Mathematics for Engineers and Scientists 4

Notes for F1.8XD2

2018

This is a one-semester course building on Mathematics for Engineers and Scientists 1, 2 and 3.

There are three main parts of the course:
(1) Laplace Transforms: Laplace transforms, inverse Laplace transforms, solving differential equations (DEs) and systems of DEs with Laplace transforms (Chapters 1 & 2).
(2) Analytic Geometry: Revision of vector algebra, scalar and vector products, lines and planes, derivatives of scalar and vector functions, directional derivatives, linear approximation of curves, tangent planes, grad, div, and curl (Chapters 3 & 4).
(3) Linear Algebra: Systems of linear euations, Gaussian elimination, vectors and matrices, matrix algebra, inverse matrices, determinants, eigenvectors and eigenvalues, applications to DEs, diagonalisation of matrices (Chapters 5, 6 & 7).

Problems for tutorials can be found at the end of each chapter.

Contents

Chapter 1

The Laplace Transform

1.1 Introduction

Laplace transforms are an interesting field of mathematics that can be used to solve problems involving differential equations and integro-differential equations. The technique is quite abstract in nature but it allows the solution of differential equations, without the need to do any integration or differentiation. The processes of integration and differentiation are replaced by algebraic manipulation, which is often considered easier to apply than concepts taken from calculus. A further advantage of the Laplace transform method for solving initial value problems that is the initial conditions are incorporated in an entirely natural way.

The main reason for considering Laplace transforms at this stage is that most engineering disciplines engage in a subject called control engineering, where Laplace transforms are used to analyse the response of engineering systems to changes in inputs to the system, whether it be a chemical reactor vessel or an auto-pilot in a plane.

Before we can attempt to solve differential equations using the Laplace transform, we need to introduce it and consider the Laplace transform and inverse Laplace transform for a number of simple functions and differential operators.

1.2 The Laplace Transform

1.2.1 Definitions and notation

The Laplace transform of a function,

$$f(t)$$

is denoted by

$$\mathcal{L}\{f(t)\} = \int_0^\infty \mathrm{e}^{-st} f(t)\,\mathrm{d}t\,. \tag{1.1}$$

Alternative notation used for the Laplace transform are:

$$\mathcal{L}\{f(t)\} = F(s) \quad \text{or } \bar{f}(s)\,.$$

From (1.1) we see that the Laplace transform consists of an improper integral (one of the integration bounds is infinite). In the first instance with an improper integral you have to focus on whether the integral has a solution or not. For example the Laplace transform of the function

$$f(t) = \mathrm{e}^{at}$$

only exists if $a - s < 0$, otherwise the Laplace transform would be

$$\int_0^\infty \mathrm{e}^{-st} f(t)\,\mathrm{d}t = \int_0^\infty \mathrm{e}^{bt}\,\mathrm{d}t$$

with $b = a - s$ positive so that the integral is unbounded.

Returning to the terminology used to describe the Laplace transform (1.1), e^{-st} is called the **kernel** of the integral. In the original function, $f(t)$, the independent variable is t, which can be considered a time variable. The function $f(t)$ is considered to exist in the **time domain**. The Laplace transform $F(s)$ exists in a **frequency domain**. s is the independent variable of the Laplace transform and strictly speaking is a complex variable although we shall for the most part only consider it to be real.

1.2.2 Laplace transforms of some simple functions

Consider the function,

$$f(t) = c \tag{1.2}$$

where c is a constant. The Laplace transform of (1.2) is given by

$$\mathcal{L}\{f(t)\} = \int_0^\infty c\mathrm{e}^{-st}\,\mathrm{d}t\,.$$

As this is an improper integral it should be considered in the limit of the upper bound being finite and the limit to infinity taken once the integral is evaluated.

$$\mathcal{L}\{f(t)\} = \int_0^\infty c\mathrm{e}^{-st}\,\mathrm{d}t = \lim_{b\to\infty}\int_0^b c\mathrm{e}^{-st}\,\mathrm{d}t$$

where

$$\int_0^b c\mathrm{e}^{-st}\,\mathrm{d}t = \left[-\frac{c}{s}\mathrm{e}^{-st}\right]_0^b = \frac{c}{s}\left(-\mathrm{e}^{-sb} - (-\mathrm{e}^0)\right) = \frac{c}{s}(1 - \mathrm{e}^{-sb})\,.$$

We can now let b tend to infinity,

$$\lim_{b\to\infty}\frac{c}{s}(1 - \mathrm{e}^{-sb}) = \frac{c}{s}\,.$$

To summarise,

$$\boxed{\mathcal{L}\{c\} = \frac{c}{s}\,.}$$

This can be considered as a **Laplace transform pair**,

$$\boxed{\left(f(t) = c\,,\quad F(s) = \frac{c}{s}\right)\,.}$$

Worked Examples (Evaluating Laplace transforms by direct integration). Let us consider another example in detail.

Question 1.1. Find the Laplace transform of t.

Solution. From the definition of the Laplace transform,

$$\mathcal{L}\{t\} = \int_0^\infty t\mathrm{e}^{-st}\,\mathrm{d}t = \lim_{b\to\infty}\int_0^b t\mathrm{e}^{-st}\,\mathrm{d}t\,.$$

Integrating by parts,

$$\int_0^b u\frac{\mathrm{d}v}{\mathrm{d}t}\,\mathrm{d}t = [uv]_0^b - \int_0^b \frac{\mathrm{d}u}{\mathrm{d}t}v\,\mathrm{d}t$$

with $u = t$ so $\mathrm{d}u/\mathrm{d}t = 1$ and $\mathrm{d}v/\mathrm{d}t = \mathrm{e}^{-st}$ so $v = -\mathrm{e}^{-st}/s$,

$$\int_0^\infty t\mathrm{e}^{-st}\,\mathrm{d}t = \left[-\frac{t}{s}\mathrm{e}^{-st}\right]_0^b + \frac{1}{s}\int_0^\infty \mathrm{e}^{-st}\,\mathrm{d}t = -\frac{b}{s}\mathrm{e}^{-bs} + \frac{1}{s}\int_0^\infty \mathrm{e}^{-st}\,\mathrm{d}t\,.$$

We could evaluate the integral on the right-hand side but we can let b tend to infinity now and avoid the integration,

$$\lim_{b\to\infty}\int_0^\infty t\mathrm{e}^{-st}\,\mathrm{d}t = \frac{1}{s}\int_0^\infty \mathrm{e}^{-st}\,\mathrm{d}t = \frac{1}{s}\mathcal{L}\{1\} = \frac{1}{s^2}\,.$$

So the Laplace transform pair is

$$\boxed{\left(f(t) = t\,,\quad F(s) = \frac{1}{s^2}\right)\,.}$$

Exercise 1.1. Use the same approach (*i.e.* one integration by parts) to show that

$$\boxed{\left(f(t) = t^2\,,\quad F(s) = \frac{2}{s^3}\right)\,.}$$

The main point of this section is not the integration process. We are interested in growing our collection of functions that we know the Laplace transform for, although we shall not do this by integrating the Laplace transform for every function we come across.

For example, inspecting the three Laplace transform pairs given above a pattern is emerging for algebraic terms. The general rule is:

General Rule for Laplace transforms of algebraic terms.

Consider the function $f(t) = t^n$.
Its Laplace transform is $F(s) = \dfrac{n!}{s^{n=1}}$.

One more example evaluated by direct integration:

Question 1.2. Evaluate the Laplace transform of the function $f(t) = \mathrm{e}^{kt}$.

Solution.

$$\mathcal{L}\{\mathrm{e}^{kt}\} = \int_0^\infty \mathrm{e}^{kt}\mathrm{e}^{-st}\,\mathrm{d}t = \lim_{b\to\infty}\int_0^b \mathrm{e}^{-(s-k)t}\,\mathrm{d}t\,,$$

where

$$\lim_{b\to\infty}\int_0^b \mathrm{e}^{-(s-k)t}\,\mathrm{d}t = \left[-\frac{1}{s-k}\mathrm{e}^{-(s-k)t}\right] = -\frac{1}{s-k}\left(\mathrm{e}^{-(s-k)b} - \mathrm{e}^0\right]_0^b\,,$$

so

$$\mathcal{L}\{\mathrm{e}^{kt}\} = \lim_{b\to\infty} -\frac{1}{s-k}\left(\mathrm{e}^{-(s-k)b} - 1\right]_0^b = \frac{1}{s-k}\,,$$

provided that $s - k > 0$, *i.e.* $s > k$.
Hence the Laplace transform pair is

$$\boxed{\left(f(t) = \mathrm{e}^{kt}\,,\quad F(s) = \frac{1}{s-k}\right)\,.}$$

As you will appreciate from the above examples, deriving Laplace transforms by direct integration is quite a tedious process. To avoid direct integration many ingenious mathematical tricks and theorems are often used to find Laplace transforms. For example consider the following worked example.

Example 1.1. Evaluate the Laplace transform of the function $f(t) = \mathrm{e}^{\mathrm{i}at}$, where a is some real number and $\mathrm{i}^2 = -1$.
Following the same steps as in Question 1.2, the solution comes to

$$\boxed{\left(f(t) = \mathrm{e}^{\mathrm{i}at}\,,\quad F(s) = \frac{1}{s-\mathrm{i}a}\right)\,.}$$

So why have we done two worked examples that are so very similar? The answer is that the second leads onto an additional useful result giving us two extra Laplace transforms. Three Laplace transforms for the price of one!

Complex numbers given in exponential form can be represented using Euler's formula

$$\mathrm{e}^{\mathrm{i}at} = \cos at + \mathrm{i}\sin at\,.$$

Reconsidering the Laplace transform of Example 1.1,

$$\mathcal{L}\{\mathrm{e}^{\mathrm{i}at}\} = \mathcal{L}\{\cos at + \mathrm{i}\sin at\} = \mathcal{L}\{\cos at\} + \mathrm{i}\mathcal{L}\{\sin at\}\,. \tag{1.3}$$

The separation of the Laplace transform into an application of the Laplace transform to the two terms is possible as the Laplace transform is a linear operator. This will be discussed further below, for now just accept it.
Again from Example 1.1,

$$\mathcal{L}\{\mathrm{e}^{\mathrm{i}at}\} = \frac{1}{s-\mathrm{i}a} = \frac{1}{s-\mathrm{i}a} \times \frac{s+\mathrm{i}a}{s+\mathrm{i}a} = \frac{s+\mathrm{i}a}{s^2+a^2}\,.$$

So, from (1.3),

$$\mathcal{L}\{\cos at\} + \mathrm{i}\mathcal{L}\{\sin at\} = \frac{s}{s^2+a^2} + \mathrm{i}\frac{a}{s^2+a^2}\,. \tag{1.4}$$

Therefore equating real and imaginary parts of (1.4) gives the following results:
Laplace Transforms of Trigonometric Functions:

The Laplace transform pair for $\cos at$:

$$f(t) = \cos at\,, \quad F(s) = \frac{s}{s^2+a^2}\,.$$

The Laplace transform pair for $\sin at$:

$$f(t) = \sin at\,, \quad F(s) = \frac{a}{s^2+a^2}\,.$$

1.2.3 Properties of the Laplace transform

The Laplace transform is a linear operator which means it has the property that

$$\mathcal{L}\{\alpha f(t) + \beta g(t)\} = \alpha\mathcal{L}\{f(t)\} + \beta\mathcal{L}\{g(t)\}$$

where α and β are constants and $f(t)$ and $g(t)$ are functions.

Question 1.3. Determine $\mathcal{L}\{3t + 2\mathrm{e}^{3t}\}$.

Solution. Using the linearity property of the Laplace transform,

$$\mathcal{L}\{3t + 2\mathrm{e}^{3t}\} = 3\mathcal{L}\{t\} + 2\mathcal{L}\{\mathrm{e}^{3t}\}\,.$$

We can now apply the Laplace transforms derived in the previous section to get

$$\mathcal{L}\{3t + 2\mathrm{e}^{3t}\} = \frac{3}{s^2} + \frac{2}{s-3}\,.$$

Question 1.4. Determine $\mathcal{L}\{5 - 3t + 4\sin 2t - 6e^{4t}\}$.

Solution. Using the linearity property of the Laplace transform,

$$\mathcal{L}\{5 - 3t + 4\sin 2t - 6e^{4t}\} = 5\mathcal{L}\{1\} - 3\mathcal{L}\{t\} + 4\mathcal{L}\{\sin 2t\} - 6\mathcal{L}\{e^{4t}\}$$

and applying the Laplace transforms derived or presented in the previous section,

$$\mathcal{L}\{5-3t+4\sin 2t-6e^{4t}\} = 5\times\frac{1}{s}-3\times\frac{1}{s^2}+4\times\frac{2}{s^2+4}-6\times\frac{1}{s-4} = \frac{5}{s}-\frac{3}{s^2}+\frac{8}{s^2+4}-\frac{6}{s-4}\,.$$

Another mathematical tool that can be used to grow our collection of Laplace transforms is called the **First Shift Theorem.**

1.2.4 First Shift Theorem

If $f(t)$ is a function having a Laplace transform $F(s)$, then the function $e^{at}f(t)$ has a Laplace transform given by

$$\boxed{\mathcal{L}\{e^{at}f(t)\} = F(s-a)\,.}$$

Sometimes $F(s-a)$ is written as $[F(s)]_{s\to s-a}$.

Here are a couple of worked examples using the first shift theorem to derive Laplace transforms.

Question 1.5. Determine $\mathcal{L}\{te^{-2t}\}$

Solution.

$$\text{For } f(t) = t\,, \quad \mathcal{L}\{f(t)\} = \mathcal{L}\{t\} = F(s) = \frac{1}{s^2}\,.$$

Then by the first shift theorem,

$$\mathcal{L}\{te^{-2t}\} = F(s)_{s\to s+2} = \left[\frac{1}{s^2}\right]_{s\to s+2} = \frac{1}{(s+2)^2}\,.$$

Question 1.6. Determine $\mathcal{L}\{e^{-3t}\sin 2t\}$

Solution.

$$\text{For } f(t) = \sin 2t\,, \quad \mathcal{L}\{f(t)\} = \mathcal{L}\{\sin 2t\} = F(s) = \frac{2}{s^2+4}\,.$$

Then by the first shift theorem,

$$\mathcal{L}\{e^{-3t}\sin 2t\} = F(s)_{s\to s+3} = \left[\frac{2}{s^2+4}\right]_{s\to s+3} = \frac{2}{(s+3)^2+4} = \frac{2}{s^2+6s+13}\,.$$

1.2.5 Table of Laplace transforms

Deriving a Laplace transform every time a function crops up is a time-consuming process. Remembering the Laplace transforms for all of the functions given above is also not a realistic proposition for most students.

Therefore in the exam the list of Laplace transforms presented in Table 1.1 is handed out.

The table together with the first shift theorem and the linearity property of the Laplace transform allows you to determine the Laplace transform of many functions.

Table 1.1: Table of Laplace transforms.

$f(t)$	$F(s)$
c	c/s
t	$1/s^2$
t^n	$n!/s^{n+1}$
e^{kt}	$1/(s-k)$
$\sin at$	$a/(s^2+a^2)$
$\cos at$	$s/(s^2+a^2)$
$t\sin at$	$2as/(s^2+a^2)^2$
$t\cos at$	$(s^2-a^2)/(s^2+a^2)^2$
$af(t)+bg(t)$	$aF(s)+bG(s)$
$f'(t)$	$sF(s)-f(0)$
$f''(t)$	$s^2F(s)-sf(0)-f'(0)$
$\delta(t-a)$	e^{-as}
$e^{at}f(t)$	$F(s-a)$
$f(t)=\begin{cases} g(t-a) & t>a \\ 0 & t<a \end{cases}$	$e^{-as}G(s)$

Table 1.1 summarises the results presented in the previous sections. In addition there are Laplace transforms for derivatives and something called the Dirac delta function. These

are useful for the solution of differential equations and will be considered in detail in the coming sections.

Exercise 1.2. Confirm by applying the first shift theorem to the Laplace transform of $t\mathrm{e}^{\mathrm{i}at}$ that

$$\mathcal{L}\{t\cos at\} = \frac{s^2 - a^2}{(s^2+a^2)^2} \quad \text{and } \mathcal{L}\{t\sin at\} = \frac{2as}{(s^2+a^2)^2}\,.$$

1.3 The Inverse Laplace Transform

If $\mathcal{L}\{f(t)\} = F(s)$ then the inverse Laplace transform, written as $\mathcal{L}^{-1}$ is $\mathcal{L}^{-1}\{F(s)\} = f(t)$.

Example 1.2. As $\mathcal{L}\{\mathrm{e}^{kt}\} = \dfrac{1}{s-k}, \quad \mathcal{L}^{-1}\left\{\dfrac{1}{s-k}\right\} = \mathrm{e}^{kt}$.

Example 1.3. As $\mathcal{L}\{\sin at\} = \dfrac{a}{s^2+a^2}, \quad \mathcal{L}^{-1}\left\{\dfrac{a}{s^2+a^2}\right\} = \sin at$.

1.3.1 Using partial fractions

The most obvious way to find an inverse transformation is to use the table of Laplace transforms, however some manipulation of ratios of polynomials of the form

$$\frac{p(s)}{q(s)} \tag{1.5}$$

is often required. For example they might be represented as partial fractions, see Appendix B if your memory needs jogging.
Similar to the Laplace transformation, the inverse Laplace transform is a linear operator, so

$$\mathcal{L}^{-1}\{\alpha F(s) + \beta G(s)\} = \alpha\mathcal{L}^{-1}\{F(s)\} + \beta\mathcal{L}^{-1}\{G(s)\}\,.$$

Using partial fractions and the linear property given above we can calculate our first inverse Laplace transforms that do not appear in the Laplace transform Table 1.1:

Question 1.7. Find

$$\mathcal{L}^{-1}\left\{\frac{1}{(s+3)(s-2)}\right\}.$$

Solution. We are ultimately going to use the table of Laplace transforms but the first thing to do is represent the Laplace transform using partial fractions:

$$\frac{1}{(s+3)(s-2)} = \frac{A}{(s+3)} + \frac{B}{(s-2)} = \frac{A(s-2)+B(s+3)}{(s+3)(s-2)}$$

so

$$A(s-2)+B(s+3) = 1\,.$$

Taking $s=2$ then $s=-3$ gives $B=1/5$ and $A=-1/5$ so

$$\mathcal{L}^{-1}\left\{\frac{1}{(s+3)(s-2)}\right\} = \frac{1}{5}\mathcal{L}^{-1}\left\{\frac{1}{(s-2)}\right\} - \frac{1}{5}\mathcal{L}^{-1}\left\{\frac{1}{(s+3)}\right\}\,.$$

By inspection of the table of Laplace transforms, the inverse Laplace transform is then

$$\mathcal{L}^{-1}\left\{\frac{1}{(s+3)(s-2)}\right\} = \frac{1}{5}\mathrm{e}^{2t} - \frac{1}{5}\mathrm{e}^{-3t}\,.$$

Question 1.8. Find

$$\mathcal{L}^{-1}\left\{\frac{s+1}{s^2(s^2+9)}\right\}\,.$$

Solution. The first thing to do is represent the Laplace transform using partial fractions,

$$\begin{aligned}\frac{s+1}{s^2(s^2+9)} &= \frac{A}{s}+\frac{B}{s^2}+\frac{Cs+D}{s^2+9} = \frac{As(s^2+9)+B(s^2+9)+Cs^3+Ds^2}{s^2(s^2+9)}\\ &= \frac{(A+C)s^3+(B+D)s^2+9As+9B}{s^2(s^2+9)}\,.\end{aligned}$$

Equating terms; 1: $9B=1$; s: $9A=1$; s^2: $B+D=0$; s^3: $A+C=0$.
These give $A=1/9$, $B=1/9$, $C=-1/9$ and $D=-1/9$. Then

$$\mathcal{L}^{-1}\left\{\frac{s+1}{s^2(s^2+9)}\right\} = \frac{1}{9}\mathcal{L}^{-1}\left\{\frac{1}{s}\right\} + \frac{1}{9}\mathcal{L}^{-1}\left\{\frac{1}{s^2}\right\} - \frac{1}{9}\mathcal{L}^{-1}\left\{\frac{s}{(s^2+9)}\right\} - \frac{1}{9}\mathcal{L}^{-1}\left\{\frac{1}{(s^2+9)}\right\}\,.$$

The first three terms are easy to evaluate using the table of Laplace transforms,

$$\frac{1}{9}\mathcal{L}^{-1}\left\{\frac{1}{s}\right\} + \frac{1}{9}\mathcal{L}^{-1}\left\{\frac{1}{s^2}\right\} - \frac{1}{9}\mathcal{L}^{-1}\left\{\frac{s}{(s^2+9)}\right\} = \frac{1}{9} + \frac{1}{9}t - \frac{1}{9}\cos 3t\,,$$

while the fourth term requires a little more work,

$$\frac{1}{9}\mathcal{L}^{-1}\left\{\frac{1}{(s^2+9)}\right\} = \frac{3}{3\times 9}\mathcal{L}^{-1}\left\{\frac{1}{(s^2+9)}\right\} = \frac{1}{27}\mathcal{L}^{-1}\left\{\frac{3}{(s^2+9)}\right\} == \frac{1}{27}\sin 3t\,,$$

the key point being multiplying the rational function by 3 and dividing by 3 to give a Laplace transform that has the same form as one of the ‘standard’ transforms.
This is a common operation in the business of finding inverse Laplace transforms
Putting this all together,

$$\mathcal{L}^{-1}\left\{\frac{s+1}{s^2(s^2+9)}\right\} = \frac{1}{9} + \frac{1}{9}t - \frac{1}{9}\cos 3t - \frac{1}{27}\sin 3t\,,$$

1.3.2 Finding inverses using the first shift theorem

The first shift theorem applied to a inverse Laplace transform says

$$\mathcal{L}^{-1}\{F(s-a)\} = \mathrm{e}^{at} f(t)\,.$$

Sometimes the alternative notation is simpler to apply:

$$\mathcal{L}^{-1}\{[F(s)]_{s\to s-a}\} = \mathrm{e}^{at} f(t)\,.$$

We will only do one example in this section. However, you will have ample opportunity to see this technique throughout the rest of these notes.

Question 1.9. Find

$$\mathcal{L}^{-1}\left\{\frac{1}{(s+2)^2}\right\}.$$

Solution.

$$\mathcal{L}^{-1}\left\{\frac{1}{(s+2)^2}\right\} = \mathcal{L}^{-1}\left\{\left[\frac{1}{s^2}\right]_{s\to s+2}\right\} = t\mathrm{e}^{-2t}\,.$$

1.3.3 Completing the square

Sometimes a quotient of the form $p(s)/q(s)$ cannot be simplified using partial fractions (without the use of complex numbers). For example consider

$$\frac{2}{s^2+6s+13}\,.$$

The quadratic term $s^2+6s+13$ has complex roots so cannot be factorised (using purely real factors) and therefore partial fractions do not offer a way forward in the next worked example. The alternative approach is to complete the square, see Appendix C if you have forgotten how to do this.

Question 1.10. Find

$$\mathcal{L}^{-1}\left\{\frac{2}{s^2+6s+13}\right\}.$$

Solution. As stated above the quotient cannot be represented as the sum of two partial fractions as the quadratic term $s^2+6s+13$ has no real roots. The answer is to complete the square:

$$s^2+6s+13 = (s+3)^2+13-3^2 = (s+3)^2+4\,.$$

Hence

$$\mathcal{L}^{-1}\left\{\frac{2}{s^2+6s+13}\right\} = \mathcal{L}^{-1}\left\{\frac{2}{(s+3)^2+4}\right\} = \mathcal{L}^{-1}\left\{\left[\frac{2}{s^2+2^2}\right]_{s\to s+3}\right\} = \mathrm{e}^{-3t}\sin 2t\,.$$

Question 1.11. Find

$$\mathcal{L}^{-1}\left\{\frac{s+7}{s^2+2s+5}\right\}.$$

Solution. Completing the square in the denominator, $s^2+2s+5=(s+1)^2+5-1=(s+1)^2+4$ so

$$\begin{aligned}\mathcal{L}^{-1}\left\{\frac{s+7}{s^2+2s+5}\right\} &= \mathcal{L}^{-1}\left\{\frac{s+7}{(s+1)^2+4}\right\}=\mathcal{L}^{-1}\left\{\frac{s+1+6}{(s+1)^2+4}\right\}\\ &= \mathcal{L}^{-1}\left\{\left[\frac{s+6}{s^2+4}\right]_{s\to s+1}\right\}=\mathcal{L}^{-1}\left\{\left[\frac{s}{s^2+4}\right]_{s\to s+1}\right\}+\mathcal{L}^{-1}\left\{\left[\frac{6}{s^2+4}\right]_{s\to s+1}\right\}.\end{aligned}$$

(We need to get all the s terms grouped with a "+1" if we are to use the first shift theorem.)

The first term is ready for inversion:

$$\mathcal{L}^{-1}\left\{\left[\frac{s}{s^2+4}\right]_{s\to s+1}\right\}=\mathrm{e}^{-t}\cos 2t\,.$$

The second term requires a little manipulation to see the inverse:

$$\mathcal{L}^{-1}\left\{\left[\frac{6}{s^2+4}\right]_{s\to s+1}\right\}=3\mathcal{L}^{-1}\left\{\left[\frac{2}{s^2+4}\right]_{s\to s+1}\right\}=3\mathrm{e}^{-t}\sin 2t\,.$$

Putting the two terms together,

$$\mathcal{L}^{-1}\left\{\frac{s+7}{s^2+2s+5}\right\}=\mathrm{e}^{-t}\cos 2t+3\mathrm{e}^{-t}\sin 2t\,.$$

Let us consider one more example, applying the same ideas to bed them in.

Question 1.12. Find

$$\mathcal{L}^{-1}\left\{\frac{2s-1}{s^2+6s+10}\right\}.$$

Solution. Again we complete the square in the quadratic term, $s^2+6s+10=(s+3)^2+10-3^2=(s+3)^2+1$. Then

$$\begin{aligned}\mathcal{L}^{-1}\left\{\frac{2s-1}{s^2+6s+10}\right\} &= \mathcal{L}^{-1}\left\{\frac{2s-1}{(s+3)^2+1}\right\}=\mathcal{L}^{-1}\left\{\frac{2(s+3)-7}{(s+3)^2+1}\right\}\\ &= 2\mathcal{L}^{-1}\left\{\frac{s+3}{(s+3)^2+1}\right\}-7\mathcal{L}^{-1}\left\{\frac{1}{(s+3)^2+1}\right\}\\ &= 2\mathcal{L}^{-1}\left\{\left[\frac{s}{s^2+1}\right]_{s\to s+3}\right\}-7\mathcal{L}^{-1}\left\{\left[\frac{1}{s^2+1}\right]_{s\to s+3}\right\}\\ &= 2\mathrm{e}^{-3t}\cos t-7\mathrm{e}^{-3t}\sin t\,.\end{aligned}$$

This process, at least initially, has to be done step by step in a slow and methodical way, otherwise silly mistakes might creep into the solution.

1.4 Problems

Problem 1.1. Compute the Laplace transforms for the following functions from first principles (*i.e.* carrying out the integrations!):

$$\begin{array}{llllll} (i) & c\,; & (ii) & t\,; & (iii) & \mathrm{e}^{kt}\,; \\ (iv) & \mathrm{e}^{\mathrm{i}at}\,; & (v) & at + b\,\mathrm{e}^{ct}\,; & (vi) & \mathrm{e}^{10t}\sin 2t\,; \end{array}$$

where a, b, c are constants and $\mathrm{i} = \sqrt{-1}$.

Problem 1.2. Use the table of Laplace transform and the first shift theorem to find the Laplace transforms of the following functions:

$$\begin{array}{llll} \text{(a)} & 5 - 3t\,; & \text{(d)} & t^2\mathrm{e}^{-4t}\,; \\ \text{(b)} & 7t^3 - 2\sin 3t\,; & \text{(e)} & (t+1)^2\,; \\ \text{(c)} & 4t\,\mathrm{e}^{-2t}\,; & \text{(f)} & \sin^2 t\,. \end{array}$$

Problem 1.3. Show that $\mathcal{L}\{\mathrm{e}^{at}f(t)\} = F(s-a)$, *i.e.*, prove the first shift theorem.

Problem 1.4. Show that $\mathcal{L}\{tf(t)\} = -\dfrac{\mathrm{d}F}{\mathrm{d}s}(s)$ *[Hint: Write $F(s)$ using the definition of Laplace transform, and differentiate the left- and right-hand sides of the expression with respect to s.]*

Problem 1.5. (a) Use the results from Problem 1.1 (*iv*) to obtain Laplace transforms of functions $\cos at$ and $\sin at$.

(b) Use the results from Problem 1.1 (*ii*) and the first shift theorem to obtain Laplace transforms of functions $t\cos at$, $t\sin at$.

Problem 1.6. (Advanced.) Show that the Laplace transform of $f(t) = t^n$ is $F(s) = \dfrac{n!}{s^{n+1}}$. *[Hint: Recursively employ integration by parts or use proof by induction.]*

Problem 1.7. Use partial fractions, the first shift theorem and the table of Laplace transforms to find the inverse Laplace transforms for the following functions

$$\begin{array}{llll} \text{(a)} & \dfrac{1}{(s+3)(s+7)}\,; & \text{(d)} & \dfrac{3s}{(s-1)(s^2-4)}\,; \\[2ex] \text{(b)} & \dfrac{2s+6}{s^2+4}\,; & \text{(e)} & \dfrac{s}{(s-1)^2(s^2+4)}\,; \\[2ex] \text{(c)} & \dfrac{4s}{(s-1)(s+1)^2}\,; & \text{(f)} & \dfrac{s}{s^2+4s+8}\,. \end{array}$$

Answers

2.(e) $2/s^3 + 2/s^2 + 1/s$

2.(f) $\dfrac{1}{2}\left(\dfrac{1}{s} - \dfrac{s}{s^2+4}\right)$.

7.(a) $\frac{1}{4}(e^{-3t} - e^{-7t})$

7.(b) $2\cos 2t + 3\sin 2t$

7.(c) $e^t - e^{-t} + 2te^{-t}$

7.(d) $-e^t + \frac{3}{2}e^{2t} - \frac{1}{2}e^{-2t}$

7.(e) $\frac{3}{25}e^t + \frac{1}{5}te^t - \frac{3}{25}\cos 2t - \frac{4}{25}\sin 2t$

7.(f) $e^{-2t}(\cos 2t - \sin 2t)$.

Chapter 2

Solution of Differential Equations Using Laplace Transforms

2.1 Laplace Transforms of Derivatives

Consider the Laplace transform of the derivative of a function,

$$\mathcal{L}\left\{\frac{\mathrm{d}f}{\mathrm{d}t}\right\} = \int_0^\infty \mathrm{e}^{-st}\frac{\mathrm{d}f}{\mathrm{d}t}\,\mathrm{d}t\,.$$

Applying integration by parts,

$$\int u\frac{\mathrm{d}v}{\mathrm{d}t}\,\mathrm{d}t = uv - \int \frac{\mathrm{d}u}{\mathrm{d}t}v\,\mathrm{d}t\,,$$

with $u = \mathrm{e}^{-st}$, so $\mathrm{d}u/\mathrm{d}t = -s\mathrm{e}^{-st}$, $v = f$ and $\mathrm{d}v/\mathrm{d}t = \mathrm{d}f/\mathrm{d}t$,

$$\int_0^\infty \mathrm{e}^{-st}\frac{\mathrm{d}f}{\mathrm{d}t}\,\mathrm{d}t = \left[\mathrm{e}^{-st}f(t)\right]_0^\infty + s\int_0^\infty \mathrm{e}^{-st}f(t)\,\mathrm{d}t = (0 - f(0)) + s\mathcal{L}\{f(t)\}\,.$$

We have
the Laplace transformation of the first derivative of a function $f(t)$:

$$\boxed{\mathcal{L}\left\{\frac{\mathrm{d}f}{\mathrm{d}t}\right\} = sF(s) - f(0)\,.}$$

Similarly the Laplace transformation of a functions second derivative can be found by a second integration by parts:
the Laplace transformation of the second derivative of a function $f(t)$:

$$\boxed{\mathcal{L}\left\{\frac{\mathrm{d}^2 f}{\mathrm{d}t^2}\right\} = s^2F(s) - sf(0) - f'(0)\,.}$$

It has to be said that in the analysis above it is assumed that the function $f(t)$ and its derivatives are sufficiently nice for the integrals to exist.
Note the Laplace transforms of derivatives are included in the table of Laplace transforms.

2.2 Constant-Coefficient Linear Differential Equations

2.2.1 Solution of first-order differential equations

Now that we have the Laplace transforms of derivatives and we can find inverse Laplace transforms, we can solve differential equations without having to integrate or differentiate anything!

Example 2.1. Consider the initial value problem,

$$\frac{\mathrm{d}y}{\mathrm{d}t} - 2y = 2\mathrm{e}^{3t}\,, \qquad y(0) = 2\,. \tag{2.1}$$

Take the Laplace transform of both sides of the differential equation:

$$\mathcal{L}\left\{\frac{\mathrm{d}y}{\mathrm{d}t} - 2y\right\} = \mathcal{L}\left\{\frac{\mathrm{d}y}{\mathrm{d}t}\right\} - 2\mathcal{L}\{y\} = \mathcal{L}\{2\mathrm{e}^{3t}\} \text{ so}$$

$$sY(s) - y(0) - 2Y(s) = sY(s) - 2 - 2Y(s) = \frac{2}{s-3}\,, \tag{2.2}$$

where $Y(s)$ denotes the Laplace transform of the function $y(t)$.
We can now reorganise the above equation (2.2) to make the Laplace transform $Y(s)$ the subject of the equation:

$$(s-2)Y(s) = 2 + \frac{2}{s-3} = \frac{2(s-3)+2}{s-3} = \frac{2(s-2)}{s-3} \qquad \text{so}$$

$$Y(s) = \frac{2(s-2)}{(s-2)(s-3)} = \frac{2}{s-3}\,.$$

Now all we have to do is find the inverse Laplace transform. In this case the inverse transform can be found directly from the table of Laplace transforms,

$$y(t) = \mathcal{L}^{-1}\{Y(s)\} = \mathcal{L}^{-1}\left\{\frac{2}{s-3}\right\} = 2\mathrm{e}^{3t}.$$

This is the solution to the initial value problem (2.1).

We have solved our first differential equation using the Laplace transform.

If you review the example given above you will see the Laplace-transform method for solving differential equations can be separated into three steps:

The Laplace Transform Method.
Step 1. Take the Laplace Transform of the given differential equation
Step 2. Make the transformed variable ($Y(s)$ above) the subject of the transformed equation.
Step 3. Apply the inverse Laplace transform to find $y(t)$.

As you can see from the Example 2.1, the initial value is included in the solution, without having to introduce a constant A and then use the initial value to find the value of A.
The Laplace transform method of solving differential equations comes into its own when considering higher-order differential equations.

2.2.2 Solution of second-order differential equations

The general inhomogeneous linear second-order constant-coefficient differential equation reads

$$a\frac{\mathrm{d}^2y}{\mathrm{d}t^2}+b\frac{\mathrm{d}y}{\mathrm{d}t}+cy=f(t)\qquad(\text{with } a\neq 0)\,.$$

Note that by using Laplace transforms, we can solve this equation without the need to separate the solution into the complementary function (the solution to the homogeneous problem (*i.e.* with $f(t)\equiv 0$) and a particular integral to extend the solution to the inhomogeneous differential equation. The solution of second-order differential equations follows a similar line to the solution of first-order differential equations using the Laplace-transform method.

Question 2.1. Solve the second-order differential equation

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2}+b\frac{\mathrm{d}y}{\mathrm{d}t}+6y=0$$

subject to the initial values

$$y(0)=1\,,\quad \frac{\mathrm{d}y}{\mathrm{d}t}(0)=0\,.$$

Solution. Apply the Laplace transform to both sides of the differential equation,

$$\mathcal{L}\left\{\frac{\mathrm{d}^2y}{\mathrm{d}t^2}\right\}+\mathcal{L}\left\{5\frac{\mathrm{d}y}{\mathrm{d}t}\right\}+\mathcal{L}\{6y\}=0\qquad\text{so}$$

$$s^2Y(s)-sy(0)-\frac{\mathrm{d}y}{\mathrm{d}t}(0)+5(sY(s)-y(0))+6Y(s)=0\,,$$

substitute the initial values into the transformed equation,

$$s^2Y(s)-s+5(sY(s)-1)+6Y(s)=0\,,$$

and reorganise the transformed equation such that $Y(s)$ is the subject of the equation,

$$(s^2+5s+6)Y(s)=s+5\,,\quad i.e.\ \ Y(s)=\frac{s+5}{s^2+5s+6}\,.$$

We now take the inverse Laplace transform of both sides,

$$y(t)=\mathcal{L}^{-1}\{Y(s)\}=\mathcal{L}^{-1}\left\{\frac{s+5}{s^2+5s+6}\right\}.\tag{2.3}$$

Considering the right-hand side of (2.3),

$$\frac{s+5}{s^2+5s+6}=\frac{s+5}{(s+2)(s+3)}=\frac{A}{s+2}+\frac{B}{s+3}=\frac{A(s+3)+B(s+2)}{(s+2)(s+3)}$$

so $A(s+3)+B(s+2)=s+5$ and putting $s=-2$ and then $s=-3$ gives $A=3$ and $B=-2$.
Then

$$y(t)=\mathcal{L}^{-1}\{Y(s)\}=\mathcal{L}^{-1}\left\{\frac{3}{s+2}+\frac{(-2)}{s+3}\right\}=3\mathcal{L}^{-1}\left\{\frac{1}{s+2}\right\}-2\mathcal{L}^{-1}\left\{\frac{1}{s+3}\right\}=3\mathrm{e}^{-2t}-2\mathrm{e}^{-3t}.$$

Let's do another worked example, this time inhomogeneous, of degenerate type.

Question 2.2. Solve the second-order differential equation

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 5\frac{\mathrm{d}y}{\mathrm{d}t} + 6y = \mathrm{e}^{-2t}, \tag{2.4}$$

subject to the initial data

$$y(0) = 1\,, \quad \frac{\mathrm{d}y}{\mathrm{d}t}(0) = -1\,.$$

Solution. Apply the Laplace transform to both sides of the differential equation,

$$\mathcal{L}\left\{\frac{\mathrm{d}^2y}{\mathrm{d}t^2}\right\} + 5\mathcal{L}\left\{\frac{\mathrm{d}y}{\mathrm{d}t}\right\} + 6\mathcal{L}\{y\} = \mathcal{L}\{\mathrm{e}^{-2t}\} \quad \text{so}$$

$$s^2Y(s) - sy(0) - \frac{\mathrm{d}y}{\mathrm{d}t}(0) + 5(sY(s) - y(0)) + 6Y(s) = \frac{1}{s+2}\,,$$

substitute the initial values into the transformed equation,

$$s^2Y(s) - s + 1 + 5(sY(s) - 1) + 6Y(s) = \frac{1}{s+2}\,,$$

and reorganise the transformed equation so that $Y(s)$ is the subject of the equation,

$$(s^2 + 5s + 6)Y(s) = (s+2)(s+3)Y(s) = \frac{1}{s+2} + s + 4 = \frac{1 + (s+4)(s+2)}{s+2}$$

$$= \frac{s^2 + 6s + 9}{s+2} = \frac{(s+3)^2}{s+2}$$

$$\text{so} \qquad Y(s) = \frac{(s+3)^2}{(s+2)^2(s+3)} = \frac{s+3}{(s+2)^2}\,.$$

We can now take the inverse Laplace transform of both sides,

$$y(t) = \mathcal{L}^{-1}\{Y(s)\} = \mathcal{L}^{-1}\left\{\frac{s+3}{(s+2)^2}\right\} = \mathcal{L}^{-1}\left\{\frac{(s+2)+1}{(s+2)^2}\right\} = \mathcal{L}^{-1}\left\{\frac{1}{s+2} + \frac{1}{(s+2)^2}\right\}$$

$$= \mathcal{L}^{-1}\left\{\frac{1}{s+2}\right\} + \mathcal{L}^{-1}\left\{\frac{1}{(s+2)^2}\right\} = \mathrm{e}^{-2t} + \mathcal{L}^{-1}\left\{\left[\frac{1}{s^2}\right]_{s\to s+2}\right\} = \mathrm{e}^{-2t} + \mathrm{e}^{-2t}t = \mathrm{e}^{-2t}(1+t)\,.$$

Question 2.3. Solve the second-order differential equation,

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 9y = 0\,,$$

subject to the initial values, $y(0) = 0$ and $\frac{\mathrm{d}y}{\mathrm{d}t}(0) = 1$.

Solution. Apply the Laplace transform to both sides of the differential equation,

$$\mathcal{L}\left\{\frac{\mathrm{d}^2y}{\mathrm{d}t^2}\right\} + \mathcal{L}\{9y\} = s^2Y(s) - sy(0) - \frac{\mathrm{d}y}{\mathrm{d}t}(0) + 9Y(s) = 0\,.$$

Substitute the initial values into the transformed equation,

$$s^2Y(s) - 1 + 9Y(s) = 0\,.$$

Reorganise the transformed equation so that $Y(s)$ is the subject of the equation,

$$s^2Y(s) + 9Y(s) = 1 \qquad \text{so} \qquad Y(s) = \frac{1}{s^2+9}\,.$$

Take the inverse Laplace transform of both sides,

$$y(t) = \mathcal{L}^{-1}\{Y(s)\} = \mathcal{L}^{-1}\left\{\frac{1}{s^2+9}\right\} = \frac{1}{3}\mathcal{L}^{-1}\left\{\frac{3}{s^2+9}\right\} = \frac{1}{3}\sin 3t\,.$$

2.2.3 An LCR circuit example

Before we move on, let us consider one last worked example. This is an initial value problem you might have seen before, but solved then using techniques considered in Mathematics for Engineers and Scientists 3.
The worked example is a model for an LCR circuit, see Fig. 2.1. An LCR circuit is one that includes a resistor, a capacitor and an inductor, connected in series with a voltage source $e(t)$.

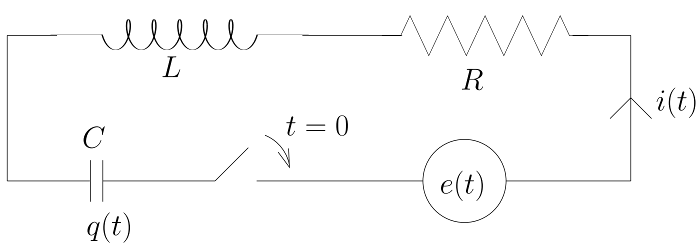


Figure 2.1: An LCR circuit.

Before closing the switch at time $t = 0$ the charge, q, on the capacitor and the resulting current, $i = \mathrm{d}q/\mathrm{d}t$, in the circuit are zero. Applying Kirchhoff's second law to the circuit gives a second-order inhomogeneous differential equation for the charge on the capacitor,

$$L\frac{\mathrm{d}^2q}{\mathrm{d}t^2} + R\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{1}{C}q = e(t)\,,$$

where

$$q(0) = 0 \quad \text{and} \quad \frac{\mathrm{d}q}{\mathrm{d}t}(0) = 0\,.$$

In the circuit equation given above the different components have the following values, $R = 160\,\Omega$, $L = 1\,\mathrm{H}$, $C = 10^{-4}\mathrm{F}$ and $e(t) = 20\,\mathrm{V}$.

Substitute the values of the electrical properties into the differential equation:

$$\frac{\mathrm{d}^2q}{\mathrm{d}t^2} + 160\frac{\mathrm{d}q}{\mathrm{d}t} + 10^4 q = 20\,.$$

Take the Laplace transforms of the differential equation:

$$\mathcal{L}\left\{\frac{\mathrm{d}^2q}{\mathrm{d}t^2}\right\} + 160\mathcal{L}\left\{\frac{\mathrm{d}q}{\mathrm{d}t}\right\} + 10^4\mathcal{L}\{q\} = s^2Q(s) - sq(0) - \frac{\mathrm{d}q}{\mathrm{d}t}(0) + 160(sQ(s) - q(0)) + 10^4Q(s)$$

$$= \mathcal{L}\{20\} = 20/s\,.$$

Substitute in the initial values:

$$(s^2 + 160s + 10^4)Q(s) = 20/s\,.$$

Make the Laplace transform of the solution of the differential equation the subject of the equation:

$$Q(s) = \frac{20}{s(s^2 + 160s + 10^4)}\,.$$

To take the inverse Laplace transform of the RHS, it needs to be represented using partial fractions. Note that $s^2 + 160s + 10^4$ doesn't have real factors so that we shall have to complete the square at some stage.

$$\frac{20}{s(s^2 + 160s + 10^4)} = \frac{A}{s} + \frac{Bs + C}{s^2 + 160s + 10^4} = \frac{A(s^2 + 160s + 10^4) + Bs^2 + Cs}{s(s^2 + 160s + 10^4)}\,.$$

Equating terms, 1: $10^4A = 20$; s: $160A + C = 0$; s^2: $A + B = 0$.
These equations have the solution $A = 1/500$, $B = -1/500$, $C = -160/500$ $(= -8/25)$ so

$$Q(s) = \frac{20}{s(s^2 + 160s + 10^4)} = \frac{1}{500}\left(\frac{1}{s} - \frac{s + 160}{s^2 + 160s + 10^4}\right)\,.$$

We now complete the square,

$$s^2 + 160s + 10000 = (s + 80)^2 + 3600 = (s + 80)^2 + 60^2.$$

Then

$$Q(s) = \frac{1}{500}\left(\frac{1}{s} - \frac{s + 160}{(s + 80)^2 + 60^2}\right) = \frac{1}{500}\left(\frac{1}{s} - \frac{s + 80}{(s + 80)^2 + 60^2} - \frac{80}{(s + 80)^2 + 60^2}\right)$$

$$= \frac{1}{500}\left(\frac{1}{s} - \frac{s + 80}{(s + 80)^2 + 60^2} - \frac{4}{3}\frac{60}{(s + 80)^2 + 60^2}\right)\,.$$

Then

$$\mathcal{L}^{-1}\left\{\frac{20}{s(s^2 + 160s + 10^4)}\right\} = \frac{1}{500}\mathcal{L}^{-1}\left\{\frac{1}{s} - \frac{s + 80}{(s + 80)^2 + 60^2} - \frac{4}{3}\frac{60}{(s + 80)^2 + 60^2}\right\}$$

$$= \frac{1}{500}\mathcal{L}^{-1}\left\{\frac{1}{s} - \left[\frac{s}{s^2 + 60^2}\right]_{s\to s+80} - \frac{4}{3}\left[\frac{60}{s^2 + 60^2}\right]_{s\to s+80}\right\} = \frac{1}{500}\left(1 - \mathrm{e}^{-80t}\cos 60t - \frac{4}{3}\mathrm{e}^{-80t}\sin 60t\right)\,.$$

Therefore the solution reads

$$q(t) = \mathcal{L}^{-1}\{Q(s)\} = \frac{1}{500}\left(1 - \mathrm{e}^{-80t}\left(\cos 60t + \frac{4}{3}\sin 60t\right)\right)\,.$$

2.3 Differential Equations and the Dirac Delta Function

So far we have solved differential equations using the Laplace-transform method without the direct application of techniques taken from calculus. It has offered a number of advantages compared to other techniques. We shall consider the application of the Laplace transform method to a class of differential equations that are not easily solved any other way.

Laplace transforms can be used to solve problems involving an impulsive force or current (*i.e.*, charge), where the impulse is delivered over a short time interval, say (t_0, t_1),

$$I(t) = \int_{t_0}^{t_1} f(t)\,\mathrm{d}t\,,$$

where $I(t)$ is the total momentum input to the system if $f(t)$ is a force.

Suppose that the applied force is given by the function

$$f_\varepsilon(t) = \begin{cases} 1/\varepsilon & \text{for } 0 < t_0 < t < t_0 + \varepsilon \\ 0 & \text{otherwise} \end{cases}\,.$$

The above function should be interpreted as a constant force $1/\varepsilon$ applied over a time interval of length ε. By construction,

$$I_\varepsilon = \int_0^\infty f_\varepsilon(t)\,\mathrm{d}t = \int_{t_0}^{t_0+\varepsilon} (1/\varepsilon)\,\mathrm{d}t = 1$$

so that the total impulse I_ε is the area under the curve $f_\varepsilon(t)$ and is independent of ε. Taking the limit $\varepsilon \to 0$,

$$f_\varepsilon(t) \to \delta(t - t_0)$$

where $\delta(t - t_0)$ is called the **Dirac delta function**. This is a peculiar name for this construct as it does not have all of the properties of a function. For this reason it is a member of a class called **generalised functions**. The Dirac delta function is zero everywhere except at $t = t_0$ where it has a singularity and is therefore undefined. The significant point is the Dirac delta function has the property

$$\int_0^\infty \delta(t - t_0)\,\mathrm{d}t = 1\,.$$

This is important as it represents an impulse of magnitude 1 acting over an infinitely short time interval. Another important property of the Dirac delta function is that

$$\int_0^\infty \delta(t - t_0) f(t)\,\mathrm{d}t = f(t_0)$$

(as long as f is continuous at t_0). This is called the **sifting property** of Dirac functions as it makes it possible to isolate a particular value of a function.

This means the Laplace transform of a Dirac delta function can be evaluated:

$$\mathcal{L}\{\delta(t-a)\} = \int_0^\infty \delta(t-a)\mathrm{e}^{-st}\,\mathrm{d}t = \mathrm{e}^{-as}.$$

Therefore in principle any differential equation involving an impulse delivered over a very short time interval can be solved using Laplace transforms.
The last construct we need before we can start solving differential equations involving Dirac delta functions is the **second shift theorem.**

The Second Shift Theorem.
Taking $a > 0$, if

$$f(t) = \begin{cases} g(t-a) & \text{for } t > a \\ 0 & \text{for } t < a \end{cases}$$

then $F(s) = \mathrm{e}^{-as}G(s)$ where $G(s) = \mathcal{L}\{g(t)\}$.

This is pretty dry stuff. It will become a little clearer when you see an application of the second shift theorem.

Question 2.4. Find the inverse Laplace transform of

$$F(s) = \frac{\mathrm{e}^{-3s}}{s^2}.$$

Solution. By inspection of the second shift theorem in this example, $a = 3$ and $G(s) = 1/s^2$ so $g(t) = \mathcal{L}^{-1}\{G(s)\} = \mathcal{L}^{-1}\{s^{-2}\} = t$.
Then the second shift theorem implies that

$$f(t) = \mathcal{L}^{-1}\{F(s)\} = \mathcal{L}^{-1}\left\{\frac{\mathrm{e}^{-3s}}{s^2}\right\} = \begin{cases} t-3 & \text{for } t > 3 \\ 0 & \text{for } t < 3 \end{cases}.$$

Another worked example to see how the second shift theorem can be used.

Question 2.5. Find the inverse Laplace transform of

$$F(s) = \mathrm{e}^{-2s}\frac{1}{s+7}.$$

Solution. Here $G(s) = 1/(s+7)$ so $g(t) = \mathcal{L}^{-1}\{G(s)\} = \mathcal{L}^{-1}\{1/(s+7)\} = \mathrm{e}^{-7t}$, while $a = 2$.
Applying the 2nd shift theorem,

$$f(t) = \mathcal{L}^{-1}\{F(s)\} = \begin{cases} \mathrm{e}^{-7(t-2)} & \text{for } t > 2 \\ 0 & \text{for } t < 2 \end{cases}.$$

We now have all of the tools in place to solve problems involving instantaneous impulses. Consider a freely vibrating system without damping. The dynamics of the system are governed by a differential equation of the form

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + \omega^2 y = 0\,.$$

(See Section 2.4 in the first semester of material.) If the system is subjected to an instantaneous impulse of magnitude b at a time $t = a$, the second-order differential equation is modified, to

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + \omega^2 y = b\delta(t - a)\,.$$

Example 2.2. Let's use Laplace transforms to solve the initial value problem

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + 4y = \delta(t - 3)\,,$$

with $y(0) = 1$ and $\dfrac{\mathrm{d}y}{\mathrm{d}t}(0) = 0$.
We follow the same solution strategy as previous examples looking at solving differential equations using the Laplace transform method.

Step 1. Take the Laplace transform of the differential equation:

$$\mathcal{L}\left\{\frac{\mathrm{d}^2 y}{\mathrm{d}t^2}\right\} + 4\mathcal{L}\{y\} = \mathcal{L}\{\delta(t-3)\} \text{ so}$$

$$s^2 Y(s) - sy(0) - \frac{\mathrm{d}y}{\mathrm{d}t}(0) + 4Y(s) = \mathrm{e}^{-3s}\,.$$

Step 2. Substitute the initial values into the equation:

$$s^2 Y(s) - s + 4Y(s) = \mathrm{e}^{-3s}\,.$$

Step 3. Make $Y(s)$ the subject of the equation:

$$(s^2 + 4)Y(s) = s + \mathrm{e}^{-3s} \qquad \text{so} \qquad Y(s) = \frac{s}{s^2+4} + \mathrm{e}^{-3s}\frac{1}{s^2+4}\,.$$

Step 4. Find the inverse Laplace transform for $Y(s)$:

$$\mathcal{L}^{-1}\{Y(s)\} = \mathcal{L}^{-1}\left\{\frac{s}{s^2+4}\right\} + \mathcal{L}^{-1}\left\{\mathrm{e}^{-3s}\frac{1}{s^2+4}\right\}\,,$$

$$\text{where} \qquad \mathcal{L}^{-1}\left\{\frac{s}{s^2+4}\right\} = \cos 2t\,.$$

To find the second inverse Laplace transform,

$$\mathcal{L}^{-1}\left\{\mathrm{e}^{-3s}\frac{1}{s^2+4}\right\},$$

we use the second shift theorem.

$$\mathcal{L}^{-1}\left\{\frac{1}{s^2+4}\right\} = \frac{1}{2}\mathcal{L}^{-1}\left\{\frac{2}{s^2+4}\right\} = \frac{1}{2}\sin 2t\,,$$

therefore the 2nd shift theorem says

$$\mathcal{L}^{-1}\left\{\mathrm{e}^{-3s}\frac{1}{s^2+4}\right\} = \begin{cases} \frac{1}{2}\sin 2(t-3) & \text{for } t>3 \\ 0 & \text{for } t<3 \end{cases}.$$

Putting these results together gives the solution,

$$y(t) = \mathcal{L}^{-1}\{Y(s)\} = \begin{cases} \cos 2t + \frac{1}{2}\sin 2(t-3) & \text{for } t>3 \\ \cos 2t & \text{for } t<3 \end{cases}.$$

See Fig.2.2 for a plot of $y(t)$.

Let's look at another example before we move on.

Question 2.6. Use Laplace transforms to solve the initial value problem

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 6\frac{\mathrm{d}y}{\mathrm{d}t} + 8y = 2\delta(t-7)\,, \quad y(0)=0\,, \quad \frac{\mathrm{d}y}{\mathrm{d}t}(0) = 6\,. \tag{2.5}$$

Solution. We follow the same solution strategy as in the previous example.
Step 1. Take the Laplace transform of the differential equation:

$$\mathcal{L}\left\{\frac{\mathrm{d}^2y}{\mathrm{d}t^2}\right\} + 6\mathcal{L}\left\{\frac{\mathrm{d}y}{\mathrm{d}t}\right\} + 8\mathcal{L}\{y\} = 2\mathcal{L}\{\delta(t-7)\} \quad \text{so}$$

$$s^2Y(s) - sy(0) - \frac{\mathrm{d}y}{\mathrm{d}t}(0) + 6(sY(s) - y(0)) + 8Y(s) = 2\mathrm{e}^{-7s}.$$

Step 2. Substitute the initial values into the equation:

$$s^2Y(s) - 6 + 6sY(s) + 8Y(s) = 2\mathrm{e}^{-7s}.$$

Step 3. Make $Y(s)$ the subject of the equation:

$$(s^2+6s+8)Y(s) = 6 + 2\mathrm{e}^{-7s} \text{ so } Y(s) = \frac{6}{s^2+6s+8} + \frac{2\mathrm{e}^{-7s}}{s^2+6s+8}.$$

Step 4. Find the inverse Laplace transform for $Y(s)$. This is a long process in this example but provided you take your time you will see the approach is the same as in the previous example.

$$\mathcal{L}^{-1}\{Y(s)\} = \mathcal{L}^{-1}\left\{\frac{6}{s^2+6s+8}\right\} + \mathcal{L}^{-1}\left\{\frac{2\mathrm{e}^{-7s}}{s^2+6s+8}\right\}.$$

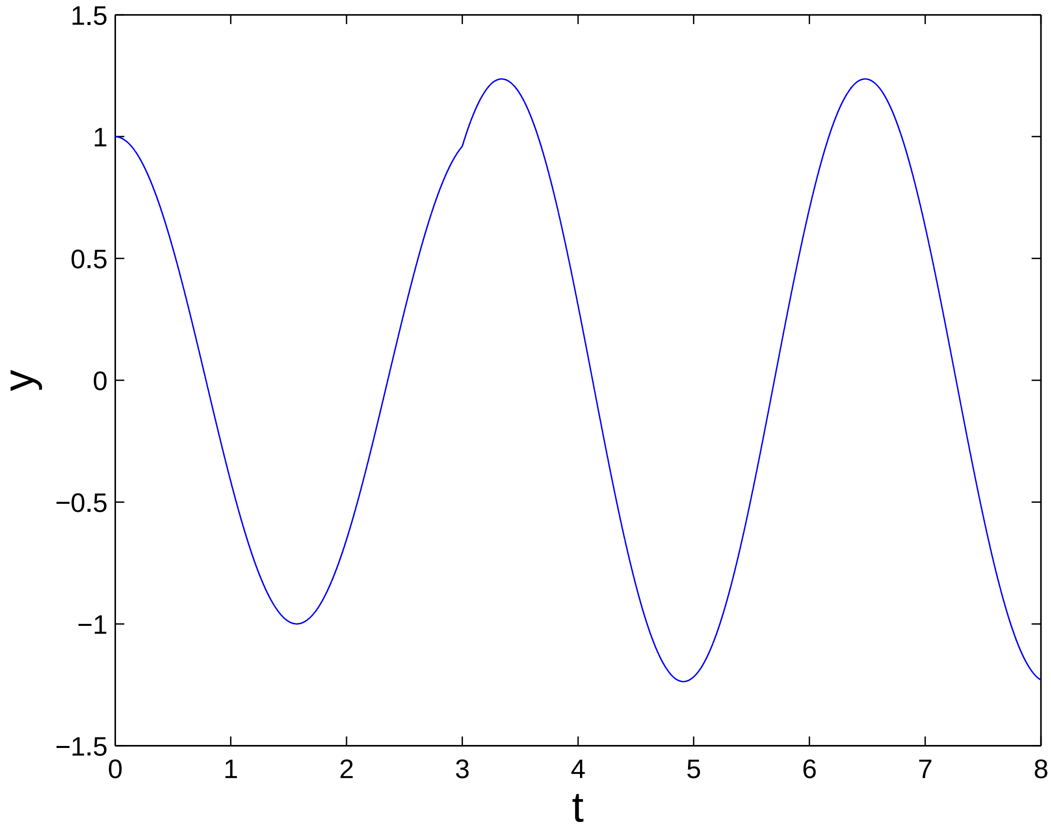


Figure 2.2: The displacement of a vibrating body subjected to an impulse at $t = 3$. (Look carefully, and you will see that there is a jump of slope at that time.)

To find the inverses we have to represent the quotients using partial fractions, where

$$s^2 + 6s + 8 = (s+2)(s+4)$$

so

$$\frac{2}{s^2+6s+8} = \frac{2}{(s+2)(s+4)} = \frac{A}{(s+2)} + \frac{B}{(s+4)} = \frac{A(s+4)+B(s+2)}{(s+2)(s+4)},$$

giving $A(s+4) + B(s+2) = 2$. Thus, by taking $s = -2$ and then $s = -4$, $A = 1$ and $B = -1$.
It follows that

$$\frac{2}{s^2+6s+8} = \frac{1}{(s+2)} - \frac{1}{(s+4)} \quad \text{and} \quad \frac{6}{s^2+6s+8} = \frac{3}{(s+2)} - \frac{3}{(s+4)} \quad \text{so}$$

$$\mathcal{L}^{-1}\left\{\frac{2}{s^2+6s+8}\right\} = \mathcal{L}^{-1}\left\{\frac{1}{(s+2)}\right\} - \mathcal{L}^{-1}\left\{\frac{1}{(s+2)}\right\} = \mathrm{e}^{-2t} - \mathrm{e}^{-4t}$$

$$\text{and} \quad \mathcal{L}^{-1}\left\{\frac{6}{s^2+6s+8}\right\} = 3\mathrm{e}^{-2t} - 3\mathrm{e}^{-4t}.$$

Then, by the second shift theorem,

$$\mathcal{L}^{-1}\left\{\frac{2\mathrm{e}^{-7s}}{s^2+6s+8}\right\} = \begin{cases} 0 & \text{for } t<7 \\ \mathrm{e}^{-2(t-7)} - \mathrm{e}^{-4(t-7)} & \text{for } t>7 \end{cases}.$$

Putting these altogether, we get

$$y(t) = \mathcal{L}^{-1}\{Y(s)\} = \begin{cases} 3\mathrm{e}^{-2t} - 3\mathrm{e}^{-4t} & \text{for } t<7 \\ 3\mathrm{e}^{-2t} - 3\mathrm{e}^{-4t} + \mathrm{e}^{-2(t-7)} - \mathrm{e}^{-4(t-7)} & \text{for } t>7 \end{cases}.$$

See Fig.2.3 for a graph of the solution.

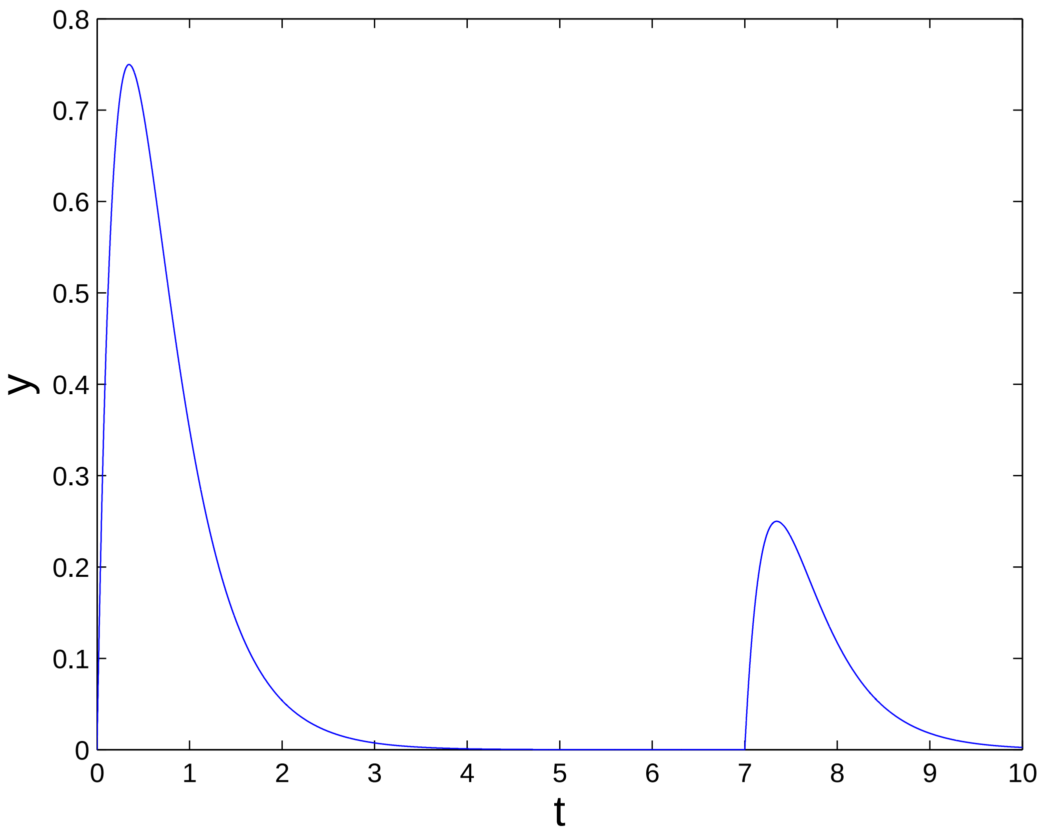


Figure 2.3: Graph of the solution of the initial value problem (2.5).

2.4 Systems of Differential Equations

The extension of the Laplace transform method for solving differential equations is natural. Recall that for one differential equation the Laplace transform method is a three-step process.

The Laplace Transform Method.
Step 1. Take the Laplace Transform of the given differential equation.
Step 2. Make the transformed variable ($Y(s)$ above) the subject of the transformed equation.
Step 3. Apply the inverse Laplace transform to find $y(t)$.

For a system of differential equations the second step is more complicated as it involves the solution of a system of (simultaneous) algebraic equations rather than one equation. By way of example consider the following initial value problem.

Question 2.7. Solve the following initial value problem using the Laplace transform method:

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = x_1 + 2x_2\,, \quad \frac{\mathrm{d}x_2}{\mathrm{d}t} = 2x_1 - 2x_2\,, \quad x_1(0) = 2\,, \quad x_2(0) = 1\,.$$

Solution. The first step is to take the Laplace transform of the differential equations:

$$\mathcal{L}\left\{\frac{\mathrm{d}x_1}{\mathrm{d}t}\right\} = \mathcal{L}\{x_1\} + 2\mathcal{L}\{x_2\} \quad \text{and} \quad \mathcal{L}\left\{\frac{\mathrm{d}x_2}{\mathrm{d}t}\right\} = 2\mathcal{L}\{x_1\} - 2\mathcal{L}\{x_2\} \quad \text{so}$$

$$sX_1 - x_1(0) = X_1 + 2X_2 \quad \text{and } sX_2 - x_2(0) = 2X_1 - 2X_2\,.$$

Substituting the initial values into the equations and reorganising to put all of the unknowns on one side of the equations,

$$(s-1)X_1 - 2X_2 = 2\,, \tag{2.6}$$

$$-2X_1 + (s+2)X_2 = 1\,. \tag{2.7}$$

This system is two equations in two unknowns, X_1 and X_2 can be solved. From (2.6)

$$X_1 = 2(X_2+1)/(s-1)\,. \tag{2.8}$$

Substituting for X_1 in (2.7),

$$-4\left(\frac{X_2+1}{s-1}\right) + (s+2)X_2 = 1$$

which can be rearranged as follows:

$$\left(-\frac{4}{s-1} + (s+2)\right)X_2 = 1 + \frac{4}{s-1} = \frac{s+3}{s-1}\,,$$

$$\therefore \left(\frac{-4+(s+2)(s-1)}{s-1}\right)X_2 = \left(\frac{s^2+s-6}{s-1}\right)X_2 = \left(\frac{(s-2)(s+3)}{s-1}\right)X_2 = \frac{s+3}{s-1}\,,$$

$$\therefore X_2 = \frac{1}{s-2}\,.$$

Having found X_2 we can substitute into (2.8) to give X_1:

$$X_1 = 2\left(1 + \frac{1}{s-2}\right)\frac{1}{s-1} = 2\left(\frac{s-1}{s-2}\right)\frac{1}{s-1} = \frac{2}{s-2}\,.$$

Now that the simultaneous equations have been solved, taking the inverse Laplace transforms gives the result:

$$x_1(t) = \mathcal{L}^{-1}\{X_1\} = \mathcal{L}^{-1}\left\{\frac{2}{s-2}\right\} = 2\mathrm{e}^{2t}\,, \quad x_2(t) = \mathcal{L}^{-1}\{X_2\} = \mathcal{L}^{-1}\left\{\frac{1}{s-2}\right\} = \mathrm{e}^{2t}\,.$$

What about second-order and inhomogeneous systems?
The method is exactly the same. A second-order system example is given below, an inhomogeneous first-order system is left as an exercise.

Question 2.8. Solve the following second-order system using the Laplace-transform method:

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2x - y = 0\,, \quad \frac{\mathrm{d}^2y}{\mathrm{d}t^2} - x + 2y = 0\,,$$

with $x(0) = 4$, $y(0) = 2$, $\dfrac{\mathrm{d}x}{\mathrm{d}t}(0) = \dfrac{\mathrm{d}y}{\mathrm{d}t}(0) = 0$.

Solution. Some of the details of the solution will be omitted as it has essentially the same steps as Question 2.7.
Take the Laplace transform of the differential equations to get

$$s^2X - x(0)s - \frac{\mathrm{d}x}{\mathrm{d}t}(0) + 2X - Y = 0\,,$$

$$s^2Y - y(0)s - \frac{\mathrm{d}y}{\mathrm{d}t}(0) - X + 2Y = 0\,.$$

Substitute for the initial values and rearrange so that the unknowns are on one side

$$(s^2 + 2)X - Y = 4s\,, \tag{2.9}$$

$$-X + (s^2 + 2)Y = 2s\,. \tag{2.10}$$

From (2.9),

$$Y = (s^2 + 2)X - 4s\,. \tag{2.11}$$

Substituting for Y in (2.10) and rearranging so that X is the subject of the equation,

$$X = \frac{4s^3 + 10s}{s^4 + 4s^2 + 3} = \frac{4s^3 + 10s}{(s^2 + 1)(s^2 + 3)}\,.$$

Partial fractions gives

$$\frac{2}{(s^2 + 1)(s^2 + 3)} = \frac{1}{(s^2 + 1)} - \frac{1}{(s^2 + 3)}$$

$$\text{so} \qquad \frac{2s}{(s^2+1)(s^2+3)} = \frac{s}{(s^2+1)} - \frac{s}{(s^2+3)} \quad \text{and}$$

$$\frac{2s^3}{(s^2+1)(s^2+3)} = s\left(\frac{s^2}{(s^2+1)} - \frac{s^2}{(s^2+3)}\right) = s\left(\frac{(s^2+1)-1}{(s^2+1)} - \frac{(s^2+3)-3}{(s^2+3)}\right)$$

$$= s\left(1 - \frac{1}{(s^2+1)} - 1 + \frac{3}{(s^2+3)}\right) = s\left(\frac{3}{(s^2+3)} - \frac{1}{(s^2+1)}\right).$$

Hence

$$X = \frac{6s}{(s^2+3)} - \frac{2s}{(s^2+1)} + \frac{5s}{(s^2+1)} - \frac{5s}{(s^2+3)} = \frac{3s}{(s^2+1)} + \frac{s}{(s^2+3)}.$$

Substituting these into (2.10) and tidying up the equation gives

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

$$= \frac{3s}{(s^2+1)} - \frac{s}{(s^2+3)}.$$

Taking the inverse Laplace transforms,

$$x(t) = \mathcal{L}^{-1}\{X(s)\} = 3\mathcal{L}^{-1}\left\{\frac{s}{(s^2+1)}\right\} + \mathcal{L}^{-1}\left\{\frac{s}{(s^2+3)}\right\} = 3\cos t + \cos\sqrt{3}t$$

and similarly

$$y(t) = \mathcal{L}^{-1}\{Y(s)\} = 3\cos t - \cos\sqrt{3}t\,.$$

There are many examples of systems of differential equations that are formulated as problems in vibration and circuit simulation. As an example application, a mathematical model for the simulation of a circuit is given here.

Example 2.3. A two-loop circuit is shown schematically in Fig. 2.4.

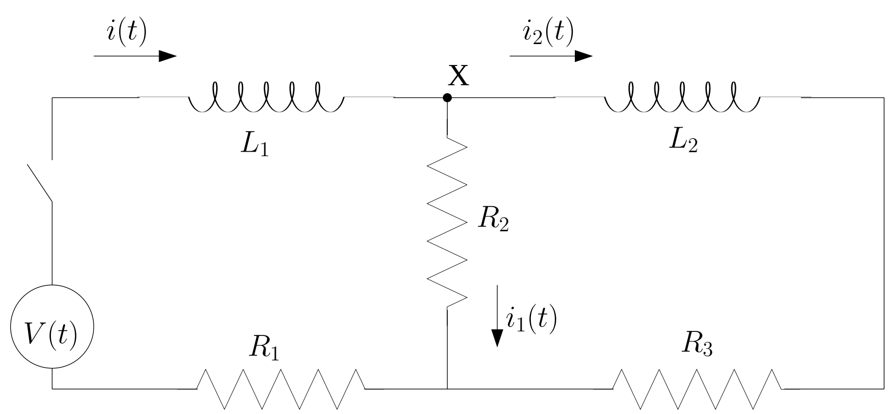


Figure 2.4: A double-loop circuit.

Applying Kirchhoff's 1st law at node X gives

$$i = i_1 + i_2\,.$$

Applying Kirchhoff's 2nd law to the left and right loops in turn then gives the differential equations

$$R_1(i_1 + i_2) + L_1\frac{\mathrm{d}}{\mathrm{d}t}(i_1 + i_2) + R_2 i_1 = V\,,$$

$$L_2\frac{\mathrm{d}i_2}{\mathrm{d}t} + R_3 i_2 - R_2 i_1 = 0\,.$$

Initially no current flows: $i_1(0) = i_2(0) = 0$.
Given that $R_1 = R_3 = 10\,\Omega$, $R_2 = 20\,\Omega$, $L_1 = L_2 = 5\,\mathrm{H}$ and $V(t) = 200\,\mathrm{V}$, we wish to find the currents i_1 and i_2 using the Laplace transform method.
We first substitute the numerical values for the constant terms into the differential equations:

$$10(i_1 + i_2) + 5\frac{\mathrm{d}}{\mathrm{d}t}(i_1 + i_2) + 20 i_1 = 200\,,$$

$$5\frac{\mathrm{d}i_2}{\mathrm{d}t} + 10 i_2 - 20 i_1 = 0\,.$$

We divide through by common factors in the differential equations to simplify them:

$$2(i_1 + i_2) + \frac{\mathrm{d}i_1}{\mathrm{d}t} + \frac{\mathrm{d}i_2}{\mathrm{d}t} + 4 i_1 = 40\,,$$

$$\frac{\mathrm{d}i_2}{\mathrm{d}t} + 2 i_2 - 4 i_1 = 0\,.$$

Next we take Laplace transforms:

$$2I_1 + 2I_2 + sI_1 - i_1(0) + sI_2 - i_2(0) + 4I_1 = 40/s\,,$$

$$sI_2 - i_2(0) + 2I_2 - 4I_1 = 0\,.$$

We then substitute in the initial values and tidy up:

$$(s+6)I_1 + (s+2)I_2 = 40/s\,, \tag{2.12}$$

$$-4I_1 + (s+2)I_2 = 0\,. \tag{2.13}$$

From (2.13)

$$I_2 = \frac{4I_1}{s+2}\,. \tag{2.14}$$

Substituting for I_2 in (2.12) and making I_1 the subject of the equation,

$$I_1 = \frac{40}{s(s+10)} = \frac{4}{s} - \frac{4}{s+10}\,,$$

on using partial fractions. From (2.14)

$$I_2 = \frac{16}{s(s+2)} - \frac{16}{(s+2)(s+10)} = \frac{8}{s} - \frac{8}{s+2} - \frac{2}{s+2} + \frac{2}{s+10} = \frac{8}{s} - \frac{10}{s+2} + \frac{2}{s+10}\,,$$

using partial fractions again.
Taking inverse Laplace transforms,

$$i_1(t) = \mathcal{L}^{-1}\{I_1\} = \mathcal{L}^{-1}\left\{\frac{4}{s}\right\} - \mathcal{L}^{-1}\left\{\frac{4}{s+10}\right\} = 4 - 4\mathrm{e}^{-10t}\,,$$

and similarly,

$$i_2(t) = \mathcal{L}^{-1}\{I_2\} = 8 - 10\mathrm{e}^{-2t} + 2\mathrm{e}^{-10t}\,.$$

2.5 Summary of Laplace transforms

These two chapters have been organised into three separate sets of theory. The first looks at finding the Laplace transforms of functions. This can be done by evaluating the improper integral definition of the Laplace transform,

$$\mathcal{L}\{f(t)\} = \int_0^\infty \mathrm{e}^{-st} f(t)\,\mathrm{d}t\,.$$

This approach is not advocated here as a simpler approach is to use a table of Laplace transforms together with the first shift theorem and the linear property of the Laplace operator.
Having mastered Laplace transforms we focused on finding inverse Laplace transforms. This is a more difficult problem as it requires a certain ability to recognise patterns. The different techniques for finding inverse Laplace transforms, such as the use of partial fractions and completing the square, are all about representing the Laplace transform in a way so that it can be matched up with functions in the table of Laplace transforms. The first shift theorem in reverse can also be used to find inverse Laplace transforms.
We looked at the application of Laplace transforms to the solution of differential equations. This involves three steps:

Step 1. Apply the Laplace transform to the differential equation.

Step 2. Rearrange the transformed equation to be explicit in the Laplace transform of the solution.

Step 3. Take the inverse Laplace transform to find the solution of the differential equation.

The Laplace transform method of solution for differential equations has a number of advantages over standard methods.

- The incorporation of initial values is simpler.
- The solution technique requires no knowledge of calculus.
- Certain problems involving near instantaneous disturbances to the modelled system can be incorporated.

The last point above involves representing instantaneous disturbances using Dirac functions and solving the differential equation using the second shift theorem.

Linear systems of differential equations are relatively easy to solve using the Laplace transform method as you can follow the three-step process given above. The second step for systems of differential equations includes the solution of simultaneous equations.

2.6 Problems

Problem 2.1. Solve the following initial value problems using Laplace tranforms:

$$\begin{array}{llll}
\text{(a)} & \dfrac{\mathrm{d}y}{\mathrm{d}t} + 3y = \mathrm{e}^{-2t}, & & y(0) = 2\,; \\
\text{(b)} & \dfrac{\mathrm{d}^2y}{\mathrm{d}t^2} + 2\dfrac{\mathrm{d}y}{\mathrm{d}t} + 5y = 1, & y(0) = 0, & \dfrac{\mathrm{d}y}{\mathrm{d}t}(0) = 0\,; \\
\text{(c)} & \dfrac{\mathrm{d}^2y}{\mathrm{d}t^2} + 4\dfrac{\mathrm{d}y}{\mathrm{d}t} + 5y = 3\mathrm{e}^{-2t}, & y(0) = 4, & \dfrac{\mathrm{d}y}{\mathrm{d}t}(0) = -7\,; \\
\text{(d)} & \dfrac{\mathrm{d}^2y}{\mathrm{d}t^2} + 8\dfrac{\mathrm{d}y}{\mathrm{d}t} + 16y = 16\sin 4t, & y(0) = -\frac{1}{2}, & \dfrac{\mathrm{d}y}{\mathrm{d}t}(0) = 1\,; \\
\text{(e)} & \dfrac{\mathrm{d}^2y}{\mathrm{d}t^2} - 2\dfrac{\mathrm{d}y}{\mathrm{d}t} + 2y = \cos t, & y(0) = 1, & \dfrac{\mathrm{d}y}{\mathrm{d}t}(0) = 0\,.
\end{array}$$

Problem 2.2. Use the second shift theorem to find the function $y(t)$ with the following Laplace transforms:

$$\text{(a)}\, Y(s) = \frac{\mathrm{e}^{-3s}}{s^4}\,; \quad \text{(b)}\, Y(s) = \mathrm{e}^{-2s}\,; \quad \text{(c)}\, Y(s) = \frac{s\mathrm{e}^{-s}}{s^2+2s+5}\,.$$

Problem 2.3. Solve the initial value problems:

$$\begin{array}{ll}
\text{(a)} & \dfrac{\mathrm{d}^2y}{\mathrm{d}t^2} + 3\dfrac{\mathrm{d}y}{\mathrm{d}t} + 2y = \delta(t-4), \quad y(0) = \dfrac{dy(0)}{dt} = 0\,; \\
\text{(b)} & \dfrac{1}{2}\dfrac{\mathrm{d}^2y}{\mathrm{d}t^2} + \dfrac{\mathrm{d}y}{\mathrm{d}t} + y = 2\delta(t-3), \quad y(0) = 1, \ \dfrac{dy(0)}{dt} = 0\,.
\end{array}$$

Problem 2.4. Application of Kirchhoff's laws for a two-loop circuit leads to the following system of differential equations for the currents i_1 and i_2

$$L_1\frac{\mathrm{d}}{\mathrm{d}t}(i_1+i_2) + R_1(i_1+i_2) + R_2 i_1 + L_3\frac{\mathrm{d}i_1}{\mathrm{d}t} = V\,,$$

$$L_2\frac{\mathrm{d}i_2}{\mathrm{d}t} + R_3 i_2 - R_2 i_1 - L_3\frac{\mathrm{d}i_1}{\mathrm{d}t} = 0\,,$$
$$i_1(0) = 1\,, \quad i_2(0) = 0\,.$$

Solve the system of equations using Laplace transforms, taking $R_1 = R_2 = R_3 = 2$, $L_1 = L_2 = L_3 = 1$ and $V = 3$.

Problem 2.5. Find the solution $x_1(t), x_2(t)$ of the system of differential equations:

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} - x_1 - x_2 &= \mathrm{e}^{-2t}; \\
\frac{\mathrm{d}x_2}{\mathrm{d}t} - 4x_1 + 2x_2 &= -2\mathrm{e}^{t};
\end{aligned}$$

with $x_1(0) = 0$ and $x_2(0) = 1$.

Answers

1.(a) $y(t) = \mathrm{e}^{-3t} + \mathrm{e}^{-2t}$,
1.(b) $y(t) = \frac{1}{5}\left(1 - \mathrm{e}^{-t}\cos 2t - \frac{1}{2}\mathrm{e}^{-t}\sin 2t\right)$,
1.(c) $y(t) = \mathrm{e}^{-2t}(\cos t + \sin t + 3)$,
1.(d) $y(t) = t\mathrm{e}^{-4t} - \frac{1}{2}\cos 4t$,
1.(e) $y(t) = \frac{1}{5}\cos t - \frac{2}{5}\sin t + \frac{4}{5}\mathrm{e}^{t}\cos t - \frac{2}{5}\mathrm{e}^{t}\sin t$.

2.

$$\text{(a)}\quad y(t) = \begin{cases} 0 & t < 3 \\ \frac{(t-3)^3}{6} & t > 3 \end{cases}\ ;\ \text{(b)}\quad y(t) = \delta(t-2)\,;$$

$$\text{(c)}\quad y(t) = \begin{cases} 0 & t < 1 \\ \mathrm{e}^{-(t-1)}(\cos 2(t-1) - \frac{1}{2}\sin 2(t-1)) & t > 1 \end{cases}.$$

3.(a)

$$y(t) = \begin{cases} 0 & t < 4 \\ \mathrm{e}^{-(t-4)} - \mathrm{e}^{-2(t-4)} & t > 4 \end{cases};$$

3.(b)

$$y(t) = \begin{cases} \mathrm{e}^{-t}(\cos t + \sin t) & t < 3 \\ \mathrm{e}^{-t}(\cos t + \sin t) + 4\mathrm{e}^{-(t-3)}\sin(t-3) & t > 3 \end{cases}.$$

4. $i_1(t) = \frac{1}{2}(1 + \mathrm{e}^{-2t})$, $i_2(t) = \frac{1}{2}(1 - \mathrm{e}^{-2t})$.

5. $x_1(t) = \frac{1}{2}(-\mathrm{e}^{-3t} + \mathrm{e}^{t})$, $x_2(t) = 2\mathrm{e}^{-3t} - \mathrm{e}^{-2t}$.

Chapter 3

Geometry

3.1 Introduction

In this and the following chapter we look at vectors and combine them with ideas taken from calculus. Our starting point is the concept of a vector and vector operations that you are likely to be familiar with already. This area of mathematics forms the basis of all CAD packages and is also useful in simulating the stress fields and fluid motion around complex shapes. A final area of application of interest to students reading for physics and electrical engineering is electromagnetism.
The vector calculus operators discussed at the end of this part make it possible to state fundamental systems of partial differential equations in a concise way. Physical laws amenable to such a treatment include the Navier Stokes equations (fluid motion) and Maxwell's Equations (electro-magnetism).

3.2 Revision of Vector Operations

A **vector** is a quantity that has two properties, a magnitude (how big it is) and a direction. This is compared to a **scalar** quantity which just has a magnitude, such as the speed of a car, or the distance travelled by a car. It is usually denoted in bold, or it may be underlined or have an arrow on top. The old favourite examples of vector quantities are forces, displacements (where you are in the world), velocities and accelerations.
Vectors are incredibly useful things as many problems can be formulated and solved in terms of vectors and vector operations, particularly if you are working in three-dimensional space (as most of us are). The reason for this is most people have great difficulty picturing a line in 3D space but are entirely comfortable with lines in 2D space. You can draw lines in two dimensions on a piece of paper for a start. The beauty of vectors is that operations in two-dimensional space extend in a completely natural way to three-dimensional space. Vectors can be represented in a number of ways, such as

$$\boldsymbol{r} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} , \quad (1,2,3) \quad \text{or} \quad \boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k} \,.$$

The three numbers represent the Cartesian co-ordinates, and $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ are **unit vectors** in the directions of the x, y and z co-ordinate axis. For the most part we will use the first bracket, *i.e.* column, notation for a vector.

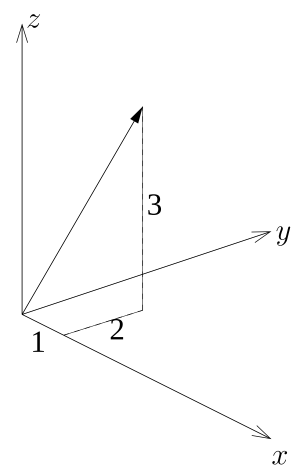


Figure 3.1: Cartesian co-ordinates of a point vector.

The magnitude of a vector $\boldsymbol{r}$ is denoted by $|\boldsymbol{r}|$ and is given by

$$|\boldsymbol{r}| = \sqrt{a^2 + b^2 + c^2}$$

where (a, b, c) are the Cartesian co-ordinates of the vector. So the vector

$$\boldsymbol{r} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \tag{3.1}$$

has magnitude $|\boldsymbol{r}| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$.

It is often convenient to calculate **unit vectors**. These are vectors with a given direction and magnitude of one. They are often written as

$$\hat{\boldsymbol{r}}\,.$$

To calculate a unit vector in the direction of a vector $\boldsymbol{r}$, we divide $\boldsymbol{r}$ by its magnitude:

$$\hat{\boldsymbol{r}} = \boldsymbol{r}/|\boldsymbol{r}|\,.$$

For example, by construction, the following vector is a unit vector:

$$\hat{\boldsymbol{r}} = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}\,.$$

3.2.1 Vector addition

A graphical representation of how vectors are added together is shown in the Fig. 3.2. Working with Cartesian co-ordinates two vectors can be added together by addition of the components of the vectors. For example if

$$\boldsymbol{r} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \text{ and } \boldsymbol{y} = \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix} \text{ then } \boldsymbol{r} + \boldsymbol{y} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix} = \begin{pmatrix} 1+2 \\ 1-3 \\ -1+4 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 3 \end{pmatrix}\,.$$

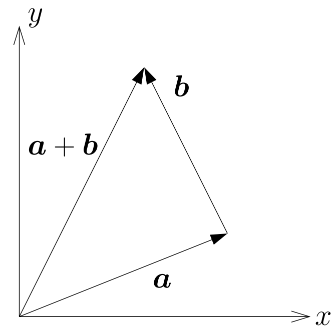


Figure 3.2: Vector addition.

3.2.2 Multiplication by a scalar

Multiplying a vector $\boldsymbol{r}$ by a scalar α gives a vector of magnitude $|\alpha||\boldsymbol{r}|$ with the same direction as $\boldsymbol{r}$ if α is positive and the opposite direction if α is negative (see Fig. 3.3). Using Cartesian co-ordinates, the components of the vector are muliplied by the scalar,

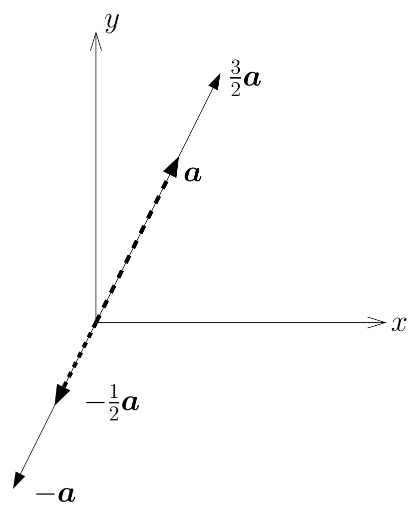


Figure 3.3: Multiplication of a vector by a scalar.

e.g., for $\alpha = 3$ and

$$\boldsymbol{r} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} , \quad \alpha\boldsymbol{r} = 3 \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \\ 9 \end{pmatrix} .$$

3.2.3 Scalar product

The extension of scalar multiplication to vectors leads to two types of “vector multiplication”: the **scalar product** and the **vector product**.
The scalar product (or dot product) takes two vectors and produces a scalar. Given two vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$, then the scalar product is written as $\boldsymbol{a} \cdot \boldsymbol{b}$.

Geometrically the scalar product is the magnitude of $\boldsymbol{b}$ multiplied by the projection of $\boldsymbol{a}$ onto $\boldsymbol{b}$, see Fig. 3.4.

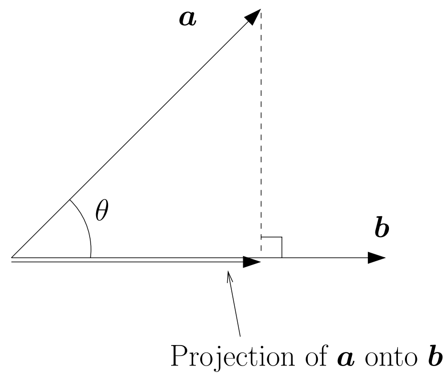


Figure 3.4: A geometric interpretation of the scalar product of $\boldsymbol{a}$ and $\boldsymbol{b}$.

The projection of $\boldsymbol{a}$ onto $\boldsymbol{b}$ is $|\boldsymbol{a}|\cos\theta$, where θ is the angle between the directions of the two vectors. If you do not see this immediately, note that in Fig. 3.4 the projection of $\boldsymbol{a}$ onto $\boldsymbol{b}$ is the side adjacent to the angle θ, and the magnitude of $\boldsymbol{a}$ is the hypotenuse. Hence

$$\boldsymbol{a}\cdot\boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\cos\theta\,. \tag{3.2}$$

If the scalar product is zero then, assuming that neither vector is the zero vector, $\cos\theta = 0$ which implies that

$$\theta = \pi/2 = 90^\circ.$$

This is very important as finding vectors that are at right angles often involves finding zero scalar products:
Given two non-zero vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ then

$$\boldsymbol{a}\cdot\boldsymbol{b} = 0 \quad \Leftrightarrow \quad \boldsymbol{a} \text{ and } \boldsymbol{b} \text{ are orthogonal.} \tag{3.3}$$

Orthogonal means at right angles. Sometimes such vectors are described as being **normal** to one another.
Another special case for the scalar product is if the scalar product of a vector is taken with itself:

$$\boldsymbol{a}\cdot\boldsymbol{a} = |\boldsymbol{a}||\boldsymbol{a}|\cos 0 = |\boldsymbol{a}|^2\,.$$

The square root of the scalar product of a vector with itself gives the vector's magnitude:

$$\sqrt{\boldsymbol{a}\cdot\boldsymbol{a}} = |\boldsymbol{a}|\,.$$

For scalar products the order of the vectors is immaterial so $\boldsymbol{a}\cdot\boldsymbol{b} = \boldsymbol{b}\cdot\boldsymbol{a}$.

In Cartesians the scalar product is given by the sum of the pair-wise products of the coordinates:

$$\boldsymbol{a}\cdot\boldsymbol{b} = a_1b_1 + a_2b_2 + a_3b_3\,, \text{ where } \boldsymbol{a} = \begin{pmatrix} a_1\\ a_2\\ a_3\end{pmatrix} \text{ and } \boldsymbol{b} = \begin{pmatrix} b_1\\ b_2\\ b_3\end{pmatrix}\,.$$

Question 3.1. Find the angle between $\boldsymbol{a} = (1, -2, 2)$ and $\boldsymbol{b} = (0, 3, -4)$.

Solution. $\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta$, so

$$\cos\theta = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|}$$

for θ the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.
Here $|\boldsymbol{a}| = \sqrt{1+4+4} = \sqrt{9} = 3$, $|\boldsymbol{b}| = \sqrt{0+9+16} = \sqrt{25} = 5$ and $\boldsymbol{a}\cdot\boldsymbol{b} = 1.0+(-2).3+2.(-4) = -14$. Hence

$$\cos\theta = \frac{-14}{15} \quad \text{so} \quad \theta = \cos^{-1}(-14/15) = \pi - \cos^{-1}(14/15)\,.$$

3.2.4 Vector product

The vector product (or cross product) of vectors $\boldsymbol{a}$, $\boldsymbol{b}$ is written

$$\boldsymbol{a} \times \boldsymbol{b}\,.$$

As the name suggests the vector product produces a vector. Geometrically the vector product is a vector with a magnitude given by the area of a parallelogram and direction perpendicular to the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, see Fig. 3.5. The magnitude of the resulting

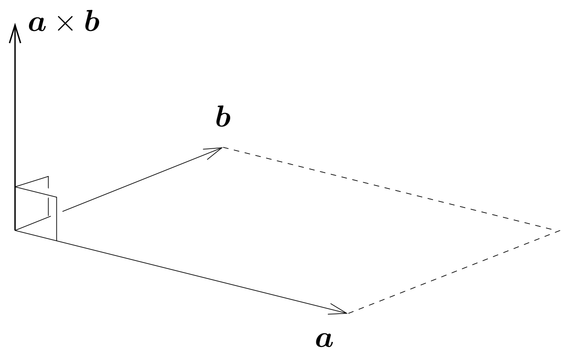


Figure 3.5: A geometric interpretation of the vector product of $\boldsymbol{a}$ and $\boldsymbol{b}$.

vector is given by the area of the parallelogram defined by $\boldsymbol{a}$ and $\boldsymbol{b}$ as shown in Fig. 3.5. Two directions are perpendicular to the plane defined by the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, vertically "up" and vertically "down". In this case the vector is vertically up because of the "right-hand rule". This says that if your right thumb represents $\boldsymbol{a}$ and right index finger $\boldsymbol{b}$, then the right-hand middle finger (arranged at right angles to thumb and index finger) gives "up", the direction of $\boldsymbol{a} \times \boldsymbol{b}$ (and an "upward" normal vector, $\boldsymbol{n}$, to the parallelogram). An expression for the vector product is then

$$\boldsymbol{a} \times \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\boldsymbol{n} \sin\theta\,,$$

where $\boldsymbol{n}$ denotes the unit vector in the vertical direction and θ is again the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.

If we consider the vector product $\boldsymbol{b}\times\boldsymbol{a}$, it has the same magnitude as $\boldsymbol{a}\times\boldsymbol{b}$ but the right-hand rule determines that the direction of the vector exactly opposite ("down" instead of "up"). This implies that the order in the vector product is important:

$$\boldsymbol{a}\times\boldsymbol{b} = -\boldsymbol{b}\times\boldsymbol{a}\,.$$

Because of the property (3.3),

$$\boldsymbol{a}\cdot(\boldsymbol{a}\times\boldsymbol{b}) = 0 \;\text{ and }\; \boldsymbol{b}\cdot(\boldsymbol{a}\times\boldsymbol{b}) = 0\,.$$

This can be a very useful property
Making use of Cartesian co-ordinates and the $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ notation, the vector product reads

$$\boldsymbol{a}\times\boldsymbol{b} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}\,, \tag{3.4}$$

where

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k} \;\text{ and }\; \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = b_1\boldsymbol{i} + b_2\boldsymbol{j} + b_3\boldsymbol{k}\,.$$

In (3.4) the vector product is given by a 3×3 determinant. Recall that this can be evaluated by breaking it down into three 2×2 determinants:

$$\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \boldsymbol{i}\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - \boldsymbol{j}\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + \boldsymbol{k}\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\,,$$

Then each 2×2 determinant can be evaluated from the formula

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc\,, \qquad \text{so}$$

$$\begin{aligned} \boldsymbol{a}\times\boldsymbol{b} &= \boldsymbol{i}(a_2b_3 - a_3b_2) - \boldsymbol{j}(a_1b_3 - a_3b_1) + \boldsymbol{k}(a_1b_2 - a_2b_1) \\ &= \boldsymbol{i}(a_2b_3 - a_3b_2) + \boldsymbol{j}(a_3b_1 - a_1b_3) + \boldsymbol{k}(a_1b_2 - a_2b_1)\,, \end{aligned}$$

or, using bracket/column notation,

$$\boldsymbol{a}\times\boldsymbol{b} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}\,.$$

You might not find this an easy formula to remember, it is then a better strategy to remember (3.4) and then evaluate the 3×3 determinant each time. Note that determinants will be looked at again in Section 6.3.

Question 3.2. Calculate the vector product $\boldsymbol{a}\times\boldsymbol{b}$ for $\boldsymbol{a} = (2,1,0)$ and $\boldsymbol{b} = (2,3,0)$. Hence find the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.

Solution.

$$\begin{aligned}
\boldsymbol{a}\times\boldsymbol{b} &= \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 2 & 1 & 0 \\ 2 & 3 & 0 \end{vmatrix} \\
&= \boldsymbol{i}\begin{vmatrix} 1 & 0 \\ 3 & 0 \end{vmatrix} - \boldsymbol{j}\begin{vmatrix} 2 & 0 \\ 2 & 0 \end{vmatrix} + \boldsymbol{k}\begin{vmatrix} 2 & 1 \\ 2 & 3 \end{vmatrix} \\
&= \boldsymbol{i}((1)(0)-(0)(3)) - \boldsymbol{j}((2)(0)-(0)(2)) + \boldsymbol{k}((2)(3)-(1)(2)) = 4\boldsymbol{k}\,.
\end{aligned}$$

Thus

$$\boldsymbol{a}\times\boldsymbol{b} = \begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix}\,.$$

As $\boldsymbol{a}\times\boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\boldsymbol{n}\sin\theta$ with $\boldsymbol{n}$ a unit vector and θ the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$, here $\boldsymbol{n} = \boldsymbol{k} = (0,0,1)$ and

$$\sin\theta = \frac{|\boldsymbol{a}\times\boldsymbol{b}|}{|\boldsymbol{a}||\boldsymbol{b}|} = \frac{4}{\sqrt{4+1}\,\sqrt{4+9}} = \frac{4}{\sqrt{5}\,\sqrt{13}} = \frac{4}{\sqrt{65}}\,.$$

Therefore the angle $= \theta = \sin^{-1}(4/\sqrt{65}) \approx 0.519 \approx 29.7°$.

Exercise 3.1. The angular momentum vector, taken about the origin, $\boldsymbol{H}$ of a particle of mass m moving with velocity $\boldsymbol{v}$ is given by

$$\boldsymbol{H} = \boldsymbol{r}\times(m\boldsymbol{v})\,,$$

where $\boldsymbol{r}$ is the displacement of the particle from the origin.
If the motion of the particle is described by it having an the angular velocity $\boldsymbol{\omega}$ about the origin, then

$$\boldsymbol{v} = \boldsymbol{\omega}\times\boldsymbol{r}\,.$$

Using the result for the vector triple product

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

where $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are vectors, show that if $\boldsymbol{r}$ is perpendicular to $\boldsymbol{\omega}$, then

$$\boldsymbol{H} = mr^2\boldsymbol{\omega}\,,$$

where $r = |\boldsymbol{r}|$.

Finally, recall the scalar triple product: $[\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}] = [\boldsymbol{b},\boldsymbol{c},\boldsymbol{a}] = [\boldsymbol{c},\boldsymbol{a},\boldsymbol{b}] = \boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = (\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{a} = \boldsymbol{b}\cdot(\boldsymbol{c}\times\boldsymbol{a}) = \ldots$.
Using the determinant formula for the vector product, (3.4), it is quite easy to show that

$$[\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}] = \boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}\,. \qquad (3.5)$$

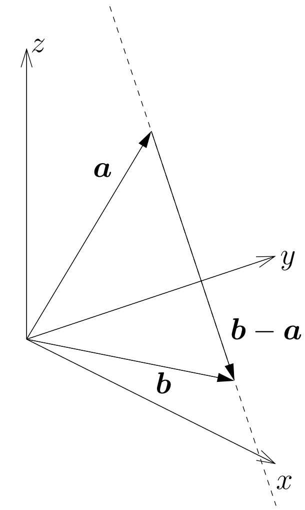


Figure 3.6: A line in 3D space passing through the points $\boldsymbol{a}$ and $\boldsymbol{b}$.

3.3 Lines in Three Dimensions

3.3.1 Parametric equation of a line

Consider two point vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, see Fig. 3.6.
Vector addition gives us $\boldsymbol{b} = \boldsymbol{a} + (\boldsymbol{b} - \boldsymbol{a})$.
In this case $\boldsymbol{b} - \boldsymbol{a}$ can be thought of as a **direction vector** The equation of the line, passing through $\boldsymbol{a}$ and in the direction $\boldsymbol{b} - \boldsymbol{a}$ is given, parametrically, by the equation

$$\boldsymbol{r} = \boldsymbol{a} + t(\boldsymbol{b} - \boldsymbol{a})\,. \tag{3.6}$$

Here t can be thought of as a free parameter, $t = 0$ gives the point $\boldsymbol{a}$, $t = 1$ gives the point $\boldsymbol{b}$. A value of t between zero and one gives a point between $\boldsymbol{a}$ and $\boldsymbol{b}$ on the line. t outside of the unit interval gives points on the line either side of $\boldsymbol{a}$ and $\boldsymbol{b}$.
A more general representation of the equation of a line using vectors is

$$\boldsymbol{r} = \boldsymbol{a} + t\boldsymbol{m}\,,$$

where $\boldsymbol{a}$ is a point on the line and $\boldsymbol{m}$ is a **direction vector** (a vector parallel to the line). The advantage of this type of representation is it is the same whether the line is in 2D or 3D space.

Question 3.3. Suppose a line passes through the point vectors

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

Derive a parametric equation of the line.

Solution. A parametric form of an equation of a line is $\boldsymbol{r} = \boldsymbol{a} + t(\boldsymbol{b} - \boldsymbol{a})$.

As

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{b} - \boldsymbol{a} = \begin{pmatrix} 1 \\ 5 \\ -1 \end{pmatrix},$$

an equation of the line in this case is

$$\boldsymbol{r} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} + t \begin{pmatrix} 1 \\ 5 \\ -1 \end{pmatrix},$$

or equivalently, $\boldsymbol{r} = (1+t)\boldsymbol{i} + (5t-2)\boldsymbol{j} + (2-t)\boldsymbol{k}$.

3.3.2 Angle between two lines

Given two vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$ say, the angle between them can be calculated using vector operations, such as the scalar product (easier than the vector product as we saw above). The angle between two lines is simply the angle between their two direction vectors (see Fig. 3.7), and can therefore be given by use of the direction vectors' scalar (or vector) product.

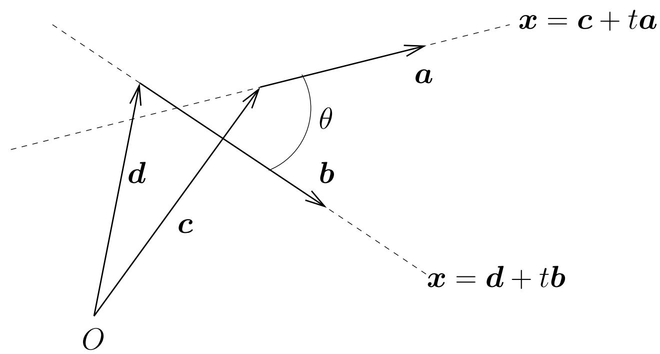


Figure 3.7: Angle between two lines.

For example if two lines are given by $\boldsymbol{r} = \boldsymbol{c} + t\boldsymbol{a}$, $\boldsymbol{r} = \boldsymbol{d} + t\boldsymbol{b}$, with $\boldsymbol{a} = (2, 1, 0)$ and $\boldsymbol{b} = (2, 3, 0)$, the angle θ between the direction vectors is given by

$$\cos\theta = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|} = \frac{(2)(2) + 1(3)}{\sqrt{5}\sqrt{13}} = \frac{7}{\sqrt{65}}.$$

The angle between the lines is then given by

$$\theta = \cos^{-1}(7/\sqrt{65}) \approx 0.519 \approx 29.7^\circ.$$

3.4 Equations of a Plane

3.4.1 Non-parametric (Cartesian) equation for a plane

The equation of a plane differs from an equation of a line as we want an equation that gives the position vector of any point on a flat surface, the plane. Any plane's orientation

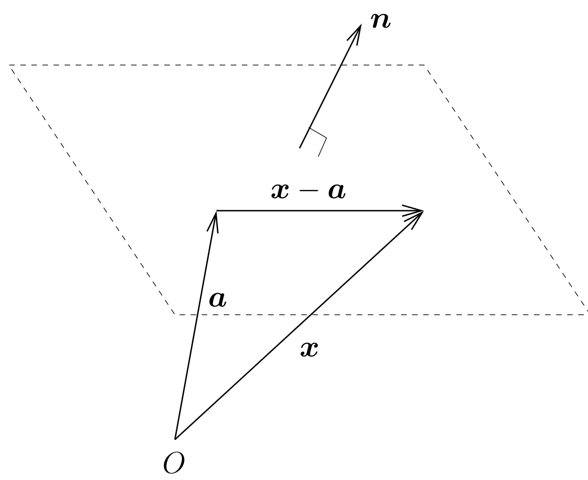


Figure 3.8: The important vectors in the definition of a plane.

is specified in terms of a vector normal to the plane, *i.e.* at right angles to its surface, here denoted by $\boldsymbol{n}$, see Fig. 3.8.
In particular, a plane is defined by a vector normal to the plane, $\boldsymbol{n}$, and a point vector in the plane, $\boldsymbol{a}$ (see Fig. 3.8). With reference to the figure, if $\boldsymbol{r}$ is any point on the plane, the vector $\boldsymbol{r}-\boldsymbol{a}$ is a direction vector in the plane, and therefore $\boldsymbol{r}-\boldsymbol{a}$ has an orientation that is perpendicular to the normal vector, $\boldsymbol{n}$. If we take a scalar product of $\boldsymbol{r}-\boldsymbol{a}$ and $\boldsymbol{n}$, then as these two vectors are at normal to each other,

$$(\boldsymbol{r}-\boldsymbol{a})\cdot\boldsymbol{n} = |\boldsymbol{r}-\boldsymbol{a}||\boldsymbol{n}|\cos\frac{\pi}{2} = 0 \quad \text{so}$$

$$\boxed{\boldsymbol{n}\cdot\boldsymbol{r} = \boldsymbol{n}\cdot\boldsymbol{a}\,. \qquad (3.7)}$$

This is the equation of a plane with normal $\boldsymbol{n}$ and having a point vector $\boldsymbol{a}$ lying on it.

One interesting property of the equation of a plane is if the normal vector is a unit vector, *i.e.* has a magnitude of one, so that the plane can be described by

$$\hat{\boldsymbol{n}}\cdot\boldsymbol{r} = d \quad \text{so}$$

with $|\hat{\boldsymbol{n}}| = 1$, then the size of right-hand side, $|d|$ in the equation above, is the perpendicular distance from the origin to the plane.
More generally, the distance h of a point $\boldsymbol{b}$ from a plane $\boldsymbol{n}\cdot\boldsymbol{r} = d$ is given by

$$h = |d - \boldsymbol{n}\cdot\boldsymbol{b}|/|\boldsymbol{n}|\,. \qquad (3.8)$$

This can be got by simple trigonometry (see Fig. 3.9). Taking $\boldsymbol{a}$ to be a point on the plane, the plane's equation is

$$\hat{\boldsymbol{n}}\cdot\boldsymbol{r} = \hat{\boldsymbol{n}}\cdot\boldsymbol{a} = d \quad \text{so}$$

$$d = \hat{\boldsymbol{n}}\cdot\boldsymbol{a}\,.$$

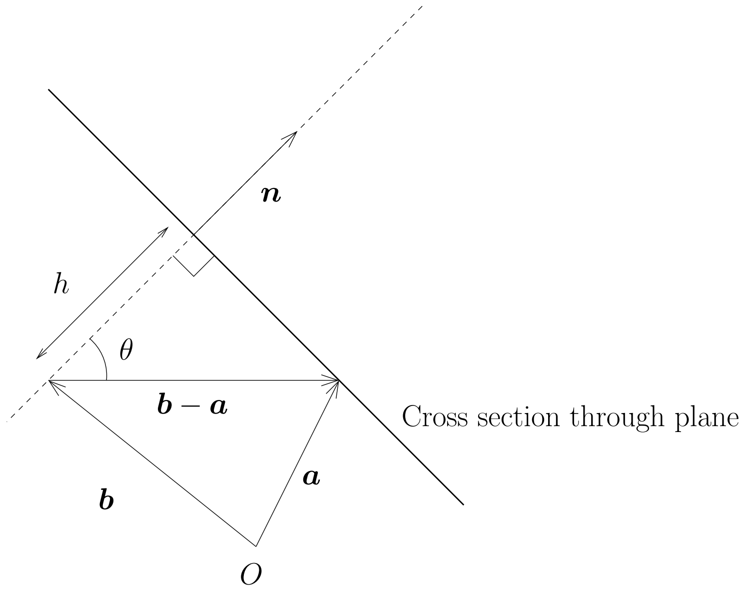


Figure 3.9: The distance h of a point $\boldsymbol{b}$ from a plane.

The perpendicular distance, h, to the plane from $\boldsymbol{b}$ is then the length of the projection of the vector from $\boldsymbol{b}$ to $\boldsymbol{a}$, $\boldsymbol{a}-\boldsymbol{b}$, onto $\boldsymbol{n}$: $h = ||\boldsymbol{a}-\boldsymbol{b}|\cos\theta|$ where θ is the angle between $\boldsymbol{a}-\boldsymbol{b}$ and $\boldsymbol{n}$. But remember that $\boldsymbol{n}\cdot(\boldsymbol{a}-\boldsymbol{b}) = |\boldsymbol{a}-\boldsymbol{b}||\boldsymbol{n}|\cos\theta$. Hence

$$h = \frac{|\boldsymbol{n}\cdot(\boldsymbol{a}-\boldsymbol{b})|}{|\boldsymbol{n}|} = \frac{|\boldsymbol{n}\cdot\boldsymbol{a}-\boldsymbol{n}\cdot\boldsymbol{b}|}{|\boldsymbol{n}|} = \frac{|d-\boldsymbol{n}\cdot\boldsymbol{b}|}{|\boldsymbol{n}|}\,.$$

Note that we would generally write the non-parametric equation of a plane, $\boldsymbol{n}\cdot\boldsymbol{r} = d$, where

$$\boldsymbol{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

as

$$n_1x + n_2y + n_3z = d\,.$$

Question 3.4. Suppose we have a plane with orientation defined by the normal vector $\boldsymbol{n} = 3\boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k}$ and a point in the plane $\boldsymbol{a} = (1, 1, 3)$. Represent the non-parametric equation of the plane in vector notation and in Cartesian form, and calculate the perpendicular distance of the plane from the origin.

Solution. The general equation for a plane in vector notation reads

$$\boldsymbol{n}\cdot\boldsymbol{r} = \boldsymbol{n}\cdot\boldsymbol{a}\,.$$

The left-hand side in this case is

$$\boldsymbol{n}\cdot\boldsymbol{r} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}\cdot\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

while the right-hand side is

$$\boldsymbol{n}\cdot\boldsymbol{a} = \begin{pmatrix}3\\-1\\2\end{pmatrix}\cdot\begin{pmatrix}1\\1\\3\end{pmatrix} = 3-1+6 = 8$$

so the plane is

$$\begin{pmatrix}3\\-1\\2\end{pmatrix}\cdot\boldsymbol{r} = \begin{pmatrix}x\\y\\z\end{pmatrix} = 8\,.$$

To convert this to its Cartesian form is straight forward; we simply calculate the scalar product on the left-hand side of the equation:

$$3x - y + 2z = 8\,. \tag{3.9}$$

In getting the distance from the origin, we might alternatively write the plane $\boldsymbol{n}\cdot\boldsymbol{r} = d$ as $\hat{\boldsymbol{n}}\cdot\boldsymbol{r} = \hat{d}$, where $\hat{\boldsymbol{n}}$ is the unit vector in the direction of $\boldsymbol{n}$: $\hat{\boldsymbol{n}} = \boldsymbol{n}/|\boldsymbol{n}|$. An equation of a plane can simply be multiplied throughout by a scalar to get an alternative form. Here we multiply throughout by $1/|\boldsymbol{n}|$ to get a new right-hand side $\hat{d} = d//|\boldsymbol{n}|$. Here $|\boldsymbol{n}| = \sqrt{9+1+4} = \sqrt{14}$. Eqn. (3.9) is then replaced by

$$\frac{3}{\sqrt{14}}x - \frac{1}{\sqrt{14}}y + \frac{2}{\sqrt{14}}z = 8/1\sqrt{14} \approx 2.138\,.$$

Note that this equation has precisely the correct form for reading off the perpendicular distance to the origin as the left-hand side is the scalar product of the general point, $\boldsymbol{r}$, with a unit vector normal to the plane.
Perpendicular distance to the origin to 4 s.f. is then 2.138.

3.4.2 Parametric representation of a plane

Given two non-parallel direction vectors, $\boldsymbol{b}$ and $\boldsymbol{c}$, parallel to a plane and a point vector in the plane, $\boldsymbol{a}$, the equation of the plane can be derived. A parametric equation of a plane can be derived as a natural extension of the parametric equation of a line,

$$\boldsymbol{r} = \boldsymbol{a} + t\boldsymbol{d}\,.$$

The parametric equation for a plane is given as

$$\boldsymbol{r} = \boldsymbol{a} + s\boldsymbol{b} + t\boldsymbol{c}\,. \tag{3.10}$$

In the equation of a line the location of a point is given by a numerical value of t once all of the vectors are defined. For the equation of a plane the location of a point in the plane is prescribed using the parameters s and ts. A schematic diagram of a plane defined in this way is given in Fig. 3.10).
Given (3.7) and (3.10) are equivalent representations of a plane, it is instructive to see how (3.10) can be converted into the general form (3.7).
The conversion is relatively straightforward. Both representations require a point in the plane, $\boldsymbol{a}$. The difference is that (3.7) requires a normal vector, $\boldsymbol{n}$, perpendicular to the

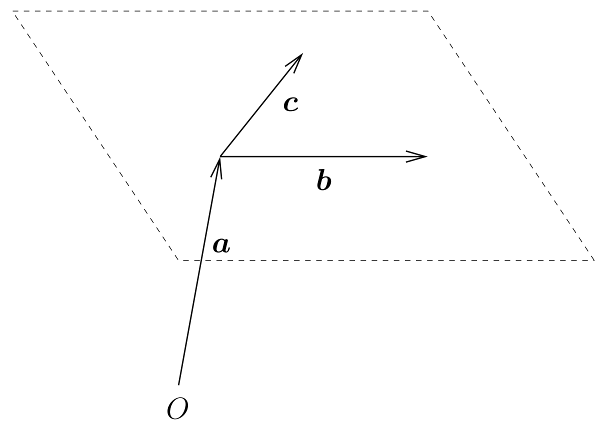


Figure 3.10: The definition of a plane by a point vector and two direction vectors.

plane and (3.10) requires two direction vectors, $\boldsymbol{b}$ and $\boldsymbol{c}$, parallel to the plane but not with each other.
The link between $\boldsymbol{n}$, and $\boldsymbol{b}$ and $\boldsymbol{c}$, is the vector product

$$\boldsymbol{n} = \boldsymbol{b} \times \boldsymbol{c} \,.$$

This gives a vector that is perpendicular to both $\boldsymbol{b}$ and $\boldsymbol{c}$, and hence it is perpendicular to the plane, giving the equation of a plane of the form

$$(\boldsymbol{b} \times \boldsymbol{c}) \cdot \boldsymbol{r} = (\boldsymbol{b} \times \boldsymbol{c}) \cdot \boldsymbol{a} \,.$$

Question 3.5. Suppose that a plane is given by

$$\boldsymbol{r} = \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 2 \\ -4 \\ -3 \end{pmatrix} + t \begin{pmatrix} -1 \\ 4 \\ 0 \end{pmatrix} .$$

Convert it into its Cartesian form and hence determine the perpendicular distance between the origin and the plane.

Solution. In the above, the point in the plane and the two direction vectors are given as

$$\boldsymbol{a} = \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix} , \quad \text{and } \boldsymbol{b} = \begin{pmatrix} 2 \\ -4 \\ -3 \end{pmatrix} \text{ and } \boldsymbol{c} = \begin{pmatrix} -1 \\ 4 \\ 0 \end{pmatrix} ,$$

respectively. Therefore a vector normal to the plane is given by

$$\boldsymbol{n} = \boldsymbol{b} \times \boldsymbol{c} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 2 & -4 & -3 \\ -1 & 4 & 0 \end{vmatrix} = \boldsymbol{i} \begin{vmatrix} -4 & -3 \\ 4 & 0 \end{vmatrix} - \boldsymbol{j} \begin{vmatrix} 2 & -3 \\ -1 & 0 \end{vmatrix} + \boldsymbol{k} \begin{vmatrix} 2 & -4 \\ -1 & 4 \end{vmatrix}$$

$$= \boldsymbol{i}(0 + 12) - \boldsymbol{j}(0 - 3) + \boldsymbol{k}(8 - 4) = 12\boldsymbol{i} + 3\boldsymbol{j} + 4\boldsymbol{k} \,.$$

Then the equation of a plane in the form given by (3.7) is

$$\boldsymbol{n} \cdot \boldsymbol{r} = \boldsymbol{n} \cdot \boldsymbol{a} = d \text{ with } d = \boldsymbol{n} \cdot \boldsymbol{a} = \begin{pmatrix} 12 \\ 3 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix} = 48 + 3 + 0 = 51\,.$$

The Cartesian form of the plane is therefore

$$12x + 3y + 4z = 51\,.$$

To calculate the perpendicular distance h from the plane to the origin we use this equation and the result $h = |d|/|\boldsymbol{n}|$, (3.8). Here $|d| = 51$ while $|\boldsymbol{n}| = \sqrt{144 + 9 + 16} = \sqrt{169} = 13$. Thus the perpendicular distance between the plane and the origin is 51/13.

3.4.3 Plane defined by three point vectors

The above analysis requires that the orientation of the plane is given by a vector perpendicular to the plane, the normal vector, $\boldsymbol{n}$. Another possibility is the plane is defined by three points, say A, B, and C, in the plane, see Fig. 3.11.

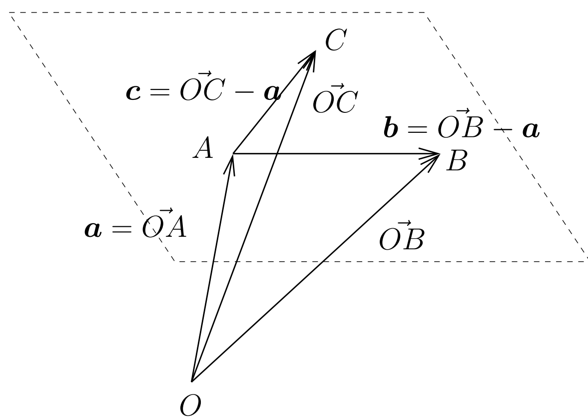


Figure 3.11: The definition of a plane by three points, A, B, C, lying on it.

To find the equation of a plane through three points A, B, C in the plane, we note that the direction vectors

$$\boldsymbol{b} = \vec{OB} - \vec{OA} \text{ and } \boldsymbol{c} = \vec{OC} - \vec{OA} \tag{3.11}$$

are parallel to the plane (see Fig. 3.11). As we have two vectors that are parallel to the plane, if we take their vector product the result is a vector that is perpendicular to the two vectors (3.11) and therefore perpendicular to the plane:

$$\boldsymbol{n} = \boldsymbol{b} \times \boldsymbol{c}\,.$$

We can now appeal to the original analysis that gave us the result for a plane given a point in the plane $\boldsymbol{a}$ and a normal vector $\boldsymbol{n}$.

Example 3.1. Suppose we have three point vectors $\boldsymbol{a} = (1, 0, 1)$, $\boldsymbol{b} = (-2, 5, 0)$ and $\boldsymbol{c} = (3, 1, 1)$ lying on a plane.
The vectors

$$\boldsymbol{b} - \boldsymbol{a} = \begin{pmatrix} -2 \\ 5 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 5 \\ -1 \end{pmatrix} \text{ and } \boldsymbol{c} - \boldsymbol{a} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

are parallel to the plane.
We can calculate their vector product to get a normal to the plane:

$$\boldsymbol{n} = (\boldsymbol{b} - \boldsymbol{a}) \times (\boldsymbol{c} - \boldsymbol{a}) = \begin{pmatrix} -3 \\ 5 \\ -1 \end{pmatrix} \times \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ -3 & 5 & -1 \\ 2 & 1 & 0 \end{vmatrix} = \boldsymbol{i} \begin{vmatrix} 5 & -1 \\ 1 & 0 \end{vmatrix} - \boldsymbol{j} \begin{vmatrix} -3 & -1 \\ 2 & 0 \end{vmatrix} + \boldsymbol{k} \begin{vmatrix} -3 & 5 \\ 2 & 1 \end{vmatrix}$$

$$= \boldsymbol{i}(0 + 1) - \boldsymbol{j}(0 + 2) + \boldsymbol{k}(-3 - 10) = \boldsymbol{i} - 2\boldsymbol{j} - 13\boldsymbol{k}\,.$$

Then the equation of the plane passing through the three point vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ is

$$\boldsymbol{n} \cdot \boldsymbol{r} = (\boldsymbol{i} - 2\boldsymbol{j} - 13\boldsymbol{k}) \cdot (x\boldsymbol{i} + y\boldsymbol{j} + x\boldsymbol{k}) = \boldsymbol{n} \cdot \boldsymbol{a} = (\boldsymbol{i} - 2\boldsymbol{j} - 13\boldsymbol{k}) \cdot (\boldsymbol{i} + \boldsymbol{k}) = 1 - 13 = -12$$

since $\boldsymbol{a} = \boldsymbol{i} + \boldsymbol{k}$ lies on the plane. Equivalently, the plane is

$$x - 2y - 13z = -12\,.$$

3.4.4 Two intersecting planes

Two planes that are not parallel will intersect. The intersection of two planes defines a line, see Fig. 3.12.

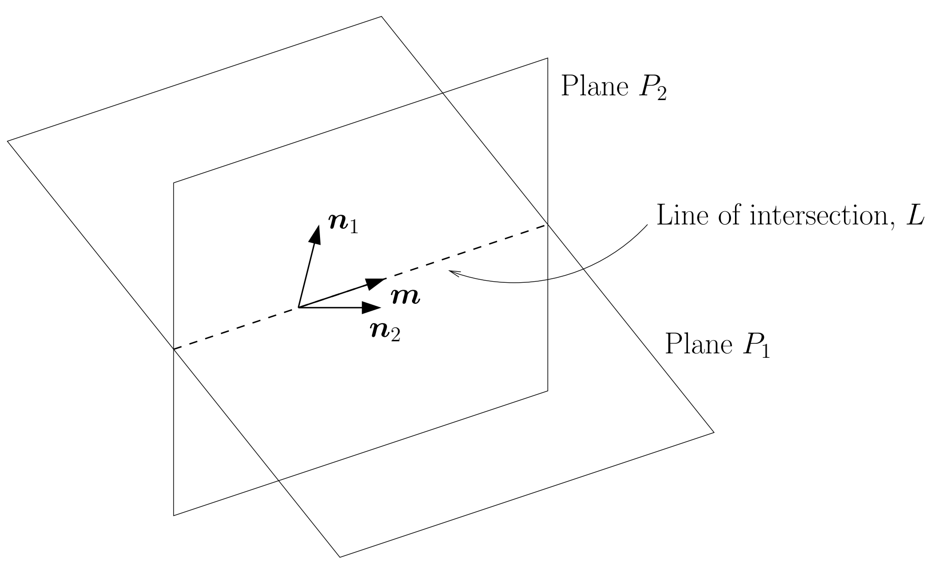


Figure 3.12: Two intersecting planes P_1, P_2 defining a line L. Normals to the planes are $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, respectively.

To calculate the line of intersection of two planes we can proceed by finding two points that are on the line. This allows us to derive a direction vector for the line and we already have a point on the line. As we have seen in Subsection 3.3.1, this is enough information to derive the equation of the line. To appreciate this more fully let us consider an example.

Example 3.2. Consider the two planes $3x + 2y - z = 3$ and $x - y + z = 1$. Let's find their line of intersection.
We look for a point on the line with $x = 0$. Specifying this gives two simultaneous equations for a point that is common to both planes:

$$2y - z = 3\,, \qquad -y + z = 1\,.$$

These are solved by $y = 4$ and $z = 5$. One point on the line of intersection is therefore $\boldsymbol{r} = \boldsymbol{a} = (0, 4, 5)$.
Similarly, trying $y = 0$, to look for another point on the line of intersection, and therefore common to both planes, gives the two simultaneous equations

$$3x - z = 3\,, \qquad x + z = 1\,.$$

These equations have solution, $x = 1$ and $z = 0$. Therefore another point on the line of intersection is $\boldsymbol{r} = \boldsymbol{b} = (1, 0, 0)$.
Now we have two points on the line of intersection we can now write down its parametric equation:

$$\boldsymbol{r} = \boldsymbol{a} + t(\boldsymbol{b} - \boldsymbol{a}) = \begin{pmatrix} 0 \\ 4 \\ 5 \end{pmatrix} + t\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 4 \\ 5 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 4 \\ 5 \end{pmatrix} + t\begin{pmatrix} 1 \\ -4 \\ -5 \end{pmatrix}\,.$$

In Cartesian form, $x = t$, $y = 4 - 4t$, $z = 5 - 5t$.

Note that it is possible to calculate a direction vector for the line using the normal vectors for the planes. The line of intersection is in both planes so is perpendicular to both planes' normals, say $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, and therefore a direction vector is given by the vector product

$$\boldsymbol{m} = \boldsymbol{n}_1 \times \boldsymbol{n}_1\,,$$

see Fig. 3.12. Using this would mean we would only have to determine one point on the line. However, the calculation of a vector product is more complicated than the solution of the second set of simultaneous equations in two unknowns to find a second point on the line.

3.4.5 Parallel planes and the angle between two planes

Whether two planes are parallel or not is determined by their orientations. In terms of vectors this is given by their normal vectors. Consider two planes defined by

$$\boldsymbol{n}_1 \cdot \boldsymbol{r} = d_1 \;\text{ and }\; \boldsymbol{n}_2 \cdot \boldsymbol{r} = d_2\,.$$

These planes are parallel if the two normal vectors, $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, are parallel. This is true if one of them is proportional to the other,

$$\boldsymbol{n}_1 = C\boldsymbol{n}_2 \,,$$

where C is some constant.

Example 3.3. Consider the two planes

$$x + 2y - z = 3 \text{ and } 12x + 24y - 12z = 5 \,.$$

These are parallel planes as $\boldsymbol{n}_1 = (1, 2, -1)$ and $\boldsymbol{n}_2 = (12, 24, -12)$ so that $\boldsymbol{n}_2 = 12\boldsymbol{n}_1$.

Two planes that are not parallel will intersect, defining a line as discussed in Subsection 3.4.4. As the condition for two planes to be parallel is dependent on their normal vectors, the angle between two intersecting planes can be calculated from the normal vectors: the angle between the planes is the same as the angle between the normals, see Fig. 3.13.

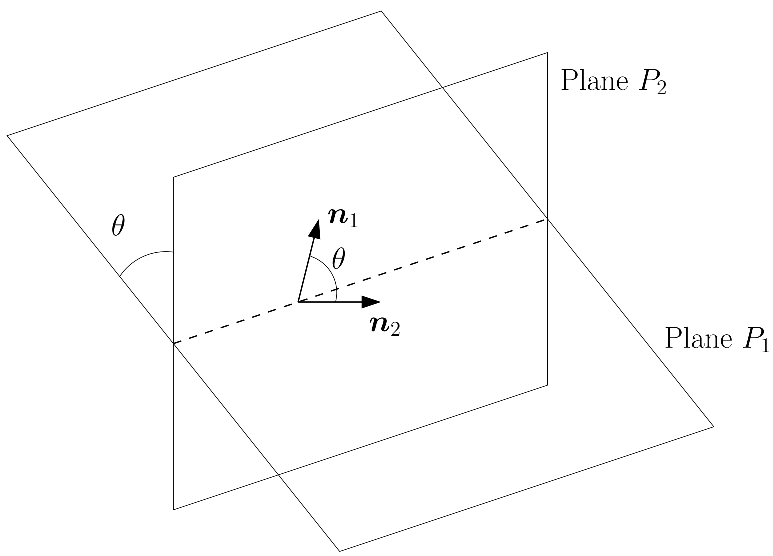


Figure 3.13: Two intersecting planes P_1, P_2, with normals to the planes are $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, respectively. The angle between the planes (and between the normals) is θ.

The angle between the two planes is given by the scalar product of the two normal vectors. In Fig. 3.13 the angle is denoted by θ. The angle between the planes can then be found by using (3.2):

$$\boldsymbol{n}_1 \cdot \boldsymbol{n}_2 = |\boldsymbol{n}_1||\boldsymbol{n}_2| \cos\theta \quad \text{so} \quad \theta = \cos^{-1}\left(\frac{\boldsymbol{n}_1 \cdot \boldsymbol{n}_2}{|\boldsymbol{n}_1||\boldsymbol{n}_2|}\right) .$$

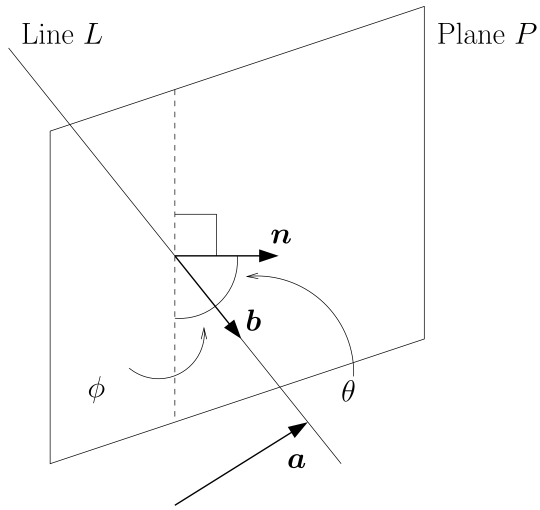


Figure 3.14: The angle ϕ between a line L and a plane P.

3.4.6 Angle between a line and a plane

The last configuration to consider is the angle between a line and a plane. This can be calculated by considering the normal vector defining the orientation of the plane and the direction vector of the line (see Fig. 3.14). If we define the plane by $\boldsymbol{n} \cdot \boldsymbol{r} = d$ and the line by $\boldsymbol{r} = \boldsymbol{a} + t\boldsymbol{b}$, we can get the angle θ between the plane's normal vector, $\boldsymbol{n}$, and the line's direction vector, $\boldsymbol{b}$, by using the scalar product:

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{n} \cdot \boldsymbol{b}}{|\boldsymbol{n}||\boldsymbol{b}|}\right) .$$

Since the normal vector is perpendicular to the plane, the angle between the plane and the line, ϕ, is given by

$$\phi = \frac{\pi}{2} - \theta\,, \qquad \text{alternatively, using degrees, } \phi = 90^\circ - \theta\,. \tag{3.12}$$

Question 3.6. A laser located at the point $\boldsymbol{a} = (2, 4, 1)$ is directed at a mirror lying in the plane given by $2x + 5y + 7z = 11$. If the laser intersects the plane at the point $\boldsymbol{b} = (1, -1, 2)$, what is the angle of incidence of the laser light with the mirror?

Solution. A direction vector for the laser beam is $\boldsymbol{a} - \boldsymbol{b} = (1, 5, -1)$, while a normal to the mirror is $(2, 5, 7)$.
Hence the angle between the normal to the mirror and the laser is given by

$$\cos\theta = \frac{\boldsymbol{n} \cdot (\boldsymbol{a} - \boldsymbol{b})}{|\boldsymbol{n}||\boldsymbol{a} - \boldsymbol{b}|} = \frac{2 + 25 - 7}{\sqrt{4 + 25 + 49}\,\sqrt{1 + 25 + 1}} = \frac{20}{\sqrt{(78)(27)}} = \frac{20}{9\sqrt{26}} .$$

Then $\theta \approx 1.120 \approx 64.16^\circ$.
(Taking a direction vector or normal vector pointing the other way would have given $\cos\theta = -20/(9\sqrt{26})$. We would then want to subtract our value of θ from π (or 180°) to get a value betweeen 0 and $\pi/2$ (0 and 90°).)
The angle between the mirror and the laser is therefore $\pi/2 - \theta \approx 0.451 \approx 25.84^\circ$.

Question 3.7. Where does the line $\boldsymbol{r} = (1, 2, -5) + t(2, -3, 1)$ intersect the plane $2x + 5y - 3z = 6$?

Solution. The equation of the line can be represented in Cartesian form as

$$x = 1 + 2t\,, \quad y = 2 - 3t\,, \quad z = -5 + t\,.$$

A point of intersection must satisfy this set of equations as well as the equation for the plane. Substituting x, y and z into the equation for the plane gives an equation for t:

$$2(1 + 2t) + 5(2 - 3t) - 3(-5 + t) = 6 \text{ so}$$

$$2 + 4t + 10 - 15t + 15 - 3t = 27 - 14t = 6 \quad \Rightarrow \quad 14t = 21 \quad \Rightarrow \quad t = 3/2\,.$$

We can now calculate the coordinates of the point of intersection:

$$x = 1 + 2t = 4\,, \quad y = 2 - 3t = -\frac{5}{2}\,, \quad z = -5 + t = -\frac{7}{2}\,.$$

3.5 Problems

Problem 3.1. What is the angle between the vectors $(2, 2, 1)$ and $(1, -1, 1)$?

Problem 3.2. Evaluate the dot and cross products of the following pairs of vectors.

(a) $(1, 2, 3),\ (4, 5, 6)$.

(b) $(1, -2, 3),\ (3, 1, 2)$.

(c) $(3, 1, 3),\ (4, 2, 1)$.

Problem 3.3. Find the parametric equation of the line passing through each of the pairs of points in the previous question.

Problem 3.4. Verify the identity $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}\,(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{c}\,(\boldsymbol{a} \cdot \boldsymbol{b})$ for the vectors $\boldsymbol{a} = (1, 2, 3)$, $\boldsymbol{b} = (1, -1, 2)$, and $\boldsymbol{c} = (3, 1, 1)$.

Problem 3.5. If $\boldsymbol{a} = (1, 2, -1)$ and $\boldsymbol{b} = (2, -3, 1)$, find a unit vector perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$.

Problem 3.6. Find parametric and non-parametric equations of the planes passing through each of the following sets of points.

(a) $(1, 2, 3),\ (4, 5, 6),\ (0, 0, 0)$.

(b) $(1, 2, 3),\ (4, 5, 6),\ (3, 2, 1)$.

Problem 3.7. Find a non-parametric equation of the plane passing through each of the following points $\boldsymbol{a}$, with the given normal vectors $\boldsymbol{n}$.

(a) $\boldsymbol{a} = (1, 2, 3),\ \boldsymbol{n} = (4, 5, 6)$.

(b) $\boldsymbol{a} = (4, 5, 6),\ \boldsymbol{n} = (3, 2, 1)$.

Problem 3.8. The vertices of a triangle are given by $A = (1, 1, -1)$, $B = (2, -1, 1)$, $C = (-1, 1, 1)$. Use vector operations to determine the angle between the sides AB and AC.

Problem 3.9. Find the angle between the planes

$$\begin{aligned} 2x + y - 2z &= 5\,, \\ 3x - 6y - 2z &= 7\,. \end{aligned}$$

Problem 3.10. Find the angle between the plane

$$2x + y - 2z = 5\,,$$

and the line

$$\begin{aligned} x &= 2 + t\,, \\ y &= 1 + 2t\,, \\ z &= 5t\,. \end{aligned}$$

Problem 3.11. For each of the two planes, find the shortest distance from the origin:
(a) The plane passing through $\boldsymbol{a} = (2, -3, 1)$ and parallel to the vectors $\boldsymbol{d}_1 = (2, 3, 4)$ and $\boldsymbol{d}_2 = (3, 2, 0)$;
(b) The plane passing through $\boldsymbol{a} = (0, 3, 6)$ and parallel to the vectors $\boldsymbol{d}_1 = (3, 2, 2)$ and $\boldsymbol{d}_2 = (-3, 1, 4)$.

Problem 3.12. A plane is given by $x + 2y + z = 6$. A second plane passes through $\boldsymbol{a} = (3, 2, 1)$ and is parallel to the first. Find the equation for the second plane and the distance between the two planes.

Problem 3.13. For the following planes and lines, determine, in each case, whether or not the line is parallel to the plane. If the plane and line are parallel, find their separation. If they are not, find the angle between the line and the plane, and also determine the point at which the line crosses the plane.
(a) $P:\ x + y + z = 2\,;\quad L:\ \boldsymbol{r} = t(-1, 2, 2)$.
(b) $P:\ x + 2y - 3z = 0\,;\quad L:\ \boldsymbol{r} = (1, 1, -1) + t(1, 1, 1)$.
(c) $P:\ -2x + y + 2z = 1\,;\quad L:\ \boldsymbol{r} = (1, 1, -2) + t(-1, 6, -4)$.
(d) $P:\ -2x + y + 2z = 1\,;\quad L:\ \boldsymbol{r} = (1, 1, -2) + t(2, 2, -1)$.

Answers

1. 1.377 radians.

2. (a) 32, $(-3, 6, -3)$. (b) 7, $(-7, 7, 7)$. (c) 17, $(-5, 9, 2)$.

3. (a) $(1, 2, 3) + t(3, 3, 3)$. (b) $(1, -2, 3) + t(-2, -3, 1)$. (c) $(3, 1, 3) + t(-1, -1, 2)$.

5. $(\frac{1}{\sqrt{59}}(-\boldsymbol{i} - 3\boldsymbol{j} - 7\boldsymbol{k})$.

6. (a) $-3x + 6y - 3z = 0$. (b) $x - 2y + z = 0$.

7. (a) $4x + 5y + 6z = 32$. (b) $3x + 2y + z = 28$.

8. 76° (approx.).

9. 79° (approx.).

10. 21° (approx.).

11. (a) $57/\sqrt{233}$. (b) 0.

12. $x + 2y + z = 8$; $\sqrt{2/3}$.

13. (a) Not parallel; 35.26°; $(-2/3, 4/3, 4/3)$. (b) Parallel; $3\sqrt{2/7}$.
13. (c) Parallel; 2. (d) Not parallel; 26.39°; $(-2, -2, -1/2)$.

Chapter 4

Vector Differentiation

4.1 Differentiation of Vectors

In many applications vector functions change with time or some other variable, for example the velocity of a aircraft:

$$\boldsymbol{v}(t) = \begin{pmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \end{pmatrix} .$$

If we wish to calculate the plane's acceleration then we differentiate each individual component with respect to time:

$$\boldsymbol{a}(t) = \begin{pmatrix} a_1(t) \\ a_2(t) \\ a_3(t) \end{pmatrix} = \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}(t) = \begin{pmatrix} \mathrm{d}v_1/\mathrm{d}t \\ \mathrm{d}v_2/\mathrm{d}t \\ \mathrm{d}v_3/\mathrm{d}t \end{pmatrix}$$

so $a_i = \mathrm{d}v_i/\mathrm{d}t$ for $i = 1, 2, 3$.

Example 4.1. Suppose we have a plane descending from 10,000 m towards an airport. (Planes tend to do this by going in circles as they lose altitude.) If the plane's location is given by the co-ordinates $x = 10000\cos(t/100)$, $y = 10000\sin(t/100)$, $z = 10000 - 5t$, we can determine its velocity.

The plane's displacement vector reads

$$\boldsymbol{r} = \begin{pmatrix} 10000\cos(t/100) \\ 10000\sin(t/100) \\ 10000 - 5t \end{pmatrix} .$$

Differentiating each component gives the velocity,

$$\boldsymbol{v} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \begin{pmatrix} -100\sin(t/100) \\ 100\cos(t/100) \\ -5 \end{pmatrix} .$$

Exercise 4.1. Given the vector function

$$\boldsymbol{r} = \begin{pmatrix} 1+t \\ t^2 \\ \frac{2}{3}t^3 \end{pmatrix},$$

find $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ and write it in the form

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = V(t)\hat{T}(t),$$

where $\hat{T}(t)$ is a unit-vector function (the "unit tangent vector").
If $\boldsymbol{r}$ is tought of as position and t as time, so that $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ is velocity, $V = |\mathrm{d}\boldsymbol{r}/\mathrm{d}t|$ is speed; $\hat{T}(t)$ gives direction of motion.

4.1.1 Differentiation of sums and products of vectors

Suppose that $\boldsymbol{a}(t)$ and $\boldsymbol{b}(t)$ are vector functions, then the following rules of differentiation apply:

$$\frac{\mathrm{d}}{\mathrm{d}t}(C\boldsymbol{a}) = C\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} \quad \text{if } C \text{ is a constant;} \tag{4.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{a}+\boldsymbol{b}) = \frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} + \frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}; \tag{4.2}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{a}\cdot\boldsymbol{b}) = \frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t}\cdot\boldsymbol{b} + \boldsymbol{a}\cdot\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}; \tag{4.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{a}\times\boldsymbol{b}) = \frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t}\times\boldsymbol{b} + \boldsymbol{a}\times\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t}. \tag{4.4}$$

These rules of differentiation are natural extensions of the rules of differentiation for scalar functions. Examples of these rules of differentiation are not be presented as you should already be sufficiently accomplished at differentiation and the calculation of vector products and scalar products.
The same rules hold replacing the ordinary derivative by partial derivatives.

4.1.2 Linear approximation of a curve in three dimensions

Before we consider linear approximations to curves in three-dimensional space, let us remind our selves of linear approximations to 1D functions, and curves in 2D.

Consider the function $f(x) = x^2$.
If we consider a truncated Taylor series about the point $x_0 = 1$,

$$f(x) \approx f(x_0) + (x - x_0)f'(x_0), \tag{4.5}$$

where, here, $f(x_0) = f(1) = 1$, $f'(x_0) = 2x_0 = 2$, so $f(x_0) + (x - x_0)f'(x_0) = 1 + 2(x - 1)$ and

$$f(x) \approx f(x_0) + (x - x_0)f'(x_0) = 2x - 1.$$

See Fig. 4.1 to compare the linear approximation with the function x^2.

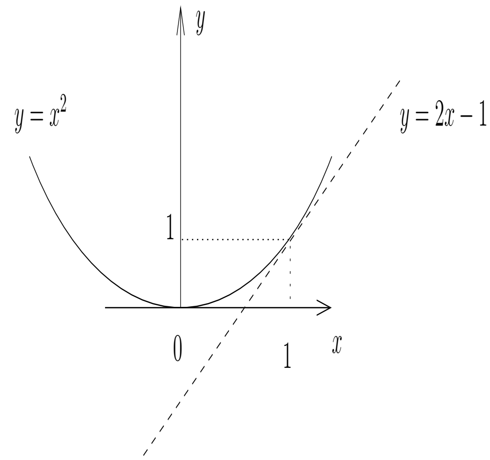


Figure 4.1: The tangent line (dashed) giving a linear approximation to the curve $y = x^2$ (solid) around $x = 1$.

Alternatively, a curve in 2D might be given parametrically as $x = x(t)$, $y = y(t)$. This is the same as $y = f(x)$ if $y(t) = f(x(t))$. A tangent to the curve at a point (x_0, y_0), given by $t = t_0$ with $x_0 = x(t_0)$, $y_0 = y(t_0)$, is then the straight line passing through (x_0, y_0) with direction vector $\begin{pmatrix} \frac{\mathrm{d}x}{\mathrm{d}t}(t_0) \\ \frac{\mathrm{d}y}{\mathrm{d}t}(t_0) \end{pmatrix}$. This can be seen to be the same as the more familiar $y = f(x_0) + f'(x_0)(x - x_0)$:
The straight line passing through (x_0, y_0) with direction vector $\begin{pmatrix} \frac{\mathrm{d}x}{\mathrm{d}t}(t_0) \\ \frac{\mathrm{d}y}{\mathrm{d}t}(t_0) \end{pmatrix}$ is given parametrically by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + s \begin{pmatrix} \frac{\mathrm{d}x}{\mathrm{d}t}(t_0) \\ \frac{\mathrm{d}y}{\mathrm{d}t}(t_0) \end{pmatrix}$$

so that

$$x = x_0 + s\frac{\mathrm{d}x}{\mathrm{d}t}(t_0)\,, \quad y = y_0 + s\frac{\mathrm{d}y}{\mathrm{d}t}(t_0)\,.$$

Eliminating s from these two equations gives

$$y = y_0 + \frac{\mathrm{d}y/\mathrm{d}t(t_0)}{\mathrm{d}x/\mathrm{d}t(t_0)}(x - x_0) = f(x_0) + f'(x_0)(x - x_0)\,,$$

on using $y_0 = f(x(t_0)) = f(x_0)$ and the chain rule:

$$f'(x) = \frac{\mathrm{d}y}{\mathrm{d}x}(x) = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}\,.$$

The same idea can be applied in three dimensions. Consider the vector function $\boldsymbol{x}(t) = (x(t), y(t), z(t))$. We are interested in a linear approximation centred at the point vector given by $t = t_0$, *i.e.* on $\boldsymbol{x}_0 = (x(t_0), y(t_0), z(t_0))$.
Let $\boldsymbol{R}(t)$ denote the linear approximation, then

$$\boldsymbol{R}(t) = \boldsymbol{x}(t_0) + (t - t_0)\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t_0)\,.$$

Alternatively, the tangent line to the curve $\boldsymbol{x} = \boldsymbol{x}(t)$ at the point $\boldsymbol{x}_0$ is given, on writing $s = t - t_0$, by

$$\boldsymbol{r} = \boldsymbol{x}_0 + s\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t_0)\,.$$

In two dimensions, since a vector in the tangent direction is $(\mathrm{d}x/\mathrm{d}t, \mathrm{d}y/\mathrm{d}t)$, and $(\mathrm{d}y/\mathrm{d}t, -\mathrm{d}x/\mathrm{d}t)\cdot(\mathrm{d}x/\mathrm{d}t, \mathrm{d}y/\mathrm{d}t) = 0$, a vector directed perpendicularly to the curve is $\boldsymbol{n} = (\mathrm{d}y/\mathrm{d}t, -\mathrm{d}x/\mathrm{d}t)$. Another normal vector is given by $\boldsymbol{n}/(\mathrm{d}x/\mathrm{d}t) = (f'(x), -1)$.

Question 4.1. Consider the trajectory of a particle as given by the vector function

$$\boldsymbol{x}(t) = \begin{pmatrix} 2t+3 \\ t^2+3t \\ t^3+2t^2 \end{pmatrix}\,. \tag{4.6}$$

Calculate a linear approximation to the particle's trajectory at $t = 1$.

Solution. When $t = 1$, $\boldsymbol{x} = (5, 4, 3)$.
Differentiating (4.6) with respect to t,

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) = \begin{pmatrix} 2 \\ 2t+3 \\ 3t^2+4t \end{pmatrix}\,.$$

Substituting $t = 1$ into this, $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(1) = (2, 5, 7)$.
The linear approximation then reads

$$\boldsymbol{x}(t) = \begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix} + (t-1)\begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 2t+3 \\ 5t-1 \\ 7t-4 \end{pmatrix}\,.$$

How good an approximation is this? Graphically it is difficult to visualise a curve or its linear approximation in 3D, although we can get a feel for the accuracy by looking at the different projections onto 2D planes normal to the coordinate axis.
Note that we might also write the tangent to the curve at $(5, 4, 3)$ as

$$\boldsymbol{x} = \begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix} + s\begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 2s+5 \\ 5s+4 \\ 7s+3 \end{pmatrix}\,.$$

4.2 Gradient of a Scalar Function

There are many physical and chemical properties that can be represented as a scalar function or, as it is sometimes called, a scalar field. For example, the temperature or chemical concentration in a particular application can be represented as a scalar field.
It is possible, given a function of more than one independent variable, to define the **gradient** vector or derivative vector.

Given a function of the form $f(x, y, z)$, then we can define its gradient as

$$\text{gradient of } f = \text{ grad } f = \nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \\ \dfrac{\partial f}{\partial y} \\ \\ \dfrac{\partial f}{\partial z} \end{pmatrix} .$$

Example 4.2. Consider the temperature function $T(x, y, z) = x^2 + 2y^2 + z^2$. According to Fourier's law of heat conduction, heat flux (power flow per unit area) is given by $\boldsymbol{q} = -k\nabla T$, for $k =$ thermal conductivity. Here the heat flux will be

$$\boldsymbol{q} = \nabla T = k \begin{pmatrix} \partial T/\partial x \\ \partial T/\partial y \\ \partial T/\partial z \end{pmatrix} = \begin{pmatrix} 2kx \\ 4ky \\ 2kz \end{pmatrix} .$$

The gradient will be useful in a following subsection on tangent planes (as might be used as approximations to surfaces in 3D).
The gradient has already appeared, in disguised form, in the lecture course in the previous semester, F18XC.

4.2.1 Directional derivatives

In the two-dimensional case, the directional derivative of a function $f(x, y)$ in the direction of an angle α, or, equivalently, in the direction of a unit vector $\hat{\boldsymbol{a}} = (\cos\theta, \sin\theta)$, where θ gives the angular direction of a vector $\boldsymbol{a}$ as measured anti-clockwise from the x direction, see Fig. 4.2, is

$$m_\alpha = D_{\hat{\boldsymbol{a}}} f = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta = \hat{\boldsymbol{a}} \cdot (\partial f/\partial x, \partial f/\partial y) = \frac{\boldsymbol{a} \cdot (\partial f/\partial x, \partial f/\partial y)}{|\boldsymbol{a}|} ,$$

where $\hat{\boldsymbol{a}} = \boldsymbol{a}/|\boldsymbol{a}|$ is the unit vector in the $\boldsymbol{a}$ direction. This then gives the directional derivative $D_{\boldsymbol{a}} f$ of f in the direction of a general vector $\boldsymbol{a}$.
In three dimensions, the directional derivative of a function $f(x, y, z)$ in the direction of a vector $\boldsymbol{a} = (a_1, a_2, a_3)$ is likewise given by

$$D_{\boldsymbol{a}} f = \hat{\boldsymbol{a}} \cdot (\partial f/\partial x, \partial f/\partial y, \partial f/\partial z) = \frac{\boldsymbol{a} \cdot (\partial f/\partial x, \partial f/\partial y, \partial f/\partial z)}{|\boldsymbol{a}|} .$$

It follows that, irrespective of the number of dimensions,

$$D_{\boldsymbol{a}} f = \frac{\boldsymbol{a} \cdot \nabla f}{|\boldsymbol{a}|} = \hat{\boldsymbol{a}} \cdot \nabla f .$$

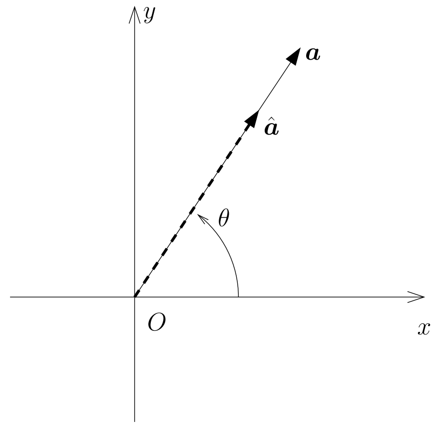


Figure 4.2: In two dimensions, a vector $\boldsymbol{a}$, its corresponding unit vector $\hat{\boldsymbol{a}}$, and angle of direction, θ.

Question 4.2. Consider the function $f(x, y, z) = xy + yz + xz$. Find the "slope" in the direction $\boldsymbol{v} = (1, 2, 3)$.

Solution. First of all we need the gradient, $\nabla f = (\partial f/\partial x, \partial f/\partial y, \partial f/\partial z)$, where $\dfrac{\partial f}{\partial x} = \dfrac{\partial}{\partial x}(xy+yz+xz) = y+z$, $\dfrac{\partial f}{\partial y} = \dfrac{\partial}{\partial y}(xy+yz+xz) = x+z$, $\dfrac{\partial f}{\partial z} = \dfrac{\partial}{\partial z}(xy+yz+xz) = x+y$, so

$$\nabla f = \begin{pmatrix} y+z \\ x+z \\ x+y \end{pmatrix} .$$

The directional derivative is then given by a scalar product:

$$D_{\boldsymbol{v}} f = \frac{\boldsymbol{v} \cdot \nabla f}{|\boldsymbol{v}|} = \frac{(1,2,3) \cdot (y+z, x+z, x+y)}{\sqrt{1+4+9}}$$

$$= \frac{(y+z) + 2(x+z) + 3(x+y)}{\sqrt{14}} = \frac{1}{\sqrt{14}}(5x + 4y + 3z) .$$

Having defined the directional derivative then it is possible to find the direction with greatest positive slope. Using the geometric definition of a scalar product,

$$D_{\boldsymbol{v}} f = \frac{\boldsymbol{v} \cdot \nabla f}{|\boldsymbol{v}|} = \frac{|\boldsymbol{v}||\nabla f| \cos\theta}{|\boldsymbol{v}|} = |\nabla f| \cos\theta ,$$

for θ the angle between the direction vector $\boldsymbol{v}$ and the gradient of f.
Therefore the maximum directional derivative occurs when $\cos\theta = 1$, that is, for $\theta = 0$ ($\boldsymbol{v}$ pointing in the direction of fastest increase of f).

The direction with the greatest slope is

$$\boldsymbol{v} = \nabla f$$

and the directional derivative in this direction is

$$|\nabla f| \,.$$

Likewise, the direction with the greatest negative slope is

$$\boldsymbol{v} = -\nabla f$$

and the directional derivative in this direction is then

$$-|\nabla f| \,.$$

Question 4.3. Consider the function given in Question 4.2, $f(x, y, z) = xy + yz + xz$. In which direction is f increasing the fastest from the point $(1, 3, -1)$?

Solution. Recall that the gradient of f is $\nabla f = (y + z, x + z, x + y)$. At the point given this is

$$\nabla f = \begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix} .$$

This (or, more simply, $(1, 0, 2)$) is the direction that the function is increasing the fastest. The magnitude of the vector gives the rate of increase of the function in this direction: $|\nabla f(1, 3, -1)| = |(2, 0, 4)| = \sqrt{4 + 16} = \sqrt{20} = 2\sqrt{5}$.

Example 4.3. Following the release of a toxic chemical the concentration field is given, at a particular time, by

$$c = C\mathrm{e}^{x^2+y^2+z^2} .$$

Due to diffusion the rate of change of the toxic chemical at any location is related to the gradient

$$\nabla c = C\nabla \mathrm{e}^{x^2+y^2+z^2} = C \begin{pmatrix} 2x\mathrm{e}^{x^2+y^2+z^2} \\ 2y\mathrm{e}^{x^2+y^2+z^2} \\ 2z\mathrm{e}^{x^2+y^2+z^2} \end{pmatrix} = 2C\mathrm{e}^{x^2+y^2+z^2} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2C\mathrm{e}^{x^2+y^2+z^2}\boldsymbol{r} \,.$$

In particular, at a point $(-1, 1, 2)$ the gradient vector is

$$\nabla c = 2C\mathrm{e}^6 \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} .$$

The direction of maximum negative slope (along which the toxic chemical diffuses) is $-\nabla c = 2C\mathrm{e}^6(1, -1, -2)$. An alternative, simpler, vector in the same direction is $(1, -1, -2)$. The slope or size of the gradient in that direction (which gives the speed at which the chemical diffuses) is $-|\nabla c| = -2C\mathrm{e}^6|(-1, 1, 2)| = -2\sqrt{6}\,\mathrm{e}^6 C$.

4.2.2 Equations for a tangent plane and normal line

A surface in three dimensions can be specified by a function as

$$f(x, y, z) = c\,,$$

where c is some constant. For example the equation

$$f(x, y, z) = 2x + 3y + 4z = 1$$

represents a plane while

$$g(x, y, z) = x^2 + y^2 + z^2 = 4$$

represents the surface of a sphere of radius 2 centred on the origin.
In this sub-section we will be deriving approximations to surfaces given in this way by a tangent plane at a particular point on the surface. This can be thought of as an extension to linear approximation of a curve by a line. This is a useful thing to do as equations of planes can be analysed using the techniques considered in Subsection 4.1.2, whereas it is often very difficult to deal with an equation for a general surface analytically.
Before we consider this problem in detail let us look at a simple surface,

$$f(x, y, z) = z = 10\,.$$

This is the equation of a plane that is parallel to the x and y axes. If we differentiate the function with respect to the three independent variables this gives the gradient $\nabla f = (0, 0, 1)$.
The individual components of the gradient give the derivatives in each of the three co-ordinate directions. So, for example, the function above does not change in the x direction so the x component of the derivative is zero. Another interpretation of the derivative vector is as a direction vector. This interpretation leads to the result that in this case the derivative direction is perpendicular, or normal, to the plane.

Important Result
This is a specific example of the general result that, given a surface, $f(\boldsymbol{r}) = c$, then the gradient evaluated at a point on the surface $\boldsymbol{r}_0 = (x_0, y_0, z_0)$, $\nabla f(x_0, y_0, z_0)$, is perpendicular to the surface at $\boldsymbol{r}_0$, see Fig. 4.3.

(It follows from noting that following any curve $\boldsymbol{r}(t) = (x(t), y(t), z(t))$ on the curve gives $\mathrm{d}f/\mathrm{d}t = \mathrm{d}\boldsymbol{r}/\mathrm{d}t \cdot \nabla f$ (by the chain rule), while $\mathrm{d}f/\mathrm{d}t = \mathrm{d}c/\mathrm{d}t = 0$, so ∇f must be normal to any curves in the surface, hence to all their tangent lines, and hence to the tangent plane.)
This is an important result as it makes it possible to calculate planes at a point on a surface.
Given a point $\boldsymbol{r}_0 = (x_0, y_0, z_0)$ on a surface $f(\boldsymbol{r}) = c$, then $\boldsymbol{n} = \nabla f(\boldsymbol{r}_0)$ is perpendicular to the surface at $\boldsymbol{r}_0$.
Recall the equation of a plane,

$$\boldsymbol{n} \cdot \boldsymbol{r} = d\,,$$

where $\boldsymbol{n}$ is a vector normal to the plane. Therefore

$$\nabla f(\boldsymbol{r}_0) \cdot \boldsymbol{r} = d\,,$$

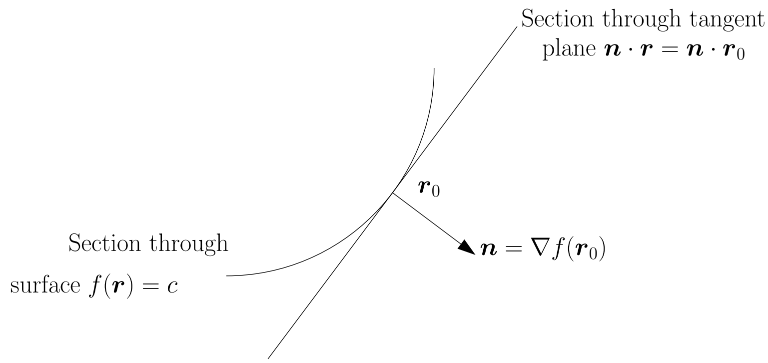


Figure 4.3: The tangent plane to a surface $f(\boldsymbol{r}) = c$ at the point $\boldsymbol{r}_0 = (x_0, y_0, z_0)$.

is the equation of the tangent plane plane. The constant d can be found by noting that the point $\boldsymbol{r}_0$ is in the plane.

Analogously, the parametric equation of a normal line to the surface at $\boldsymbol{r}_0$ is

$$\boldsymbol{r} = \boldsymbol{r}_0 + t\nabla f(\boldsymbol{r}_0)\,,$$

i.e., the gradient at $\boldsymbol{r}_0$ is the direction vector of the normal line passing through $\boldsymbol{r}_0$. Let us see this put into action.

Example 4.4. Consider the surface

$$f(x, y, z) = x^2 + 2y^2 + 3z^2 = 10$$

(the surface of an ellipsoid).
Suppose that we want to derive an equation of a plane that approximates the surface at the point $(1, \sqrt{3}, 1)$.
Applying the gradient,

$$\nabla f = \begin{pmatrix} 2x \\ 4y \\ 6z \end{pmatrix} = \begin{pmatrix} 2 \\ 4\sqrt{3} \\ 6 \end{pmatrix}$$

at $(1, \sqrt{3}, 1)$.
The equation of the plane is then

$$\begin{pmatrix} 2 \\ 4\sqrt{3} \\ 6 \end{pmatrix} \cdot \boldsymbol{r} = d$$

or, equivalently, $2x + 4\sqrt{3}\,y + 6z = d$.
To find d we note that $\boldsymbol{r} = (1, \sqrt{3}, 1)$ lies on the plane so

$$d = \begin{pmatrix} 2 \\ 4\sqrt{3} \\ 6 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \sqrt{3} \\ 1 \end{pmatrix} = 2 + 12 + 6 = 20\,.$$

The equation of the normal plane is therefore

$$\begin{pmatrix} 2 \\ 4\sqrt{3} \\ 6 \end{pmatrix} \cdot \boldsymbol{r} = 20\,.$$

We could tidy it up a little further as 2 is common to all terms and could be removed. The plane can then be written as $x + 2\sqrt{3}\,y + 3z = 10$.

Using the normal vector above, or, more simply, $\boldsymbol{n} = (1, 2\sqrt{3}, 3)$, the equation of the normal line passing through $\boldsymbol{r}_0$ can be written as

$$\boldsymbol{r} = \begin{pmatrix} 1 \\ \sqrt{3} \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 2\sqrt{3} \\ 3 \end{pmatrix} .$$

4.3 Introduction to div and curl

This section introduces the vector operators **div** and **curl**. Vector calculus is a very useful field of mathematics to engineers and scientists. We only give the briefest introduction.

In the previous section we considered the calculus of multi-dimensional scalar functions as these are useful for characterising scalar fields such as the temperature distribution in a thermal system or the concentration distribution following the release of a toxic chemical.

We defined the gradient of a multi-dimensional scalar function $f(x, y, z)$ as a vector function grad f:

$$\text{grad } f = \nabla f = \begin{pmatrix} \partial f/\partial x \\ \partial f/\partial y \\ \partial f/\partial z \end{pmatrix} .$$

Vector functions as well as arising as gradients of multi-dimensional scalar functions are of interest in their own right to engineers and scientists as they allow us to characterise vector quantities such as forces, velocities and accelerations. A vector function takes the form

$$\boldsymbol{f}(x, y, z) = \begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} ,$$

where $f_x = f_1$, $f_y = f_2$ and $f_z = f_3$ represent the components of the vector function parallel to the x, y and z axes.

When considering derivatives of a vector function there are two possibilities. These arise in the same way that "multiplication" of two vectors can take two forms, the scalar or dot product and the vector or cross product. The divergence operator is related to the scalar product, it is defined now.

Definition of the divergence operator
Given a vector function

$$\boldsymbol{f}(x, y, z) = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix},$$

the divergence of $\boldsymbol{f}$ is defined to be

$$\text{div } \boldsymbol{f} = \nabla \cdot \boldsymbol{f} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \partial f_1/\partial x + \partial f_2/\partial y + \partial f_3/\partial z\,.$$

Note that div takes a vector function and produces a scalar one whereas grad takes a scalar function and produces a vector one.

Example 4.5. A compressible fluid with varying density $\rho(\boldsymbol{r}, t)$ flows with velocity $\boldsymbol{v} = (v_1, v_2, v_3)$. To ensure that mass is conserved, the "continuity equation",

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0\,,$$

has to hold.
Suppose that, at some time, $\boldsymbol{q} = \rho\boldsymbol{v} = (2x - y^2, x^2 + 3z, 4y - z^2)$. An evaluation of div $\boldsymbol{q}$ tells us how fast the density is decreasing.

$$\nabla \cdot \boldsymbol{q} = \frac{\partial q_1}{\partial x} + \frac{\partial q_2}{\partial y} + \frac{\partial q_3}{\partial z} = \frac{\partial}{\partial x}(2x - y^2) + \frac{\partial}{\partial y}(x^2 + 3z) + \frac{\partial}{\partial z}(4y - z^2) = 2 + 0 - 2z = 2 - 2z\,.$$

Hence $\partial\rho/\partial t = -\text{div}\boldsymbol{q} = 2z - 2$.

Question 4.4. Evaluate the divergence of the vector function $\boldsymbol{f} = (3x^2y, z, x^2)$

Solution.

$$\nabla \cdot \boldsymbol{f} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \cdot \begin{pmatrix} 3x^2y \\ z \\ x^2 \end{pmatrix} = \frac{\partial}{\partial x}(3x^2y) + \frac{\partial z}{\partial y} + \frac{\partial x^2}{\partial z} = 6xy + 0 + 0 = 6xy\,.$$

The other derivative operator for vector functions, based on the vector product, is the curl, defined here:

Definition of curl
Given a vector function,

$$\boldsymbol{f}(x,y,z) = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix},$$

the curl of $\boldsymbol{f}$ is defined to be

$$\text{curl } \boldsymbol{f} = \nabla \times \boldsymbol{f} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \times \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ f_1 & f_2 & f_3 \end{vmatrix}.$$

Example 4.6. Curl can be thought of as giving a measure of how a vector field turns. Suppose that a fluid rotates as a rigid body with angular velocity $\boldsymbol{\omega}$ about an axis through the origin O. Then the velocity $\boldsymbol{v}$ of the fluid at a point with position vector $\boldsymbol{r} = (x,y,z)$ is given by

$$\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \omega_1 & \omega_2 & \omega_3 \\ x & y & z \end{vmatrix} = (\omega_2 z - \omega_3 y)\boldsymbol{i} + (\omega_3 x - \omega_1 z)\boldsymbol{j} + (\omega_1 y - \omega_2 x)\boldsymbol{k}.$$

Then

$$\begin{aligned} \nabla \times \boldsymbol{v} &= \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \omega_2 z - \omega_3 y & \omega_3 x - \omega_1 z & \omega_1 y - \omega_2 x \end{vmatrix} = \left(\frac{\partial}{\partial y}(\omega_1 y - \omega_2 x) - \frac{\partial}{\partial z}(\omega_3 x - \omega_1 z)\right)\boldsymbol{i} \\ &+ \left(\frac{\partial}{\partial z}(\omega_2 z - \omega_3 y) - \frac{\partial}{\partial x}(\omega_1 y - \omega_2 x)\right)\boldsymbol{j} + \left(\frac{\partial}{\partial x}(\omega_3 y - \omega_1 z) - \frac{\partial}{\partial y}(\omega_2 z - \omega_3 y)\right)\boldsymbol{k} \\ &= 2\omega_1 \boldsymbol{i} + 2\omega_2 \boldsymbol{j} + 2\omega_3 \boldsymbol{k} = 2\boldsymbol{\omega}. \end{aligned}$$

The curl of the velocity is twice the angular velocity.

Question 4.5. Find the curl of the vector function

$$\boldsymbol{f} = \begin{pmatrix} 2x - y^2 \\ x^2 + 3z \\ 4y - z^2 \end{pmatrix}.$$

Solution.

$$\text{curl } \boldsymbol{f} = \nabla \times \boldsymbol{f} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \times \begin{pmatrix} 2x - y^2 \\ x^2 + 3z \\ 4y - z^2 \end{pmatrix} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ 2x - y^2 & x^2 + 3z & 4y - z^2 \end{vmatrix}$$

$$
\begin{aligned}
&= \boldsymbol{i}\begin{vmatrix} \partial/\partial y & \partial/\partial z \\ x^2+3z & 4y-z^2 \end{vmatrix} - \boldsymbol{j}\begin{vmatrix} \partial/\partial x & \partial/\partial z \\ 2x-y^2 & 4y-z^2 \end{vmatrix} + \boldsymbol{k}\begin{vmatrix} \partial/\partial x & \partial/\partial y \\ 2x-y^2 & x^2+3z \end{vmatrix} \\
&= \boldsymbol{i}\left(\frac{\partial}{\partial y}(4y-z^2) - \frac{\partial}{\partial z}(x^2+3z)\right) \\
&- \boldsymbol{j}\left(\frac{\partial}{\partial x}(4y-z^2) - \frac{\partial}{\partial z}(2x-y^2)\right) \\
&+ \boldsymbol{k}\left(\frac{\partial}{\partial x}(x^2+3z) - \frac{\partial}{\partial y}(2x-y^2)\right) \\
&= (4-3)\boldsymbol{i} - (0+0)\boldsymbol{j} + (2x+2y)\boldsymbol{k} = \boldsymbol{i} + 2(x+y)\boldsymbol{k}\,.
\end{aligned}
$$

In some circumstances the div and curl of a physical property can have physical meaning, for example in the vector field of the velocity $\boldsymbol{v}$ of an incompressible fluid, mass conservation can be expressed as

$$\text{div } \boldsymbol{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z} = 0\,.$$

This can be read as the net flow of fluid into a control volume is zero.
Similarly a velocity field for a fluid where the curl is zero, curl $\boldsymbol{v} = \boldsymbol{0}$, means the fluid has no rotational motion.

The vector operators div and curl and the scalar operator grad can be combined. For example,

$$\text{div(grad } f) = \nabla \cdot (\nabla f) = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \cdot \begin{pmatrix} \partial f/\partial x \\ \partial f/\partial y \\ \partial f/\partial z \end{pmatrix} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

This is also written as $\nabla^2 f$.
Provided $f(\boldsymbol{r})$ has continuous second-order derivatives it is possible to show that

$$\text{curl grad } f = \boldsymbol{0}\,.$$

Similarly it is possible to show that

$$\text{div curl } \boldsymbol{f} = 0\,.$$

An advantage of working with the vector and scalar operators is that they provide short hand for systems of partial differential equations. For example, in a solid material the mechanism for heat transfer is conduction, this leads to a partial differential equation that the temperature field $T(x, y, z, t)$ satisfies. The relevant differential equation without using vector and scalar operators reads

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right)\,,$$

where ρ is the density, c is the specific heat and k is the thermal conductivity. Using vector and scalar operators the differential equation reads

$$\rho c \frac{\partial T}{\partial t} = \text{div } (k \text{ grad } T) = \nabla \cdot (k \nabla T)\,.$$

One final example is the system of equations in electromagnetism, Maxwell's equations. Using vector and scalar operators these are, in a vacuum (without any electric charge or current),

$$\nabla \cdot \boldsymbol{H} = 0\,, \quad \nabla \cdot \boldsymbol{E} = 0\,, \quad \varepsilon_0 \boldsymbol{E} = \nabla \times \boldsymbol{H}\,, \quad \boldsymbol{H} = -\mu_0 \nabla \times \boldsymbol{E}\,.$$

Here ε_0 is the electric permittivity of free space, μ_0 is the magnetic permeability of a vacuum, $\boldsymbol{E}$ is the electric field and $\boldsymbol{H}$ is the magnetic field.

4.4 Summary of Vector Geometry

This part began with how vector operations, such as the scalar or vector product could be used to define useful geometric objects, such as a line in three-dimensional space,

$$\boldsymbol{r} = \boldsymbol{a} + t\boldsymbol{b}\,,$$

or give the non-parametric equation of a plane,

$$\boldsymbol{n} \cdot \boldsymbol{r} = d\,.$$

The equation of the plane can take a few different forms depending on what information is available. For example if you have three points in the plane, $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$, then the normal vector defining the orientation of the plane must be constructed by deriving two direction vectors parallel to the plane and using the vector product,

$$\boldsymbol{n} = (\boldsymbol{a} - \boldsymbol{b}) \times (\boldsymbol{b} - \boldsymbol{c})\,.$$

Further relationships between planes and lines, for example the line of intersection between planes or the angle between two intersecting planes, could then be determined without the need to visualise them in three-dimensional space.

Once comfortable with the analysis of planes and lines, the calculus of vectors was introduced. This made it possible to approximate three-dimensional curves by a linear approximation and made it possible to construct planes as local approximations to three-dimensional surfaces, *i.e.* $\nabla f(\boldsymbol{r}_0) \cdot \boldsymbol{r} = d$ is the tangent (approximating) plane to the surface $f(\boldsymbol{r}) = c$ at $\boldsymbol{r}_0$ on taking $d = \nabla f(\boldsymbol{r}_0) \cdot \boldsymbol{r}_0$.

We also found that for a scalar function f we could calculate the directional derivative in any direction defined by a vector $\boldsymbol{v}$:

$$D_{\boldsymbol{v}} f = \frac{\boldsymbol{v} \cdot \nabla f}{|\boldsymbol{v}|} = \frac{\boldsymbol{v} \cdot (\partial f/\partial x, \partial f/\partial y, \partial f/\partial z)}{|\boldsymbol{v}|}\,.$$

Moreover, the direction of the maximum slope is $\boldsymbol{v} = \nabla f$.

The last part of the block introduced the vector operators div and curl.

4.5 Problems

Problem 4.1. If $\boldsymbol{a} = 5t^2\boldsymbol{i} + t\boldsymbol{j} - t^3\boldsymbol{k}$ and $\boldsymbol{b} = \sin(t)\boldsymbol{i} - \cos(t)\boldsymbol{j}$, find

(a) $\dfrac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{a}\cdot\boldsymbol{b})$, (b) $\dfrac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{a}\times\boldsymbol{b})$, (c) $\dfrac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{a}\cdot\boldsymbol{a})$.

Problem 4.2. Using the laws given in the lectures, show that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{a}\cdot\boldsymbol{b}\times\boldsymbol{c}) = \boldsymbol{a}\cdot\boldsymbol{b}\times\frac{d\boldsymbol{c}}{dt} + \boldsymbol{a}\cdot\frac{d\boldsymbol{b}}{dt}\times\boldsymbol{c} + \frac{d\boldsymbol{a}}{dt}\cdot\boldsymbol{b}\times\boldsymbol{c}.$$

Problem 4.3. A particle moves along a curve $\boldsymbol{r}(t) = (x(t), y(t), z(t))$, $x = t$, $y = t^2$, $z = 2t^3/3$.

(a) Find its velocity ($\mathrm{d}\boldsymbol{r}/\mathrm{d}t$) and acceleration ($\mathrm{d}^2\boldsymbol{r}/\mathrm{d}t^2$) at time $t = 1$.

(b) Find the linear approximation of the particle trajectory, velocity and acceleration at $t = 1$

Problem 4.4. Calculate the gradient of the following scalar fields.

(a) $f(x, y, z) = x^2y^3z^4$.

(b) $f(x, y, z) = \sin x \cos(yz)$.

(c) $f(x, y, z) = x^2yz + xy^2z$.

Problem 4.5. Calculate the divergence and curl of the following vector fields.

(a) $\boldsymbol{f}(x, y, z) = (x^2y^2, 2xyz, z^2)$.

(b) $\boldsymbol{f}(x, y, z) = (x^2y^2, y^2z^2, x^2z^2)$.

(c) $\boldsymbol{f}(x, y, z) = (\sin x, \cos y, \tan z)$.

Problem 4.6. Suppose that two vector fields are given by $\boldsymbol{g}(\boldsymbol{r}) = \boldsymbol{r} = (x, y, z)$ and $\boldsymbol{h}(\boldsymbol{r}) = \boldsymbol{\omega}\times\boldsymbol{r}$, where $\boldsymbol{\omega}$ is a constant vector. Show that:

(a) $\nabla\cdot\boldsymbol{g} = 3$; $\nabla\times\boldsymbol{g} = (0, 0, 0)$.

(b) $\nabla\cdot\boldsymbol{h} = 0$; $\nabla\times\boldsymbol{h} = 2\boldsymbol{\omega}$.

Problem 4.7. For each of the scalar functions in Problem 4.4, calculate the directional derivative $\boldsymbol{a}.\nabla f$, at the point $\boldsymbol{r} = (\pi, \pi, 1)$, $\boldsymbol{a} = (1, 2, 2)$.

Problem 4.8. Find the equations for the normal line and tangent plane for the following surfaces, at the points indicated.

(a) $xy + yz + zx = 11$, at $(1, 2, 3)$.

(b) $x^2 + 4y^2 - z^2 = 0$, at $(3, 2, 5)$.

(c) $x^2 + xy + y^2 = 3$, at $(-1, -1, 1)$.

Problem 4.9. Compute the directional derivative at position $(2, 1, -1)$ of the vector field $\phi = x^2yz^3$ in the direction $(1, 0, 0)$. In which direction is the directional derivative at $(2, 1, -1)$ a maximum?

Problem 4.10. Show that if $\boldsymbol{f}(x, y, z) = (yz, xz, xy)$ then $\nabla \times f = 0$, and find a simple scalar field $\phi(x, y, z)$ such that that $\boldsymbol{f} = \operatorname{grad}\phi$. Can you find another $\phi(x, y, z)$ such that $\boldsymbol{f} = \operatorname{grad}\phi$ still holds?

Problem 4.11 (Advanced). Show that if the scalar field $\phi(x, y, z)$ is a solution of Laplace's equation $\nabla \cdot (\nabla\phi) = 0$, then for $\operatorname{grad}\phi$ we have that $\nabla \times \operatorname{grad}\phi = 0$ and $\nabla \cdot \operatorname{grad}\phi = 0$.

Problem 4.12 (Advanced). Suppose that $\phi(x, y, z)$ is a scalar field, and that $\boldsymbol{f}(x, y, z)$ is a vector field. Prove the following:

1. $\operatorname{curl}(\operatorname{grad}(\phi)) = \mathbf{0}$;
2. $\operatorname{curl}(\operatorname{curl}(\boldsymbol{f})) = \operatorname{grad}(\operatorname{div}(\boldsymbol{f})) - \nabla^2\boldsymbol{f}$;
3. $\nabla \times \boldsymbol{f} = \boldsymbol{i} \times \dfrac{\partial \boldsymbol{f}}{\partial x} + \boldsymbol{j} \times \dfrac{\partial \boldsymbol{f}}{\partial y} + \boldsymbol{k} \times \dfrac{\partial \boldsymbol{f}}{\partial z}$.

Answers

3. (b) $(t, -1 + 2t, -4/3 + 2t)$, $(1, 2t, -2 + 4t)$, $(0, 2, 4t)$.

5. (a) $2xy^2 + 2xz + 2z$, $-2xy\boldsymbol{i} + (2yz - 2x^2y)\boldsymbol{k}$. (b) $2xy^2 + 2yz^2 + 2x^2z$, $-2y^2z\boldsymbol{i} - 2xz^2\boldsymbol{j} - 2x^2y\boldsymbol{k}$. (c) $\cos x - \sin y + \sec^2 z$, $\mathbf{0}$.

7. $\frac{8\pi^4}{3}(1 + \pi)$. (b) $\frac{1}{3}$. (c) $\frac{\pi^2}{3}(9 + 4\pi)$.

8. (a) $(1, 2, 3) + u(5, 4, 3)$, $5x + 4y + 3z = 22$. (b) $(3, 2, 5) + u(6, 16, -10)$, $3x + 8y - 5z = 0$. (c) $(-1, -1, 1) + u(-3, -3, 0)$, $x + y = -2$.

10. xyz (+ constant).

Chapter 5

Systems of Linear Equations

5.1 Linear Equations and Elementary Row Operations

Physical and engineering applications
When modelling many physical and engineering problems, we are often left with a system of algebraic equations for unknown quantities x_1, x_2, x_3, …, x_n, say. These unknown quantities may represent components of modes of oscillation in structures for example, or more generally[1]:

Structures: stresses and moments in complicated structures;

Hydraulic networks: hydraulic head at junctions and the rate of flow (discharge) for connecting pipes;

Circuits: electrical currents in circuits;

Surveying: error adjustments *via* least squares method;

Curve fitting: determining the coefficients of the best polynomial approximation;

General networks: abstract network problems – global communication systems;

Finite difference schemes: nodal values in the numerical implementation of a finite-difference scheme for solving differential equation boundary value problems;

Finite element method: elemental values in the numerical implementation of a finite element method for solving boundary value problems – useful in arbitrary geometrical configurations;

Nonlinear cable analysis: bridges, structures;

Production totals: factories, companies;

[1] see `www.nr.com` and `www.ulib.org`

Airline scheduling: flight/ticket availability.

In any of these contexts, the system of algebraic equations that we must solve will in many cases be linear or at least can be well approximated by a linear system of equations. Linear algebraic equations are characterised by the property that no variable is raised to a power other than one or is multiplied by any other variable. The question is: is there a systematic procedure for solving such systems?

We thus aim to study systems of m linear equations in n unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \ldots &= \ldots \\ \ldots &= \ldots \\ \ldots &= \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m\,. \end{aligned} \tag{5.1}$$

Here $x_1, x_2, \ldots, x_n$ are the unknowns and a_{ij}s are coefficients and b_js are constants. Given such a system we would like to have answers to the following questions:

- Does it have a solution? What conditions are required on the coefficients a_{ij} and the constants b_j for the system to have a solution?
- How many solutions?
- How do we find them?

If the constants $b_1 = b_2 = \cdots = b_m = 0$ then the system (5.1) is said to be **homogeneous**. A homogeneous system always has at least one solution, namely $x_1 = x_2 = \cdots = x_n = 0$; this is the **trivial solution**.

Question 5.1. Solve the equation

$$ax = b \tag{5.2}$$

Solution. There are 3 cases:
1. If $a \neq 0$, then for any b (5.2) has a unique solution

$$x = \frac{b}{a}\,.$$

For example the equation

$$3x = 12$$

has unique solution

$$x = \frac{12}{3} = 4\,.$$

2. If $a = 0$ but $b \neq 0$, then (5.2) becomes

$$0 \times x = b\,.$$

Since b is not zero, no value of x will make this statement true, so (5.2) has no solutions in this case.

3. If both $a = 0$ and $b = 0$, then (5.2) becomes

$$0 \times x = 0\,.$$

This is true for any x. So (5.2) has infinitely many solutions in this case.

We shall see that the three cases in the above example exactly mirror the general case.

The coefficients a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ present on the left-hand of the system of equations (5.1) can be arranged into a square table

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \tag{5.3}$$

called an $m \times n$ matrix, more specifically, the **coefficient matrix**. Here m is the number of rows in the matrix (the number of equations) and n is the number of columns (the number of unknowns). A number positioned in the i-th row and the j-th column is called the (ij) matrix entry. For the matrix at hand the i-th row thus corresponds to the i-th equation while the j-th column corresponds to the j-th variable x_j.
The coefficients b_i, $i = 1, \ldots, m$ on the right-hand sides of the equations are arranged into a column

$$\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

We put all of the data specifying the system (5.1) into an $m \times (n+1)$ **augmented matrix** by adding to the matrix $\mathbf{A}$ one extra column on the right. To emphasise that the coefficients b_i are of a different nature than the coefficients a_{ij} we may separate the extra column by a vertical bar:

$$(\mathbf{A}|\boldsymbol{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right).$$

Each row of the augmented matrix represents an equation of system (5.1). To facilitate the solution of system (5.1), we initially define two matrix **row operations**:

1. $\boldsymbol{R}_j \to \boldsymbol{R}_j + k\boldsymbol{R}_i$ means: add k times row i to row j of the matrix;

2. $\boldsymbol{R}_i \to k\boldsymbol{R}_i$ means: multiply row i by $k \neq 0$.

Question 5.2. Solve the system of linear equations

$$\begin{aligned} x + 2y - 3z &= 4\,, \\ x + 3y + z &= 11\,, \\ 2x + 5y - 4z &= 13\,. \end{aligned}$$

Solution. Translated into a matrix formulation, this becomes

$$(\mathbf{A}|\boldsymbol{b}) = \left(\begin{array}{ccc|c} 1^* & 2 & -3 & 4 \\ 1 & 3 & 1 & 11 \\ 2 & 5 & -4 & 13 \end{array}\right).$$

We use the first row to kill the rest of the first column. This corresponds to eliminating x from all but the first equation. The top left entry of the matrix, which is used to drive all the entries directly below it to zero, is called a **pivot** and is marked by an asterisk. We apply

$$\boldsymbol{R}_2 \to \boldsymbol{R}_2 - \boldsymbol{R}_1\,, \qquad \boldsymbol{R}_3 \to \boldsymbol{R}_3 - 2\boldsymbol{R}_1\,,$$

to obtain

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1^* & 4 & 7 \\ 0 & 1 & 2 & 5 \end{array}\right).$$

The process of driving all the entries below a pivot to zero is called a **down-sweep**. Next we use the second enry on the second row as a pivot to kill the elements below it. We do this by applying $\boldsymbol{R}_3 \to \boldsymbol{R}_3 - \boldsymbol{R}_2$

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & 4 & 7 \\ 0 & 0 & -2 & -2 \end{array}\right).$$

We can further simplfy the matrix by multiplying the last row by $-1/2$:

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & 4 & 7 \\ 0 & 0 & 1^* & 1 \end{array}\right).$$

We may now reverse the strategy and use the third row to kill elements in the third column. This process is called an **up-sweep**. After applying

$$\boldsymbol{R}_2 \to \boldsymbol{R}_2 - 4\boldsymbol{R}_3\,, \qquad \boldsymbol{R}_1 \to \boldsymbol{R}_1 + 3\boldsymbol{R}_3,$$

we obtain

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 7 \\ 0 & 1^* & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

We next use the second row to kill elements in the second column by applying $\boldsymbol{R}_1 \to \boldsymbol{R}_1 - 2\boldsymbol{R}_2$

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

The unique solution of the system can now be read off the rightmost column:

$$x = 1\,, \quad y = 3\,, \quad z = 1\,.$$

The process that we employed to solve the above problems is called **Gaussian elimination**. In it one first carries out a sequence of down-sweeps using pivots on the diagonal successively starting with the first row. One next scales the last row and uses it for up-sweeping. Then one scales the second row and uses it for up-sweeping, *etc.* until one obtains a unit matrix in the left block of the augmented matrix. The right-most column of the resulting matrix contains the solution. This algorithm assumes that all pivots encountered in the course of its execution are non-zero.

Question 5.3. Solve the system of linear equations

$$\begin{aligned} x + y + z &= 1 \\ 3x + 4y - z &= 2 \\ -2x + y - z &= 8\,. \end{aligned}$$

Solution. The augmented matrix for this problem is

$$\left(\begin{array}{ccc|c} 1^* & 1 & 1 & 1 \\ 3 & 4 & -1 & 2 \\ -2 & 1 & -1 & 8 \end{array}\right).$$

After down-sweeping using the top left entry as a pivot we obtain

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1^* & -4 & -1 \\ 0 & 3 & 1 & 10 \end{array}\right).$$

down-sweeping using the second row we get

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -4 & -1 \\ 0 & 0 & 13 & 13 \end{array}\right). \tag{5.4}$$

After scaling the last row and then up-sweeping, we get

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

Finally up-sweeping in the second column yields

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

and the unique solution is $x = -3$, $y = 3$, $z = 1$.

As an alternative to scaling the rows (to make all the pivots equal to one) and up-sweeping following down-sweeping, we can carry out the alternative (but equivalent) process of **back substitution**. Once down-sweeping is complete, the final row (last equation) is used to determine the last variable. This is then substituted into the final version of the second last equation to find the second last variable. Both are put into equation got from the third last row, giving the third last variable, and so on.

Example 5.1. We can finish off Question 5.3 using back substitution.
After down-sweeping is complete, we have the augmented matrix (5.4). The third row implies

$$13z = 13 \qquad \text{so } z = 1\,.$$

The second row now gives

$$y - 4z = y - 4 = -1\,,$$

on using $z = 1$, so $y = 3$.
Then the first row leads to

$$x + y + z = x + 3 + 1 = x + 4 = 1\,,$$

on substituting for the values of z and y, so $x = -3$.

Question 5.4. Solve the system of linear equations

$$\begin{aligned} x + 2y + 3z &= 1\,, \\ 4x + 5y + 6z &= 2\,, \\ 7x + 8y + 9z &= 1\,. \end{aligned}$$

Solution. The augmented matrix for this problem is

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \\ 7 & 8 & 9 & 1 \end{array}\right)\,.$$

Down-sweeping brings this matrix to

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -2 \\ 0 & 0 & 0 & -2 \end{array}\right)\,.$$

From here on we are unable to proceed with up-sweeping because the entry we normally use as a pivot is zero. However, we do not have to proceed because the last equation now reads $0 \times x + 0 \times y + 0 \times z = -2$ which cannot be satisfied for any values of x,y, z. The system therefore has no solutions.

Question 5.5. Solve the system of linear equations

$$\begin{aligned} x + 2y - 3z &= 6\,, \\ 2x - y + 4z &= 2\,, \\ 4x + 3y - 2z &= 14\,. \end{aligned}$$

Solution. The augmented matrix for this system is

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 6 \\ 2 & -1 & 4 & 2 \\ 4 & 3 & -2 & 14 \end{array}\right).$$

Down-sweeping yields

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 6 \\ 0 & 1^* & -2 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

The last row corresponds to the equation

$$0 \times x + 0 \times y + 0 \times z = 0$$

which is true for any values of x, y, z. To simplify the system further we do the up-sweep using the entry on the second row. We obtain

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

The first two rows correspond to the equations

$$\begin{aligned} x + z &= 2\,, \\ y - 2z &= 2\,, \end{aligned}$$

respectively. While the value of z remains undetermined (arbitrary), given such a value, the corresponding values of x and y are obtained from

$$x = 2 - z\,, \quad y = 2 + 2z\,.$$

The **general solution** of our system is

$$(x, y, z) = (2 - z, 2 + 2z, z)\,,$$

where z is arbitrary and is called a **free variable**. There are thus infinitely many solutions. A **particular solution** is obtained by fixing the value of z. For example $z = 3$, $x = -1$, $y = 8$ is a particular solution.

5.2 Gaussian Elimination: General Case

Sometimes pivots do not appear where we naively expect them to be and rows may need to be interchanged. Consider the following problem:

Question 5.6. Solve the system of linear equations represented by the augmented matrix

$$\left(\begin{array}{ccc|c} 1^* & 2 & 3 & 1 \\ 2 & 4 & 5 & 1 \\ -1 & 1 & 2 & 4 \end{array}\right).$$

Solution. After down-sweeping from the first pivot, we obtain

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 0^* & -1 & -1 \\ 0 & 3 & 5 & 5 \end{array}\right).$$

We are blocked from using the natural pivot by a zero. However, we can interchange rows 2 and 3 and use a new pivot.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 3^* & 5 & 5 \\ 0 & 0 & -1 & -1 \end{array}\right).$$

Remaining down-sweeps change nothing. After scalings and up-sweeps we obtain

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

and the system has a unique solution $(x, y, z) = (-2, 0, 1)$.

We see that it is convenient to introduce one more row operation: $\boldsymbol{R}_i \leftrightarrow \boldsymbol{R}_j$ which interchanges the rows i and j.

Question 5.7. Solve the system of linear equations represented by the augmented matrix

$$\left(\begin{array}{ccc|c} 1^* & 2 & 3 & 1 \\ 2 & 4 & 5 & 1 \\ -1 & -2 & 2 & 4 \end{array}\right).$$

Solution. All that is changed from the previous example is the entry $A_{3,2}$ but it makes a big difference. After down-sweeping the first column we get

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 0^* & -1 & -1 \\ 0 & 0 & 5 & 5 \end{array}\right)$$

and there is nothing further down in the second column. If we can't move down we move right:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 0 & -1^* & -1 \\ 0 & 0 & 5 & 5 \end{array}\right).$$

We down-sweep column 3 to obtain

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

Scaling and up-sweeping using the entry $A_{2,3}$ as a pivot we obtain

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

There are infinitely many solutions. We can express the pivot variables (those multiplied by the units) in terms of the non-pivot ones:

$$x = -2y - 2\,, \quad z = 1\,.$$

The value of y is arbitrary.

A matrix is said to be in **echelon form** if

1. Any all-zero rows are below all other rows; and
2. The first non-zero entry of any row (called the **pivot entry**) is strictly further right than the first non-zero entry of any row above it

If $(\mathbf{A}'|\boldsymbol{b}')$ is the augmented matrix obtained by using the Gaussian elimination just after the last down-sweep, then the matrix $\mathbf{A}'$ is in echelon form. Here is an example of an augmented matrix after the last down-sweep:

$$\left(\begin{array}{cccccc|c} \boxed{1} & 2 & 1 & 1 & 4 & 1 & 2 \\ 0 & 0 & \boxed{2} & -2 & 6 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & \boxed{3} & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The coefficient part $\mathbf{A}'$ is in echelon form. The pivot entries are put in boxes.

A matrix is said to be in **reduced echelon** form if all the following are true

1. It is in echelon form.
2. The first non-zero entry of any row is 1. This is called a **pivotal 1**.
3. All entries above a pivotal 1 are 0.

If $(\mathbf{A}''|\boldsymbol{b}'')$ is the final augmented matrix obtained by Gaussian elimination (as described above), then $\mathbf{A}''$ is in the reduced echelon form. In the above example, the augmented matrix, after scalings and up-sweeps, is finally in the form

$$\left(\begin{array}{cccccc|c} \boxed{1} & 2 & 0 & 2 & 1 & 0 & 1 \\ 0 & 0 & \boxed{1} & -1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

It is in reduced echelon form (as its coefficient part). The pivotal 1's are put into boxes.

Consider now a general system of m equations for n unknowns. The coefficient matrix $\mathbf{A}$ has dimensions $m \times n$. At the end of Gaussian elimination it has the form

$$\left(\begin{array}{cccccccccccc|c} 0 & \dots & 0 & \boxed{1} & \boxed{} & 0 & \boxed{} & 0 & \boxed{} & 0 & \boxed{} & & d_1 \\ 0 & \dots & 0 & 0 & \dots & \boxed{1} & \boxed{} & 0 & \boxed{} & 0 & \boxed{} & & d_2 \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & \boxed{1} & \boxed{} & 0 & \boxed{} & & d_3 \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots & \boxed{} & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 & \dots & \boxed{1} & \boxed{} & & d_l \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & & d_{l+1} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & & d_m \end{array}\right).$$

The first l rows contain pivotal 1's while the last $m-l$ rows have coefficients zero. There are the following possibilities :

1. If at least one of the entries d_j with $l+1 \le j \le m$ is non-zero, the system has no solutions.

2. If $d_j = 0$ for all $l+1 \le j \le m$ and $l < n$, the system has infinitely many solutions. The non-pivotal variables can be taken as free (undertermined) variables. All pivotal variables are expressed *via* (only) the non-pivotal ones.

3. If $d_j = 0$ for all $l+1 \le j \le m$ and $l = n$ the system has a unique solution: $x_i = d_i$, $i = 1, \dots, n$.

Question 5.8. Determine how many solutions there is to the system of equations

$$\begin{aligned} x + 2y - z &= 0\,, \\ 2x + 5y + 2z &= 0\,, \\ x + 4y + 7z &= 0\,, \\ x + 3y + 3z &= 0\,. \end{aligned}$$

Solution. Down-sweeping, we obtain

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 2 & 5 & 2 & 0 \\ 1 & 4 & 7 & 0 \\ 1 & 3 & 3 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \\ 0 & 1 & 4 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

We see that in the echelon form there are two equations for 3 unknowns ($l = 2 < n = 3$) and the system has infinitely many solutions.

Question 5.9. Solve the system

$$\begin{aligned} x + 3y - 5z + w &= 4\,, \\ 2x + 5y - 2z + 4w &= 6\,. \end{aligned}$$

Solution. Down-sweeping, we obtain

$$\left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 2 & 5 & -2 & 4 & 6 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 0 & -1 & 8 & 2 & -2 \end{array}\right).$$

Scaling the second row and up-sweeping, we obtain

$$\left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 0 & 1 & -8 & -2 & 2 \end{array}\right) \to \left(\begin{array}{cccc|c} \boxed{1} & 0 & 19 & 7 & -2 \\ 0 & \boxed{1} & -8 & -2 & 2 \end{array}\right).$$

The coefficient matrix is now in reduced echelon form. We have $l = 2 < n = 4$ and the system has infinitely many solutions. The pivotal variables are x and y and the general solution is

$$(x, y, z, w) = (-2 - 19z - 7w, 2 + 8z + 2w, z, w)\,.$$

The non-pivotal variables w, z are arbitrary.

The number of non-zero rows in the echelon form of matrix $\mathbf{A}$ is called the **rank** of $\mathbf{A}$, and is denoted by rank($\mathbf{A}$). (It gives a measure of the "size" of $\mathbf{A}$.)

Question 5.10. Find the rank of

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 3 & 4 & 4 & 7 \\ -2 & 1 & 2 & -3 \\ 5 & 3 & 4 & 6 \\ 4 & 5 & 3 & 13 \end{pmatrix}.$$

Solution. To find the rank we bring the matrix to echelon form by down-sweeping:

$$\begin{pmatrix} 1 & 1 & 1 & 2 \\ 3 & 4 & 4 & 7 \\ -2 & 1 & 2 & -3 \\ 5 & 3 & 4 & 6 \\ 4 & 5 & 3 & 13 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 4 & 1 \\ 0 & -2 & -1 & -4 \\ 0 & 1 & -1 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -2 & 4 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -2 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix is now in echelon form. There are 3 non-zero rows so $\text{rank}(\mathbf{A}) = 3$.

Note: For a system of m equations in n unknowns, so that the coefficient matrix $\mathbf{A}$ is $m \times n$ and the augmented matrix $(\mathbf{A}|\boldsymbol{b})$ is $m \times (n+1)$,

- If $\text{rank}(\mathbf{A}) < \text{rank}(\mathbf{A}|\boldsymbol{b})$, there is no solution.
- If $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}|\boldsymbol{b}) = n$, there is a unique solution.
- If $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}|\boldsymbol{b}) < n$, there are infinitely many solutions.

(The rank r of an $m \times n$ matrix $\mathbf{A}$ is the same as that of its transpose $\mathbf{A}^T$, and satisfies $r \leq$ the smaller of m and n.)

5.3 Geometric Interpretation

For two unknowns, the system of equations

$$a_{11}x + a_{12}y = b_1\,, \tag{5.5}$$

represents a line in the plane (assuming that a_{11} and a_{12} aren't both zero). Taking a second equation,

$$a_{21}x + a_{22}y = b_2\,, \tag{5.6}$$

gives a second line, and having the two equations holding together, (5.5) and (5.6), leads to the crossing point of the lines, assuming that the lines are not parallel, *i.e.* that $a_{11}/a_{12} \neq a_{21}/a_{22}$, or, equivalently, $a_{11}a_{22} \neq a_{21}a_{12}$ (see Fig. 5.1). If the lines are parallel, $a_{11}a_{22} = a_{21}a_{12}$, *e.g.* for $x + y = 1$, $2x + 2y = 1$, the lines, generally, do not intersect and then (5.5) and (5.6) have no solution, see Fig. 5.2. If $a_{11}/a_{12} = a_{21}/a_{22} = b_1/b_2$, the lines are coincident and there are infinitely many solutions (*e.g.* $x + y = 1$ and $2x + 2y = 2$).

In three dimensions,

$$a_{11}x + a_{12}y + a_{13}z = b_1 \tag{5.7}$$

represents a plane. Combining this with the equation of a second plane,

$$a_{21}x + a_{22}y + a_{23}z = b_2 \tag{5.8}$$

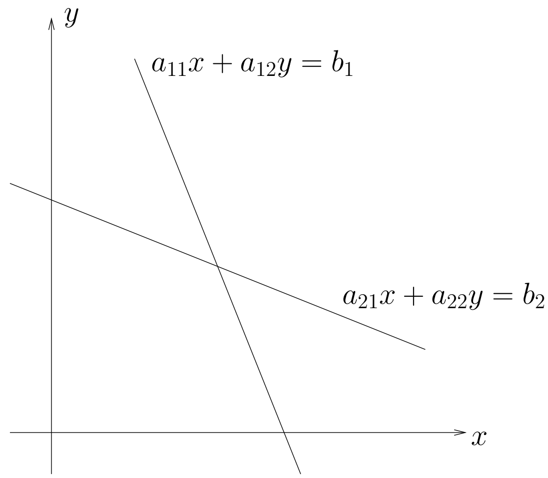


Figure 5.1: Crossing of two non-parallel lines ($a_{11}a_{22} \neq a_{21}a_{12}$).

gives the planes' line of intersection (see Subsection 3.4.4). If we solve (5.7) and (5.8), we can find x and y in terms of the free parameter z, say, and this would then be one form of the equation of the line of intersection.

Now adding a third equation, for a third plane,

$$a_{31}x + a_{32}y + a_{33}z = b_3\,, \tag{5.9}$$

the solution of (5.7), (5.8) and (5.9) (assuming it exists) gives the point where all three planes intersect. This is where the line of intersection crosses the third plane.

If there is no solution, a plane must be parallel to the line of intersection of the other two.

If there are infinitely many solutions, the line of intersection lies on the third plane.

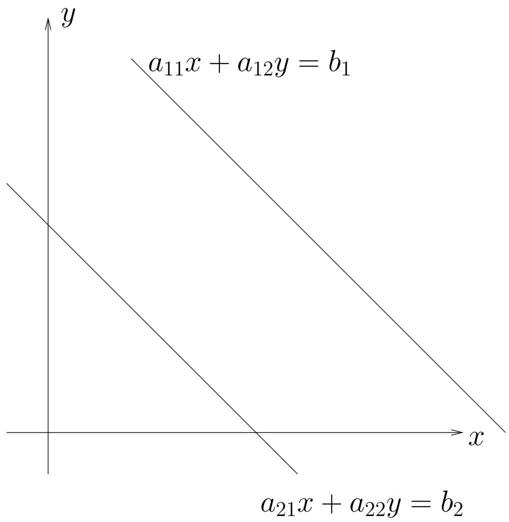


Figure 5.2: Two parallel lines ($a_{11}a_{22} = a_{21}a_{12}$).

5.4 Problems

Problem 5.1. Find how many solutions the following system has:

$$\begin{aligned} 8x_1 + 6x_2 &= 1\,, \\ 20x_1 + 15x_2 &= 0\,. \end{aligned}$$

[Use the quantities $a_{11}a_{22} - a_{12}a_{21}$ and $a_{22}b_1 - a_{12}b_2$ appearing in lectures.]

Problem 5.2. For the system of linear equations

$$\begin{aligned} x_1 + 2x_2 - 3x_3 &= -1\,, \\ 3x_1 - x_2 + 2x_3 &= 7\,, \\ 5x_1 + 3x_2 - 4x_3 &= 2\,, \end{aligned}$$

find the augmented matrix. Simplify it by performing the sequence of row operations

$$R_2 \to R_2 - 3R_1\,, \quad R_3 \to R_3 - 5R_1\,, \quad R_3 \to R_3 - R_2\,.$$

How many solutions does this system have?

Problem 5.3. For the system of linear equations

$$\begin{aligned} x - z &= 2\,, \\ -2x + y + 4z &= -5\,, \\ 2x - 3y - 3z &= 8\,, \end{aligned}$$

find the augmented matrix. Simplify it by down-sweeping in each column. Find the solution by back substitution or by up-sweeping.

Problem 5.4. For the system

$$\begin{aligned} x + 2y + z + 3w &= 1\,, \\ 2x + 5y + 2z + 5w &= 17\,, \\ -x - 2y - 2w &= 4\,, \end{aligned}$$

find the augmented matrix. Simplify it by down-sweeping in each column. Find the general solution using back substitution or up-sweeping and taking z to be a free variable.

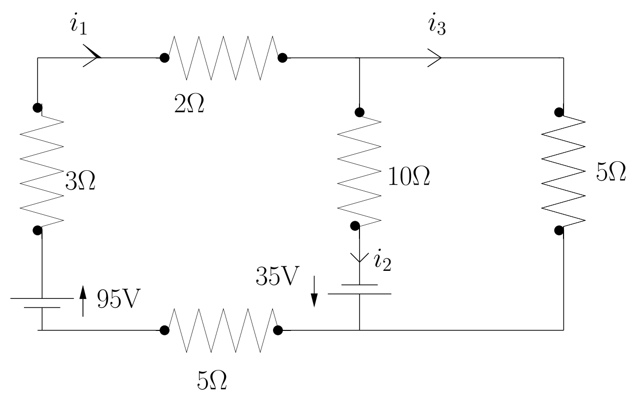
Problem 5.5. For the system with augmented matrix

$$\left(\begin{array}{cccc|c} 1 & 2 & -2 & 3 & 2 \\ 2 & 4 & -3 & 4 & 5 \\ 5 & 10 & -8 & 11 & 12 \end{array}\right),$$

simplify the matrix by down-sweeping in each column. Find the general solution using back substitution (or up-sweeping).

Problem 5.6. For the electric circuit appearing below, Kirchhoff's laws give

$$\begin{aligned} i_2 + i_3 &= i_1\,, \\ 10i_1 + 5i_3 &= 95\,, \\ -5i_3 + 10i_2 &= 35\,. \end{aligned}$$



Find the unknown currents i_1, i_2 and i_3 using Gaussian elimination.

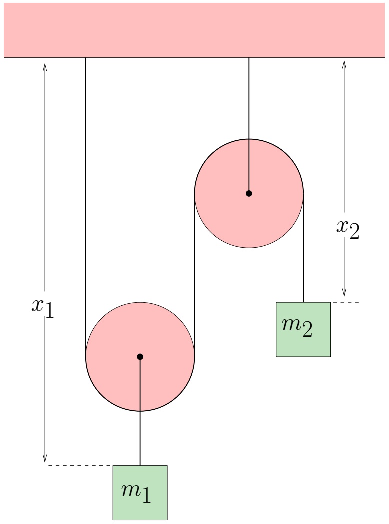
Problem 5.7. Consider the following chemical process:

$$(x_1)\mathrm{CO} + (x_2)\mathrm{CO_2} + (x_3)\mathrm{H_2} \to (x_4)\mathrm{CH_4} + (x_5)\mathrm{H_2O}\,. \tag{5.10}$$

Find the smallest positive integer values of x_1, x_2, x_3, x_4, x_5 so that (5.10) balances. *[Note that the equations for conservation of C, O and H atoms read, respectively,*

$$\begin{aligned} x_1 + x_2 &= x_4\,, \\ x_1 + 2x_2 &= x_5\,, \\ 2x_3 &= 4x_4 + 2x_5\,.] \end{aligned}$$

Problem 5.8. Two weights of mass $m_1 = m_2 = 3$ kg are arranged with light ropes and light pulleys as shown below.



Newton's laws and the equation giving constancy of rope length give

$$\begin{aligned} 3a_1 + 2T &= 3g\,, \\ 3a_2 + T &= 3g\,, \\ 2a_1 + a_2 &= 0\,, \end{aligned}$$

where $a_1 = \ddot{x}_1$ and $a_2 = \ddot{x}_2$ are the downward accelerations, T is the tension in the rope linking the first pulley and the second mass, and g is acceleration due to gravity. Use Gaussian elimination to find a_1, a_2 and T in terms of g.

Problem 5.9. What condition should be placed on the parameter a so that the system

$$\begin{aligned} x + 2y + 3z &= 0\,, \\ 4x + 9y + (a+12)z &= 2\,, \\ -2x - 9y + (10-3a)z &= -10\,, \end{aligned}$$

has infinitely many solutions?

Problem 5.10. What condition should be placed on the parameters a and b so that the system

$$\begin{aligned} x + 2y &= -a\,, \\ -3x - 6y &= 4a - 2b\,, \\ 2x + 7y &= 1 - 2a\,, \end{aligned}$$

has a unique solution?

Problem 5.11. Find the rank of the matrix

$$\mathbf{B} = \begin{pmatrix} 1 & 3 & -2 & 5 & 4 \\ 1 & 4 & 1 & 3 & 5 \\ 1 & 4 & 2 & 4 & 3 \\ 2 & 7 & -3 & 6 & 13 \end{pmatrix}.$$

Would the problem $\mathbf{B}\boldsymbol{x} = \mathbf{0}$ have non-trivial (meaning non-zero) solutions?
For $\mathbf{A}$ the matrix given by the first four columns of $\mathbf{B}$ and $\boldsymbol{b}$ the 4-vector formed by the last column of $\mathbf{B}$, does the problem $\mathbf{A}\boldsymbol{y} = \boldsymbol{b}$, have (a) no solution, (b) a unique solution, or (c) infinitely many solutions.

Answers

3. $x = 11/5$, $y = -7/5$, $z = 1/5$.

4. $(x, y, z, w) = (-54 + 4z, 20 - z, z, 5 - z)$

5. $(x_1, x_2, x_3, x_4) = (4 - 2x_2 + x_4, x_2, 2x_4 + 1, x_4)$.

6. $i_1 = 8$ A, $i_2 = 5$ A, $i_3 = 3$ A.

7. 1, 1, 7, 2, 3

8. $a_1 = -g/5$, $a_2 = 2g/5$, $T = 9g/5$.

9. -8.

10. $a = 2b$.

11. 3, yes, (c).

Chapter 6

Matrices

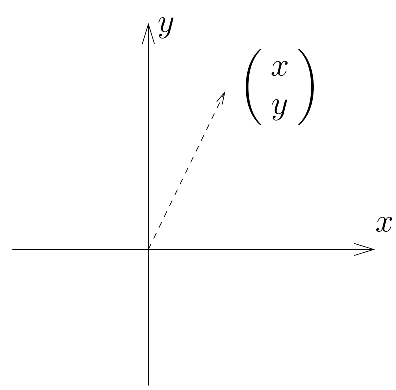
6.1 Vectors and Matrices

In this section we explain how to define an action of a matrix on a vector which will allow us to recast the system of linear equations in terms of that action. We will further define operations with matrices such as matrix multiplication and the inverse of a matrix.
Recall the following notions:

- $\mathbb{R}^2$ denotes the set of all ordered pairs of real numbers

$$\begin{pmatrix} x \\ y \end{pmatrix}$$

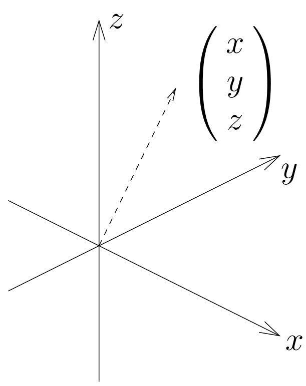
which are called **two-dimensional vectors** and can be depicted as



- $\mathbb{R}^3$ denotes the set of all ordered triples of real numbers

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

called **3-dimensional vectors** and depicted as



- $\mathbb{R}^n$ denotes the set of ordered n-tuples

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

which are called ***n*-dimensional vectors** or ***n*-vectors**. $\mathbb{R}^n$ is called ***n*-dimensional space**.

The n-vectors have the following properties

1. We can add any two n-vectors by adding their components:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

2. We can multiply any vector by a number λ by multiplying each component:

$$\lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

3. There is a special n-vector called the **zero vector** whose every component is a zero:

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Let $\mathbf{A} = (A_{ij})$ be an $m \times n$ matrix and let

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

be an n-vector or a point in $\mathbb{R}^n$. We can define an action of the matrix $\mathbf{A}$ on the vector $\boldsymbol{x}$ to be a vector denoted $\mathbf{A}\boldsymbol{x}$ with components

$$\mathbf{A}\boldsymbol{x} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} . \tag{6.1}$$

Example 6.1. Let

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & 4 & -2 \\ 0 & 2 & 5 & 1 \\ 0 & 1 & -3 & 0 \end{pmatrix}$$

and let

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

be a point in $\mathbb{R}^4$. Then

$$T_A(\boldsymbol{x}) := \mathbf{A}\boldsymbol{x} = \begin{pmatrix} 24 \\ 10 \\ -11 \end{pmatrix}$$

is a point in $\mathbb{R}^3$.

Since the action $\mathbf{A}\boldsymbol{x}$ is defined for any vector $\boldsymbol{x} \in \mathbb{R}^n$ we say that $\mathbf{A}$ gives rise to a **linear transformation**:

$$T_A : \mathbb{R}^n \to \mathbb{R}^m$$

by mapping

$$\boldsymbol{x} \mapsto \mathbf{A}\boldsymbol{x} .$$

We now make two observations

1. For any $m \times n$ matrix $\mathbf{A}$ we always have

$$T_A(\mathbf{0}) = \mathbf{0}$$

   where the $\mathbf{0}$ on the left-hand side is the zero vector in $\mathbb{R}^n$ and the $\mathbf{0}$ on the right-hand side is the zero vector in $\mathbb{R}^m$.

2. For some matrices $\mathbf{A}$ the linear transformation T_A is many-to-one.

Example 6.2. For the matrix **A** as in Example 6.1 and

$$\boldsymbol{x}_1 = \begin{pmatrix} -12 \\ 7 \\ 1 \\ -17 \end{pmatrix}, \qquad \boldsymbol{x}_2 = \begin{pmatrix} 22 \\ 1 \\ -1 \\ 5 \end{pmatrix},$$

we have

$$T_A(\boldsymbol{x}_1) = T_A(\boldsymbol{x}_2) = \begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix} = \boldsymbol{y}$$

and also many other vectors from $\mathbb{R}^4$ are mapped onto $\boldsymbol{y}$. This happens because T_A "crushes" $\mathbb{R}^4$ into $\mathbb{R}^3$.

A system (5.1) of m linear equations for n unknowns in matrix notation can be written as

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}\,. \tag{6.2}$$

A linear transformation associated with **A** acts as

$$T_A : \mathbb{R}^n \to \mathbb{R}^m$$

so that equations (6.2) can be recast as

$$T_A(\boldsymbol{x}) = \boldsymbol{b}$$

and the problem of solving (6.2) is equivalent to finding all vectors in $\mathbb{R}^n$ which T_A maps onto $\boldsymbol{b}$ – a specific vector in $\mathbb{R}^m$.

Suppose now we have an $m \times n$ matrix **A** and an $n \times k$ matrix **B**. The matrix **B** has n rows and k columns. Each column in matrix **B** can be considered to be a vector from $\mathbb{R}^n$. Hence we can map each column in **B** into a column of m elements (a vector in $\mathbb{R}^m$) using the rule (6.1). The resulting columns are then arranged into an $m \times k$ matrix called a product of matrices **A** and **B** and written **AB**. (The order is important!) Thus we have defined a **matrix multiplication**, $\mathbf{C} = \mathbf{AB}$, so that the (ij) entry in the product matrix is given by the formula

$$C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{in}B_{nj}\,. \tag{6.3}$$

Question 6.1. Find the matrix product **AB** for the matrices

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ 4 & 1 & 1 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 7 & -1 \\ 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Solution. The result is a 2×2 matrix

$$\mathbf{AB} = \begin{pmatrix} (1\times 7 - 1\times 2 + 0\times 0) & (1\times(-1) - 1\times 0 + 0\times 1) \\ (4\times 7 + 1\times 2 + 1\times 0) & (4\times(-1) + 1\times 0 + 1\times 1) \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ 30 & -3 \end{pmatrix}.$$

Note that in general not any two matrices can be multiplied. The number of columns in the first matrix must match the number of rows in the second.

Example 6.3. In the above question one can also define **BA**:

$$\mathbf{BA} = \begin{pmatrix} 7 & -1 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 4 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -8 & -1 \\ 2 & -2 & 0 \\ 4 & 1 & 1 \end{pmatrix}.$$

6.2 Inverse Matrices

Let **A** be an $n \times n$ matrix. Its **inverse**, denoted by $\mathbf{A}^{-1}$, is an $n \times n$ matrix such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}\,,$$

where **I** is the $n \times n$ identity matrix:

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

For example one can check that

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$$

and therefore

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}.$$

(Multiplying a matrix (or vector) by the identity matrix leaves the matrix unchanged – assuming the products are defined: for **I** the $m \times m$ identity matrix, **A** an $l \times m$ matrix and **B** an $m \times n$ matrix, then $\mathbf{AI} = \mathbf{A}$ and $\mathbf{IB} = \mathbf{B}$.)

Recall that a system of linear equations (5.1) can be written in a matrix form

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}\,. \tag{6.4}$$

Suppose further that $n = m$, then **A** is an $n \times n$ matrix, $\boldsymbol{x}$ is an $n \times 1$ matrix (*i.e.* an n-vector) and $\boldsymbol{b}$ is an $n \times 1$ matrix (another n-vector). If the system has a unique solution, then Gaussian elimination (with up-sweeping and scaling) will bring its augmented matrix $(\mathbf{A}|\boldsymbol{b})$ to the form $(\mathbf{A}'|\boldsymbol{b}') = (\mathbf{I}|\boldsymbol{b}')$ where **I** is the $n \times n$ identity matrix and $\boldsymbol{b}'$ is the solution $\boldsymbol{x} = \boldsymbol{b}'$. On the other hand, if we apply the inverse matrix $\mathbf{A}^{-1}$ to both sides of our equation (6.4),

$$\mathbf{A}^{-1}\mathbf{A}\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b}$$

and therefore

$$\mathbf{I}\boldsymbol{x} = \boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b}\,.$$

This means that $\boldsymbol{b}' = \mathbf{A}^{-1}\boldsymbol{b}$. Solving the system for different right-hand sides $\boldsymbol{b}$ gives us information on the entries of the inverse matrix $\mathbf{A}^{-1}$. Thus if $\boldsymbol{b}$ is a column vector with 1 on the i-th row and zeroes everywhere else, $\boldsymbol{b} = \boldsymbol{e}_i$, the i-th **elementary column vector** (in three dimensions, $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ are the three unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$), $\mathbf{A}^{-1}\boldsymbol{b}$ gives the i-th column of the matrix $\mathbf{A}^{-1}$. Solving the system for all such elementary column vectors with $i = 1, 2, \ldots, n$ gives the entire matrix $\mathbf{A}^{-1}$. We can do it all at once by applying the same row operations to a large augmented matrix $(\mathbf{A}|\mathbf{I})$ where the second block is an $n \times n$ identity matrix (the n elementary column vectors put together).

Question 6.2. Use Gaussian elimination to find the inverse for the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}.$$

Solution. We start with an augmented matrix

$$\left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array}\right),$$

which we can treat as a single 2×4 matrix. Down-sweeping in the first column, we obtain

$$\left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & 1/2 & -5/2 & 1 \end{array}\right).$$

Multiplying the second row by 2 and up-sweeping the second column, we obtain

$$\left(\begin{array}{cc|cc} 2 & 0 & 6 & -2 \\ 0 & 1 & -5 & 2 \end{array}\right).$$

Finally multiplying the first row by 1/2 brings our matrix to the form

$$\left(\begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1 & -5 & 2 \end{array}\right)$$

and we see that when the left block became the identity matrix the right block became the inverse matrix to the original coefficient matrix $\mathbf{A}$.

Warning: Not every square matrix has an inverse! However, it can be proved that if system (6.4) has a unique solution then $\mathbf{A}^{-1}$ exists for its matrix of coefficients $\mathbf{A}$.

Question 6.3. Find $\mathbf{B}^{-1}$ for

$$\mathbf{B} = \begin{pmatrix} 6 & 11 & 5 \\ 18 & 34 & 15 \\ 12 & 25 & 11 \end{pmatrix} .$$

Solution. We start with the matrix

$$(\mathbf{B}|\mathbf{I}) = \left(\begin{array}{ccc|ccc} 6 & 11 & 5 & 1 & 0 & 0 \\ 18 & 34 & 15 & 0 & 1 & 0 \\ 13 & 25 & 11 & 0 & 0 & 1 \end{array}\right)$$

and carry out the usual down-sweeping,

$$\to \left(\begin{array}{ccc|ccc} 6 & 11 & 5 & 1 & 0 & 0 \\ 0 & 1 & 0 & -3 & 1 & 0 \\ 0 & 3 & 1 & -2 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 6 & 11 & 5 & 1 & 0 & 0 \\ 0 & 1 & 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 7 & -3 & 1 \end{array}\right) ,$$

up-sweeping,

$$\to \left(\begin{array}{ccc|ccc} 6 & 0 & 5 & 34 & -11 & 0 \\ 0 & 1 & 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 7 & -3 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 6 & 0 & 0 & -1 & 4 & -5 \\ 0 & 1 & 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 7 & -3 & 1 \end{array}\right) ,$$

and finally scaling,

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/6 & 2/3 & -5/6 \\ 0 & 1 & 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 7 & -3 & 1 \end{array}\right) .$$

Hence

$$\mathbf{B}^{-1} = \begin{pmatrix} -1/6 & 2/3 & -5/6 \\ -3 & 1 & 0 \\ 7 & -3 & 1 \end{pmatrix} .$$

6.3 Determinants

A **determinant** of a 2×2 matrix $\mathbf{A}$ is the number

$$\det(\mathbf{A}) := \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc\,. \tag{6.5}$$

In Question 6.2, $\det(\mathbf{A}) = 2 \times 3 - 1 \times 5 = 1$. A determinant of a 3×3 matrix is a number defined by

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \times \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \times \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \times \det \begin{pmatrix} d & e \\ g & h \end{pmatrix} ,$$

where the determinants of 2×2 matrices are computed according to formula (6.5).

Example 6.4.

$$\det\begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix} = 1 \times \det\begin{pmatrix} -5 & 3 \\ -6 & 4 \end{pmatrix} - (-3) \times \det\begin{pmatrix} 3 & 3 \\ 6 & 4 \end{pmatrix} + 3 \times \det\begin{pmatrix} 3 & -5 \\ 6 & -6 \end{pmatrix}$$
$$= (-5 \times 4 - 3 \times (-6)) + 3(3 \times 4 - 3 \times 6) + 3(3 \times (-6) - (-5) \times 6)$$
$$= -2 + 3 \times (-6) + 3 \times 12 = 16\,.$$

In general a determinant of an $n \times n$ matrix $\mathbf{A} = (A_{ij})$ can be defined recursively as

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{j+1} A_{1j} \times \det(A^{(1,j)})\,, \tag{6.6}$$

where $\mathbf{A}^{(1,j)}$ is an $(n-1) \times (n-1)$ matrix obtained from $\mathbf{A}$ by removing from $\mathbf{A}$ the 1st row and the j-th column.

We will accept without proof the following theorem :

Theorem 6.1. *The homogeneous system of n equations for n unknowns: $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ has a non-trivial solution if and only if* $\det(\mathbf{A}) = 0$.

Question 6.4. Given a system of linear equations

$$\begin{aligned} x + 2y &= 0\,, \\ -x + z &= 0\,, \\ 2x - 3y = az &= 0\,, \end{aligned}$$

find for what values of parameter a it has a non-trivial solution.

Solution. Computing

$$\det\begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \\ 2 & -3 & a \end{pmatrix} = 7 - 2a$$

and setting it to zero, we conclude that the system has a non-trivial solution if and only if $a = 7/2$.

One can also prove that if two matrices $\mathbf{A}$ and $\mathbf{B}$ can be turned into each other by elementary row operations, their determinants are proportional to each other. Thus if $\det(\mathbf{A}) = 0$ then $\det(\mathbf{B}) = 0$ and *vice versa*. This implies in particular that:

- if a matrix $\mathbf{A}$ has a zero row its determinant vanishes;
- if a matrix $\mathbf{A}$ has two rows proportional, its determinant vanishes;
- if a matrix $\mathbf{A}$ has rows which are **linearly dependent** (*i.e.* one row can be written as a sum of others, or a non-trivial sum of rows vanishes), its determinant vanishes.

(The same comments apply to columns.)
(If the row operations are restricted to adding or subtracting a multiple of one to or from another, avoiding swaps and scalings, the value of the determinant is unchanged. The same applies to manipulating columns.)

The following theorem relates matrix multiplication and determinants.

Theorem 6.2. *For any two $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$,* $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$.

Question 6.5. For the matrices $\mathbf{A}$ and $\mathbf{B}$ below check that $\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det(\mathbf{A})\det(\mathbf{B})$

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} -1 & 0 \\ 2 & 5 \end{pmatrix}$$

Solution. We have

$$\mathbf{AB} = \begin{pmatrix} 3 & 10 \\ 9 & 20 \end{pmatrix}, \qquad \mathbf{BA} = \begin{pmatrix} -1 & -2 \\ -3 & 24 \end{pmatrix}$$

and the determinants are $\det(\mathbf{A}) = 6$, $\det(\mathbf{B}) = -5$, $\det(\mathbf{AB}) = \det(\mathbf{BA}) = -30$ so that indeed $\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det(\mathbf{A})\det(\mathbf{B})$.

Determinants and inverse matrices

Determinants can be used to compute entries of an inverse matrix $\mathbf{A}^{-1}$ directly through the entries of the original matrix $\mathbf{A}$.

Given an $n \times n$ matrix $\mathbf{A}$, define the (i, j)-th **minor**, denoted by M_{ij}, to be the determinant of the $(n-1) \times (n-1)$ matrix which results from deleting row i and column j of $\mathbf{A}$.

Example 6.5. For the matrix

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

we find, for instance,

$$M_{12} = \det \begin{pmatrix} 0 & 3 \\ -1 & 1 \end{pmatrix} = 3\,, \quad M_{31} = \det \begin{pmatrix} -2 & 0 \\ 2 & 3 \end{pmatrix} = -6\,.$$

The entries of the inverse matrix $\mathbf{A}^{-1}$ can becomputed by using the minors using the following formula:

$$A^{-1}_{ij} = \frac{(-1)^{i+j} M_{ji}}{\det(\mathbf{A})}\,. \tag{6.7}$$

Note that the order of indices for the minor is interchanged compared with the order of indices in the inverse matrix. In practice it is convenient first to compute the matrix

$C_{ij} = (-1)^{i+j} M_{ij}$ which is called the **cofactor matrix** of matrix $\mathbf{A}$, and then take its **transpose**, that is, interchange the rows and columns and finally divide the result by $\det(\mathbf{A})$. In general transpose of a matrix $\mathbf{A}$ is denoted by $\mathbf{A}^T$ and means that $A^T_{ij} = A_{ji}$. For example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

With this notation we have

$$\mathbf{A}^{-1} = (1/\det(\mathbf{A}))\mathbf{C}^T.$$

Formula (6.7) assumes that $\det(\mathbf{A}) \neq 0$, so that we can divide by that quantity. Having $\det(\mathbf{A}) \neq 0$ is a necessary and sufficient condition for the inverse matrix to $\mathbf{A}$ to exist.

Example 6.6. For an arbitrary 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

the above formula gives

$$\boxed{\mathbf{A}^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.}$$

(This can be worth remembering!)

Question 6.6. Compute the inverse of matrix given in Example 6.4, using formula (6.7).

Solution. We first find the cofactor matrix

$$\mathbf{C} = \begin{pmatrix} +\det\begin{pmatrix} -5 & 3 \\ -6 & 4 \end{pmatrix} & -\det\begin{pmatrix} 3 & 3 \\ 6 & 4 \end{pmatrix} & +\det\begin{pmatrix} 3 & -5 \\ 6 & -6 \end{pmatrix} \\ -\det\begin{pmatrix} -3 & 3 \\ -6 & 4 \end{pmatrix} & +\det\begin{pmatrix} 1 & 3 \\ 6 & 4 \end{pmatrix} & -\det\begin{pmatrix} 1 & -3 \\ 6 & -6 \end{pmatrix} \\ +\det\begin{pmatrix} -3 & 3 \\ -5 & 3 \end{pmatrix} & -\det\begin{pmatrix} 1 & 3 \\ 3 & 3 \end{pmatrix} & +\det\begin{pmatrix} 1 & -3 \\ 3 & -5 \end{pmatrix} \end{pmatrix}.$$

We obtain

$$\mathbf{C} = \begin{pmatrix} -2 & 6 & 12 \\ -6 & -14 & -12 \\ 6 & 6 & 4 \end{pmatrix}.$$

Taking the transpose of this matrix and dividing it by 16 we obtain

$$\mathbf{A}^{-1} = \frac{1}{16} \begin{pmatrix} -2 & -6 & 6 \\ 6 & -14 & 6 \\ 12 & -12 & 4 \end{pmatrix}.$$

6.4 Problems

Problem 6.1. For a matrix $\mathbf{A}$ and vectors $\boldsymbol{x}_1 \in \mathbb{R}^3$ and $\boldsymbol{x}_2 \in \mathbb{R}^5$, find vectors $\boldsymbol{y} = \mathbf{A}\boldsymbol{x}_1 \in \mathbb{R}^5$ and $\boldsymbol{z} = 2\boldsymbol{y} - 3\boldsymbol{x}_2 \in \mathbb{R}^5$, with

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -2 \\ 7 & 1/2 & -1 \\ -1/3 & 0 & 2 \\ 2 & 2 & 2 \\ 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{x}_1 = \begin{pmatrix} 6 \\ -2 \\ 1 \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1/3 \\ 11 \end{pmatrix}.$$

Problem 6.2. (a) The matrix performing clockwise rotations about the x axis in 3-dimensional space has the form

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

where θ is the angle of rotation. Use this matrix to calculate the components of vector

$$\boldsymbol{v} = \begin{pmatrix} 1 \\ -2.5 \\ 0.3 \end{pmatrix}$$

after clockwise rotation by $\pi/3$ about the x axis.
(b) Check, using matrix multiplication, the identity

$$\mathbf{R}_x(\pi/6)\mathbf{R}_x(\pi/6) = \mathbf{R}_x(\pi/3)\,.$$

[Can you guess what $\mathbf{R}_x(\theta_1)\mathbf{R}_x(\theta_2)$ is in general?]

Problem 6.3. Find $\mathbf{CD}$, $\mathbf{DC}$ and $(\mathbf{CD})^2$ using matrix multiplication, where

$$\mathbf{C} = \begin{pmatrix} 3 & -1 & 2 \\ 0 & 4 & 5 \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} 11 & 2 \\ -1 & -3 \\ 0 & 1 \end{pmatrix}.$$

Problem 6.4. Use Gaussian elimination to find the inverse of

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

Problem 6.5. Find $\mathbf{B}^{-1}$ and $\mathbf{C}^{-1}$, by any method, for

$$\mathbf{B} = \begin{pmatrix} 1 & 4 \\ -1 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} 0 & \frac{3}{2} \\ -\frac{1}{3} & 2 \end{pmatrix}.$$

Problem 6.6. Compute the determinants $\det(\mathbf{D})$, $\det(\mathbf{E})$ of the following matrices:

$$\mathbf{D} = \begin{pmatrix} 11 & -3 \\ 5 & 2 \end{pmatrix}, \qquad \mathbf{E} = \begin{pmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{pmatrix}.$$

Problem 6.7. Check, by computing a determinant, that the homogeneous system of linear equations

$$\begin{aligned} x - z &= 0\,, \\ 2x + 7y - 3z &= 0\,, \\ 5x + 14y - 7z &= 0\,, \end{aligned}$$

has a non-trivial solution.

Problem 6.8. Find for what values of parameter a the following system of linear equations (for the unknowns x, y) has a non-trivial solution:

$$\begin{aligned} ax + 12y &= 0\,, \\ 3x + ay &= 0\,. \end{aligned}$$

Problem 6.9. Using the general properties of determinants, compute the determinant

$$\begin{vmatrix} 1 & 77 & 0 & 1 \\ 2 & 0 & 5 & -1 \\ 0 & -77 & 0 & 1 \\ 3 & 154 & 0 & 0 \end{vmatrix}.$$

Problem 6.10. Find the cofactor matrix $\mathbf{C}$ and hence the inverse $\mathbf{A}^{-1}$ for the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{pmatrix}$$

using the method of minors. *[The determinant is* $\det(\mathbf{A}) = -12$.*]*

Answers

1.

$$\boldsymbol{y} = \begin{pmatrix} 4 \\ 40 \\ 0 \\ 10 \\ -2 \end{pmatrix}, \quad \boldsymbol{z} = \begin{pmatrix} 5 \\ 83 \\ 0 \\ 19 \\ -37 \end{pmatrix}.$$

2. $(1, -0.99, 2.31)^T$.

3.

$$\mathbf{CD} = \begin{pmatrix} 34 & 11 \\ -4 & -7 \end{pmatrix}, \qquad \mathbf{DC} = \begin{pmatrix} 33 & -3 & 32 \\ -3 & -11 & -17 \\ 0 & 4 & 5 \end{pmatrix}, \qquad (\mathbf{CD})^2 = \begin{pmatrix} 1112 & 297 \\ -108 & 5 \end{pmatrix}.$$

4.

$$\mathbf{A}^{-1} = \begin{pmatrix} 27 & -16 & 6 \\ 8 & -5 & 2 \\ -5 & 3 & -1 \end{pmatrix}.$$

5.

$$\mathbf{B}^{-1} = \begin{pmatrix} 1/3 & -2/3 \\ 1/6 & 1/6 \end{pmatrix}, \qquad \mathbf{C}^{-1} = \begin{pmatrix} 4 & -3 \\ 2/3 & 0 \end{pmatrix}.$$

6. $\det(\mathbf{D}) = 37$, $\det(\mathbf{E}) = -12$.

8. $a = \pm 6$.

9. -1540.

10. The cofactor matrix is

$$\mathbf{C} = \begin{pmatrix} 12 & -8 & 6 \\ -6 & 2 & -3 \\ 12 & -4 & 0 \end{pmatrix}, \quad \mathbf{A}^{-1} = -\frac{1}{12} \begin{pmatrix} 12 & -6 & 12 \\ -8 & 2 & -4 \\ 6 & -3 & 0 \end{pmatrix}.$$

Chapter 7

Eigenvalues and Eigenvectors

7.1 Introduction

In this chapter we study "eigenvalue" problems. These arise in many situations, for example: calculating the natural frequencies of oscillation of a vibrating system; finding principal axes of stress and strain; calculating oscillations of an electrical circuit; image processing; data mining (web search engines); *etc.*

Differential equations: solving arbitrary order linear differential equations analytically;

Vibration analysis: calculating the natural frequencies of oscillation of a vibrating system – bridges, cantilevers;

Principal axes of stress and strain: mechanics;

Dynamic stability: linear stability analysis;

Column buckling: lateral deflections – modes of buckling;

Electrical circuit: oscillations, resonance;

Principle component analysis: extracting the salient features of a mass of data;

Markov Chains: transition matrices;

Data mining: web search engines – analysis of fixed point problems;

Image processing: fixed point problems again;

Quantum mechanics: quantised energy levels.

Suppose that $\mathbf{A}$ is a square $n \times n$ matrix then we can ask if there are any non-zero vectors such that $\mathbf{A}$ just stretches when it acts on them:

$$\mathbf{A}\boldsymbol{x} = \lambda \boldsymbol{x}$$

where λ is a number. For $\lambda = 0$ the corresponding vectors are said to belong to the null space of $\mathbf{A}$. For $\lambda = 1$ the corresponding vector is called a fixed point because it does not change under the action of $\mathbf{A}$.

Example 7.1. Let

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix} \qquad \boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

then

$$\mathbf{A}\boldsymbol{x} = \begin{pmatrix} 4 \\ 4 \\ 8 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = 4\boldsymbol{x}\,.$$

If

$$\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}\,.$$

for some vector $\boldsymbol{x}$ then such a number λ is called an **eigenvalue** of $\mathbf{A}$, and the corresponding $\boldsymbol{x}$ is called an **eigenvector**.

A matrix can have more than one eigenvector and eigenvalue. In the previous example the matrix $\mathbf{A}$ was shown to have an eigenvalue $\lambda = 4$ with eigenvector

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

One can check that $\lambda = -2$ is also an eigenvalue of $\mathbf{A}$ with corresponding eigenvectors

$$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

that is, $\mathbf{A}\boldsymbol{x}_1 = -2\boldsymbol{x}_1$ and $\mathbf{A}\boldsymbol{x}_2 = -2\boldsymbol{x}_2$. This example also shows that there can be more than one eigenvector corresponding to the same eigenvalue. In fact, for the above example, one can check that any linear combination

$$\boldsymbol{y} = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2$$

where c_1, c_2 are constants (not both equal to zero), is also an egenvector of $\mathbf{A}$ with eigenvalue -2.

The problem of finding the eigenvalues of an $n \times n$ matrix $\mathbf{A}$ is equivalent to finding for what values of λ the equation $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ has a non-trivial solution. This system of equations can be rewritten as

$$(\mathbf{A} - \lambda\mathbf{I})\boldsymbol{x} = \mathbf{0}\,,$$

where $\mathbf{I}$ is the $n \times n$ identity matrix. For this homogeneous system to have a non-trivial solution gives a condition on the values of λ.

Question 7.1. Find the eigenvalues of matrix

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

Solution. Consider the matrix

$$\mathbf{A} - \lambda I = \begin{pmatrix} 1-\lambda & 2 \\ 1 & -\lambda \end{pmatrix}.$$

Down-sweeping in the first column we obtain

$$\begin{pmatrix} 1-\lambda & 2 \\ 0 & -\lambda - \frac{2}{1-\lambda} \end{pmatrix}.$$

This matrix has a non-trivial null space if the lower right entry is zero, that is, if

$$\lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1) = 0\,.$$

Therefore $\mathbf{A}$ has exactly 2 eigenvalues, $\lambda = 2$ and $\lambda = -1$.

The presence of a variable λ in the matrix made Gaussian elimination messier than usual. (Even worse, our computation has been incomplete as we tacitly assumed that $\lambda - 1 \neq 0$. This assumption effectively excluded the case $\lambda = 1$ from our computations and one needs to check separately that 1 is not an eigenvalue of $\mathbf{A}$.)

A more straightforward way to compute eigenvalues is by using determinants.

By Theorem 6.1, λ is an eigenvalue of matrix $\mathbf{A}$ if and only if

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0\,. \tag{7.1}$$

The expression $\chi_A(\lambda) := \det(\mathbf{A} - \lambda I)$ is a polynomial of degree n called the **characteristic polynomial** of $\mathbf{A}$. Equation (7.1) is called the **characteristic equation** of $\mathbf{A}$. Thus the eigenvalues of $\mathbf{A}$ coincide with the roots of its characteristic polynomial.

Question 7.2. Find the eigenvalues of matrix

$$\mathbf{C} = \begin{pmatrix} 8 & -12 & 5 \\ 15 & -25 & 11 \\ 24 & -42 & 19 \end{pmatrix}.$$

Solution. The characteristic equation reads

$$\begin{aligned} 0 &= \chi_C(\lambda) = \det(\mathbf{C} - \lambda\mathbf{I}) = \det \begin{pmatrix} 8-\lambda & -12 & 5 \\ 15 & -25-\lambda & 11 \\ 24 & -42 & 19-\lambda \end{pmatrix} \\ &= (8-\lambda)[(-25-\lambda)(19-\lambda) - (-42)(11)] - (-12)[15(19-\lambda) - 24 \times 11] \\ &\quad +5[15 \times (-42) - 24(-25-\lambda)] = -\lambda^3 + 2\lambda^2 + \lambda - 2\,. \end{aligned}$$

The polynomial factorises to give

$$(\lambda - 1)(\lambda + 1)(\lambda - 2) = 0$$

so that the eigenvalues are

$$\lambda_1 = 1\,, \qquad \lambda_2 = -1\,, \qquad \lambda_3 = 2\,.$$

To find the eigenvalues and the corresponding eigenvectors of an $n \times n$ matrix $\mathbf{A}$ we proceed as follows.

1. Find the characteristic polynomial $\chi_A(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$.
2. Find the distinct roots $\lambda_1, \lambda_2, \ldots, \lambda_k$ of the characteristic equation $\chi_A(\lambda) = 0$. These are the eigenvalues of $\mathbf{A}$.
3. For each root λ_i, $i = 1, 2, \ldots, k$, solve the homogeneous equation $(\mathbf{A} - \lambda_i\mathbf{I})\boldsymbol{x} = \mathbf{0}$ for

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

The solutions are the corresponding eigenvectors.

Question 7.3. Find the eigenvalues and eigenvectors of

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Solution. (Step 1) The characteristic polynomial is

$$\chi_A(\lambda) = \det \begin{pmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{pmatrix} = -\lambda(\lambda - 1)(\lambda - 2)\,.$$

(Step 2) The above polynomial has roots

$$\lambda_1 = 0\,, \qquad \lambda_2 = 1\,, \qquad \lambda_3 = 2\,.$$

These are the eigenvalues of $\mathbf{A}$.
(Step 3) For $\lambda_1 = 0$ the homogeneous equation $(\mathbf{A} - \lambda_1\mathbf{I})\boldsymbol{x} = \mathbf{0}$ reads

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

Solving this system gives a corresponding eigenvector (or any multiple thereof)

$$\boldsymbol{x}_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

<u>For $\lambda_2 = 1$</u> we have to solve the system $(\mathbf{A} - \mathbf{I})\boldsymbol{x} = \mathbf{0}$ which reads

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

Solving this gives an eigenvector (or any scalar multiple thereof) corresponding to λ_2:

$$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

<u>For $\lambda_3 = 2$</u> we have to solve the system $(\mathbf{A} - 2\mathbf{I})\boldsymbol{x} = \mathbf{0}$ which reads

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Solving this gives an eigenvector (or any scalar multiple thereof) corresponding to λ_3:

$$\boldsymbol{x}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

7.2 Algebraic and Geometric Multiplicity of Eigenvalues

Note that in Question 7.3, for example, each eigenvalue is a simple root of the characteristic equation; in other words each eigenvalue only appears "once". On the other hand, in Example 7.1, the characteristic equation for $\mathbf{A}$ is

$$(4 - \lambda)(2 + \lambda)^2 = 0$$

so that $\lambda_1 = 4$ is a single root, and appears once as a solution of the equation, whereas $\lambda_2 = -2$ is a double root, and appears twice as a solution of the equation. We say that $\lambda_1 = 4$ is an eigenvalue of **algebraic multiplicity** 1 and that $\lambda_2 = -2$ is an eigenvalue of algebraic multiplicity 2. More generally, if the characteristic polynomial for a matrix $\mathbf{A}$ has a factor $(\lambda_i - \lambda)^j$, λ_i will be an eigenvalue of algebraic multiplicity j.

The number of distinct (independent) eignevectors corresponding to an eigenvalue is termed its **geometric multiplicity**. The geoemetric multiplicity of an eigenvalue must be at least 1. In Example 7.1 we saw that there were 2 independent eigenvectors corresponding to the eigenvalue of algebraic multiplicity 2. In general an eigenvalue's geoemetric multiplicity is no greater than its algebraic multiplicity:

$$1 \leq \text{geoemetric multiplicity} \leq \text{algebraic multiplicity}. \tag{7.2}$$

Question 7.4. Find the eigenvalues and their algebraic and geoemetric multiplicities for

$$\mathbf{A} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}, \ \mathbf{B} = \begin{pmatrix} 5 & 4 & -2 \\ -2 & 11 & -1 \\ 4 & -4 & 11 \end{pmatrix}, \ \text{and } \mathbf{C} = \begin{pmatrix} 9 & 3 & 6 \\ -6 & 3 & 0 \\ -3 & 0 & 15 \end{pmatrix}.$$

Solution.

The matrix A. The characteristic equation is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \det \begin{pmatrix} \lambda - 9 & 0 & 0 \\ 0 & \lambda - 9 & 0 \\ 0 & 0 & \lambda - 9 \end{pmatrix} = (\lambda - 9)^3 = 0\,.$$

The above polynomial has the triple root

$$\lambda = 9$$

so that $\mathbf{A}$ has only one eigenvalue, $\lambda = 9$, of algebraic multiplicity three.
To now find the eigenvectors, we solve the homogeneous equation $(\mathbf{A} - \lambda \mathbf{I})\boldsymbol{x} = \mathbf{0}$, so

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Attempting to solve this system leaves x, y and z undetermined; the general eigenvector is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

There are three independent eignevectors :

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

or, equivalently, the general eigenvector can be written in terms of three arbitrary constants: $\lambda = 9$ is an eigenvalue of geometric multiplicity three.

The matrix B. The characteristic equation is

$$\det(\lambda\mathbf{I}-\mathbf{B}) = \det\begin{pmatrix} \lambda-5 & -4 & 2 \\ 2 & \lambda-11 & 1 \\ -4 & 4 & \lambda-11 \end{pmatrix}$$

$$= (\lambda-5)(\lambda^2-22\lambda+121)+16+16-4(\lambda-5)+8(\lambda-11)+8(\lambda-11)$$
$$= \lambda^3-27\lambda^2+231\lambda-605+32+12\lambda+20-176 = \lambda^3-27\lambda^2+243\lambda-729$$
$$= \lambda^3-3\times 9\lambda^2+3\times 9^2\lambda-9^3 = (\lambda-9)^3 = 0\,.$$

Again there is the triple root

$$\lambda = 9$$

so that $\mathbf{B}$ also has only one eigenvalue, $\lambda = 9$, of algebraic multiplicity three.
To now find the eigenvectors, we solve the homogeneous equation $(\mathbf{B}-\lambda\mathbf{I})\boldsymbol{x} = \mathbf{0}$, so

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

Subtracting twice the second row from the first and adding twice the second row to the third gives

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

The second equation then gives $-2x+2y-z = 0$ so $z = 2y-2x$ with x and y undetermined; the general eigenvector is

$$\begin{pmatrix} x \\ y \\ 2y-2x \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

There are two independent eignevectors :

$$\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix},$$

or, equivalently, the general eigenvector can be written in terms of two arbitrary constants: $\lambda = 9$ is an eigenvalue of geometric multiplicity two.

The matrix C. The characteristic equation is

$$\det(\lambda\mathbf{I}-\mathbf{C}) = \det\begin{pmatrix} \lambda-9 & -3 & -6 \\ 6 & \lambda-3 & 0 \\ 3 & 0 & \lambda-15 \end{pmatrix}$$

$$= (\lambda-9)(\lambda^2-18\lambda+45)+18(\lambda-15)+18(\lambda-3) = (\lambda-9)(\lambda^2-18\lambda+45)+36(\lambda-9)$$
$$= (\lambda-9)(\lambda^2-18\lambda+81) = (\lambda-9)^3 = 0\,.$$

Once again there is the triple root

$$\lambda = 9$$

so that $\mathbf{C}$ also has only one eigenvalue, $\lambda = 9$, of algebraic multiplicity three.
To now find the eigenvectors, we solve the homogeneous equation $(\mathbf{C} - \lambda\mathbf{I})\boldsymbol{x} = \mathbf{0}$, so

$$\begin{pmatrix} 0 & 3 & 6 \\ -6 & -6 & 0 \\ -3 & 0 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

Dividing the first and third rows by three and the second by minus six recasts the problem as

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

then adding the third row to the second gives

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

and subtracting the second row from the first leads to

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

The second equation then gives $y = -2z$ while the third gives $x = 2z$; the general eigenvector is

$$\begin{pmatrix} 2z \\ -2z \\ z \end{pmatrix} = z \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} .$$

There is just one (independent) eignevector :

$$\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} ,$$

or, equivalently, the general eigenvector can be written in terms of one arbitrary constant: $\lambda = 9$ is an eigenvalue of geometric multiplicity one.

Note from this question that for an eigenvalue λ of an $n \times n$ matrix $\mathbf{A}$,

$$\text{geometric multiplicity} = \text{number of independent eigenvectors}$$

$$= \text{number of parameters in the expression for the general eigenvector}$$

$$= n - \text{rank of } (\mathbf{A} - \lambda\mathbf{I})\,.$$

From the form of the characteristic equation, it can also be observed, again for an $n \times n$ matrix $\mathbf{A}$, that n equals the sum of the algebraic multiplicities of the eigenvalues of $\mathbf{A}$.

7.3 Practical Application: Mass-Spring Systems

Question 7.5. Two identical simple pendula oscillate in the plane as shown in Figure 7.1. Both pendula consist of light rods of length $\ell = 10$ and are suspended from the same ceiling a distance $L = 15$ apart, with equal masses $m = 1$ attached to their ends. The angles the pendula make to the downward vertical are θ_1 and θ_2, and they are coupled through the spring shown which has stiffness coefficient $k = 1$. The spring has unstretched length $L = 15$. Assume that the acceleration due to gravity $g = 10$. Describe the system's dynamics by differential equations and find their general solutions.

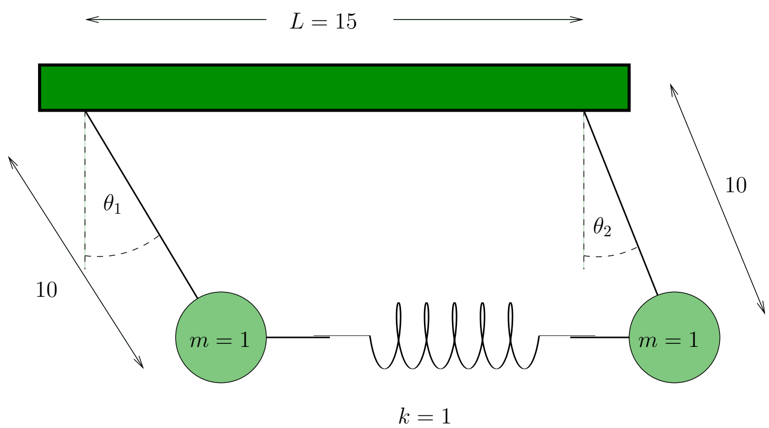


Figure 7.1: Simple coupled pendula system.

Solution. Assuming that the oscillations of the spring remain small in amplitude, so that $|\theta_1| \ll 1$ and $|\theta_2| \ll 1$, by applying Newton's second law and Hooke's law one finds that the coupled pendula system gives rise to the system of differential equations

$$\frac{\mathrm{d}^2\boldsymbol{\theta}}{\mathrm{d}t^2} = \mathbf{A}\boldsymbol{\theta}, \quad \text{where} \quad \mathbf{A} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}, \tag{7.3}$$

and

$$\boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$$

is the vector of unknown angles for each of the pendula shown in Figure 7.1.

We further look for a solution of the form $\boldsymbol{\theta}(t) = \mathrm{Re}\{\boldsymbol{v}\mathrm{e}^{\mathrm{i}\omega t}\}$ for a constant vector $\boldsymbol{v}$. Substituting $\boldsymbol{\theta}(t)$ into (7.3) and dividing both sides by $\mathrm{e}^{i\omega(t)}$ we obtain that the system of differential equations (7.3) reduces to solving the eigenvalue problem

$$(\mathbf{A} + \omega^2\mathbf{I})\boldsymbol{v} = \mathbf{0}\,. \tag{7.4}$$

We obtain using the characteristic equation method explained in the previous section that matrix $\mathbf{A}$ has two distinct eigenvalues $\lambda_1 = -1$ and $\lambda = -3$. Since $\lambda = -\omega^2$ we obtain the characteristic frequencies $\omega_1 = 1$ and $\omega_2 = \sqrt{3}$ (it is usual to take frequencies to be positive, as was done in F18XC). An eigenvector corresponding to $\lambda_1 = -1$ is

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and for $\lambda_2 = -3$ an eigenvector is

$$\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Taking the real combinations of the complex exponents:

$$\cos(\omega t) = \frac{1}{2}(\mathrm{e}^{i\omega t} + \mathrm{e}^{-i\omega t}), \qquad \sin(\omega t) = \frac{1}{2i}(\mathrm{e}^{i\omega t} - \mathrm{e}^{-i\omega t}),$$

we obtain a general solution in the form

$$\begin{pmatrix} \theta_1(t) \\ \theta_2(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} (C_1 \cos(t) + C_2 \sin(t)) + \begin{pmatrix} 1 \\ -1 \end{pmatrix} (C_3 \cos(\sqrt{3}t) + C_3 \sin(\sqrt{3}t))$$

where C_i, $i = 1, 2, 3, 4$ are arbitrary constants which can be fixed by initial conditions.

7.4 Diagonalisation

An $n \times n$ matrix $\tilde{\mathbf{A}}$ is said to be **similar** to an $n \times n$ matrix $\mathbf{A}$ if

$$\tilde{\mathbf{A}} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$$

for some (non-singular) $n \times n$ matrix $\mathbf{X}$. This transformation, which gives $\tilde{\mathbf{A}}$ from $\mathbf{A}$, is called a **similarity transformation**.

Diagonalisation transformation

If an $n \times n$ matrix $\mathbf{A}$ has a set of n linearly independent eigenvectors, $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$, ..., $\boldsymbol{x}^{(n)}$, then if we set $X = [\boldsymbol{x}^{(1)}|\boldsymbol{x}^{(2)}|\ldots|\boldsymbol{x}^{(n)}]$ (*i.e.* the matrix whose columns are the eigenvectors of $\mathbf{A}$), we have

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \mathbf{D},$$

where $\mathbf{D}$ is the $n \times n$ diagonal matrix

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix},$$

i.e. the matrix whose diagonal entries are the eigenvalues of the matrix $\mathbf{A}$ and whose all other entries are zero. Note that we have

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1}.$$

(Also note that if our matrix has some eigenvalue whose algebraic multiplicity is **greater** than its geometric multiplicity, the total number of independent eigenvectors will be less than n and $\mathbf{A}$ will not be diagonalisable.)

Symmetric Matrices

A real square matrix $\mathbf{A}$ is said to be **symmetric** if transposition leaves it unchanged, *i.e.* $\mathbf{A}^T = \mathbf{A}$.

Theorem 7.1. *If $\mathbf{A}$ is a real symmetric $n \times n$ matrix, then its eigenvalues, $\lambda_1, \ldots, \lambda_n$, are all real.*
The matrix has corresponding eigenvectors $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$, ..., $\boldsymbol{x}^{(n)}$, all linearly independent (even if there aren't n distinct eigenvalues, so that $\lambda_i = \lambda_j$ for some i and j).
Moreover, any two eigenvectors $\boldsymbol{x}^{(i)}$, $\boldsymbol{x}^{(j)}$ correspondinging to two distinct eigenvalues $\lambda_i \neq \lambda_j$ are orthogonal to each other: $\boldsymbol{x}^{(i)} \cdot \boldsymbol{x}^{(j)} = 0$.
In particular, all real symmetric matrices are diagonalisable.

Example 7.2. The matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

is symmetric. The eigenvalues of $\mathbf{A}$ are 1, 3 and 6 – all real. The corresponding eigenvectors are

$$\boldsymbol{x}^{(1)} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{x}^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{x}^{(3)} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

One can easily check that these vectors are mutually orthogonal. For example,

$$\boldsymbol{x}^{(1)} \cdot \boldsymbol{x}^{(2)} = x_1^{(1)} x_1^{(2)} + x_2^{(1)} x_2^{(2)} + x_3^{(1)} x_3^{(2)} = (-2) \times 0 + 1 \times 0 + 0 \times 1 = 0\,.$$

Note that as usual the eigenvectors are defined up to rescaling. We can use this freedom to pick eigenvectors all to have length one (all to be unit vectors), so they satisfy $\boldsymbol{x}^{(i)} \cdot \boldsymbol{x}^{(i)} = 1$. To do that we divide each eigenvector by its length

$$\boldsymbol{x}^{(i)} \to \frac{1}{\sqrt{\boldsymbol{x}^{(i)} \cdot \boldsymbol{x}^{(i)}}}\, \boldsymbol{x}^{(i)}\,.$$

Thus the properly normalised eigenvectors satisfy

$$\boldsymbol{x}^{(i)} \cdot \boldsymbol{x}^{(j)} = \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } i = j \end{cases}.$$

These relations imply that for the matrix $\mathbf{X} = [\boldsymbol{x}^{(1)}|\boldsymbol{x}^{(2)}|\ldots|\boldsymbol{x}^{(m)}]$ whose columns are the eigenvectors of $\mathbf{A}$, one has

$$\mathbf{X}^{-1} = \mathbf{X}^T\,.$$

In general matrices for which the inverse coincides with the transposed matrix are called **orthogonal matrices**.

Question 7.6. Find a matrix $\mathbf{X}$ which diagonalises the matrix $\mathbf{A}$ from the previous example *via* a similarity transformation.

Solution. Normalised eigenvectors with length one are

$$\boldsymbol{x}^{(1)} = \frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{x}^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{x}^{(3)} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Hence

$$\mathbf{X} = \begin{pmatrix} -\frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix}.$$

Now we can check that

$$\mathbf{X}^{-1} = \mathbf{X}^T = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \end{pmatrix},$$

and furthermore

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \end{pmatrix}\begin{pmatrix} 2 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} -\frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

Diagonalisation can be used to obtain large powers of a given matrix. Note that

$$\mathbf{A}^n = (\mathbf{X}\mathbf{D}\mathbf{X}^{-1})(\mathbf{X}\mathbf{D}\mathbf{X}^{-1})\ldots(\mathbf{X}\mathbf{D}\mathbf{X}^{-1}) = \mathbf{X}\mathbf{D}^n\mathbf{X}^{-1}$$

and it is straightforward to take the n-th power of a diagonal matrix just by taking the n-th power of each diagonal entry.

Question 7.7. Find $\mathbf{A}^n$ for the matrix from the previous example.

Solution. We have

$$\begin{aligned}\mathbf{A}^n = \mathbf{X}\mathbf{D}^n\mathbf{X}^{-1} &= \begin{pmatrix} -\frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3^n & 0 \\ 0 & 0 & 6^n \end{pmatrix}\begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{5}(4+6^n) & \frac{2}{5}(-1+6^n) & 0 \\ \frac{2}{5}6^n & \frac{4}{5}6^n & 0 \\ 0 & 0 & 3^n \end{pmatrix}.\end{aligned}$$

The need to evaluate matrices to arbitrary powers often arises in connection with differential equations.

7.5 Systems of Linear Differential Equations

We have already seen how one (second-order) coupled system could be tackled in Question 7.5. Other oscillatory problems like this proceed as in the folowing question.

Question 7.8. Find the general solution to

$$\frac{\mathrm{d}^2\boldsymbol{x}}{\mathrm{d}t^2} + \mathbf{B}\boldsymbol{x} = \mathbf{0}\,, \tag{7.5}$$

where $\boldsymbol{x}$ is a two-dimensional vector and $\mathbf{B} = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$.

Solution. First look for a special solution in complex form, $\boldsymbol{x} = \boldsymbol{v}\mathrm{e}^{\mathrm{i}\omega t}$ with $\boldsymbol{v}$ a constant vector.
Differentiating $\boldsymbol{x}$ twice, we get $\mathrm{d}^2\boldsymbol{x}/\mathrm{d}t^2 = -\omega^2\boldsymbol{v}\mathrm{e}^{\mathrm{i}\omega t}$. Then substituting into (7.5) gives

$$-\omega^2\boldsymbol{v}\mathrm{e}^{\mathrm{i}\omega t} + \mathbf{B}\boldsymbol{v}\mathrm{e}^{\mathrm{i}\omega t} = \mathbf{0}$$

$$\text{so} \qquad \mathbf{B}\boldsymbol{v} = \omega^2\boldsymbol{v}\,.$$

Hence $\lambda = \omega^2$ is an eigenvalue of $\mathbf{B}$ with $\boldsymbol{v}$ a corresponding eigenvector.

The characteristic equation here gives

$$0 = \det(\lambda\mathbf{I} - \mathbf{B}) = \begin{vmatrix} \lambda - 3 & 1 \\ 1 & \lambda - 3 \end{vmatrix} = (\lambda-3)^2 - 1 = (\lambda-3-1)(\lambda-3+1) = (\lambda-4)(\lambda-2)$$

so $\lambda = \lambda_1 = 4$ with $\omega = 2$, or $\lambda = \lambda_2 = 2$ with $\omega = \sqrt{2}$ (taking ω positive).

For $\boldsymbol{v}_1 = (a,b)$ an eigenvalue corresponding to $\lambda_1 = 4$,

$$(\mathbf{B} - \lambda_1\mathbf{I})\boldsymbol{v}_1 = \left(\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} - \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}\right)\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -a-b \\ -a-b \end{pmatrix} = \mathbf{0}$$

so $a = -b$ and an eigenvector is $\boldsymbol{v}_1 = (1,-1)$.

For $\boldsymbol{v}_2 = (a,b)$ an eigenvalue corresponding to $\lambda_2 = 2$,

$$(\mathbf{B} - \lambda_2\mathbf{I})\boldsymbol{v}_12 = \left(\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\right)\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a-b \\ b-a \end{pmatrix} = \mathbf{0}$$

so $a = b$ and an eigenvector is $\boldsymbol{v}_2 = (1,1)$.
These then give special solutions in complex form $\boldsymbol{v}_1\mathrm{e}^{\mathrm{i}\omega_1 t} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}\mathrm{e}^{2\mathrm{i}t}$ and $\boldsymbol{v}_2\mathrm{e}^{\mathrm{i}\omega_2 t} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\mathrm{e}^{\sqrt{2}\,\mathrm{i}t}$.
To get real solutions, we take real and imaginary parts (remembering that $\mathrm{Re}\{\mathrm{e}^{\mathrm{i}\theta}\} = \cos\theta$ and $\mathrm{Im}\{\mathrm{e}^{\mathrm{i}\theta}\} = \sin\theta$.) We then have independent solutions

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}\cos 2t \;\text{ and }\; \begin{pmatrix} 1 \\ -1 \end{pmatrix}\sin 2t$$

and also

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cos(\sqrt{2}\,t) \text{ and } \begin{pmatrix} 1 \\ 1 \end{pmatrix} \sin(\sqrt{2}\,t)\,.$$

(The first pair of solutions, coming from $\lambda_1 = 4$, has $x_1 = x_2$ – the components move together – and such motion is termed **in phase**. The second pair of solutions, coming from $\lambda_1 = 2$, has $x_1 = -x_2$ – the components move opposite to each other – and such motion is termed **out of phase**.)

The general solution is then got by combining all four solutions:

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} (a_1 \cos 2t + a_2 \sin 2t) + \begin{pmatrix} 1 \\ 1 \end{pmatrix} (b_1 \cos(\sqrt{2}\,t) + b_2 \sin(\sqrt{2}\,t))\,.$$

More generally, we can get systems of first-order equations of the form

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \mathbf{A}\boldsymbol{x} \tag{7.6}$$

with $\boldsymbol{x} = (x_1(t), x_2(t), \ldots, \boldsymbol{x}_n(t))$ and $\mathbf{A}$ a constant $n \times n$ matrix. (Inhomogeneous versions, with some sort of forcing terms, are also possible but are not considered in this course.)

For simplicity we shall only consider case where there are n linearly independent eigenvectors for the $n \times n$ matrix $\mathbf{A}$.

Example 7.3. Let's first try solving the vector differential equation

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \mathbf{A}\boldsymbol{x}$$

with $\mathbf{A}$ as in Question 7.3.

We start by looking for a simple solution of the form $\boldsymbol{x} = \boldsymbol{v}\mathrm{e}^{\lambda t}$ with $\boldsymbol{v}$ a constant non-zero vector. Differentiating $\boldsymbol{x}$ and substituting into the original equation, we get

$$\lambda \boldsymbol{v}\mathrm{e}^{\lambda t} = \mathbf{A}\boldsymbol{v}\mathrm{e}^{\lambda t} \qquad \text{so} \qquad \mathbf{A}\boldsymbol{v} = \lambda \boldsymbol{v}\,.$$

Hence λ is an eigenvalue with $\boldsymbol{v}$ a corresponding eigenvector. From Question 7.3, we have:

$$\lambda_1 = 0\,,\ \boldsymbol{v}_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \ ; \quad \lambda_2 = 1\,,\ \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \ ; \quad \lambda_3 = 2\,,\ \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\,.$$

There are then three independent special solutions:

$$\boldsymbol{v}_1\mathrm{e}^{\lambda_1 t} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \ ; \quad \boldsymbol{v}_2\mathrm{e}^{\lambda_2 t} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mathrm{e}^{t}\,; \quad \boldsymbol{v}_3\mathrm{e}^{\lambda_3 t} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \mathrm{e}^{2t}\,.$$

The general solution is given by taking a linear combination:

$$\boldsymbol{x} = a_1\boldsymbol{v}_1\mathrm{e}^{\lambda_1 t} + 2\boldsymbol{v}_2\mathrm{e}^{\lambda_2 t} + \boldsymbol{v}_3\mathrm{e}^{\lambda_3 t} = a_1 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mathrm{e}^{t} + a_3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \mathrm{e}^{2t}\,.$$

Alternatively, the individual components are $x_1 = a_2\mathrm{e}^{t}$, $x_2 = a_1 + a_3\mathrm{e}^{2t}$ and $x_3 = a_3\mathrm{e}^{2t} - a_1$.

Question 7.9. Suppose that $\boldsymbol{x}(t) = (x_1, x_2, x_3)^{\mathrm{T}}$ satisfies $\mathrm{d}\boldsymbol{x}/\mathrm{d}t = \mathbf{A}\boldsymbol{x}$ with

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

Find the general solution to the system of ODEs.

Solution. Looking for a special solution of the form $\boldsymbol{x} = \boldsymbol{v}\mathrm{e}^{\lambda t}$ gives, on substituting this into the equation, $\lambda\boldsymbol{v} = \mathbf{A}\boldsymbol{v}$, *i.e.* λ is an eigenvalue of $\mathbf{A}$ and $\boldsymbol{v}$ is an associated eigenvector.
We can find that the eigenvalues of $\mathbf{A}$ are $\lambda_1 = 5$ and $\lambda_2 = -3$, and that corresponding eigenvectors are : $\boldsymbol{v}_1 = (-1, -2, 1)$ for $\lambda = 5$; and $\boldsymbol{v}_2 = (3, 0, 1)^{\mathrm{T}}$ and $\boldsymbol{v}_3 = (-2, 1, 0)^{\mathrm{T}}$ for $\lambda = -3$.
Independent solutions to the system are then

$$\boldsymbol{v}_1\mathrm{e}^{5t}, \qquad \boldsymbol{v}_2\mathrm{e}^{-3t} \quad \text{and} \quad \boldsymbol{v}_3\mathrm{e}^{-3t},$$

so the general solution, in vector form, is

$$\boldsymbol{x} = \alpha\begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}\mathrm{e}^{5t} + \beta\begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}\mathrm{e}^{-3t} + \gamma\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}\mathrm{e}^{-3t},$$

for arbitrary constants α, β and γ.
Looking at the individual components,

$$\begin{array}{rcrrr} x_1 & = & -\alpha\mathrm{e}^{5t} & +3\beta\mathrm{e}^{-3t} & -2\gamma\mathrm{e}^{-3t}, \\ x_2 & = & -2\alpha\mathrm{e}^{5t} & & +\gamma\mathrm{e}^{-3t}, \\ x_3 & = & \alpha\mathrm{e}^{5t} & +\beta\mathrm{e}^{-3t}. & \end{array}$$

To determine the constants for an initial value problem, a system of linear equations has to be solved (as in Chap. 5).

Example 7.4. Suppose that at $t = 0$, $\boldsymbol{x}$, as in Qu. 7.9, is given by $x_1 = 3$, $x_2 = -5$, $x_3 = 3$. The constants α, β and γ must then satisfy

$$-\alpha + 3\beta - 2\gamma = 3\,, \quad -2\alpha + \gamma = -5\,, \quad \alpha + \beta = 3\,.$$

Solving these by Gaussian elimination gives $\alpha = 2$, $\beta = 1$ and $\gamma = -1$. Then

$$x_1 = -2\mathrm{e}^{5t} + 3\mathrm{e}^{-3t} + 2\mathrm{e}^{-3t}, \quad x_2 = -4\mathrm{e}^{5t} - \mathrm{e}^{-3t}, \quad \text{and} \quad x_3 = 2\mathrm{e}^{5t} + \mathrm{e}^{-3t}.$$

Systems of linear first-order ordinary differential equations with constant coefficients are normally equivalent to a single ODE: an $n \times n$ system can be usually be written as an nth-order ODE, and an nth-order ODE can always be written as an $n \times n$ system.

Example 7.5. Let us try to solve

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + 2\frac{\mathrm{d}y}{\mathrm{d}t} - 3y = 0\,.$$

We can first write $x_1 = y$ and $x_2 = \mathrm{d}y/\mathrm{d}t$. Then

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = x_2 \quad \text{and} \quad \frac{\mathrm{d}x_2}{\mathrm{d}t} = \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = 3y - 2\frac{\mathrm{d}y}{\mathrm{d}t} = 3x_1 - 2x_2\,.$$

In matrix form the system is then

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

The characteristic equation for the matrix is

$$0 = \begin{vmatrix} \lambda & -1 \\ -3 & \lambda + 2 \end{vmatrix} = \lambda(\lambda+2) - 3 = \lambda^2 + 2\lambda - 3 = (\lambda+3)(\lambda-1)$$

so the eigenvalues are $\lambda_1 = -3$ and $\lambda_2 = 1$.
An eigenvector corresponding to the first is then given by

$$\begin{pmatrix} -3 & -1 \\ -3 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -3a - b \\ -3a - b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

so an eigenvector will be $\boldsymbol{v}_1 = (1, -3)^{\mathrm{T}}$.
An eigenvector corresponding to the second is given by

$$\begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a - b \\ -3a + 3b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

so an eigenvector will be $\boldsymbol{v}_2 = (1, 1)^{\mathrm{T}}$.
The general solution in vector form can be written as

$$\boldsymbol{x} = \alpha\begin{pmatrix} 1 \\ -3 \end{pmatrix}\mathrm{e}^{-3t} + \beta\begin{pmatrix} 1 \\ 1 \end{pmatrix}\mathrm{e}^{t}$$

so the general solution for the original problem is

$$y = x_1 = \alpha\mathrm{e}^{-3t} + \beta\mathrm{e}^{t}.$$

Question 7.10. Use eigenvalues and eigenvectors to solve

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 6\frac{\mathrm{d}y}{\mathrm{d}t} + 13y = 0\,, \quad y(0) = -1\,,\ \frac{\mathrm{d}y}{\mathrm{d}t}(0) = 11\,.$$

Solution. Write $x_1 = y$ and $x_2 = \mathrm{d}y/\mathrm{d}t$. Then

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = x_2 \quad \text{and} \quad \frac{\mathrm{d}x_2}{\mathrm{d}t} = \frac{\mathrm{d}^2y}{\mathrm{d}t^2} = -13y - 6\frac{\mathrm{d}y}{\mathrm{d}t} = -13x_1 - 6x_2\,.$$

In matrix form the system is then

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -13 & -6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

The characteristic equation for the matrix is

$$0 = \begin{vmatrix} \lambda & -1 \\ 13 & \lambda+6 \end{vmatrix} = \lambda(\lambda+6) + 13 = \lambda^2 + 6\lambda + 13\,.$$

Therefore the eigenvalues are $\lambda = \frac{1}{2}(-6 \pm \sqrt{36-52}\,) = \frac{1}{2}(-6 \pm \sqrt{16}\,\mathrm{i}) = -3 \pm 2\mathrm{i}$.
The corresponding eigenvectors are given by

$$\begin{pmatrix} -3\pm 2\mathrm{i} & -1 \\ 13 & 3\pm 2\mathrm{i} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} (-3\pm 2\mathrm{i})a - b \\ 13a + (3\pm 2\mathrm{i})b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

so the eigenvectors can be taken to be $(1, -3 \mp 2\mathrm{i})^{\mathrm{T}}$.
The general solution in vector form can be written as

$$\boldsymbol{x} = \alpha\begin{pmatrix} 1 \\ -3-2\mathrm{i} \end{pmatrix}\mathrm{e}^{(-3+2\mathrm{i})t} + \beta\begin{pmatrix} 1 \\ -3+2\mathrm{i} \end{pmatrix}\mathrm{e}^{(-3-2\mathrm{i})t}$$

or equivalently

$$\boldsymbol{x} = \begin{pmatrix} \alpha\mathrm{e}^{(-3+2\mathrm{i})t} + \beta\mathrm{e}^{(-3-2\mathrm{i})t} \\ (-3-2\mathrm{i})\alpha\mathrm{e}^{(-3+2\mathrm{i})t} + (-3+2\mathrm{i})\beta\mathrm{e}^{(-3-2\mathrm{i})t} \end{pmatrix}.$$

Using $\mathrm{e}^{(-3\pm 2\mathrm{i})t} = \mathrm{e}^{3t}(\cos 2t \pm \mathrm{i}\sin 2t)$:

$$\begin{aligned} x_1 &= \mathrm{e}^{-3t}(\alpha(\cos 2t + \mathrm{i}\sin 2t) + \beta(\cos 2t - \mathrm{i}\sin 2t) \\ &= \mathrm{e}^{-3t}((\alpha+\beta)\cos 2t + \mathrm{i}(\alpha-\beta)\sin 2t) \\ &= \mathrm{e}^{-3t}(C_1\cos 2t + C_2\sin 2t)\,, \end{aligned}$$

on writing $C_1 = (\alpha+\beta)$ and $C_2 = \mathrm{i}(\alpha-\beta)$;

$$\begin{aligned} x_2 &= \mathrm{e}^{-3t}[(-3-2\mathrm{i})\alpha(\cos 2t + \mathrm{i}\sin 2t) + (-3+2\mathrm{i})\beta(\cos 2t - \mathrm{i}\sin 2t)] \\ &= \mathrm{e}^{-3t}[\alpha((-3\cos 2t + 2\sin 2t) - \mathrm{i}(3\sin 2t + 2\cos 2t)) + \beta((-3\cos 2t + 2\sin 2t) + \mathrm{i}(3\sin 2t + 2\cos 2t) \\ &= \mathrm{e}^{-3t}[(\alpha+\beta)(2\sin 2t - 3\cos 2t) - \mathrm{i}(\alpha-\beta)(3\sin 2t + 2\cos 2t)] \\ &= \mathrm{e}^{-3t}(C_1(2\sin 2t - 3\cos 2t) - C_2(3\sin 2t + 2\cos 2t))\,. \end{aligned}$$

Now

$$\boldsymbol{x} = \mathrm{e}^{-3t}\begin{pmatrix} C_1\cos 2t + C_2\sin 2t \\ C_1(2\sin 2t - 3\cos 2t) - C_2(3\sin 2t + 2\cos 2t) \end{pmatrix}$$

where we also know that $t = 0$, $y = x_1 = -1$ and $\mathrm{d}y/\mathrm{d}t = x_2 = 11$ so
$\begin{pmatrix} C_1 \\ -3C_1 + 2C_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 11 \end{pmatrix}$. Therefore $C_1 = -1$ and $C_2 = 4$. These give

$$y = x_1 = \mathrm{e}^{-3t}(4\sin 2t - \cos 2t)\,.$$

7.6 Problems

Problem 7.1. Find the eigenvalues and eigenvectors for the matrices

$$\mathbf{A} = \begin{pmatrix} 8 & 3 \\ -10 & -3 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 3 & 5 \\ 5 & 3 \end{pmatrix}.$$

Problem 7.2. Use the results of the previous question to find a general solution to the system of equations

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= 8x + 3y\,, \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= -10x - 3y\,. \end{aligned}$$

Problem 7.3. Compute the eigenvalues and eigenvectors of the following 3×3 matrices:

$$\mathbf{D} = \begin{pmatrix} 3 & 5 & 7 \\ 5 & 3 & -7 \\ 0 & 0 & 2 \end{pmatrix}, \qquad \mathbf{E} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

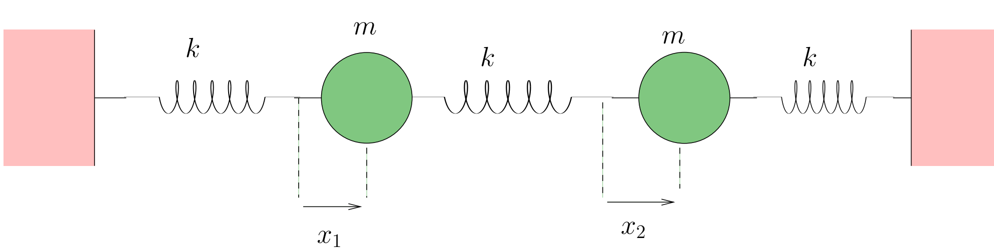
Problem 7.4. Find the general solution to the system of differential equations

$$\frac{\mathrm{d}^2\boldsymbol{x}}{\mathrm{d}t^2} = \mathbf{B}\boldsymbol{x}\,,$$

where

$$\boldsymbol{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} -2 & \sqrt{2} \\ \sqrt{2} & -3 \end{pmatrix}.$$

Problem 7.5. Consider two equal masses m connected by three springs with stiffness coefficients k as indicated in the diagram:



Newton's equations of motion describing this system are

$$\begin{aligned} \left(\frac{m}{k}\right)\frac{\mathrm{d}^2x_1}{\mathrm{d}t^2} &= -2x_1 + x_2\,, \\ \left(\frac{m}{k}\right)\frac{\mathrm{d}^2x_2}{\mathrm{d}t^2} &= x_1 - 2x_2\,, \end{aligned}$$

where x_1, x_2 are the displacements of the two bodies from their equilibrium positions. Assume that $m/k = 1$. Rewrite the above equations in matrix form. Find the natural vibration frequences. Write the general solution to the equations of motion.

Problem 7.6. In quantum physics the outcome of a measurement will correspond to the eigenvalues of an operator. The spins of a particle in the x, y and z directions are described by the three Pauli spin matrices

$$\mathbf{S}_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{S}_y = \frac{1}{2}\begin{pmatrix} 0 & \mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix}, \quad \mathbf{S}_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit. Calculate the values of any measurements of spins along the x, y and z directions, which are given by the corresponding eigenvalues of $\mathbf{S}_x$, $\mathbf{S}_y$, $\mathbf{S}_z$.

Problem 7.7. Determine the algebraic and geometric multiplicities of the eigenvalues for the matrix

$$\mathbf{E} = \begin{pmatrix} 2 & 75 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Problem 7.8. For the matrix

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

find the diagonalisation transformation.

Problem 7.9. For the matrix

$$\mathbf{C} = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

find the diagonalisation transformation.

Problem 7.10. For the matrix

$$\mathbf{D} = \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix},$$

calculate $\mathbf{D}^{54}$.

Answers

1. For $\mathbf{A}$: $\lambda_1 = 2$, $\boldsymbol{x}_1 = \alpha\begin{pmatrix} 1 \\ -2 \end{pmatrix}$; $\lambda_2 = 3$, $\boldsymbol{x}_2 = \alpha\begin{pmatrix} -3 \\ 5 \end{pmatrix}$.

For $\mathbf{B}$: $\lambda_1 = 8$, $\boldsymbol{x}_1 = \alpha\begin{pmatrix} 1 \\ 1 \end{pmatrix}$; $\lambda_2 = -2$, $\boldsymbol{x}_2 = \alpha\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

2.

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = C_1 \mathrm{e}^{2t}\begin{pmatrix} 1 \\ -2 \end{pmatrix} + C_2 \mathrm{e}^{3t}\begin{pmatrix} -3 \\ 5 \end{pmatrix}.$$

3. For $\mathbf{D}$: $\lambda_1 = -2$, $\boldsymbol{x}_1 = \alpha \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. $\lambda_2 = 2$, $\boldsymbol{x}_2 = \alpha \begin{pmatrix} 7 \\ -7 \\ 4 \end{pmatrix}$; $\lambda_3 = 8$, $\boldsymbol{x}_3 = \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

For $\mathbf{E}$: $\lambda_1 = 0$, $\boldsymbol{x}_1 = \alpha \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$; $\lambda_2 = 2$, $\boldsymbol{x}_2 = \alpha_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

4. $\boldsymbol{x}(t) = \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix} (C_1 \cos(t) + C_2 \sin(t)) + \begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix} (C_3 \cos(2t) + C_4 \sin(2t))$.

5. The natural frequences are $\omega_1 = 1$, $\omega_2 = \sqrt{3}$. The general solution is

$$(A_1 \cos(t) + A_2 \sin(t)) \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (A_3 \cos(\sqrt{3}\,t) + A_4 \sin(\sqrt{3}\,t)) \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

6. Each of the 3 matrices has eigenvalues 1/2 and −1/2..

7. $\lambda_1 = 2$ with algebraic multiplicity 3 and geometric multiplicity 2.

8. $\mathbf{X}^{-1}\mathbf{B}\mathbf{X} = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$ where $\mathbf{X} = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$, $\quad \mathbf{X}^{-1} = \frac{1}{7} \begin{pmatrix} 3 & 1 \\ -1 & 2 \end{pmatrix}$.

9.

$$\mathbf{X}^{-1}\mathbf{C}\mathbf{X} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

10.

$$\mathbf{D}^{54} = \frac{1}{2} \begin{pmatrix} (6^{54} + 4^{54}) & (6^{54} - 4^{54}) \\ (6^{54} - 4^{54}) & (6^{54} + 4^{54}) \end{pmatrix}.$$

Appendix A

Useful Formulæ for F1.8XD2

Standard Derivatives :

$F(x)$	$F'(x)$
x^n	nx^{n-1}
$\sin ax$	$a\cos ax$
$\cos ax$	$-a\sin ax$
$\tan ax$	$a\sec^2 ax$
e^{ax}	$a\mathrm{e}^{ax}$
$\ln x$	$1/x$
$\cosh ax$	$a\sinh ax$
$\sinh ax$	$a\cosh ax$
$F(ax+b)$	$aF'(ax+b)$
$u(x)v(x)$	$u'v+uv'$
$\dfrac{u(x)}{v(x)}$	$\dfrac{vu'-uv'}{v^2}$
$u(v(x))$	$\dfrac{\mathrm{d}u}{\mathrm{d}v}\dfrac{\mathrm{d}v}{\mathrm{d}x}$

Standard Integrals :

$f(x)$	$\int f(x)\,\mathrm{d}x$
$x^n \ (n \neq -1)$	$x^{n+1}/(n+1)$
$\sin ax$	$-\dfrac{\cos ax}{a}$
$\cos ax$	$\dfrac{\sin ax}{a}$
$\tan x$	$-\ln\cos x$
$1/(1+x^2)$	$\tan^{-1} x$
$\ln x$	$x\ln x - x$
e^{ax}	e^{ax}/a
$1/x$	$\ln x$
$u(x)$	$\displaystyle\int u(x(y))\frac{\mathrm{d}x}{\mathrm{d}y}\,\mathrm{d}y$

Integration by Parts :

$$\int_a^b u(x)\frac{\mathrm{d}v}{\mathrm{d}x}\,\mathrm{d}x = [u(x)v(x)]_a^b - \int_a^b \frac{\mathrm{d}u}{\mathrm{d}x}v(x)\,\mathrm{d}x$$

Trigonometrical Formulæ :

$$\sin^2 A + \cos^2 A = 1\,, \qquad \sec^2 A = \tan^2 A + 1\,,$$

$$\sin(A+B) = \sin A\cos B + \cos A\sin B\,, \quad \sin(A-B) = \sin A\cos B - \cos A\sin B$$

$$\cos(A+B) = \cos A\cos B - \sin A\sin B\,, \quad \cos(A-B) = \cos A\cos B + \sin A\sin B$$

$$\sin 2A = 2\sin A\cos A\,, \quad \cos 2A = 2\cos^2 A - 1 = 1 - 2\sin^2 A$$

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

$$\cos A\cos B = \frac{1}{2}(\cos(A-B) + \cos(A+B))$$

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

$$\sin A + \sin B = 2\sin\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right)$$

$$\sin A - \sin B = 2\sin\left(\frac{A-B}{2}\right)\cos\left(\frac{A+B}{2}\right)$$

$$\cos A + \cos B = 2\cos\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right)$$

$$\cos A - \cos B = -2\sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)$$

Laplace transforms :

$f(t)$	$F(s)$
c	c/s
t	$1/s^2$
t^n	$n!/s^{n+1}$
e^{kt}	$1/(s-k)$
$\sin at$	$a/(s^2+a^2)$
$\cos at$	$s/(s^2+a^2)$
$t\sin at$	$2as/(s^2+a^2)^2$
$t\cos at$	$(s^2-a^2)/(s^2+a^2)^2$
$af(t)+bg(t)$	$aF(s)+bG(s)$
$f'(t)$	$sF(s)-f(0)$
$f''(t)$	$s^2F(s)-sf(0)-f'(0)$
$\delta(t-a)$	e^{-as}
$\mathrm{e}^{at}f(t)$	$F(s-a)$
$f(t)=\begin{cases} g(t-a) & t>a \\ 0 & t<a \end{cases}$	$\mathrm{e}^{-as}G(s)$

Appendix B

Partial Fractions

In the same way you can add algebraic fractions together you can also separate them into combinations of algebraic fractions, such terms are called partial fractions.

Example B.1. Consider

$$\frac{7x+10}{x^2+3x+2}.$$

Suppose we want to convert this into the sum of partial fractions.
We first factorise the denominator:

$$\frac{7x+10}{x^2+3x+2} = \frac{7x+10}{(x+2)(x+1)}.$$

Secondly, we assume a solution with unknown coefficients (say A and B):

$$\frac{7x+10}{(x+2)(x+1)} = \frac{A}{x+2} + \frac{B}{x+1}.$$

Next, multiply both sides of the equation by the denominator $(x+2)(x+1)$ and tidy up:

$$7x+10 = A(x+1) + B(x+2)\,. \tag{B.1}$$

Now find A and B.
Since (B.1) is true for all values of x it must hold when x is -1 and then -2.
Taking $x=-1$, $-7+10 = (-1+1)A + (-1+2)B = B$ so $B=3$.
Taking $x=-2$, $-14+10 = (-2+1)A + (-2+2)B = -A$ so $A=4$. Hence

$$\frac{7x+10}{x^2+3x+2} = \frac{4}{x+2} + \frac{3}{x+1}.$$

Where you have an algebraic fraction like

$$\frac{x-2}{(x+1)^2},$$

so that the denominator has a repeated root, the above approach does not work.

In this case you have to represent the algebraic fraction as the sum of two partial fractions in the form

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

Now multiplying both sides by $(x+1)^2$ gives

$$x - 2 = A(x+1) + B\,. \tag{B.2}$$

You can now equate terms to get two simultaneous equations in A and B to solve:

$$x: \quad 1 = A; \qquad\qquad 1: -2 = A + B\,.$$

Hence $A = 1$ and then $B = -3$. (Note that you could instead take $x = -1$ in (B.2) to get $B = -3$ immediately.) Then

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

You can also get algebraic fractions where the denominator includes quadratic terms $(ax^2 + bx + c)$ that cannot be factorised into two linear factors (without using complex numbers) as the roots are complex. Such cases give rise to a partial fractions of the form

$$\frac{Ax+B}{ax^2+bx+c}$$

where A and B are real constants.

Example B.2. We convert the algebraic fraction

$$\frac{7x}{(x^2+x+1)(x-2)}$$

into partial fractions.
Note that $x^2 + x + 1$ has complex roots so cannot be factorised using real factors. The partial fraction representation is given as:

$$\frac{7x}{(x^2+x+1)(x-2)} = \frac{Ax+B}{x^2+x+1} + \frac{C}{x-2}.$$

Now we have the partial fraction representation of the algebraic fraction we find the unknowns, A, B and C, in the same way as the other examples.
Step 1. Recombine the partial fractions over a common denominator:

$$\frac{Ax+B}{x^2+x+1} + \frac{C}{x-2} = \frac{(Ax+B)(x-2)+C(x^2+x+1)}{(x^2+x+1)(x-2)}.$$

Step 2. Multiply out the factors in the numerator:

$$\frac{(Ax+B)(x-2)+C(x^2+x+1)}{(x^2+x+1)(x-2)} = \frac{Ax^2-2Ax+Bx-2B+Cx^2+Cx+C}{(x^2+x+1)(x-2)}.$$

Step 3. Tidy up:

$$\frac{7x}{(x^2+x+1)(x-2)} = \frac{(A+C)x^2 + (-2Ax + B + C)x - 2B + C}{(x^2+x+1)(x-2)}.$$

Step 4. Equate terms:

$$x^2: \quad A + C = 0\,; \tag{B.3}$$

$$x: \quad -2A + B + C = 7\,; \tag{B.4}$$

$$1: \quad -2B + C = 0\,. \tag{B.5}$$

Step 5. Solve system of equations to give A, B and C.
We have three equations in three unknowns but the approach is the same as the other worked examples. From (B.3), $C = -A$ and substituting into the two remaining equations gives

$$-3A + B = 7\,, \tag{B.6}$$

$$A + 2B = 0\,. \tag{B.7}$$

From (B.7), $A = -2B$ and substituting for A in (B.6) gives

$$-3(-2B) + B = 7B = 7 \qquad \text{so} \qquad B = 1\,.$$

Back substituting, $A = -2B = -2$ and $C = -A = 2$.
Putting this all together gives the final result:

$$\frac{7x}{(x^2+x+1)(x-2)} = \frac{1-2x}{x^2+x+1} + \frac{2}{x-2}.$$

(Again, some work can be saved by seeing what you get by trying $x = 2$.)

Some general rules for partial fractions are:
Given a, b, c, l, m, n, p, q, r and s (*i.e.* they are numbers) then:

$$\frac{mx+n}{(px+q)(rx+s)} = \frac{A}{px+q} + \frac{B}{rx+s};$$

$$\frac{mx+n}{(px+q)^2} = \frac{A}{px+q} + \frac{B}{(px+q)^2};$$

$$\frac{lx^2+mx+n}{(ax^2+bx+c)(px+q)} = \frac{Ax+B}{ax^2+bx+c} + \frac{C}{px+q};$$

$$\frac{lx^2+mx+n}{(px+q)^2(rx+s)} = \frac{A}{px+q} + \frac{B}{(px+q)^2} + \frac{C}{rx+s}.$$

In all cases, you can multiply through by the denominator on the left-hand side of the equation, tidy up and compare terms in the top line. (Work is often saved by seeing what $x = -q/p$ (or $-s/r$) gives.)

Appendix C

Completing the Square

In any quadratic function you can complete the square. Consider the function

$$x^2 - 3x + 2\,. \tag{C.1}$$

You can complete the square by recalling that, for any number a,

$$(x+a)^2 = x^2 + 2ax + a^2$$

so that, subtracting a^2 from both sides and writing $b = 2a$ (*i.e.* $a = b/2$),

$$x^2 + bx = \left(x + \frac{b}{2}\right)^2 - \frac{b^2}{4}\,.$$

Thus, to complete the square in (C.1), where $b = -3$, we add and subtract $3^2 = 9/4$ to get

$$x^2 - 3x + 2 = \left(x + \frac{3}{2}\right)^2 - \frac{9}{4} + 2 = \left(x + \frac{3}{2}\right)^2 - \frac{1}{4}\,.$$

Once you have completed a square you can find the roots of a quadratic equation, although in the Laplace transform method that is not the point of the procedure.