A Short Course in Mathematics

for Engineering and Science Students

Simon A. Mathias
Department of Engineering
Durham University

July 27, 2023

Citation: Mathias, S. A. (2023) A Short Course in Mathematics for Engineering and Science Students. Durham University. https://samathiasuk.github.io/MATHSnotes/MathiasMathsBook.pdf

Preface

This textbook is designed to help students understand how to formulate and solve partial differential equations (PDE). It also provides revision material for many basic mathematical concepts including calculus, series expansion, complex numbers, partial differentiation, vectors and matrices.

The book starts with an introduction to differentiation and integration. Complex numbers are used to explain the difference between trigonometric and hyperbolic functions. Various techniques for solving ordinary differential equations (ODE) are introduced and explained. Maclaurin's power series are used to derive series expansions for many well known functions including exponential, sine and cosine. The concepts of partial and total differentiation are explained. It is shown how to solve PDEs using techniques for solving ODEs through similarity transforms and separation of variables. The latter leads on to the development of the Fourier transform and Fourier series. Basic concepts associated with matrices are introduced and used to derive fundamental engineering concepts including the Mohr circle and Hookes' law. The concept of vector calculus is introduced and used to derive the Cauchy momentum equation and associated wave equations.

This book is mostly derived from lecture notes, originally developed by the author, for geophysics students at Durham University. The author is grateful for associated contributions from Prof Neil Goulty and Prof Jim McElwaine. The author also acknowledges that many of the questions in the problem sheets were obtained from Stroud's Engineering Mathematics and Kreyszig's Advanced Engineering Mathematics.

Simon Mathias, Durham University 27/07/2023

Contents

1

Differentiation and integration

1.1 Learning outcomes

You should be able to:

- Describe the purpose of calculus.
- Determine algebraically the gradients of non-linear functions.
- Show $d(e^x)/dx = e^x$.
- Prove and apply the product rule.
- Understand and apply the chain rule.
- Convert a table of derivatives to a table of integrals.
- Evaluate definite and indefinite integrals.
- Apply the method of integration by substitution.
- Apply the method of integration by partial fractions.

1.2 Differentiation - the gradient at a point

The origin of calculus (differentiation and integration) largely dates back to Isaac Newton (1643-1727), who, during the time of the Great Plague (1665-1666), was pondering about the calculation of gradients. Gradients relate to numerous important physical properties (velocity on a distance-time graph, acceleration on a velocity time graph, etc.). For straight lines such as $y = mx + c$, the gradient of the line is obviously m. But what about the gradients of curved lines, for example, as shown in Fig. 1.1?

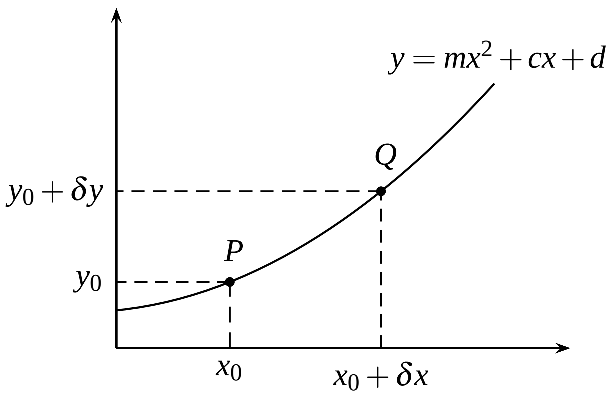


Figure 1.1: Approximate gradient of a curved line.

Indeed, we can approximate the gradient between the two points, P and Q, as $\delta y/\delta x$. But what about the gradient at a single point? We can approach this result by looking at the gradient between P and Q when P and Q are moved infinitesimally close together, i.e. making $\delta x \to 0$. But as $\delta x \to 0$, it can also be seen that $\delta y \to 0$. What does it mean if we say

$$\lim_{\delta x \to 0} \frac{\delta y}{\delta x} = \frac{0}{0} \quad ?$$

Newton set out to solve this conundrum (in a way similar to) as follows:

Consider the function $y = mx^2 + cx + d$

Assume point P is at (x_0, y_0) and point Q is at $(x_0 + \delta x, y_0 + \delta y)$. The linear gradient between P and Q is found from $\delta y / \delta x$. What is the slope, $\delta y / \delta x$ as a function of x_0 and δx?

At $x = x_0 + \delta x$, $y = y_0 + \delta y$.

$\therefore y_0 + \delta y = m(x_0 + \delta x)^2 + c(x_0 + \delta x) + d$

and of course $y_0 = mx_0^2 + cx_0 + d$.

$\therefore \delta y = m(x_0 + \delta x)^2 - mx_0^2 + c(x_0 + \delta x) - cx_0$

$\therefore \delta y = m(x_0^2 + 2\delta x x_0 + \delta x^2) - mx_0^2 + c\delta x$

$\therefore \delta y = m(2\delta x x_0 + \delta x^2) + c\delta x$

and $\div$ both sides by δx we have

$$\frac{\delta y}{\delta x} = m(2x_0 + \delta x) + c$$

Now we write the derivative $\dfrac{dy}{dx}$, which represents the slope of the line at a discrete point.

As we move points Q and P infinitesimally close together and setting $x_0 = x$ we have

$$\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x} = 2mx + c$$

So for $y = mx^2 + cx + d$

$$\frac{dy}{dx} = 2mx + c$$

From the previous result, it can be appreciated that:

y	dy/dx
1	0
x	1
x^2	$2x$
x^3	$3x^2$
x^n	nx^{n-1}

But the key result here is the identification of the "derivative"

$$\boxed{\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x}}$$

which is the gradient of a function y at an infinitesimal point, x.

Although it is known that Newton formulated his calculus theory around 1665, he did not formally publish his results until 1693. However, prior to that, the German mathematician, Gottfried Leibniz (1646-1716), independently established and published the theory of calculus in 1684. Both Leibniz and Newton were members of the Royal Society of London, and after a long and rigorous investigation, the Royal Society authoritatively ruled that Newton was the originator of the theory. Nevertheless, the form of calculus notation used today originates from the Leibniz manuscript as opposed to the apparently more cumbersome form of Newton.

1.3 The derivative of e^x

Consider the exponential series

$$y = e^x \equiv 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \tag{1.1}$$

Find $\dfrac{dy}{dx}$

Consider the point $(y + \delta y, x + \delta x)$

$$\begin{aligned} y + \delta y &= e^{x+\delta x} = 1 + x + \delta x + \frac{(x+\delta x)^2}{2!} + \frac{(x+\delta x)^3}{3!} + \ldots \\ &= 1 + x + \delta x + \frac{x^2}{2!} + x\delta x + \frac{\delta x^2}{2!} + \frac{x^3 + 3x^2\delta x + 3x\delta x^2 + \delta x^3}{3!} \end{aligned}$$

Collecting terms of δx

$$y + \delta y = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \delta x\left(1 + x + \frac{x^2}{2!}\right) + \frac{\delta x^2}{2}(1 + 2x) + \frac{\delta x^3}{3}\ldots$$

Now recalling Eq. (1.1)

$$\delta y = \delta x\left(1 + x + \frac{x^2}{2!}\right) + \frac{\delta x^2}{2}(1 + 2x) + \frac{\delta x^3}{3}\ldots$$

Dividing both sides by δx

$$\frac{\delta y}{\delta x} = 1 + x + \frac{x^2}{2!} + \frac{\delta x}{2}(1 + 2x) + \frac{\delta x^2}{3}\ldots$$

and

$$\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x} = 1 + x + \frac{x^2}{2!} + \ldots = e^x$$

$$\boxed{\text{So} \quad \frac{d(e^x)}{dx} = e^x}$$

1.4 The Product and Quotient Rules

Now consider a function of the form $y = u(x)v(x)$

Consider again the point $(y+\delta y, x+\delta x)$

$$u = u(x) \qquad\qquad v = v(x)$$

$$u+\delta u = u(x+\delta x) \qquad\qquad v+\delta v = v(x+\delta x)$$

$$y = y(x) = uv$$

$$y+\delta y = y(x+\delta x)$$

$$\therefore \quad y+\delta y = (u+\delta u)(v+\delta v)$$

$$\therefore \quad uv+\delta y = uv+v\delta u+u\delta v+\delta u\delta v$$

$$\frac{\delta y}{\delta x} = v\frac{\delta u}{\delta x}+u\frac{\delta v}{\delta x}+\delta u\frac{\delta v}{\delta x}$$

Now it can be said that (for u and v are continuous functions)

$$\lim_{\delta x\to 0}\delta u = 0 \qquad \lim_{\delta x\to 0}\delta v = 0$$

and therefore $$\frac{dy}{dx} = \lim_{\delta x\to 0}\frac{\delta y}{\delta x} = u\frac{dv}{dx}+v\frac{du}{dx}$$

Using this approach we can arrive at two important rules:

The Product rule: $$\frac{d}{dx}(uv) = u\frac{dv}{dx}+v\frac{du}{dx}$$

$$\boxed{\text{The Quotient rule:} \quad \frac{d}{dx}\left(\frac{u}{g}\right) = \frac{1}{g^2}\left(g\frac{du}{dx} - u\frac{dg}{dx}\right)}$$

1.5 The Chain Rule

What happens when you get something like $y = 3(x+3)^7$?

One way to deal with this is to make a substitution.

Say $g = x+3 \Rightarrow y = 3g^7$

So $\dfrac{dy}{dg} = 21g^6$

But $\dfrac{dy}{dx} = ?$

Recall that $\dfrac{dy}{dx} = \lim\limits_{\delta x \to 0} \dfrac{\delta y}{\delta x}$

Now $\dfrac{\delta y}{\delta x} = \dfrac{\delta y}{\delta x}\dfrac{\delta g}{\delta x}\dfrac{\delta x}{\delta g} = \dfrac{\delta g}{\delta x}\dfrac{\delta y}{\delta g}$

from which it follows that $\boxed{\dfrac{dy}{dx} = \dfrac{dg}{dx}\dfrac{dy}{dg}}$ where $y = f(g(x))$.

So for the case above $\dfrac{dy}{dx} = \dfrac{dg}{dx} \times 21g^6 = 21(x+3)^6$.

1.6 Integration - the opposite of differentiation

Integration is simply the opposite of differentiation, i.e.

$$\int \frac{dy}{dx} dx = y$$

We can therefore use the derivatives we previously established to derive a table of corresponding integrals as shown in Table 1.1.

Table 1.1: Table of derivatives and integrals.

y	dy/dx	y	$\int y dx$
x^n	nx^{n-1}	x^n	$(n+1)^{-1}x^{n+1}+C$
e^x	e^x	e^x	e^x+C
$\sin x$	$\cos x$	$\cos x$	$\sin x+C$
$\cos x$	$-\sin x$	$\sin x$	$-\cos x+C$
$\tan x$	$\sec^2 x$	$\sec^2 x$	$\tan x+C$
$\ln x$	x^{-1}	x^{-1}	$\ln x+C$

1.7 The integration constant

Consider $\frac{d}{dx}(x^4+2) = 4x^3 \quad \therefore \int 4x^3 dx = x^4+2$

Now consider $\frac{d}{dx}(x^4-5) = 4x^3 \quad \therefore \int 4x^3 dx = x^4-5$

Note that we always get different answers.

So we should say $\int 4x^3 dx = x^4+C$

where C is the unknown integration constant and requires additional information.

How can we find the integration constant?

Consider the integral

$$F = \int 8x^3 + 6x^2 dx = \frac{8}{4}x^4 + \frac{6}{3}x^3 + C = 2x^4 + 2x^3 + C$$

We can only find C if we have knowledge of F at a given point x.

Eg. $F = 3$ when $x = 2$

$$\therefore \quad 3 = 2 \times 16 + 2 \times 8 + C$$

$$\therefore \quad C = 3 - 32 - 16 = -45$$

1.8 Definite and indefinite integral

The integrals we have previously discussed are often referred to as “indefinite integrals”, typically denoted as

$$\int f(x)dx$$

This is as opposed to “definite integrals, which are bound by “limits”, in this case a and b (see Fig. 1.2).

$$\int_a^b f(x)dx$$

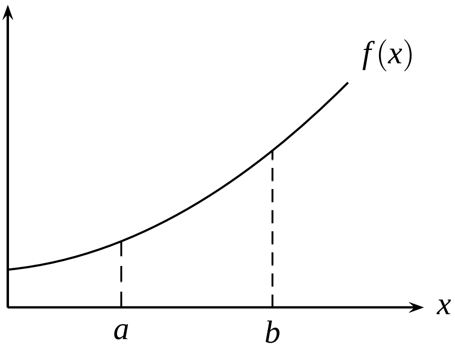


Figure 1.2: Schematic diagram illustrating the meaning of the limits a and b.

Note that the value of the definite integral, in this case, corresponds to the area bounded by $y = f(x)$, $x = a$, $x = b$ and $y = 0$.

An example of definite integral evaluation is shown below

$$F = \int_a^b 8x^3 + 6x^2 dx = \left[2x^4 + 2x^3\right]_a^b = 2(b^4 + b^3 - a^4 - a^3)$$

1.9 Integration by substitution

Consider integrals such as

$$\int \sin(2x+3)dx \quad \text{and} \quad \int (4x+2)^5 dx$$

The best way to deal with these is by integration by substitution.

Recall the chain rule $\dfrac{dy}{dx} = \dfrac{dg}{dx}\dfrac{dy}{dg}$

Consider $\int f(g(x))dx = \int f(g(x))\frac{dg}{dx}\frac{dx}{dg}dx$

From which it follows that $\boxed{\int f(g(x))dx = \int f(g(x))\frac{dx}{dg}dg}$

Challenge 1.1 Evaluate the integral $\int \sin(2x+3)dx$

Let $g = 2x+3$ such that $\frac{dg}{dx} = 2$ and $\frac{dx}{dg} = \frac{1}{2}$

It follows that

$$\int \sin(2x+3)dx = \int \frac{\sin g}{2}dg = -\frac{\cos g}{2} + C = -\frac{\cos(2x+3)}{2} + C$$

Challenge 1.2 Evaluate the integral $\int (4x+2)^5 dx$

Let $g = 4x+2$ such that $\frac{dg}{dx} = 4$ and $\frac{dx}{dg} = \frac{1}{4}$

It follows that

$$\int (4x+2)^5 dx = \int \frac{g^5}{4}dg = \frac{g^6}{24} + C = \frac{(4x+2)^6}{24} + C$$

1.10 Integration by partial fractions

$$\boxed{\text{Partial fractions:} \quad \frac{C}{ab} = \frac{A}{a} + \frac{B}{b}}$$

Now consider something like $F = \int \frac{7x+8}{2x^2+11x+5}dx$

The denominator can be split such that $F = \int \frac{7x+8}{(2x+1)(x+5)} dx$

and splitting into partial fractions: $F = \int \frac{A}{2x+1} + \frac{B}{x+5} dx$

But what are A and B?

Well, it can be seen that: $\frac{7x+8}{(2x+1)(x+5)} = \frac{A}{2x+1} + \frac{B}{x+5}$

$$\therefore 7x+8 = (x+5)A + (2x+1)B$$

Find a value of x that eliminates A or B.

Choose $x = -5 \quad \Rightarrow 7 \times (-5) + 8 = (2 \times (-5) + 1)B$

$$\therefore 8 - 35 = -9B \quad \Rightarrow B = \frac{27}{9} = 3$$

Choose $x = -0.5 \quad \Rightarrow 8 - 3.5 = 4.5A \quad \Rightarrow A = 1$

Check: $7x+8 = (x+5) + 3(2x+1) = 7x+8 \quad \checkmark$

So $\quad F = \int \frac{7x+8}{2x^2+11x+5} dx = \int \frac{1}{2x+1} + \frac{3}{x+5} dx$

Now let us consider the general problem of

$$G = \int \frac{1}{ax+b} dx$$

$$g = ax+b \quad \frac{dg}{dx} = a \quad \frac{dx}{dg} = \frac{1}{a}$$

$$\therefore G = \int \frac{1}{g} dx = \int \frac{1}{g} \frac{dx}{dg} dg$$

$$\therefore G = \frac{1}{a}\int \frac{1}{g}dg = \frac{\ln g}{a} + C = \frac{\ln(ax+b)}{a} + C$$

Hence: $$\boxed{\int \frac{1}{ax+b}dx = \frac{\ln(ax+b)}{a} + C}$$

So $$F = \int \frac{1}{2x+1} + \frac{3}{x+5}dx = \frac{\ln(2x+1)}{2} + 3\ln(x+5) + C$$

1.11 Problem sheet

Problem 1.1 (see Worked Solution 1.1)

Determine algebraically, from first principals, the gradient of the graph of $y = 6x^2 + 5$ at the point P where $x = 0.5$.

Problem 1.2 (see Worked Solution 1.2)

If $y = 2x^4 + 3x^3 - 7x^2 + 2x - 6$, give expressions for:

a) $\frac{dy}{dx}$ b) $\frac{d^2y}{dx^2}$ c) $\frac{d^3y}{dx^3}$

Problem 1.3 (see Worked Solution 1.3)

Differentiate the following functions using the product rule:

a) $y = x^4 \sin x$ b) $y = e^x \sin x$

Problem 1.4 (see Worked Solution 1.4)

Differentiate the following functions using the chain rule:

a) $y = (2x-3)^5$ b) $y = \cos(3x^2 + x)$

Problem 1.5 (see Worked Solution 1.5)

Starting from $\dfrac{d(e^x)}{dx} = e^x$ show that $\dfrac{d(\ln x)}{dx} = \dfrac{1}{x}$

Problem 1.6 (see Worked Solution 1.6)

Starting from $\dfrac{d(\sin x)}{dx} = \cos x$ and $\dfrac{d(\cos x)}{dx} = -\sin x$

show that $\dfrac{d(\tan x)}{dx} = \sec^2 x$

Problem 1.7 (see Worked Solution 1.7)

Using the table of integrals developed in class, evaluate the following integrals

a) $\displaystyle\int_4^5 x^3 dx$ b) $\displaystyle\int_0^{\pi/2} 3\cos x dx$ c) $\displaystyle\int \frac{4}{x} dx$ d) $\displaystyle\int 2\sec^2 x dx$

Problem 1.8 (see Worked Solution 1.8)

Using the method of integration by substitution, evaluate the following

a) $\displaystyle\int 3e^{4x+1} dx$ b) $\displaystyle\int (5x+2)^{-1} dx$

Problem 1.9 (see Worked Solution 1.9)

Using the method of integration by partial fractions, evaluate the following

a) $\displaystyle\int \frac{4x+15}{x^2+7x+12} dx$ b) $\displaystyle\int \frac{8x+15}{2x^2+9x+9} dx$

1.12 Worked solutions

Worked Solution 1.1 (see Problem 1.1)

$y = 6x^2 + 5$

$y_0 + \delta y = 6(x_0 + \delta x)^2 + 5$

$6x_0^2 + 5 + \delta y = 6x_0^2 + 12x_0\delta x + 6\delta x^2 + 5$

$$\frac{\delta y}{\delta x} = 12x_0 + 6\delta x$$

$$\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x} = 12x$$

At $x = 0.5$, $\dfrac{dy}{dx} = 6$

Worked Solution 1.2 (see Problem 1.2)

$y = 2x^4 + 3x^3 - 7x^2 + 2x - 6$

a) $\dfrac{dy}{dx} = 8x^3 + 9x^2 - 14x + 2$ b) $\dfrac{d^2y}{dx^2} = 24x^2 + 18x - 14$

c) $\dfrac{d^3y}{dx^3} = 48x + 18$

Worked Solution 1.3 (see Problem 1.3)

a) $y = x^4 \sin x$

$$\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx} = x^4 \cos x + 4x^3 \sin x$$

b) $y = e^x \sin x$

$$\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx} = e^x \cos x + e^x \sin x = e^x(\cos x + \sin x)$$

Worked Solution 1.4 (see Problem 1.4)

a) $y = (2x-3)^5$

$$g = 2x-3, \qquad y = g^5, \qquad \frac{dy}{dg} = 5g^4$$

$$\frac{dy}{dx} = \frac{dg}{dx}\frac{dy}{dg} = 2 \times 5g^4 = 10(2x-3)^4$$

b) $y = \cos(3x^2+x)$

$$g = 3x^2+x, \qquad y = \cos g, \; \frac{dy}{dg} = -\sin g$$

$$\frac{dy}{dx} = \frac{dg}{dx}\frac{dy}{dg} = -(6x+1)\sin(3x^2+x)$$

Worked Solution 1.5 (see Problem 1.5)

Given $\dfrac{de^x}{dx} = e^x$

$$y = \ln x, \qquad e^y = x, \qquad \frac{dx}{dy} = e^y$$

and $\dfrac{d\ln x}{dx} = \dfrac{dy}{dx} = \dfrac{1}{e^y} = \dfrac{1}{x}$

Worked Solution 1.6 (see Problem 1.6)

$$y = \tan x = \frac{\sin x}{\cos x}$$

$$\frac{dy}{dx} = \frac{1}{g^2}\left(g\frac{du}{dx} - u\frac{dg}{dx}\right) = \frac{1}{\cos^2 x}(\cos^2 x + \sin^2 x) = \sec^2 x$$

recall that $\boxed{\cos^2 x + \sin^2 x = 1}$

Worked Solution 1.7 (see Problem 1.7)

a) $\int_4^5 x^3 dx = \left[\frac{x^4}{4}\right]_4^5 = \frac{5^4 - 4^4}{4} = 92.25$

b) $\int_0^{\pi/2} 3\cos x dx = [3\sin x]_0^{\pi/2} = 3 - 0 = 3$

c) $\int \frac{4}{x} dx = 4\ln x + C$

d) $\int 2\sec^2 x dx = 2\tan x + C$

Worked Solution 1.8 (see Problem 1.8)

a) $F = \int 3e^{4x+1} dx$

$$g = 4x + 1, \qquad \frac{dg}{dx} = 4, \frac{dx}{dg} = \frac{1}{4}$$

$$F = \int 3e^g dx = 3\int e^g \frac{dg}{dx}\frac{dx}{dg} dx = 3\int e^g \frac{dx}{dg} dg$$

$$\therefore F = \frac{3}{4}\int e^g dg = \frac{3}{4}e^{4x+1} + C$$

b) $\int (5x+2)^{-1} dx$

$$g = 5x + 2, \qquad \frac{dg}{dx} = 5, \frac{dx}{dg} = \frac{1}{5}$$

$$F = \int g^{-1} dx = \int g^{-1}\frac{dg}{dx}\frac{dx}{dg} dx = \int g^{-1}\frac{dx}{dg} dg$$

$$\therefore F = \frac{1}{5}\int g^{-1} dg = \frac{\ln(5x+2)}{5} + C$$

Worked Solution 1.9 (see Problem 1.9)

a) $F = \int \frac{4x + 15}{x^2 + 7x + 12} dx$

$$\frac{4x+15}{x^2+7x+12} = \frac{4x+15}{(x+3)(x+4)} = \frac{A}{x+3} + \frac{B}{x+4}$$

$\therefore 4x+15 = (x+4)A + (x+3)B$

$x = -3, \quad A = -12+15, \quad A = 3$

$x = -4, \quad -1 = -B, \quad B = 1$

Check: $\qquad 4x+15 = 3x+12+x+3 \quad \checkmark$

So $F = \displaystyle\int \frac{3}{x+3} + \frac{1}{x+4} dx = 3\ln(x+3) + \ln(x+4) + C$

b) $\displaystyle\int \frac{8x+15}{2x^2+9x+9} dx$

$$\frac{8x+15}{2x^2+9x+9} = \frac{8x+15}{(2x+3)(x+3)} = \frac{A}{2x+3} + \frac{B}{x+3}$$

$\therefore 8x+15 = (x+3)A + (2x+3)B$

$x = -3, \quad -24+15 = -3B, \quad B = 3$

$x = -1.5, \quad -12+15 = 1.5A, \quad A = 2$

Check: $\qquad 8x+15 = 2x+6+6x+9 \quad \checkmark$

So $F = \displaystyle\int \frac{2}{2x+3} + \frac{3}{x+3} dx = \ln(2x+3) + 3\ln(x+3) + C$

2

Complex numbers

2.1 Learning outcomes

You should be able to:

- Understand what a complex number is.
- Understand that $i^4 = i^{-4} = 1$.
- Add, subtract, multiply and divide complex numbers.
- Draw an Argand diagram.
- Show why complex numbers can be presented in polar, exponential and trigonometric forms.
- Use the above knowledge to prove trigonometric formulae and derivatives.

2.2 NIRIC

Let us start by considering NIRIC:

Natural numbers	$1, 2, 3, 4$, etc.
Integers	$-3, -2, -1, 0, 1, 2, 3$, etc.
Rational numbers	$7/34, 13/71, 2/3$, etc.
Irrational numbers	$\pi = 3.14159\ldots, e = 2.7182\ldots$, etc.
Complex numbers	$6 + i3, 1.789 - i2.36, \pi - i2\pi$, etc.

Note that a 'complex' is a mixture. A complex number is a mixture of real and imaginary numbers. Imaginary numbers are denoted by the prefix i (or sometimes j).

So what is an imaginary number?

$$\boxed{i = \sqrt{-1}}$$

2.3 When do complex numbers come about?

Lets start with quadratic equations:

$$\begin{array}{l|l} x^2 - 1 = 0 & x^2 + 1 = 0 \\ x^2 = 1 & x^2 = -1 \\ x = \pm\sqrt{1} & x = \pm\sqrt{-1} \\ x = \pm 1 & x = \pm i \end{array}$$

Now recall the general quadratic formula

$$ax^2 + bx + c = 0$$

where x is found from

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

and consider

$x^2+2x-3=0$	$x^2+2x+3=0$
$x=\dfrac{-2\pm\sqrt{4+12}}{2}$	$x=\dfrac{-2\pm\sqrt{4-12}}{2}$
$x=-1\pm 2$	$x=-1\pm i\sqrt{2}$
$x=1$ or 3	$x=-1+i\sqrt{2}$ or $-1-i\sqrt{2}$

The right-hand-side is a set of complex results.

2.4 Powers of i

$$
\begin{aligned}
&i=\sqrt{-1}\\
&i^2=(\sqrt{-1})^2=-1\\
&i^3=i^2\times i=-i\\
&i^4=i^2\times i^2=1\\
&i^5=i^4\times i=i\\
&i^{10}=(i^4)^2\times i^2=-1
\end{aligned}
$$

What about negative integer powers of i?

$$i^{-1}=\frac{1\times i}{i\times i}=\frac{i}{-1}=-i$$

$$i^{-2}=\frac{1}{-1}=-1$$

$$i^{-3}=\frac{1}{i^2\times i}=\frac{1\times i}{i^2\times i\times i}=\frac{i}{i^4}=i$$

$$i^{-4}=1$$

The important thing to remember is that $i^4=1$. Therefore we can handle really high orders of i very easily. For example $i^{-126}=(i^4)^{-31}i^{-2}=1\times -1=-1$.

2.5 Addition, subtraction and multiplication

Consider

$$z_1 = a + ib \quad \text{and} \quad z_2 = c + id$$

It follows that

$$z_1 + z_2 = a + c + i(b + d)$$

$$z_1 - z_2 = a - c + i(b - d)$$

$$\begin{aligned} z_1 \times z_2 \quad &= (a + ib)(c + id) \\ &= ac + ibc + iad + i^2 bd \\ &= ac - bd + i(bc + ad) \end{aligned}$$

2.6 Division

Now consider

$$z_1 \div z_2 = \frac{a + ib}{c + id}$$

which is fine, except that it is generally considered good form not to have imaginary numbers in both the nominator and the denominator. One way to eliminate i from the denominator is to multiply both top and bottom by the so-called 'conjugate'.

Consider the complex number $a + ib$. The conjugate of this is $a - ib$. The conjugate of a complex number is the same as the complex number of concern, except that the sign of the imaginary part is changed.

Now we will multiply the top and bottom of the above equation by the conjugate of the bottom

$$
\begin{aligned}
z_1 \div z_2 &= \frac{(a+ib)\times(c-id)}{(c+id)\times(c-id)} \\
&= \frac{ac+ibc-iad-i^2bd}{c^2+icd-icd-i^2d^2} \\
&= \frac{ac+bd+i(bc-ad)}{c^2+d^2}
\end{aligned}
$$

All imaginary numbers are now at the top!

2.7 Argand diagrams

Considering again a complex number of the form $z = a + ib$. It is possible to represent this in a two-dimensional plot in Fig. 2.1. Diagrams such as in Fig. 2.1 are often referred to as Argand diagrams, after Jean-Robert Argand (1806).

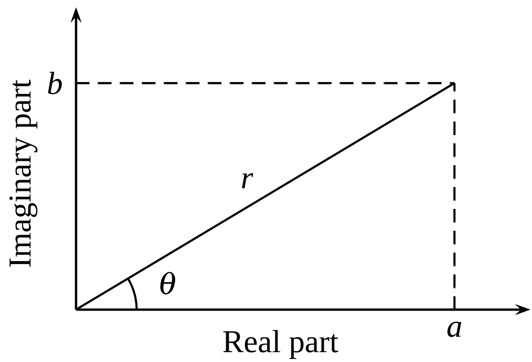


Figure 2.1: Example of an Argand diagram.

2.8 Polar form of a complex number

From the Argand diagram above it can be easily appreciated that another way of representing a complex number is by its radial distance from the origin, r, and the angle of orientation, θ. It can be easily seen that

$$\begin{aligned} r^2 &= a^2 + b^2 \\ \tan\theta &= b/a \end{aligned} \tag{2.1}$$

The polar form of a complex number is generally written as

$$z = r\angle\theta$$

Sometimes people refer to the 'modulus' of a complex number

$$|z| = \sqrt{a^2 + b^2} \equiv r$$

and the argument of a 'complex' number

$$\arg z = \arctan\left(\frac{b}{a}\right) \equiv \theta$$

Note that arctan is a way of writing inverse tan.

2.9 Trigonometric form of a complex number

Further examination of Eq. (2.1) and the diagram above also reveals that

$$\begin{aligned} a &= r\cos\theta \\ b &= r\sin\theta \end{aligned}$$

It follows that another way of writing complex numbers is in a trigonometric form

$$\boxed{z = r\cos\theta + ir\sin\theta} \tag{2.2}$$

2.10 Exponential form of a complex number

Consider again the exponential series expansion

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \ldots$$

and set $x = i\theta$.

$$e^{i\theta} = 1 + i\theta + \frac{i^2\theta^2}{2!} + \frac{i^3\theta^3}{3!} + \frac{i^4\theta^4}{4!} + \frac{i^5\theta^5}{5!} + \ldots$$

Recalling that $i^2 = -1$ then leads to

$$e^{i\theta} = 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} + \ldots$$

Collecting real and imaginary parts we then have

$$e^{i\theta} = \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \ldots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \ldots\right)$$

Now it turns out that

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \ldots \quad \text{and} \quad \sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \ldots$$

It follows that

$$\boxed{e^{i\theta} = \cos\theta + i\sin\theta} \tag{2.3}$$

which, on comparison with Eq. (2.2), reveals that complex numbers can also be written in an exponential form such as

$$\boxed{z = re^{i\theta}}$$

Eq. (2.3) above is often referred to as Euler's formula.

2.11 More on Euler's formula

Let us again consider the Euler formula, Eq. (2.3)

$$e^{i\theta} = \cos\theta + i\sin\theta$$

it follows that

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

and consequently that

$$\begin{aligned} e^{i\theta} + e^{-i\theta} &= \cos\theta + i\sin\theta + \cos\theta - i\sin\theta \\ &= 2\cos\theta \end{aligned}$$

Similarly

$$\begin{aligned} e^{i\theta} - e^{-i\theta} &= \cos\theta + i\sin\theta - \cos\theta + i\sin\theta \\ &= 2i\sin\theta \end{aligned}$$

From which it follows:

$$\boxed{\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{i2}}$$

2.12 An amazing equation

Again consider Euler's formula

$$e^{i\theta} = \cos\theta + i\sin\theta$$

Bringing everything to the left-hand-side

$$e^{i\theta} - \cos\theta - i\sin\theta = 0$$

Setting $\theta = \pi$ then leads to

$$e^{i\pi} + 1 = 0$$

Interestingly, raising the irrational quantity, e, to the power of the imaginary and irrational quantity, $i\pi$, leads to the a real and rational quantity, -1. It is amazing to find so many key concepts in mathematics locked up in just one simple equation $(e, i, \pi, 1, 0)$.

2.13 Problem sheet

Problem 2.1 (see Worked Solution 2.1)

Simplify the following

a) i^5 b) i^{12} c) i^{-18} d) i^{37}

Problem 2.2 (see Worked Solution 2.2)

Express in the form $a + ib$

a) $(2 - i6)(3 + i2)$ b) $(i - 2)^2$ c) $\dfrac{5 + i2}{6 - i}$ d) $\dfrac{2 + i6}{3 + i2}$

Problem 2.3 (see Worked Solution 2.3)

Sketch the Argand diagrams and express in polar, exponential and trigonometric form

a) $2+i6$ b) $3-i7$

Problem 2.4 (see Worked Solution 2.4)

Express in the form $a+ib$

a) $6\,[\cos(\pi/5)+i\sin(\pi/5)]$ b) $6\angle 37^{o}$ c) $e^{2+i\pi/3}$

Problem 2.5 (see Worked Solution 2.5)

Given that $\dfrac{d(e^x)}{dx}=e^x$ and $e^{ix}=\cos x+i\sin x$ show that

a) $\dfrac{d(\sin x)}{dx}=\cos x$ b) $\dfrac{d(\cos x)}{dx}=-\sin x$

Problem 2.6 (see Worked Solution 2.6)

Starting from $e^{ix}=\cos x+i\sin x$, show that

a) $\cos(A+B)=\cos A\cos B-\sin A\sin B$

b) $\sin(A+B)=\sin A\cos B+\cos A\sin B$

2.14 Worked solutions

Worked Solution 2.1 (see Problem 2.1)

a) $i^5=i^4\times i=i$

b) $i^{12}=(i^4)^3=1$

c) $i^{-18} = (i^{-4})^4 \times i^{-2} - 1$

d) $i^{37} = (i^4)^9 \times i = i$

Worked Solution 2.2 (see Problem 2.2)

a) $(2-i6)(3+i2) = 6-i18+i4+12 = 18-i14$

b) $(i-2)^2 = i^2-4i+4 = 3-4i$

c) $\dfrac{5+i2}{6-i} = \dfrac{(5+i2)(6+i)}{(6-i)(6+i)} = \dfrac{30+i12+i5-2}{36+1} = \dfrac{28+i17}{37}$

d) $\dfrac{2+i6}{3+i2} = \dfrac{(2+i6)(3-i2)}{(3+i2)(3-i2)} = \dfrac{6+i18-i4-12i^2}{9+4} = \dfrac{18+i14}{13}$

Worked Solution 2.3 (see Problem 2.3)

a) $2+i6$

$\theta = \arctan(6/2) = \arctan(3) = 1.249 \text{ rad } (= 71.56^o)$

$r = \sqrt{4+36} = \sqrt{40} = 6.3246$

$\therefore 2+i6 = \sqrt{40}\angle 71.6^o = \sqrt{40}e^{i1.25} = \sqrt{40}[\cos 71.6^o + i\sin 71.6^o]$

b) $3-i7$

$\theta = \arctan(-7/3) = -1.166 \text{ rad } (= -66.80^o)$

$r = \sqrt{9+49} = \sqrt{58} = 7.616$

$\therefore 3+i7 = \sqrt{58}\angle -66.8^o = \sqrt{58}e^{-i1.17} = \sqrt{58}[\cos 66.8^o - i\sin 66.8^o]$

Worked Solution 2.4 (see Problem 2.4)

a) $6\left[\cos(\pi/5)+i\sin(\pi/5)\right]=6(0.809+i0.588)=4.85+i3.53$

b) $6\angle 37^o=6(\cos 37^o+i\sin 37^o)=4.79+i3.61$

c) $e^{2+i\pi/3}=e^2e^{i\pi/3}=e^2\left[\cos(\pi/3)+i\sin(\pi/3)\right]=3.69+i6.40$

Worked Solution 2.5 (see Problem 2.5)

Given $e^{ix}=\cos x+i\sin x$, we have $e^{-ix}=\cos x-i\sin x$.

It follows that $e^{ix}+e^{-ix}=2\cos x$ and $e^{ix}-e^{-ix}=i2\sin x$

from which we have $\cos x=\dfrac{e^{ix}+e^{-ix}}{2}$ and $\sin x=\dfrac{e^{ix}-e^{-ix}}{i2}$

a) $\dfrac{d(\sin x)}{dx}=\dfrac{ie^{ix}-(-i)e^{-ix}}{i2}=\dfrac{e^{x}+e^{-x}}{2}=\cos x$

b) $\dfrac{d(\cos x)}{dx}=\dfrac{ie^{ix}-ie^{-ix}}{2}=\dfrac{i^2e^{ix}-i^2e^{-ix}}{i2}=-\sin x$

Worked Solution 2.6 (see Problem 2.6)

From $e^{ix}=\cos x+i\sin x$ we have:

$e^{iA}=\cos A+i\sin A$

$e^{iB}=\cos B+i\sin B$

$e^{i(A+B)}=\cos(A+B)+i\sin(A+B)$

and $e^{i(A+B)}=e^{iA}e^{iB}$, therefore it can be said

$e^{i(A+B)}=(\cos A+i\sin A)(\cos B+i\sin B)$

$\therefore e^{i(A+B)}=\cos A\cos B+i\sin A\cos B+i\cos A\sin B+i^2\sin A\sin B$

$$\therefore e^{i(A+B)} = \cos A\cos B - \sin A\sin B + i(\sin A\cos B + \cos A\sin B)$$

$$\therefore \cos(A+B) + i\sin(A+B) = \cos A\cos B - \sin A\sin B \\ +i(\sin A\cos B + \cos A\sin B)$$

Equating real and imaginary parts

a) $\cos(A+B) = \cos A\cos B - \sin A\sin B$

b) $\sin(A+B) = \sin A\cos B + \cos A\sin B$

3

Hyperbolic functions

3.1 Learning outcomes

You should be able to:

- Define hyperbolic functions in terms of exponentials.
- Relate hyperbolic functions to trigonometric functions.
- Sketch plots of cosh, sinh and tanh.
- Derive expressions for inverse cosh, sinh and tanh.
- Obtain hyperbolic formulae analogous to trigonometric formulae.
- Derive expressions for the derivatives and integrals of cosh, sinh and tanh.

3.2 What are hyperbolic functions?

We talked earlier about

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \text{and} \quad \sin x = \frac{e^{ix} - e^{-ix}}{i2}$$

What happens when the arguments for sin and cos become imaginary?

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

$$\sin ix = \frac{e^{i^2 x} - e^{-i^2 x}}{i2} = \frac{e^{-x} - e^{x}}{i2}$$

Introducing the hyperbolic functions:

$$\boxed{\cosh x = \frac{e^{x} + e^{-x}}{2} \qquad \sinh x = \frac{e^{x} - e^{-x}}{2}}$$

$$\tanh x = \frac{\sinh x}{\cosh x} \qquad \coth x = \frac{1}{\tanh x}$$

$$\mathrm{sech} x = \frac{1}{\cosh x} \qquad \mathrm{cosech} x = \frac{1}{\sinh x}$$

3.3 Relationship with trigonometric functions

Given

$$\cos ix = \frac{e^{x} + e^{-x}}{2} \quad \text{and} \quad \sin ix = \frac{-1}{i}\left(\frac{e^{x} - e^{-x}}{2}\right)$$

it follows that

$$\boxed{\cosh x = \cos ix \qquad \sinh x = -i \sin ix \qquad \tanh x = -i \tan ix}$$

and

$$\boxed{\cos ix = \cosh x \quad \sin ix = i \sinh x \quad \tan ix = i \tanh x}$$

It is worth to compare this to

$$\cos(-x) = \cos x \quad \sin(-x) = -\sin x \quad \tan(-x) = -\tanh x$$

3.4 So what do they look like?

First recall the trigonometric functions, $\cos x$, $\sin x$ and $\tan x$ (see Figs. 3.1 and 3.2).

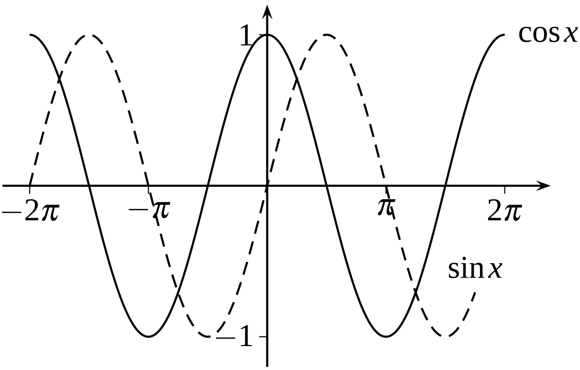


Figure 3.1: Plots of $\cos x$ and $\sin x$.

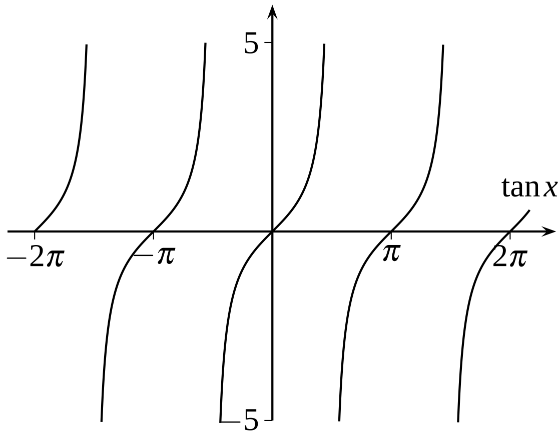


Figure 3.2: Plot of $\tan x$.

Now lets look at the hyperbolic functions:

$$\cosh x = \frac{e^x + e^{-x}}{2} \quad \text{and} \quad \sinh x = \frac{e^x - e^{-x}}{2}$$

First consider the limit as $x \to 0$.

$$\lim_{x\to 0} e^x = 1, \quad \lim_{x\to 0} e^{-x} = 1$$

It follows that:

$$\lim_{x\to 0} \cosh x = \frac{1+1}{2} = 1, \quad \lim_{x\to 0} \sinh x = \frac{1-1}{2} = 0$$

Now lets study the limit as $x \to \infty$.

$$\lim_{x\to\infty} e^x = \infty, \quad \lim_{x\to\infty} e^{-x} = 0 \tag{3.1}$$

It follows that

$$\lim_{x\to\infty} \cosh x = \lim_{x\to\infty} \sinh x = \infty \tag{3.2}$$

Similarly it can said that

$$\lim_{x\to-\infty} \cosh x = \lim_{x\to-\infty} -\sinh x = \infty \tag{3.3}$$

Plots of $\cosh x$ and $\sinh x$ are shown in Fig. 3.3. Note the complete loss of periodicity!

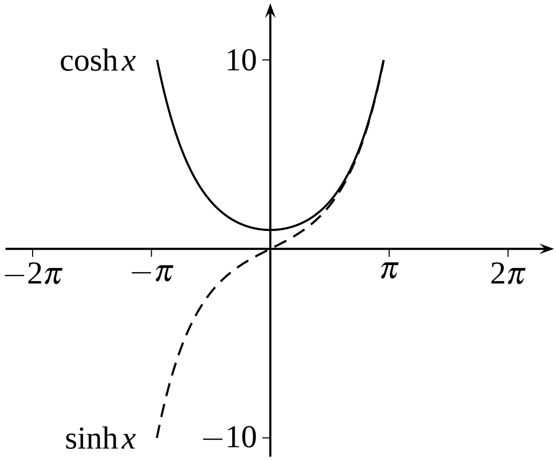


Figure 3.3: Plots of $\cosh x$ and $\sinh x$.

As for $\tanh x$. Recall that

$$\tanh x = \frac{\sinh x}{\cosh x}$$

and from Eqs. (3.1), (3.2) and (3.3) we have

$$\lim_{x\to 0} \tanh x = 0, \quad \lim_{x\to\infty} \tanh x = 1, \quad \lim_{x\to-\infty} \tanh x = -1 \tag{3.4}$$

A plot of $\tanh x$ is shown in Fig. 3.4.

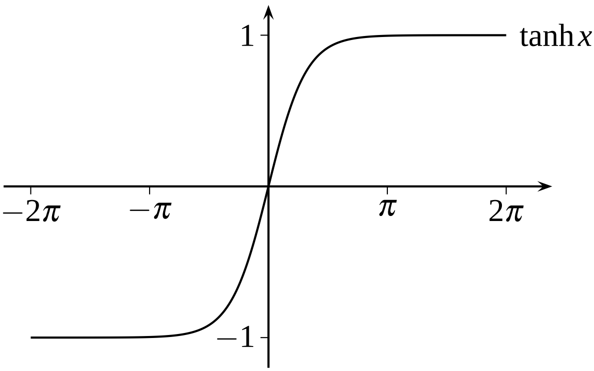


Figure 3.4: Plot of $\tanh x$.

3.5 Problem sheet

Problem 3.1 (see Worked Solution 3.1)

Consider:

a) $\cos^2 x + \sin^2 x = 1$

b) $\cos(A+B) = \cos A \cos B - \sin A \sin B$

c) $\sin(A+B) = \sin A \cos B + \cos A \sin B$

Derive equivalent results for hyperbolic functions.

Problem 3.2 (see Worked Solution 3.2)

Starting from: $\cosh x = \dfrac{e^x + e^{-x}}{2}$, $\sinh x = \dfrac{e^x - e^{-x}}{2}$

and $\tanh x = \dfrac{\sinh x}{\cosh x}$, show that the associated inverses are:

a) $\operatorname{arccosh} x = \ln\left(x \pm \sqrt{x^2 - 1}\right)$

b) $\operatorname{arcsinh} x = \ln\left(x \pm \sqrt{x^2 + 1}\right)$

c) $\operatorname{arctanh} x = \dfrac{1}{2}\ln\left(\dfrac{1+x}{1-x}\right)$

Problem 3.3 (see Worked Solution 3.3)

Derive expressions for a) the derivatives and b) the integrals of

(i) $\cosh x$ (ii) $\sinh x$ (iii) $\tanh x$

3.6 Worked solutions

Worked Solution 3.1 (see Problem 3.1)

Starting with the identities $\cosh x = \cos ix$ and $\sinh x = -i\sin ix$ it follows that

$\cosh ix = \cos x$

$\sinh ix = -i\sin -i^2 x = -i\sin x$

$\therefore i\sinh ix = -i^2 \sin x = \sin x$

a) $1 = \cos^2 z + \sin^2 z = \cosh^2 iz + (i\sinh iz)^2 = \cosh^2 iz - \sinh^2 iz$

Substituting $x = iz$ it follows that $\boxed{\cosh^2 x - \sinh^2 x = 1}$

b) $\cos(a+b) = \cos a \cos b - \sin a \sin b$

$\cosh(ia+ib) = \cosh ia \cosh ib - i^2 \sinh ia \sinh ib \therefore \cosh(ia+ib) = \cosh ia \cosh ib + \sinh ia \sinh ib$

Substituting $A = ia$ and $B = ib$ it follows that

$$\boxed{\cosh(A+B) = \cosh A \cosh B + \sinh A \sinh B}$$

c) $\sin(a+b) = \sin a \cos b + \cos a \sin b$

$i \sinh(ia+ib) = i \sin ia \cosh ib + i \cosh ia \sinh ib \therefore \sinh(ia+ib) = \sinh ia \cosh ib + \cosh ia \sinh ib$

Substituting $A = ia$ and $B = ib$ it follows that

$$\boxed{\sinh(A+B) = \sinh A \cosh B + \cosh A \sinh B}$$

Worked Solution 3.2 (see Problem 3.2)

a) $x = \cosh y$ and $y = \operatorname{arccosh} x$

It can also be said that $2x = 2\cosh y = e^y + e^{-y}$

Making the substitution $z = e^y$, this simplifies to $2x = z + z^{-1}$

from which it follows that $z^2 - 2xz + 1 = 0$

Solving for z then leads to $z = (2x \pm \sqrt{4x^2 - 4})/2$

from which it follows that $e^y = x \pm \sqrt{x^2 - 1}$

So finally $\boxed{\text{arccosh}x = \ln\left(x \pm \sqrt{x^2-1}\right)}$

b) $x = \sinh y$ and $y = \text{arcsinh}x$

It can also be said that $2x = 2\sinh y = e^y - e^{-y}$

Making the substitution $z = e^y$, this simplifies to $2x = z - z^{-1}$

from which it follows that $z^2 - 2xz - 1 = 0$

Solving for z then leads to $z = (2x \pm \sqrt{4x^2+4})/2$

from which it follows that $e^y = x \pm \sqrt{x^2+1}$

So finally $\boxed{\text{arcsinh}x = \ln\left(x \pm \sqrt{x^2+1}\right)}$

c) $x = \tanh y$ and $y = \text{arctanh}x$

It can also be said that $x = \dfrac{\sinh x}{\cosh x} = \dfrac{e^y - e^{-y}}{e^y + e^{-y}}$

Making the substitution $z = e^y$, this simplifies to $x(z + z^{-1}) = z - z^{-1}$

from which it follows that $z^2 = \dfrac{1+x}{1-x}$

and consequently $z = \sqrt{\dfrac{1+x}{1-x}}$

and finally $$\boxed{\operatorname{arctanh} x = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right)}$$

Worked Solution 3.3 (see Problem 3.3)

(i) Starting with $\cosh x = \dfrac{e^x + e^{-x}}{2}$

a) $\dfrac{d(\cosh x)}{dx} = \dfrac{e^x - e^{-x}}{2} = \sinh x$

b) $\displaystyle\int \cosh x dx = \frac{e^x - e^{-x}}{2} + C = \sinh x + C$

(ii) Starting with $\sinh x = \dfrac{e^x - e^{-x}}{2}$

a) $\dfrac{d(\sinh x)}{dx} = \dfrac{e^x + e^{-x}}{2} = \cosh x$

b) $\displaystyle\int \sinh x dx = \frac{e^x + e^{-x}}{2} + C = \cosh x + C$

(iii) Starting with $\tanh x = \dfrac{\sinh x}{\cosh x}$ and invoking the quotient rule

a) $\dfrac{d(\tanh x)}{dx} = \dfrac{1}{g^2}\left(g\dfrac{du}{dx} - u\dfrac{dg}{dx}\right) = \dfrac{1}{\cosh^2 x}\left(\cosh^2 x - \sinh^2 x\right)$

and invoking $\cosh^2 x - \sinh^2 x = 1$ leads to $\dfrac{d(\tanh x)}{dx} = \operatorname{sech}^2 x$

b) $\displaystyle\int \tanh x dx = \int \frac{\sinh x}{\cosh x} dx$

$u = \cosh x$ and $\dfrac{du}{dx} = \sinh x$

$$\therefore \int \tanh x dx = \int \frac{\sinh x}{u} \frac{dx}{du} du = \int \frac{1}{u} du = \ln u + C$$

$$\therefore \int \tanh x dx = \ln(\cosh x) + C$$

4

More differentiation

4.1 Learning outcomes

You should be able to:

- Derive equations for tangents and normals for points on curves.
- Mathematically locate maxima, minima and inflection points.
- Apply the principle of logarithmic differentiation.
- Differentiate inverse trigonometric functions.
- Differentiate inverse hyperbolic functions.

4.2 Tangents and normals

It is often desirable to determine algebraic expressions for the tangent and normal that pass through a point, P, situated on the function $y = f(x)$ at (x_1, y_1) (see Fig. 4.1).

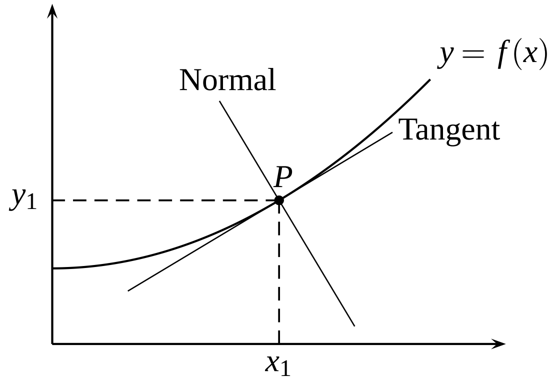


Figure 4.1: Schematic diagram showing a normal and tangent.

The tangent, y_t, at point P can be described by the linear equation

$$y_t = m_t x + c_t$$

It can be easily be realized that the slope, m_t, is defined by

$$m_t = \left.\frac{dy}{dx}\right|_{x=x_1}$$

We also know that at $x = x_1$, $y = y_1$ from which it follows that

$$y_1 = m_t x_1 + c_t$$

$$c_t = y_1 - m_t x_1$$

and

$$y_t = m_t(x - x_1) + y_1$$

The tangent, y_n, at point P can be described by the linear equation

$$y_n = m_n x + c_n$$

and recalling that $y = y_1$ at $x = x_1$, it can be said that

$$y_n = m_n(x - x_1) + y_1$$

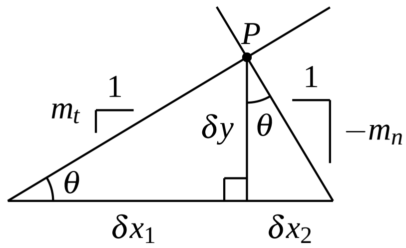


Figure 4.2: Schematic diagram showing a normal and tangent.

Now consider the geometry shown in Fig. 4.2. The gradients of the tangent and normal, m_t and m_n are defined by

$$m_t = \frac{\delta y}{\delta x_1}, \quad m_n = -\frac{\delta y}{\delta x_2}$$

But notice that

$$\tan\theta = \frac{\delta y}{\delta x_1} = \frac{\delta x_2}{\delta y}$$

from which it follows that

$$m_n = -\frac{1}{m_t}$$

and consequently

$$y_n = \frac{1}{m_t}(x_1 - x) + y_1$$

4.3 Location of maxima, minima and inflection

It has already been discussed how differentiation can be used to find gradients. Differentiation can also be used to find minima, maxima and inflection points.

Fig. 4.3 shows some results from differentiating

$$y = 4x - x^2 + x^3/15 \tag{4.1}$$

Note that y' and y'' are the first and second derivatives with respect to x.

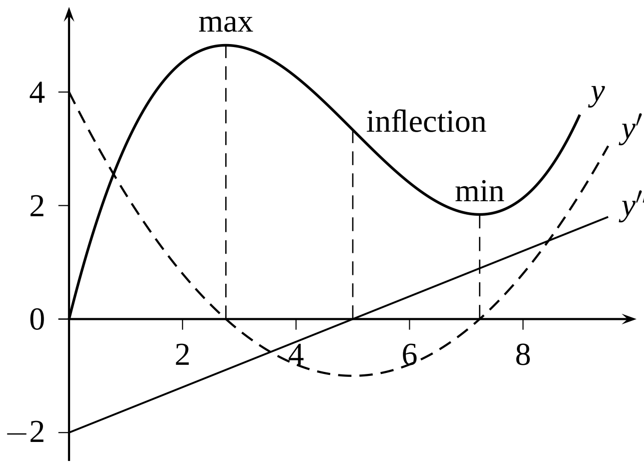


Figure 4.3: Schematic diagram illustrating the meaning of maxima, minima and inflection.

Maxima are points where the gradient is changing from positive to negative. Minima are points where the gradient is changing from negative to positive. As the gradient changes sign it must pass through zero. Therefore, maxima and minima are referred to as stationary points as they refer to points where the gradient is exactly zero.

It follows that maxima and minima points of a function $y = f(x)$ can be found where the first derivatives are zero. Consider again Eq. (4.1). Differentiating with respect to x leads to

$$y' \equiv \frac{dy}{dx} = 4 - 2x + 3x^2/15 \tag{4.2}$$

setting $y' = 0$ leads to the quadratic equation

$$0 = 4 - 2x + 3x^2/15$$

It follows that stationary points exist at

$$x = 5 \pm \sqrt{5} = 2.7639 \text{ or } 7.2361$$

As stated earlier, maxima and minima are points where the gradient is changing from positive to negative and negative to positive, respectively. It follows that the gradient of the gradient (i.e. the second derivative, y'') should be negative at maxima and positive at minima.

Consider the derivative of Eq. (4.2)

$$y'' \equiv \frac{d^2y}{dx^2} = -2 + 6x/15 \tag{4.3}$$

It can now be seen that

$$y''(x = 2.76) = -0.89 \text{ and } y''(x = 7.24) = 0.89$$

from which it can be concluded that $x = 2.76$ is the maximum and $x = 7.24$ is the minimum (compare with the figure above).

Another point of interest is the point of inflection. This exists where the gradient is stationary. It follows that this is found where $y'' = 0$. Consider Eq. (4.3). Setting $y'' = 0$ and solving for x leads to $x = 5$.

4.4 Logarithmic differentiation

We have seen earlier how products of two functions can be easily differentiated using the product rule. With more than two functions we can use nested applications of the product rule. For example, for $y = uvw$ we can say:

$$\begin{aligned}\frac{d(uvw)}{dx} &= vw\frac{du}{dx} + u\frac{d(vw)}{dx} \\ &= vw\frac{du}{dx} + uw\frac{dv}{dx} + uv\frac{dw}{dx} \\ &= uvw\left(\frac{1}{u}\frac{du}{dx} + \frac{1}{v}\frac{dv}{dx} + \frac{1}{w}\frac{dw}{dx}\right)\end{aligned}$$

However, this can become quite tedious with more functions in the products, especially when some are denominators. A more general method is to take logs on both side.

Consider again:

$$y = uvw$$

Taking logs of both sides leads to

$$\ln y = \ln(uvw) = \ln u + \ln v + \ln w \tag{4.4}$$

Now consider

$$y = \ln f \tag{4.5}$$

Applying the chain rule

$$\frac{dy}{dx} = \frac{dy}{dx}\frac{df}{dx}\frac{dx}{df} = \frac{dy}{df}\frac{df}{dx} \tag{4.6}$$

Now from Eq. (4.5), we have that

$$\frac{dy}{df} = \frac{1}{f}$$

and from Eq. (4.6), we have

$$\boxed{\frac{d(\ln f)}{dx} = \frac{1}{f}\frac{df}{dx}} \tag{4.7}$$

from which it follows that differentiating Eq. (4.4) with respect to x leads to

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{u}\frac{du}{dx} + \frac{1}{v}\frac{dv}{dx} + \frac{1}{w}\frac{dw}{dx}$$

and recalling that $y = uvw$ we have

$$\frac{d(uvw)}{dx} = uvw\left(\frac{1}{u}\frac{du}{dx} + \frac{1}{v}\frac{dv}{dx} + \frac{1}{w}\frac{dw}{dx}\right)$$

Derivation of the quotient rule

Now consider

$$y = \frac{u}{v}$$

Taking logs on both sides

$$\ln y = \ln u - \ln v$$

Differentiating both sides and recalling Eq. (4.7) yields

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{u}\frac{du}{dx} - \frac{1}{v}\frac{dv}{dx}$$

and recalling that $y = u/v$ leads to

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{u}{v}\left(\frac{1}{u}\frac{du}{dx} - \frac{1}{v}\frac{dv}{dx}\right)$$

which on rearranging, finally yields the quotient rule

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{1}{v^2}\left(v\frac{du}{dx} - u\frac{dv}{dx}\right)$$

Some worked examples

Challenge 4.1 Differentiate $\dfrac{x^2 \sin x}{\cos 2x}$ with respect to x.

Let $y = \dfrac{x^2 \sin x}{\cos 2x}$

Taking logs on both sides

$$\ln y = 2\ln x + \ln(\sin x) - \ln(\cos 2x)$$

Differentiating both sides leads to

$$\frac{1}{y}\frac{dy}{dx} = \frac{2}{x} + \frac{\cos x}{\sin x} + \frac{2\sin 2x}{\cos 2x}$$

and therefore

$$\frac{dy}{dx} = \frac{x^2 \sin x}{\cos 2x}\left(\frac{2}{x} + \cot x + 2\tan 2x\right)$$

Challenge 4.2 Differentiate $\dfrac{e^{4x}}{x^3 \cosh 2x}$ with respect to x.

Let $y = \dfrac{e^{4x}}{x^3 \cosh 2x}$

Taking logs on both sides

$$\ln y = 4x - 3\ln(x) - \ln(\cosh 2x)$$

Differentiating both sides leads to

$$\frac{1}{y}\frac{dy}{dx} = 4 - \frac{3}{x} - 2\tanh 2x$$

and therefore

$$\frac{dy}{dx} = \frac{e^{4x}}{x^3 \cosh 2x}\left(4 - \frac{3}{x} - 2\tanh 2x\right)$$

4.5 Differentiating inverse functions

In Chapter 1 we looked at how to obtain the derivative of $\ln x$. Recall

$$y = \ln x \quad \Rightarrow \quad x = e^y$$

Differentiating with respect to y

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

Multiplying both sides by dy/dx and dividing both sides by x then yields

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

Hence it is found that

$$\frac{d(\ln x)}{dx} = \frac{1}{x}$$

Note that $\ln x$ is the inverse of e^x. Exactly the same procedure can be used for finding derivatives for many other inverse functions.

Consider
$$y = \arccos x \quad \Rightarrow \quad x = \cos y$$
Differentiating with respect to y
$$\frac{dx}{dy} = -\sin y$$
Multiplying both sides by dy/dx and dividing both sides by $\sin x$ then yields
$$\frac{dy}{dx} = \frac{d(\arccos x)}{dx} = \frac{-1}{\sin y}$$
But it is possible to simplify this further. Recall that
$$\cos^2 y + \sin^2 y = 1$$
It follows that
$$\sin y = \pm\sqrt{1-\cos^2 y}$$
Furthermore, in this case, $\cos y = x$. Therefore we can say
$$\frac{d(\arccos x)}{dx} = \frac{\mp 1}{\sqrt{1-x^2}}$$

4.6 Problem sheet

Problem 4.1 (see Worked Solution 4.1)

Derive the derivatives of:

a) $\arccos x$ b) $\arcsin x$ c) $\arctan x$

d) $\operatorname{arccosh} x$ e) $\operatorname{arcsinh} x$ f) $\operatorname{arctanh} x$

Problem 4.2 (see Worked Solution 4.2)

Differentiate the following:

a) $\dfrac{e^{3x}\cos 2x}{\sin x}$ b) $\dfrac{x^3\cos(x+2)}{e^{3x}}$ c) $\ln\left(\dfrac{1+x^2}{1-x^2}\right)$

d) $\arctan\left(\dfrac{e^x\sqrt{1-x^2}}{\cosh 3x}\right)$

Problem 4.3 (see Worked Solution 4.3)

a) Locate, mathematically, the maxima, minima and inflection points of

$y = x^3 - 6x^2 + 2x + 4$

b) Derive equations for the tangent and normal at the inflection point.

c) Present your results as a sketched graph.

4.7 Worked solutions

Worked Solution 4.1 (see Problem 4.1)

a) $y = \arccos x$ therefore $x = \cos y$ and $\dfrac{dx}{dy} = -\sin y$

Recall that $\cos^2 x + \sin^2 x = 1$

$$\therefore \frac{d(\arccos x)}{dx} \equiv \frac{dy}{dx} = \frac{-1}{\sin y} = \frac{\mp 1}{\sqrt{\sin^2 y}} = \frac{\mp 1}{\sqrt{1-\cos^2 y}} = \frac{\mp 1}{\sqrt{1-x^2}}$$

b) $y = \arcsin x$ therefore $x = \sin y$ and $\dfrac{dx}{dy} = \cos y$

$$\therefore \frac{d(\arcsin x)}{dx} \equiv \frac{dy}{dx} = \frac{1}{\cos y} = \frac{\pm 1}{\sqrt{\cos^2 y}} = \frac{\pm 1}{\sqrt{1-\sin^2 y}} = \frac{\pm 1}{\sqrt{1-x^2}}$$

c) $y = \arctan x$ therefore $x = \tan y$ and $\dfrac{dx}{dy} = \sec^2 y$

Now $\cos^2 y + \sin^2 y = 1$ therefore $1 + \tan^2 y = \sec^2 y$

$$\therefore \frac{d(\arctan x)}{dx} \equiv \frac{dy}{dx} = \frac{1}{\sec^2 y} = \frac{1}{1+\tan y^2} = \frac{1}{1+x^2}$$

d) $y = \mathrm{arccosh} x$ therefore $x = \cosh y$ and $\dfrac{dx}{dy} = \sinh y$

Recall that $\cosh^2 x - \sinh^2 x = 1$

$$\therefore \frac{d(\mathrm{arccosh} x)}{dx} \equiv \frac{dy}{dx} = \frac{\pm 1}{\sqrt{\sinh^2 y}} = \frac{\pm 1}{\sqrt{\cosh^2 y - 1}} = \frac{\pm 1}{\sqrt{x^2-1}}$$

e) $y = \mathrm{arcsinh} x$ therefore $x = \sinh y$ and $\dfrac{dx}{dy} = \cosh y$

$$\therefore \frac{d(\mathrm{arcsinh} x)}{dx} \equiv \frac{dy}{dx} = \frac{\pm 1}{\sqrt{\cosh^2 y}} = \frac{\pm 1}{\sqrt{1+\sinh^2 y}} = \frac{\pm 1}{\sqrt{1+x^2}}$$

f) $y = \mathrm{arctanh} x$ therefore $x = \tanh y$ and $\dfrac{dx}{dy} = \mathrm{sech}^2 y$

Now $\cosh^2 y - \sinh^2 y = 1$ therefore $1 - \tanh^2 y = \mathrm{sech}^2 y$

$$\therefore \frac{d(\mathrm{arctanh} x)}{dx} \equiv \frac{dy}{dx} = \frac{1}{\mathrm{sech}^2 y} = \frac{1}{1-\tanh^2 y} = \frac{1}{1-x^2}$$

Worked Solution 4.2 (see Problem 4.2)

a) $y = \dfrac{e^{3x}\cos 2x}{\sin x}$

$\ln y = \ln \dfrac{e^{3x}\cos 2x}{\sin x} = 3x + \ln(\cos 2x) - \ln(\sin x)$

$\dfrac{d(\ln y)}{dx} = \dfrac{1}{y}\dfrac{dy}{dx} = 3 - \dfrac{2\sin 2x}{\cos 2x} - \dfrac{\cos x}{\sin x} = 3 - 2\tan 2x - \cot x$

$\therefore \dfrac{dy}{dx} = \dfrac{e^{3x}\cos 2x(3 - 2\tan 2x - \cot x)}{\sin x}$

b) $y = \dfrac{x^3\cos(x+2)}{e^{3x}}$

$\ln y = 3\ln x + \ln[\cos(x+2)] - 3x$

$\dfrac{1}{y}\dfrac{dy}{dx} = \dfrac{3}{x} - \tan(x+2) - 3$

$\therefore \dfrac{dy}{dx} = \dfrac{x^3\cos(x+2)}{e^{3x}}\left(\dfrac{3}{x} - \tan(x+2) - 3\right)$

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

$\dfrac{dy}{dx} = \dfrac{2x}{1+x^2} + \dfrac{2x}{1-x^2} = \dfrac{2x - 2x^3 + 2x + 2x^3}{1-x^4} = \dfrac{4x}{1-x^4}$

d) $y = \arctan\left(\dfrac{e^x\sqrt{1-x^2}}{\cosh 3x}\right) = \arctan u$

$u = \dfrac{e^x\sqrt{1-x^2}}{\cosh 3x}$

$\ln u = x + \dfrac{1}{2}\ln(1-x^2) - \ln(\cosh 3x)$

$$\frac{du}{dx} = \frac{e^x\sqrt{1-x^2}}{\cosh 3x}\left(1 - \frac{x}{1-x^2} - 3\tanh 3x\right)$$
$$\frac{dy}{du} = \frac{1}{1+u^2}$$
$$\frac{dy}{dx} = \frac{du}{dx}\frac{dx}{du}\frac{dy}{dx} = \frac{du}{dx}\frac{dy}{du}$$
$$\therefore \frac{dy}{dx} = \frac{u}{1+u^2}\left(1 - \frac{x}{1-x^2} - 3\tanh 3x\right)$$

Worked Solution 4.3 (see Problem 4.3)

a) $y = x^3 - 6x^2 + 2x + 4$

$$\frac{dy}{dx} = 3x^2 - 12x + 2 \quad \Rightarrow 0 = x^2 - 4x + \frac{2}{3} \quad \Rightarrow x = 2 \pm \sqrt{4 - 2/3}$$
$$\frac{d^2y}{dx^2} = 6x - 12 \quad \Rightarrow 0 = x - 2 \quad \Rightarrow x = 2$$
$$\left.\frac{d^2y}{dx^2}\right|_{x=2+\sqrt{4-2/3}} = 6(2 + \sqrt{4-2/3}) - 12 = 10.95 \text{ (min)}$$
$$\left.\frac{d^2y}{dx^2}\right|_{x=2-\sqrt{4-2/3}} = 6(2 - \sqrt{4-2/3}) - 12 = -10.95 \text{ (max)}$$

From which it follows that:

Maximum: $x = 2 - \sqrt{4 - 2/3}$

Minimum: $x = 2 + \sqrt{4 - 2/3}$

Inflection: $x = 2$

b) Tangent at inflection: $y_{tan} = \left.\frac{dy}{dx}\right|_{x=2}(x-2) + y(x=2)$

$\therefore y_{tan} = (3 \times 4 - 12 \times 2 + 2)(x - 2) + 2^3 - 6 \times 2^2 + 2 \times 2 + 4$

$\therefore y_{tan} = -10(x - 2) - 8$ and $y_{nor} = 0.1(x - 2) - 8$

5

More integration

5.1 Learning outcomes

You should be able to:

- List previous differentiation results as integrals.
- Understand that $\int f'/f\,dx = \ln f + C$.
- Understand that $\int f' f\,dx = f^2/2 + C$.
- Understand and apply integration by parts.
- Integrate formulae containing trigonometric functions.
- Integrate formulae containing hyperbolic functions.

5.2 How is our table integrals looking?

We have now established an extensive list of derivatives and corresponding integrals (see Table 5.1).

Table 5.1: Table of derivatives and integrals.

y	dy/dx	y	$\int y dx$
x^n	nx^{n-1}	x^n	$(n+1)^{-1}x^{n+1}+C$
e^x	e^x	e^x	e^x+C
$\ln x$	x^{-1}	x^{-1}	$\ln x+C$
$\sin x$	$\cos x$	$\cos x$	$\sin x+C$
$\cos x$	$-\sin x$	$\sin x$	$-\cos x+C$
$\tan x$	$\sec^2 x$	$\sec^2 x$	$\tan x+C$
$\sinh x$	$\cosh x$	$\cosh x$	$\sinh x+C$
$\cosh x$	$\sinh x$	$\sinh x$	$\cosh x+C$
$\tanh x$	$\text{sech}^2 x$	$\text{sech}^2 x$	$\tanh x+C$
$\arcsin x$	$(1-x^2)^{-1/2}$	$(1-x^2)^{-1/2}$	$\arcsin x+C$
$\arccos x$	$-(1-x^2)^{-1/2}$	$-(1-x^2)^{-1/2}$	$\arccos x+C$
$\arctan x$	$(1+x^2)^{-1}$	$(1+x^2)^{-1}$	$\arctan x+C$
arcsinh x	$(1+x^2)^{-1/2}$	$(1+x^2)^{-1/2}$	arcsinh x $+C$
arccosh x	$(x^2-1)^{-1/2}$	$(x^2-1)^{-1/2}$	arccosh x $+C$
arctanh x	$(1-x^2)^{-1}$	$(1-x^2)^{-1}$	arctanh x $+C$
		$\tan x$	$\ln\sec x+C$
		$\tanh x$	$\ln\cosh x+C$

5.3 Integrals of the form $\int f'/f$

Recall the integration of $\tan x$.

$$F = \int \tan x dx = \int \frac{\sin x}{\cos x} dx$$

Our strategy involved making the substitution $f = \cos x$ such that

$$F = \int \frac{\sin x}{f} dx$$

Applying the chain rule then leads to

$$F = \int \frac{\sin x}{f} \frac{dx}{df} df$$

Of course we know that $df/dx = -\sin x$. Therefore it can be said that

$$F = -\int \frac{1}{f} df = -\ln f + C = \ln \sec x + C$$

What we have benefited from is the fact that $\sin x$ is a linear function of the derivative of $\cos x$. Namely, $\sin x = -d\cos x/dx$.

A more general way of highlighting this advantage is as follows. Consider

$$\int \frac{f'}{f} dx$$

where recall that $f' = df/dx$.

Applying the chain rule it can be seen that

$$\boxed{\int \frac{f'}{f} dx = \int \frac{1}{f} df = \ln f + C} \tag{5.1}$$

So look out for when a function can be separated out into the form $f'f$. For example:

$$F = \int \frac{2x+3}{x^2+3x-5} dx$$

Note that

$$f = x^2 + 3x - 5 \text{ and } f' = 2x + 3$$

Therefore from Eq. (5.1)

$$F = \ln(x^2 + 3x + 5) + C$$

5.4 Integrals of the form $\int f'f$

Another related form is

$$\boxed{\int f' f dx = \int f' f \frac{dx}{df} dx = \int f df = \frac{f^2}{2} + C}$$

So look out for when a function can be separated out into the form f'/f. For example:

$$F = \int (2x+3)(x^2+3x-5)dx = \frac{(x^2+3x-5)^2}{2} + C$$

5.5 Integration by parts

Recall the product rule

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

Integrating both sides with respect to x leads to

$$uv = \int u\frac{dv}{dx}dx + \int v\frac{du}{dx}dx$$

and rearranging gives

$$\boxed{\int u\frac{dv}{dx}dx = uv - \int v\frac{du}{dx}dx} \tag{5.2}$$

The above result can be particularly useful for integrating products. Application of Eq. (5.2) is often referred to as “integration by parts”.

For example:

$$F = \int x^2 \ln x dx$$

The first step is to choose which part is u and which is v'. If we set $u = x^2$ we to integrate $\ln x$. Alternatively, if we set $u = \ln x$, we have to integrate x^2. The latter is obviously preferable. So, setting $u = \ln x$ and $v' = x^2$ and noting that $u' = 1/x$ and $v = x^3/3$ we have, from Eq. (5.2)

$$F = \frac{x^3 \ln x}{3} - \int \frac{x^2}{3} dx = \frac{x^3}{3}\left(\ln x - \frac{1}{3}\right) + C$$

5.6 Integrating trigonometric and hyperbolic functions

We already know the integrals for $\cos x$ and $\sin x$, but how do we deal with problems such as below?

$$F = \int \sin^2 x dx \tag{5.3}$$

We need to apply our trigonometric and hyperbolic identities:

$$\cos^2 x + \sin^2 x = 1$$

$$\cos(A+B) = \cos A \cos B - \sin A \sin B$$
$$\sin(A+B) = \cos A \sin B + \sin A \cos B$$
$$\cosh^2 x - \sinh^2 x = 1$$
$$\cosh(A+B) = \cosh A \cosh B + \sinh A \sinh B$$
$$\sinh(A+B) = \cosh A \sinh B + \sinh A \cosh B$$

Using the above we can see that for Eq. (5.3) we can benefit from setting $A = B = x$ to get

$$\cos 2x = \cos^2 x - \sin^2 x = 1 - 2\sin^2 x = -1 + 2\cos^2 x$$

from which it follows that $\sin^2 x = (1-\cos 2x)/2$. Therefore we can say that

$$\int \sin^2 x dx = \int \frac{1-\cos 2x}{2} dx = \frac{x}{2} - \frac{\sin 2x}{4} + C$$

Similarly we can say that

$$\int \cos^2 x dx = \int \frac{1+\cos 2x}{2} dx = \frac{x}{2} + \frac{\sin 2x}{4} + C$$

Challenge 5.1 Evaluate $\int \sin^3 x dx$

Let

$$F = \int \sin^3 x dx = \int \sin x \sin^2 x dx = \int \sin x dx - \int \sin x \cos^2 x dx$$

In the same way that we can say

$$\int f f' dx = \frac{f^2}{2} + C$$

it can also be said that

$$\int f^2 f' dx = \frac{f^3}{3} + C$$

Therefore we have

$$F = -\cos x + \frac{\cos^3 x}{3} + C \tag{5.4}$$

Challenge 5.2 Evaluate $\int \sinh ax \cosh bx dx$

Let

$$F = \int \sinh ax \cosh bx dx$$

Recall that

$$\sinh(A+B) = \cosh A \sinh B + \sinh A \cosh B$$

It follows that

$$\sinh(A-B) = -\cosh A \sinh B + \sinh A \cosh B$$

and

$$\sinh(A+B) + \sinh(A-B) = 2\sinh A \cosh B$$

Therefore we can say that

$$F = \frac{1}{2}\int \sinh[(a+b)x] + \sinh[(a-b)x] dx$$

which is easily evaluated as

$$F = \frac{\cosh[(a+b)x]}{2(a+b)} + \frac{\cosh[(a-b)x]}{2(a-b)} + C$$

5.7 Problem sheet

Problem 5.1 (see Worked Solution 5.1)

Integrate the following

a) $\cos^3 x$ b) $\cos^5 x$ c) $\dfrac{\sin 2x}{1+\cos^2 x}$

Problem 5.2 (see Worked Solution 5.2)

Integrate the following

a) $\cosh^3 x$ b) $\sinh^5 x$ c) $\sinh 2x \cosh 3x$

Problem 5.3 (see Worked Solution 5.3)

Integrate the following

a) $x^2 e^{3x}$ b) $e^{3x}\sin x$ c) $x^3\ln(x+4)$

Problem 5.4 (see Worked Solution 5.4)

Integrate the following

a) $\dfrac{3x^2+5}{x^3+5x-7}$ b) $\dfrac{\cos x}{1+\sin x}$ c) $(x^2-3x+6)(2x-3)$

You will benefit from remembering:

$\cos^2 x+\sin^2 x=1$

$\cos(A+B)=\cos A\cos B-\sin A\sin B$

$\sin(A+B)=\sin A\cos B+\cos A\sin B$

5.8 Worked solutions

Worked Solution 5.1 (see Problem 5.1)

a) $F=\displaystyle\int\cos^3 x dx=\int\cos x\cos^2 x dx=\int\cos x(1-\sin^2 x)dx$

$u=\sin x$ and $\dfrac{du}{dx}=\cos x$

$\therefore F=\displaystyle\int\cos x(1-\sin^2 x)\frac{1}{\cos x}du=\int(1-u^2)du$

$$\therefore F = u - \frac{u^3}{3} + C = \sin x - \frac{\sin^3 x}{3} + C$$

b) $F = \int \cos^5 x dx = \int \cos x \cos^4 x dx = \int \cos x (1 - \sin^2 x)^2 dx$

$u = \sin x \qquad \text{and} \qquad \frac{du}{dx} = \cos x$

$$\therefore F = \int \cos x (1 - \sin^2 x)^2 \frac{1}{\cos x} du = \int 1 - 2u^2 + u^4 du$$

$$\therefore F = u - \frac{2u^3}{3} + \frac{u^5}{5} + C = \sin x - \frac{2\sin^3 x}{3} + \frac{\sin^5 x}{5} + C$$

c) $F = \int \frac{\sin 2x}{1 + \cos^2 x} dx$

$\cos(A + B) = \cos A \cos B - \sin A \sin B$

$\therefore \cos 2x = \cos^2 x - \sin^2 x \quad \Rightarrow \quad 1 + \cos^2 x = (3 + \cos 2x)/2$

$$\therefore F = \int \frac{2\sin 2x}{3 + \cos 2x} dx$$

$u = \cos 2x \qquad \text{and} \qquad \frac{du}{dx} = -2\sin 2x$

$$\therefore F = \int -\frac{1}{3+u} du = -\ln(3 + u) + C = -\ln[3 + \cos 2x] + C$$

Another way is as follows:

Now $\sin(A + B) = \sin A \cos B + \cos A \sin B$

$\therefore \sin 2x = 2 \sin x \cos x$

If $u = \cos^2 x$, $\dfrac{du}{dx} = -2\sin x \cos x = -\sin 2x$

$$\therefore F = -\int \frac{1}{(1+u)} du = -\ln(1+\cos^2 x) + D$$

Recall that $1+\cos^2 x = (3+\cos 2x)/2$

$$\therefore -\ln(1+\cos^2 x) + D = -\ln(3+\cos 2x) + \ln 2 + D$$

Worked Solution 5.2 (see Problem 5.2)

a) $F = \int \cosh^3 x dx = \int \cosh x \cosh^2 x dx = \int \cosh x(1+\sinh^2 x) dx$

$u = \sinh x$ and $\dfrac{du}{dx} = \cosh x$

$$F = \int (1+u^2) du = u + \frac{u^3}{3} + C = \sinh x + \frac{\sinh^3 x}{3} + C$$

b) $F = \int \sinh^5 x dx = \int \sinh x \sinh^4 x dx = \int \sinh x(\cosh^2 x - 1)^2 dx$

$u = \cosh x$ and $\dfrac{du}{dx} = \sinh x$

$$F = \int (1-2u^2+u^4) du = u - \frac{2u^3}{3} + \frac{u^5}{5} + C$$

$$\therefore F = \cosh x - \frac{2\cosh^3 x}{3} + \frac{\cosh^5 x}{5} + C$$

c) $F = \int \sinh 2x \cosh 3x dx$

$$\sinh(A+B) = \sinh A \cosh B + \cosh A \sinh B$$

$$\therefore \sinh(A-B) = \sinh A \cosh B - \cosh A \sinh B$$

$$\therefore \sinh(A+B)+\sinh(A-B)=2\sinh A\cosh B$$

$$\therefore F=\int\frac{\sinh 5x-\sinh x}{2}dx=\frac{\cosh 5x}{10}-\frac{\cosh x}{2}+C$$

Worked Solution 5.3 (see Problem 5.3)

a) $F=\int x^2e^{3x}dx$

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

$$\therefore F=\frac{x^2e^{3x}}{3}-\int\frac{2xe^{3x}}{3}dx=\frac{x^2e^{3x}}{3}-\frac{2xe^{3x}}{9}+\int\frac{2e^{3x}}{9}dx$$

$$\therefore F=\frac{x^2e^{3x}}{3}-\frac{2xe^{3x}}{9}+\frac{2e^{3x}}{27}+C=\frac{e^{3x}}{3}\left(x^2-\frac{2x}{3}+\frac{2}{9}\right)+C$$

b) $F=\int e^{3x}\sin x dx$

$$F=-e^{3x}\cos x+3\int e^{3x}\cos x dx$$

$$F=-e^{3x}\cos x+3e^{3x}\sin x-9\int e^{3x}\sin x dx$$

$$F=-e^{3x}\cos x+3e^{3x}\sin x-9F$$

$$\therefore F=\frac{e^{3x}(3\sin x-\cos x)}{10}+C$$

c) $F=\int x^3\ln(x+4)dx$

$$u=x+4 \qquad \frac{du}{dx}=1 \qquad x^3=(u-4)^3$$

$$\therefore F = \int (u-4)^3 \ln u du$$

$$\therefore F = \frac{(u-4)^4}{4} \ln u - A$$

where

$$A = \int \frac{(u-4)^4}{4u} du$$

Now $(u-4)^4 = u^4 - 16u^3 + 96u^2 - 256u + 256$

and $\dfrac{(u-4)^4}{4u} = \dfrac{u^3}{4} - 4u^2 + 24u - 64 + \dfrac{64}{u}$

$$\therefore A = \frac{u^4}{16} - \frac{4u^3}{3} + 12u^2 - 64u + 64 \ln u + C$$

Worked Solution 5.4 (see Problem 5.4)

a) $\displaystyle\int \frac{3x^2+5}{x^3+5x-7} dx = \ln(x^3+5x-7) + C$

b) $\displaystyle\int \frac{\cos x}{1+\sin x} dx = \ln(1+\sin x) + C$

c) $\displaystyle\int (x^2-3x+6)(2x-3) dx = \frac{(x^2-3x+6)^2}{2} + C$

6

First-order differential equations

6.1 Learning outcomes

You should be able to:

- Formulate differential equations.
- Solve first-order differential equations using direct integration.
- Solve first-order differential equations using separation of variables.
- Solve homogenous first-order differential equations using $y = vx$.

6.2 The order of differential equations

A differential equation is a relationship between an independent variable, x, a dependent variable, y, and one or more derivatives of y with respect to x, e.g.

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

The order of a differential equation is given by the highest derivative involved in the equation:

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

is a first-order equation

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

is a second-order equation

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

is a third-order equation, and so on.

Another useful distinction is whether an equation is linear or non-linear. A linear differential equation is linear in the dependent variable, e.g.

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

Note that although the equation above is non-linear in x, it is linear in y, which is the dependent variable. Therefore the above equation is considered to be a linear equation. It is linear in y because none of the terms contain powers of y other than y^0 and y^1.

In contrast, the following equations are non-linear in y:

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

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

$$\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2\sin y = x^2$$

Linear equations are generally a lot easier to deal with.

6.3 Formation of differential equations

Differential equations are often formed to describe various quantitative situations. Mathematically, they can be formed by continuous differentiation of particular functions.

Challenge 6.1 Determine the underlying ordinary differential equation associated with $y = Ax^2 + Bx$.

Let

$$y = Ax^2 + Bx \tag{6.1}$$

Differentiating with respect to x leads to

$$\frac{dy}{dx} = 2Ax + B \tag{6.2}$$

and again

$$\frac{d^2y}{dx^2} = 2A \tag{6.3}$$

Substituting Eq. (6.3) into Eq. (6.2) yields

$$B = \frac{dy}{dx} - x\frac{d^2y}{dx^2} \tag{6.4}$$

which on substitution into Eq. (6.1) along with Eq. (6.3) yields the second-order differential equation

$$\frac{x^2}{2}\frac{d^2y}{dx^2} - x\frac{dy}{dx} + y = 0 \tag{6.5}$$

Challenge 6.2 Determine the underlying ordinary differential equation associated with $y = Ae^x$.

Let

$$y = Ae^x \tag{6.6}$$

Differentiating with respect to x leads to

$$\frac{dy}{dx} = Ae^x$$

which on substitution back into Eq. (6.6) leads to the first-order differential equation

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

Challenge 6.3 Determine the underlying ordinary differential equation associated with $y = Ae^x + Be^{-x}$.

Let

$$y = Ae^x + Be^{-x} \tag{6.7}$$

Differentiating with respect to x leads to

$$\frac{dy}{dx} = Ae^x - Be^{-x}$$

and differentiating again leads to

$$\frac{d^2y}{dx^2} = Ae^x + Be^{-x}$$

which on substitution back into Eq. (6.7) yields the second-order differential equation

$$\frac{d^2y}{dx^2} = y$$

Now consider

$$y = C\cosh x + D\sinh x \tag{6.8}$$

Differentiating with respect to x leads to

$$\frac{dy}{dx} = C\sinh x + D\cosh x$$

and differentiating again leads to

$$\frac{d^2y}{dx^x} = C\cosh x + D\sinh x$$

which on substitution back into Eq. (6.8) also yields the second-order differential equation

$$\frac{d^2y}{dx^2} = y \tag{6.9}$$

So Eqs. (6.7) and (6.8) are both general solutions to the differential equation in Eq. (6.9).

From the above, we have seen that it is relatively straightforward to formulate the differential equation associated with a given solution. In what follows, we will examine various techniques for reversing this process, i.e., deriving solutions to differential equations. This term we will focus on first-order differential equations. Deriving solutions to second-order equations is reserved for later in the following term.

6.4 Direct integration

The simplest technique is direct integration.

Challenge 6.4 Solve the ordinary differential equation

$$\frac{dy}{dx} = 6x^2 - 6x + 7$$

Integrating both sides with respect to x

$$\int \frac{dy}{dx} dx = \int 6x^2 - 6x + 7 dx$$

yields the general solution

$$y = 2x^3 - 3x^2 + 7x + C$$

Challenge 6.5 Solve the ordinary differential equation

$$2x\frac{dy}{dx} = 6x^3 + 5$$

Make the derivative the subject of the formula

$$\frac{dy}{dx} = 3x^2 + \frac{5}{2x}$$

Integrating both sides with respect to x

$$\int \frac{dy}{dx} dx = \int 3x^2 + \frac{5}{2x} dx$$

yields the general solution

$$y = x^3 + \frac{5}{2}\ln x + C$$

6.5 Separation of variables

All pretty easy until you end up with something like

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

Integrating both sides with respect to x leads to

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

So what can we do with the y term in the denominator?

The trick is to separate the variables on to either side and apply the chain rule.

Challenge 6.6 Solve the ordinary differential equation

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

First separate all the y-factors to the LHS and the x-factors to the RHS such that

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

Integrating both sides with respect to x then leads to

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

Consideration of the chain rule allows the above to simplify to get (recall integration by substitution)

$$\int (y+1) dy = \int 3x^2 dx$$

Evaluating the integrals on both sides then leads to

$$\frac{y^2}{2} + y = x^3 + C$$

which is a quadratic. Applying the quadratic formula finally yields

$$y = -1 \pm \sqrt{1 + 2(x^3 + C)}$$

Challenge 6.7 Solve the ordinary differential equation

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

First separate all the y-factors to the LHS and the x-factors to the RHS such that

$$\frac{1}{y-1}\frac{dy}{dx} = x - 1$$

Integrating both sides with respect to x then leads to

$$\int \frac{1}{y-1}\frac{dy}{dx}dx = \int (x-1)dx$$

Evaluating the integrals on both sides then leads to

$$\ln(y-1) = \frac{x^2}{2} - x + C$$

and solving for y we have

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

where $A = e^C$.

Challenge 6.8 Solve the ordinary differential equation

$$\frac{dy}{dx} = \frac{y-1}{x-1}$$

First separate all the y-factors to the LHS and the x-factors to the RHS such that

$$\frac{1}{y-1}\frac{dy}{dx} = \frac{1}{x-1}$$

Integrating both sides with respect to x then leads to

$$\int \frac{1}{y-1}\frac{dy}{dx}dx = \int \frac{1}{x-1}dx$$

Evaluating the integrals on both sides then leads to

$$\ln(y-1) = \ln(x-1) + C = \ln[A(x-1)]$$

where $A = e^C$.

Solving for y we have

$$y = A(x-1) + 1$$

Challenge 6.9 Solve the ordinary differential equation

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

First separate all the y-factors to the LHS and the x-factors to the RHS such that

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

Integrating both sides with respect to x then leads to

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

which expands to get

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

Evaluating the integrals then yields

$$\ln y + \frac{1}{y} = \ln x - \frac{1}{x} + C$$

Note that here it is not possible to get an explicit solution for y or x so evaluation would need to be by an iterative process. This is a common problem with solutions to non-linear differential equations.

6.6 Homogenous equations

But what happens if you end up with something like

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

It is not possible to separate out the x-factors and y-factors.

But look what happens when we set $y = vx$.

$$\frac{d(vx)}{dx} = \frac{x^2 + v^2x^2}{vx^2} = \frac{1 + v^2}{v}$$

Applying the product rule we have

$$\boxed{\frac{d(vx)}{dx} = v + x\frac{dv}{dx}}$$

Therefore we can say that

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

which is now separable.

Separating v-factors and x-factors to the LHS and RHS, respectively, and integrating both sides with respect to x then leads to

$$\int v dv = \int \frac{1}{x} dx$$

and consequently

$$\frac{v^2}{2} = \ln x + C$$

Reversing the substitution finally yields

$$y^2 = 2x^2(\ln x + C)$$

So why does this work? It works because the equation is homogenous. Homogeneous in what respect? In terms of degrees. Recall the example we just solved

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

Consider either x or y as degrees. All the terms in both the nominator and denominator (i.e., x^2, y^2 and xy) have the same number of degrees. In this respect, the above equation is homogenous. So when we replace y with vx, all the x's cancel out. Consequently we can subsequently separate variables as above.

Challenge 6.10 Solve the ordinary differential equation

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

We can see that all the terms in both the nominator and denominator have the same number of degrees, hence the equation is homogenous. Making the substitution $y = vx$ leads to

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

hence it can be said that

$$\frac{dv}{dx} = 0$$

Integrating with respect to x then leads to

$$v = C$$

and

$$y = Cx$$

Challenge 6.11 Solve the ordinary differential equation

$$(x^3 - xy^2)\frac{dy}{dx} = xy^2 - y^3$$

Solving for the derivative leads to

$$\frac{dy}{dx} = \frac{xy^2 - y^3}{x^3 - xy^2}$$

Again we can see that all the terms in both the nominator and denominator have the same number of degrees, hence the equation is homogenous. Making the substitution $y = vx$ leads to

$$v + x\frac{dv}{dx} = \frac{v^2 - v^3}{1 - v^2}$$

and

$$x\frac{dv}{dx} = \frac{v(v-1)}{1-v^2} = \frac{-v}{1+v}$$

Separating variables and integrating with respect to x

$$\int \frac{1}{v} + 1 dv = -\int \frac{1}{x} dx$$

which evaluates to yield

$$\ln v + v = -\ln x + C$$

Reversing the substitution finally leads to

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

6.7 Problem sheet

Problem 6.1 (see Worked Solution 6.1)

Derive the differential equations for the following functions

a) $y = A\ln x$ b) $y = A\cos x + B\sin x$ c) $y = Ax + Bx^2 + Cx^3$

Problem 6.2 (see Worked Solution 6.2)

Solve the following differential equations using separations of variables

a) $\dfrac{dy}{dx} = \dfrac{1+2y}{2+x}$ b) $xy\dfrac{dy}{dx} = \dfrac{x^3-1}{y-3}$ c) $\dfrac{\sin x}{1+y}\dfrac{dy}{dx} = \cos x$

d) $y\tan x\dfrac{dy}{dx} = (4+y^2)\sec^2 x$

Problem 6.3 (see Worked Solution 6.3)

Solve the following homogenous differential equations

a) $(x^2+y^2)\dfrac{dy}{dx} = xy$ b) $2x^2\dfrac{dy}{dx} = x^2+y^2$ c) $\dfrac{dy}{dx} = \dfrac{xy-y^2}{x^2+xy}$

6.8 Worked solutions

Worked Solution 6.1 (see Problem 6.1)

a) $y = A\ln x$

Differentiating both sides we have $\dfrac{dy}{dx} = \dfrac{A}{x}$

From the original equation we have $A = \dfrac{y}{\ln x}$

$$\therefore \frac{dy}{dx} = \frac{y}{x \ln x}$$

b) $y = A\cos x + B\sin x$

$$\frac{dy}{dx} = -A\sin x + B\cos x$$

$$\frac{d^2y}{dx^2} = -A\cos x - B\sin x$$

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

c) $y = Ax + Bx^2 + Cx^3$

$$\frac{dy}{dx} = A + 2Bx + 3Cx^2$$

$$\frac{d^2y}{dx^2} = 2B + 6Cx$$

$$\frac{d^3y}{dx^3} = 6C$$

$$C = \frac{1}{6}\frac{d^3y}{dx^3}, B = \frac{1}{2}\frac{d^2y}{dx^2} - \frac{x}{2}\frac{d^3y}{dx^3}, A = \frac{dy}{dx} - x\frac{d^2y}{dx^2} + \frac{x^2}{2}\frac{d^3y}{dx^3}$$

$$\therefore y = x\frac{dy}{dx} - x^2\frac{d^2y}{dx^2} + \frac{x^3}{2}\frac{d^3y}{dx^3} + \frac{x^2}{2}\frac{d^2y}{dx^2} - \frac{x^3}{2}\frac{d^3y}{dx^3} + \frac{x^3}{6}\frac{d^3y}{dx^3}$$

which reduces to $y = x\dfrac{dy}{dx} - \dfrac{x^2}{2}\dfrac{d^2y}{dx^2} + \dfrac{x^3}{6}\dfrac{d^3y}{dx^3}$

Worked Solution 6.2 (see Problem 6.2)

a) $\dfrac{dy}{dx} = \dfrac{1+2y}{2+x}$

Rearranging leads to $\dfrac{1}{1+2y}\dfrac{dy}{dx} = \dfrac{1}{2+x}$

Integrating both sides with respect to x

$$\int \frac{1}{1+2y}dy = \int \frac{1}{2+x}dx$$
$$\therefore \frac{\ln(1+2y)}{2} = \ln(2+x)+C$$

or we can say $\ln(1+2y) = \ln[e^{2C}(2+x)^2]$

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

b) $xy\dfrac{dy}{dx} = \dfrac{x^3-1}{y-3}$

$$\therefore \int y(y-3)dy = \int \frac{x^3-1}{x}dx$$
$$\therefore \frac{y^3}{3} - \frac{3y^2}{2} = \frac{x^3}{3} - \ln x + C$$

c) $\dfrac{\sin x}{1+y}\dfrac{dy}{dx} = \cos x$

$$\therefore \int \frac{1}{1+y}dy = \int \cot x dx$$
$$\therefore \ln(1+y) = \int \frac{\cos x}{\sin x}\frac{1}{\cos x}d\sin x = \ln(\sin x) + C$$
$$\therefore y = e^C \sin x - 1$$

d) $y\tan x\dfrac{dy}{dx} = (4+y^2)\sec^2 x$

$\therefore \displaystyle\int \frac{y}{4+y^2}dy = \int \frac{\sec^2 x}{\tan x}dx$

$\therefore \dfrac{\ln(4+y^2)}{2} = \ln\tan x + C$

$\therefore y^2 = e^{2C}\tan^2 x - 4$

Worked Solution 6.3 (see Problem 6.3)

a) $(x^2+y^2)\dfrac{dy}{dx} = xy$

$\therefore \dfrac{dy}{dx} = \dfrac{xy}{x^2+y^2}$

Set $y = vx$ such that we have $\dfrac{d(vx)}{dx} = \dfrac{vx^2}{x^2+v^2x^2} = \dfrac{v}{1+v^2}$

Applying the product rule we have $\dfrac{d(vx)}{dx} = v + x\dfrac{dv}{dx}$

So it follows that $v + x\dfrac{dv}{dx} = \dfrac{v}{1+v^2}$

$\therefore x\dfrac{dv}{dx} = \dfrac{v}{1+v^2} - v = \dfrac{v-v-v^3}{1+v^2} = -\dfrac{v^3}{1+v^2}$

Rearranging and integrating with respect to x

$$\int \frac{1+v^2}{v^3}dv = -\int \frac{1}{x}dx$$

$\therefore -\dfrac{1}{2v^2} + \ln v = -\ln x + C$

$\therefore -\dfrac{x^2}{2y^2} + \ln\left(\dfrac{y}{x}\right) = -\ln x + C$

b) $2x^2\dfrac{dy}{dx} = x^2 + y^2$

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

$$\therefore v + x\frac{dv}{dx} = \frac{x^2+v^2x^2}{2x^2} = \frac{1+v^2}{2}$$

$$\therefore x\frac{dv}{dx} = \frac{1-2v+v^2}{2} = \frac{(1-v)^2}{2}$$

$$\therefore \int \frac{2}{(1-v)^2}dv = \int \frac{1}{x}dx$$

$$\therefore \frac{2}{1-v} = \ln x + C$$

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

c) $\dfrac{dy}{dx} = \dfrac{xy-y^2}{x^2+xy}$

$$\therefore v + x\frac{dv}{dx} = \frac{v-v^2}{1+v}$$

$$\therefore x\frac{dv}{dx} = \frac{v-v^2}{1+v} - v = \frac{-2v^2}{1+v}$$

$$\therefore \int \frac{-1-v}{2v^2}dv = \int \frac{1}{x}dx$$

$$\therefore \frac{1}{2v} - \frac{\ln v}{2} = \ln x + C$$

7

More first-order differential equations

7.1 Learning outcomes

You should be able to:

- Show how to use integration factors to solve linear FO differential equations.
- Show how to use integration factors to solve non-linear Bernoulli equations.
- Distinguish between general solutions and particular solutions.

7.2 Integrating factor approach to linear equations

Consider the linear equation

$$\frac{dy}{dx} + 3y = e^{5x} \tag{7.1}$$

Obviously we can't solve this using any of the previous methods.

Again we will apply another trick. Multiply both sides by a factor $F = e^{3x}$.

$$e^{3x}\frac{dy}{dx} + 3e^{3x}y = e^{8x}$$

Can the above equation be further simplified? Yes! From the product rule we have

$$\frac{d(e^{3x}y)}{dx} = e^{3x}\frac{dy}{dx} + 3e^{3x}y$$

It follows that

$$\frac{d(e^{3x}y)}{dx} = e^{8x}$$

Finally we can integrate both sides with respect to x to yield

$$\int \frac{d(e^{3x}y)}{dx}dx = \int e^{8x}dx$$

from which we have

$$e^{3x}y = \frac{e^{8x}}{8} + C$$

and consequently

$$y = \frac{e^{5x}}{8} + Ce^{-3x}$$

In this example, we have multiplied the equation by the integration factor e^{3x}. Why did we use e^{3x}?

This form was chosen specifically to allow the two terms in the LHS of the original equation

$$\frac{dy}{dx} + 3y$$

to be merged, i.e.

$$\left(\frac{dy}{dx} + 3y\right) e^{3x} = \frac{d(e^{3x}y)}{dx}$$

So is there a general form of this fortuitous multiplicative factor, F?

The above equation is in fact a special case of the general form

$$\boxed{\frac{dy}{dx} + Py = Q} \tag{7.2}$$

where P and Q are functions of x. The integration factor, F, is a function such that the following equation is satisfied

$$F\left(\frac{dy}{dx} + Py\right) = \frac{d(Fy)}{dx}$$

Application of the product rule

$$\frac{d(Fy)}{dx} = F\frac{dy}{dx} + y\frac{dF}{dx}$$

we have that

$$F\frac{dy}{dx} + y\frac{dF}{dx} = F\frac{dy}{dx} + FPy$$

and consequently

$$\frac{dF}{dx} = FP$$

Separating the F-factors and P-factors yields

$$\frac{1}{F}\frac{dF}{dx} = P$$

Integrating with respect to x we have

$$\ln F = \int P dx$$

and consequently

$$\boxed{F = e^{\int P dx}} \tag{7.3}$$

Returning back to Eq. (7.2).

$$F\left[\frac{dy}{dx} + Py\right] = \frac{d(Fy)}{dx} = FQ$$

Therefore

$$\boxed{y = F^{-1}\int FQdx} \tag{7.4}$$

Recall Eq. (7.1). In this case

$$P = 3, \quad Q = e^{5x}, \quad F = e^{\int P dx} = e^{3x}$$

and so

$$y = e^{-3x}\int e^{3x}e^{5x}dx = e^{-3x}\left(\frac{e^8 x}{8} + C\right) = \frac{e^{5x}}{8} + Ce^{-3x}$$

Challenge 7.1 Solve the ordinary differential equation

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

After some rearranging we get

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

In this case

$$P = x^{-1}, \quad Q = x^2, \quad F = e^{\int P dx} = e^{\ln x} = x$$

and so

$$y = x^{-1} \int x^3 dx = x^{-1} \left(\frac{x^4}{4} + C \right) = \frac{x^3}{4} + \frac{C}{x}$$

Challenge 7.2 Solve the ordinary differential equation

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

After some rearranging we get

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

In this case

$$P = (x+1)^{-1}, \quad Q = x+1, \quad F = e^{\int P dx} = e^{\ln(x+1)} = x+1$$

and so

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

therefore

$$y = \frac{(x+1)^2}{3} + \frac{C}{x+1}$$

Challenge 7.3 Solve the ordinary differential equation

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

After some rearranging we get

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

In this case

$$P = \frac{-x}{1-x^2}, \quad Q = \frac{1}{1-x^2}, \quad F = e^{\int P dx} = e^{\int -x(1-x^2)^{-1} dx}$$

Making the substitution $u = 1 - x^2$ (where $u' = -2x$)

$$\int \frac{-x}{1-x^2} dx = \int \frac{-x}{uu'} du = \int \frac{1}{2u} du = \ln u^{1/2}$$

and so

$$F = e^{\int P dx} = e^{\ln[(1-x^2)^{1/2}]} = (1-x^2)^{1/2}$$

Finally we can say that

$$y = (1-x^2)^{-1/2} \int \frac{(1-x^2)^{1/2}}{1-x^2} dx = (1-x^2)^{-1/2} \int (1-x^2)^{-1/2} dx$$

and therefore

$$y = \frac{\arcsin x + C}{(1-x^2)^{1/2}}$$

7.3 Solving Bernoulli equations

Consider non-linear equations of the form

$$\boxed{\frac{dy}{dx} + Ry = Sy^n}$$

otherwise known as the Bernoulli equation. Can we use the integration-factor method to solve this?

Divide both sides by y^n

$$y^{-n}\frac{dy}{dx}+Ry^{1-n}=S$$

and substitute $z=y^{1-n}$. Note that from the chain rule

$$\frac{dz}{dx}=\frac{d(y^{1-n})}{dx}=\frac{dy}{dx}\frac{dx}{dy}\frac{d(y^{1-n})}{dx}=\frac{dy}{dx}\frac{d(y^{1-n})}{dy}=(1-n)y^{-n}\frac{dy}{dx}$$

Therefore we can say

$$\frac{dz}{dx}+(1-n)Rz=(1-n)S$$

Substituting $P=(1-n)R$ and $Q=(1-n)S$ we see that

$$\frac{dz}{dx}+Pz=Q$$

hence we can use integration factors as before.

Challenge 7.4 Solve the ordinary differential equation

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

Divide both sides by y^2

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

Substitute $z=y^{-1}$ (note that $z'=-y'y^{-2}$)

$$\frac{dz}{dx}-\frac{z}{x}=-x$$

In this case

$$P=-x^{-1},\quad Q=-x,\quad F=e^{\int P\,dx}=e^{-\ln x}=x^{-1}$$

and so

$$z = -x \int dx = x(-x + C) = -x^2 + xC$$

and

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

Challenge 7.5 Solve the ordinary differential equation

$$x^2 y - x^3 \frac{dy}{dx} = y^4 \cos x$$

Rearrange into the general form

$$\frac{dy}{dx} - \frac{y}{x} = -\frac{y^4}{x^3} \cos x$$

Divide both sides by y^4

$$y^{-4} \frac{dy}{dx} - \frac{y^{-3}}{x} = -\frac{\cos x}{x^3}$$

Substitute $z = y^{-3}$ (note that $z' = -3y' y^{-4}$)

$$\frac{dz}{dx} + \frac{3z}{x} = \frac{3\cos x}{x^3}$$

In this case

$$P = \frac{3}{x}, \quad Q = \frac{3\cos x}{x^3}, \quad F = e^{\int P dx} = e^{3\ln x} = x^3$$

and so

$$z = x^{-3} \int 3\cos x dx = \frac{3\sin x + C}{x^3}$$

and

$$y = \frac{x}{(3\sin x + C)^{1/3}}$$

7.4 Problem sheet

Problem 7.1 (see Worked Solution 7.1)

Solve the following linear differential equations

a) $x\dfrac{dy}{dx} - y = x^3 + 3x^2 - 2x$

b) $\dfrac{dy}{dx} + y\tan x = \sin x$

c) $x\dfrac{dy}{dx} - y = x^3 \cos x$ given that $y = 0$ when $x = \pi$.

d) $(1+x^2)\dfrac{dy}{dx} + 3xy = 5x$ given that $y = 2$ when $x = 1$.

Problem 7.2 (see Worked Solution 7.2)

Solve the following Bernoulli equations

a) $\dfrac{dy}{dx} + y = xy^3$

b) $\dfrac{dy}{dx} + y = y^4 e^x$

c) $\dfrac{dy}{dx} - 2y\tan x = y^2 \tan^2 x$

d) $\dfrac{dy}{dx} + y\tan x = y^3 \sec^4 x$

7.5 Worked solutions

Worked Solution 7.1 (see Problem 7.1)

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

First, rearrange to the standard form $y' + Py = Q$

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

Find the integration factor, F, such that $F\left(\frac{dy}{dx} + Py\right) = \frac{d(Fy)}{dx}$

$$\therefore F = e^{\int Pdx} = e^{\int -x^{-1}dx} = e^{-\ln x} = x^{-1}$$

Multiplying both sides of the differential equation by F

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

which reduces to $\frac{d(y/x)}{dx} = x + 3 - \frac{2}{x}$

It follows that $y = x\left(\frac{x^2}{2} + 3x - 2\ln x + C\right)$

b) $\frac{dy}{dx} + y\tan x = \sin x$

$$\therefore F = e^{\int Pdx} = e^{\int \tan x dx} = e^{\ln\sec x} = \sec x$$

$$\therefore \frac{d(y\sec x)}{dx} = \sec x \sin x = \tan x$$

$$\therefore y = \cos x\,(\ln\sec x + C)$$

c) $x\frac{dy}{dx} - y = x^3\cos x$

$$\therefore \frac{dy}{dx} - \frac{y}{x} = x^2\cos x$$

$$\therefore F = e^{\int -x^{-1}dx} = x^{-1}$$

$$\therefore \frac{d(y/x)}{dx} = x\cos x$$

$$\therefore y = x\left(x\sin x - \int \sin x dx\right)$$

$$\therefore y = x^2 \sin x + x\cos x + Cx$$

Given that $y = 0$ when $x = \pi$, $0 = -\pi + C\pi$, $\therefore C = 1$

$$\therefore y = x(x\sin x + \cos x + 1)$$

d) $(1+x^2)\dfrac{dy}{dx} + 3xy = 5x$

$$\therefore \frac{dy}{dx} + \frac{3xy}{1+x^2} = \frac{5x}{1+x^2}$$

$$\therefore F = e^{\int 3x(1+x^2)^{-1}dx}$$

Setting $u = 1+x^2$, $u' = 2x$

$$\therefore F = e^{\int 3/(2u)du} = u^{3/2} = (1+x^2)^{3/2}$$

$$\therefore \frac{d}{dx}\left[y(1+x^2)^{3/2}\right] = 5x(1+x^2)^{1/2}$$

$$\therefore y(1+x^2)^{3/2} = \int 5x(1+x^2)^{1/2}dx = \int \frac{5u^{1/2}}{2}du = \frac{5u^{3/2}}{3} + C$$

$$\therefore y = (1+x^2)^{-3/2}\left[\frac{5(1+x^2)^{3/2}}{3} + C\right] = \frac{5}{3} + C(1+x^2)^{-3/2}$$

Given that $y = 2$ when $x = 1$, $2 = 5/3 + 2^{-3/2}C$

$$\therefore C = 2^{3/2}/3 = 0.9428$$

Worked Solution 7.2 (see Problem 7.2)

a) $\dfrac{dy}{dx}+y=xy^3$

Divide both sides by y^3

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

Set $z=y^{-2}$.

$$\frac{dz}{dx}=\frac{dz}{dy}\frac{dy}{dx}=-2y^{-3}\frac{dy}{dx}$$

$$\therefore \frac{-1}{2}\frac{dz}{dx}+z=x$$

$$\therefore \frac{dz}{dx}-2z=-2x$$

$$\therefore F=e^{\int -2dx}=e^{-2x}$$

$$\therefore \frac{d(ze^{-2x})}{dx}=-2xe^{-2x}$$

$$\therefore ze^{-2x}=xe^{-2x}-\int e^{-2x}=e^{-2x}\left(x+\frac{1}{2}\right)+C$$

$$\therefore y^{-2}\equiv z=x+\frac{1}{2}+Ce^{2x}$$

b) $\dfrac{dy}{dx}+y=y^4e^x$

$$y^{-4}\frac{dy}{dx}+y^{-3}=e^x$$

Set $z=y^{-3}$

$$\therefore \frac{dz}{dx}=\frac{dz}{dy}\frac{dy}{dx}=-3y^{-4}\frac{dy}{dx}$$

$$\therefore \frac{dz}{dx} - 3z = -3e^x$$

$$\therefore F = e^{\int -3dx} = e^{-3x}$$

$$\therefore \frac{d(ze^{-3x})}{dx} = -3e^{-2x}$$

$$\therefore ze^{-3x} = \frac{3e^{-2x}}{2} + C$$

$$\therefore y^{-3} \equiv z = \frac{3e^x}{2} + Ce^{3x}$$

c) $\dfrac{dy}{dx} - 2y\tan x = y^2 \tan^2 x$

$$y^{-2}\frac{dy}{dx} - 2y^{-1}\tan x = \tan^2 x$$

Set $z = y^{-1}$

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

$$\therefore \frac{dz}{dx} + 2z\tan x = -\tan^2 x$$

$$\therefore F = e^{\int 2\tan x dx} = e^{2\ln\sec x} = \sec^2 x$$

$$\therefore \frac{d(z\sec^2 x)}{dx} = -\sec^2 x\tan^2 x$$

$$\therefore z\sec^2 x = -\int \sec^2 x\tan^2 x dx$$

Substitute $u = \tan x$, noting that $u' = \sec^2 x$

$$\therefore z\sec^2 x = -\int u^2 du = -\frac{u^3}{3} + C$$

$$\therefore y^{-1} = z = -\frac{\sin^2 x \tan x}{3} + C\cos^2 x$$

d) $\dfrac{dy}{dx} + y\tan x = y^3 \sec^4 x$

$$\therefore y^{-3}\frac{dy}{dx} + y^{-2}\tan x = \sec^4 x$$

Set $z = y^{-2}$

$$\therefore \frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = -2y^{-3}\frac{dy}{dx}$$

$$\therefore \frac{dz}{dx} - 2z\tan x = -2\sec^4 x$$

$$\therefore F = e^{\int -2\tan x dx} = e^{2\ln\cos x} = \cos^2 x$$

$$\therefore \frac{d(z\cos^2 x)}{dx} = -2\sec^2 x$$

$$\therefore z\cos^2 x = -2\tan x + C$$

$$\therefore y^{-2} \equiv z = \sec^2 x(C - 2\tan x)$$

8

Second-order differential equations

8.1 Learning outcomes

You should be able to:

- Find the general solutions of homogenous, linear, constant coefficient, second-order differential equations.
- Write solutions in terms of exponential, hyperbolic and trigonometric functions as appropriate.
- Explain why the general solution, for when the roots of the auxiliary equation are equal, takes the form $y = (A + Bx)e^{mx}$.

8.2 Solutions can be additions of alternative solutions

Consider the generalised linear, constant coefficient, differential equation

$$a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$$

where a, b and c are constant coefficients.

The above equation is said to be homogenous when $f(x)=0$, such that

$$a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0 \tag{8.1}$$

This week, we will focus on the homogenous form.

Let $y=u$ and $y=v$ be two solutions of the equation.

It follows that

$$a\frac{d^2u}{dx^2}+b\frac{du}{dx}+cu=0 \text{ and } a\frac{d^2v}{dx^2}+b\frac{dv}{dx}+cv=0$$

Adding these together leads to

$$a\frac{d^2(u+v)}{dx^2}+b\frac{d(u+v)}{dx}+c(u+v)=0$$

which indicates (comparing to our original equation) that $y=u+v$.

This seems a trivial point but it is necessary to show that new solutions to differential equations can be obtained by adding alternative solutions together.

8.3 A general solution in terms of exponentials

Consider Eq. (8.1) with $a = 0$ leads to

$$b\frac{dy}{dx} + cy = 0$$

rearranging leads to

$$\frac{1}{y}\frac{dy}{dx} = -\frac{c}{b}$$

Integrating both sides with respect to x

$$y = Ae^{-(c/b)x} \tag{8.2}$$

where A is an integration constant.
Let us try $y = Ae^{mx}$ as solution to our original second-order differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

where m is yet to be defined.

Substituting Eq. (8.2) into the differential equation leads to

$$am^2Ae^{mx} + bmAe^{mx} + cAe^{mx} = 0$$

which reduces to obtain

$$am^2 + bm + c = 0$$

The above equation is often referred to as the auxiliary equation.

It follows that

$$m = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Let

$$m_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad m_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \tag{8.3}$$

We have shown that Eq. (8.1) has two solutions

$$y = Ae^{m_1 x} \text{ and } y = Be^{m_2 x}$$

But neither of the above are general solutions to Eq. (8.1) because they have only one constant. A general solution to a second-order differential equation must have two integration constants. Such a solution can be obtained by adding the two above equations. Indeed, the general solution to Eq. (8.1) is

$$y = Ae^{m_1 x} + Be^{m_2 x} \tag{8.4}$$

where m_1 and m_2 are found from Eq. (8.3).

8.4 Hyperbolic form

So for the equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

the general solution is

$$y = A\exp\left[\left(\frac{-b + \sqrt{b^2 - 4ac}}{2a}\right)x\right] + B\exp\left[\left(\frac{-b - \sqrt{b^2 - 4ac}}{2a}\right)x\right]$$

Recalling that

$$\cosh x = \frac{e^x + e^{-x}}{2} \text{ and } \sinh x = \frac{e^x - e^{-x}}{2}$$

it follows that

$$y = e^{-\frac{bx}{2a}} \left\{ C \cosh\left[\left(\frac{\sqrt{b^2-4ac}}{2a}\right)x\right] + D \sinh\left[\left(\frac{\sqrt{b^2-4ac}}{2a}\right)x\right]\right\}$$

8.5 Trigonometric form

Again another way to write the previous equation is

$$y = e^{-\frac{bx}{2a}} \left\{ C \cosh\left[\left(\frac{i\sqrt{4ac-b^2}}{2a}\right)x\right] + D \sinh\left[\left(\frac{i\sqrt{4ac-b^2}}{2a}\right)x\right]\right\}$$

Recalling that

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \text{ and } \sin x = \frac{e^{ix} - e^{-ix}}{i2}$$

which leads us to

$$\cosh ix = \cos x \text{ and } \sinh ix = i \sin x$$

$$\therefore y = e^{-\frac{bx}{2a}} \left\{ C \cos\left[\left(\frac{\sqrt{4ac-b^2}}{2a}\right)x\right] + E \sin\left[\left(\frac{\sqrt{4ac-b^2}}{2a}\right)x\right]\right\}$$

Note that when $b^2 < 4ac$ the roots of the auxiliary equation are complex (recall Eq. (8.3)). When the roots are complex it is better to write the general solution in the above trigonometric form.

Challenge 8.1 Solve the ordinary differential equation

$$\frac{d^2y}{dx^2} + 3\frac{dy}{dx} + 2y = 0$$

Substituting $y = Ae^{mx}$ it is found that

$$m^2 + 3m + 2 = 0 = (m+2)(m+1)$$

So $m = -2$ or -1

$$\therefore y = Ae^{-2x} + Be^{-x}$$

Another way to write this is

$$y = e^{-3x/2}(Ae^{-x/2} + Be^{x/2})$$

$$\therefore y = e^{-3x/2}\left[C\cosh(x/2) + D\sinh(x/2)\right]$$

Challenge 8.2 Solve the ordinary differential equation

$$\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + 6y = 0$$

Substituting $y = Ae^{mx}$ it is found that

$$m^2 + 5m + 6 = 0 = (m+2)(m+3)$$

So $m = -2$ or -3

$$\therefore y = Ae^{-2x} + Be^{-3x}$$

Another way to write this is

$$y = e^{-5x/2}(Ae^{x/2} + Be^{-x/2})$$

$$\therefore y = e^{-5x/2}\left[C\cosh(x/2) + D\sinh(x/2)\right]$$

Challenge 8.3 Solve the ordinary differential equation

$$\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + 10y = 0$$

Substituting $y = Ae^{mx}$ it is found that

$$m^2 - 2m + 10 = 0$$

$$\therefore m = \frac{2 \pm \sqrt{4-40}}{2} = 1 \pm i3$$

Because of complex roots, we should use the trigonometric form

$$\therefore y = e^x(A\cos 3x + B\sin 3x)$$

Challenge 8.4 Solve the ordinary differential equation

$$\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 9y = 0$$

Substituting $y = Ae^{mx}$ it is found that

$$m^2 + 4m + 9 = 0$$

$$\therefore m = \frac{-4 \pm \sqrt{16-36}}{2} = -2 \pm i\sqrt{5}$$

Because of complex roots, we should use the trigonometric form

$$\therefore y = e^{-2x}(A\cos\sqrt{5}x + B\sin\sqrt{5}x)$$

Challenge 8.5 Solve the ordinary differential equation

$$\frac{d^2y}{dx^2} + 6\frac{dy}{dx} + 9y = 0$$

Substituting $y = Ae^{mx}$ it is found that

$$m^2 + 6m + 9 = 0$$

It follows that

$$(m+3)(m+3) = 0$$

So $m = -3$

$$\therefore y = Ae^{-3x}$$

But there is only one constant. So this is not the general solution.

8.6 Equal roots to the auxiliary equation

The situation that occurs in Example 3 happens because the auxiliary equation has equal roots. Referring back to Eq. (8.3), it can be seen that for our general equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

the roots to the auxiliary equation

$$am^2 + bm + c = 0$$

are equal when

$$b^2 = 4ac$$

To obtain a general solution with two constants for the special case when $b^2 = 4ac$, it is necessary to find an alternative solution to the basic exponential form given in Eq. (8.2).

Let us consider a second solution of the form

$$y = u(x)e^{mx}$$

Substituting into Eq. (8.1) leads to

$$a\frac{d^2}{dx^2}(ue^{mx}) + b\frac{d}{dx}(ue^{mx}) + cue^{mx} = 0 \tag{8.5}$$

The first and second-order derivatives can be expanded using product rule

$$\frac{d}{dx}(ue^{mx}) = (mu + u')e^{mx}$$

$$\frac{d^2}{dx^2}(ue^{mx}) = \frac{d}{dx}\left(\frac{d}{dx}(ue^{mx})\right) = \frac{d}{dx}\left((mu + u')e^{mx}\right)$$

and consequently

$$\frac{d^2}{dx^2}(ue^{mx}) = (m^2u + 2mu' + u'')e^{mx}$$

Substituting into Eq. (8.5) leads to

$$a(m^2u + 2mu' + u'') + b(mu + u') + cu = 0$$

Collecting derivatives of common order

$$au'' + (2am + b)u' + (am^2 + bm + c)u = 0$$

Recall, the case we are interested in is when $b^2 = 4ac$. In this case, from Eq. (8.3), $m = -b/2a$, which on substitution into the above equation leads to

$$au'' + \left(c - \frac{b^2}{4a}\right)u = 0$$

But of course, $b^2 = 4ac$, so

$$u'' = 0$$

which on integration leads to

$$u = A + Bx$$

Therefore a general solution to Eq. (8.1) for this case is seen to be

$$y = (A + Bx)e^{mx} \tag{8.6}$$

8.7 Another way to study the equal roots case

Another way of looking at this is as follows. Consider the equation

$$y = Ae^{mx} + Be^{nx}$$

where $n = m + \delta m$.

Factorising the e^{mx} term leads to

$$y = (A + Be^{\delta mx})e^{mx}$$

Recall the exponential series

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

It follows that

$$e^{\delta mx} = 1 + \delta mx + \frac{(\delta mx)^2}{2!} + \frac{(\delta mx)^3}{3!} + \ldots$$

Now if $\delta m \ll 1$

$$e^{\delta mx} \approx 1 + \delta mx \tag{8.7}$$

and consequently

$$y = [A + B(1 + \delta mx)]e^{mx}$$

Setting $C = A + B$ and $D = B\delta m$ then yields

$$y = (C + Dx)e^{mx}$$

So it can be said that

$$\lim_{m \to n} \{Ae^{mx} + Be^{nx}\} = (C + Dx)e^{mx} \tag{8.8}$$

Challenge 8.6 Solve the ordinary differential equation

$$\frac{d^2y}{dx^2} + 6\frac{dy}{dx} + 9y = 0$$

Substituting $y = Ae^{mx}$ it is found that

$$m^2 + 6m + 9 = 0$$

It follows that

$$(m+3)(m+3) = 0$$

So both roots are $m = -3$

$$\therefore y = (A + Bx)e^{-3x}$$

Challenge 8.7 Solve the ordinary differential equation

$$\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 4y = 0$$

Substituting $y = Ae^{mx}$ it is found that

$$m^2 + 4m + 4 = 0$$

It follows that

$$(m+2)(m+2) = 0$$

So both roots are $m = -2$

$$\therefore y = (A + Bx)e^{-2x}$$

8.8 Problem sheet

Problem 8.1 (see Worked Solution 8.1)

Find the general solution of $\dfrac{d^2y}{dx^2} - 5\dfrac{dy}{dx} + 6y = 0$

Problem 8.2 (see Worked Solution 8.2)

Find the particular solution of $\dfrac{d^2y}{dx^2} + 7\dfrac{dy}{dx} + 6y = 0$

given that $y(0) = 0$ and $y'(0) = 5$.

Problem 8.3 (see Worked Solution 8.3)

Solve $\dfrac{d^2y}{dx^2} + 6\dfrac{dy}{dx} + 9y = 0$

Problem 8.4 (see Worked Solution 8.4)

Solve $\dfrac{d^2y}{dx^2} + 9y = 0$ given that $y(0) = 2$ and $y(\pi/6) = 3$.

Problem 8.5 (see Worked Solution 8.5)

Solve $\dfrac{d^2y}{dx^2} - 9y = 0$ given that $y(0) = 0$ and $y'(0) = 6$.

Problem 8.6 (see Worked Solution 8.6)

Solve $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} + 8y = 0$

Problem 8.7 (see Worked Solution 8.7)

Solve $3\dfrac{d^2y}{dx^2} - 6\dfrac{dy}{dx} + 4y = 0$

given that $y(0) = 1$ and $y'(\pi/2\sqrt{3}) = 0$.

8.9 Worked solutions

Worked Solution 8.1 (see Problem 8.1)

$$\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 6y = 0$$

Try $y = Ae^{mx}$

$$\therefore m^2Ae^{mx} - 5mAe^{mx} + 6Ae^{mx} = 0$$

$$\therefore m^2 - 5m + 6 = 0$$

$$\therefore m = \frac{5 \pm \sqrt{25-24}}{2} = 2 \text{ or } 3$$

$$\therefore y = Ae^{2x} + Be^{3x}$$

Worked Solution 8.2 (see Problem 8.2)

$$\frac{d^2y}{dx^2} + 7\frac{dy}{dx} + 6y = 0$$

$$\therefore m^2 + 7m + 6 = 0$$

$$\therefore m = \frac{-7 \pm \sqrt{49-24}}{2} = -6 \text{ or -1}$$

$$\therefore y = Ae^{-6x} + Be^{-x}$$

$$\therefore \frac{dy}{dx} = -6Ae^{-6x} - Be^{-x}$$

Given that $y(0) = 0$, it can be seen that $A = -B$

Given that $y'(0) = 5$, we get $5 = -6A - B = -6A + A = -5A$

$\therefore A = -1$ and $B = 1$

$\therefore y = e^{-x} - e^{-6x}$

Worked Solution 8.3 (see Problem 8.3)

$$\frac{d^2y}{dx^2} + 6\frac{dy}{dx} + 9y = 0$$

$$\therefore m^2 + 6m + 9 = 0$$

$$\therefore m = \frac{-6 \pm \sqrt{36-36}}{2} = -3$$

Because there is only one root, $y = (A + Bx)e^{-3x}$

Worked Solution 8.4 (see Problem 8.4)

$$\frac{d^2y}{dx^2} + 9y = 0$$

$$\therefore m^2 + 9 = 0$$

$$\therefore m = \pm i3$$

$$\therefore y = Ae^{i3x} + Be^{-i3x}$$

$$\therefore y = C\cos 3x + D\sin 3x$$

Given that $y(0) = 2$, it can be said that $C = 2$

Given that $y(\pi/6) = 3$, it can be said that $D = 3$

$$\therefore y = 2\cos 3x + 3\sin 3x$$

Worked Solution 8.5 (see Problem 8.5)

$$\frac{d^2y}{dx^2} - 9y = 0$$

$\therefore m^2 - 9 = 0$ and consequently $m = \pm 3$

$$\therefore y = Ae^{3x} + Be^{-3x}$$

$$\therefore y = C\cosh 3x + D\sinh 3x$$

$$\therefore \frac{dy}{dx} = 3C\sinh 3x + 3D\cosh 3x$$

Given that $y(0) = 0$, it can be seen that $C = 0$.

Given that $y'(0) = 6$, it can be seen that $6 = 3D$ and $D = 2$.

$$\therefore y = 2\sinh 3x$$

Worked Solution 8.6 (see Problem 8.6)

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + 8y = 0$$

$$\therefore m^2 + 2m + 8 = 0$$

$$\therefore m = \frac{2 \pm \sqrt{4-32}}{2} = -1 \pm i\sqrt{7}$$

$$\therefore y = Ae^{(-1+i\sqrt{7})x} + Be^{(-1-i\sqrt{7})x} = e^{-x}\left(Ae^{i\sqrt{7}x} + Be^{-i\sqrt{7}x}\right)$$

$$\therefore y = e^{-x}\left(C\cos\sqrt{7}x + D\sin\sqrt{7}x\right)$$

Worked Solution 8.7 (see Problem 8.7)

$$3\frac{d^2y}{dx^2} - 6\frac{dy}{dx} + 4y = 0$$

$\therefore 3m^2 - 6m + 4 = 0$

$$\therefore m = \frac{6 \pm \sqrt{36-48}}{6} = 1 \pm i/\sqrt{3}$$

$$\therefore y = e^x \left[A\cos(x/\sqrt{3}) + B\sin(x/\sqrt{3})\right]$$

$$\therefore \frac{dy}{dx} = e^x \left[(A + B/\sqrt{3})\cos(x/\sqrt{3}) + (B - A/\sqrt{3})\sin(x/\sqrt{3})\right]$$

Given that $y(0) = 1$ it can be seen that $A = 1$.

Given that $y'(\pi/2\sqrt{3}) = 0$ it can be seen that

$$\therefore 0 = (1 + B/\sqrt{3})\cos(\pi/6) + (B - 1/\sqrt{3})\sin(\pi/6)$$

$$\therefore 0 = \frac{\sqrt{3}}{2} + \frac{B}{2} + \frac{B}{2} - \frac{1}{2\sqrt{3}} \text{ from which } B = -\frac{1}{\sqrt{3}}$$

$$\therefore y = e^x \left[\cos(x/\sqrt{3}) - \frac{1}{\sqrt{3}}\sin(x/\sqrt{3})\right]$$

9

More second-order differential equations

9.1 Learning outcomes

You should be able to:

- Explain why the general solution of a non-homogenous, linear, constant coefficient, second-order differential equation is equal to the solution to the homogenous equivalent (the complementary function) plus the so-called particular integral.
- Appropriately speculate about the form of particular integrals for given non-homogenous equations.

9.2 Complementary functions and particular integrals

Consider the generalised linear, constant coefficient, differential equation

$$ay'' + by' + cy = f(x) \tag{9.1}$$

where a, b and c are constant coefficients.

Recall from the previous lecture that solutions to differential equations can be added together to create new solutions. Consider two functions, y_h and y_p, which satisfy

$$ay_h'' + by_h' + cy_h = 0 \tag{9.2}$$

$$ay_p'' + by_p' + cy_p = f(x) \tag{9.3}$$

Adding above equation leads to

$$a\frac{d^2(y_h + y_p)}{dx^2} + b\frac{d(y_h + y_p)}{dx} + c(y_h + y_p) = f(x)$$

By comparison with Eq. (9.1), it can be seen that $y = y_h + y_p$.

Note that Eq. (9.2) is a homogenous version of Eq. (9.1). We know how to get a general solution to Eq. (9.2). Given that such a solution will have the two necessary unknown integration constants, we are free to find solutions to Eq. (9.3) that are not general. Particular solutions of Eq. (9.3) can be generalised by adding general solutions of Eq. (9.2).

The general solution to Eq. (9.2) is known as the complementary function. The particular solution to Eq. (9.3) is known as the particular integral. The general solution to Eq. (9.1) is equal to the complementary function plus the particular integral.

Challenge 9.1 Solve the ordinary differential equation

$$y'' - 5y' + 6y = 24$$

First let us find the complementary function (CF)

$y_h'' - 5y_h' + 6y_h = 0$

The auxiliary equation is $m^2 - 5m + 6 = 0$

Therefore $(m-2)(m-3) = 0$

It follows that $y_h = Ae^{2x} + Be^{3x}$

Now lets obtain the particular integral (PI)

$y_p'' - 5y_p' + 6y_p = 24$

What form of y_p would satisfy the above equation?

Clearly, $y_p = 0$ would not satisfy it. How about $y_p = C$?

If $y_p = C$, then $y_p' = 0$ and $y_p'' = 0$. It follows that $6C = 24$, therefore $y_p = C = 4$.

The general solution to the original equation is therefore

$y = y_h + y_p = Ae^{2x} + Be^{3x} + 4$

Challenge 9.2 Solve the ordinary differential equation

$$y'' - 5y' + 6y = x^2$$

Recall that the CF is $y_h = Ae^{2x} + Be^{3x}$

For the PI, try $y_p = Cx^2 + Dx + E$

So $y_p' = 2Cx + D$ and $y_p'' = 2C$

$\therefore 2C - 5(2Cx + D) + 6(Cx^2 + Dx + E) = x^2$

Collect terms of x and x^2

$2C - 5D + 6E + (6D - 10C)x + (6C - 1)x^2 = 0$

Equating coefficients of x and x^2

$6C - 1 = 0$, $6D - 10C = 0$ and $2C - 5D + 6E = 0$

It follows that $C = 1/6$, $D = 5/18$ and $E = 19/108$

Therefore $y = Ae^{2x} + Be^{3x} + \dfrac{x^2}{6} + \dfrac{5x}{18} + \dfrac{19}{108}$

Challenge 9.3 Solve the ordinary differential equation

$$y'' - 5y' + 6y = 2\sin 4x$$

Recall that the CF is $y_h = Ae^{2x} + Be^{3x}$

For the PI, try $y_p = C\cos 4x + D\sin 4x$

So $y_p' = -4C\sin 4x + 4D\cos 4x$

and $y_p'' = -16C\cos 4x - 16D\sin 4x$

Substituting these back into the original equation

$-16C\cos 4x - 16D\sin 4x + 20C\sin 4x - 20D\cos 4x$

$+6C\cos 4x + 6D\sin 4x = 2\sin 4x$

Equating coefficients of $\cos 4x$: $-16C - 20D + 6C = 0$ therefore $C = -2D$

Equating coefficients of $\sin 4x$: $-16D + 20C + 6D = 2$

It follows that $-16D - 40D + 6D = 2$ and therefore $D = -1/25$ and $C = 2/25$

Therefore $y_p = \dfrac{2}{25}\cos 4x - \dfrac{1}{25}\sin 4x$

and consequently $y = Ae^{2x} + Be^{3x} + \dfrac{2}{25}\cos 4x - \dfrac{1}{25}\sin 4x$

9.3 Solution by undetermined coefficients

In the above examples, we have been applying the so-called method of undetermined coefficients.

The particular integral (PI), y_p, of the differential equation

$ay'' + by' + cy = f(x)$

is found by consideration of the following set of rules.

1) Basic Rule: If $f(x)$ is one of the functions in the first column of Table 9.1, choose the corresponding function y_p in the second column of Table 9.1 and determine its undetermined coefficients by substituting y_p and its derivatives back into the differential equation.

2) Modification Rule: If a term in your choice of y_p is a solu-

tion of the homogenous equation (i.e., the differential equation with $f(x) = 0$), then multiply your choice of y_p by x. If your choice of y_p multiplied by x is also a solution of the homogenous equation, then multiply by x^2 and so on.

3) Sum Rule: If $f(x)$ is a sum of functions listed in several lines of the first column of Table 9.1, then choose y_p to be a sum of the functions in the corresponding lines of the second column of Table 9.1.

Table 9.1: Recommended functional forms for particular integrals.

Term in $f(x)$	Choice for y_p
αe^{kx}	Ce^{kx}
αx^n $(n = 0, 1, \ldots)$	$C_n x^n + C_{n-1}x^{n-1} + \ldots + C_1 x + C_0$
$\alpha \cos kx$	$C \cos kx + D \sin kx$
$\alpha \sin kx$	$C \cos kx + D \sin kx$
$\alpha e^{\lambda x} \cos kx$	$e^{\lambda x}(C \cos kx + D \sin kx)$
$\alpha e^{\lambda x} \sin kx$	$e^{\lambda x}(C \cos kx + D \sin kx)$

Challenge 9.4 Solve the ordinary differential equation

$$y'' + 4y = 8x^2$$

To find the CF, solve $y_h'' + 4y_h = 0$

The auxiliary equation is $m^2 + 4 = 0$

Therefore the CF is $y_h = A \sin 2x + B \cos 2x$

Using the "Basic Rule", considering the $8x^2$ in the RHS (right-hand-side) of the differential equation, from the above table it can be seen that the PI will take the form

$y_p = Cx^2 + Dx + E$ and of course, $y_p'' = 2C$

Substituting back into the original differential equation leads to

$2C + 4(Cx^2 + Dx + E) = 8x^2$

Equating coefficients of x^2, x and x^0 it can be seen that $4C = 8$, $4D = 0$ and $2C + 4E = 0$

Therefore $C = 2$, $D = 0$ and $E = -1$.

Consequently it can be said that $y_p = 2x^2 - 1$

So the general solution is:

$y = A\sin 2x + B\cos 2x + 2x^2 - 1$

Challenge 9.5 Solve the ordinary differential equation

$$y'' - 3y' + 2y = e^x$$

To find the CF, solve $y_h'' - 3y_h' + 2y_h = 0$

The auxiliary equation is $m^2 - 3m + 2 = 0 \Rightarrow (m-1)(m-2) = 0$.

Therefore the CF is $y_h = Ae^x + Be^{2x}$

Using the "Basic Rule", considering the e^x in the RHS of the differential equation, from the above table let us choose $y_p = Ce^x$. Substituting back into the differential equation leads to

$Ce^x - 3Ce^x + 2Ce^x \neq e^x$.

Indeed $y_p \neq Ce^x$. But recall the “Modification Rule”. Therefore try $y_p = Cxe^x$. Substituting back into the differential equation leads to $(x+2)Ce^x - 3(x+1)Ce^x + 2xCe^x = e^x$

$\therefore (x+2)C - 3(x+1)C + 2xC = 1$

Equating coefficients of x and x^0

$C - 3C + 2C = 0$ and $2C - 3C = 1$

from which it is seen that $C = -1$.

Therefore the general solution is

$y = Ae^x + Be^{2x} - xe^x$

Challenge 9.6 Solve the ordinary differential equation

$$y'' - 2y' + y = e^x + x$$

To find the CF, solve $y_h'' - 2y_h' + y_h = 0$

The auxiliary equation is $m^2 - 2m + 1 = 0 \Rightarrow (m-1)(m-1) = 0$.

Therefore the CF is $y_h = (A + Bx)e^x$

Using the “Sum Rule”, considering the $e^x + x$ term in the RHS of the differential equation, try $y_p = Ce^x + Dx + E$.

But a term in our choice of y_p is one of the solutions to the homogenous equation. From the “Modification Rule”, multiply that term by x. Therefore try $y_p = Cxe^x + Dx + E$.

But still a term in our choice of y_p is one of the solutions to the homogenous equation. From the "Modification Rule", multiply by that term by x again. Therefore try $y_p = Cx^2e^x + Dx + E$.

$y_p' = C(x^2 + 2x)e^x + D$ and $y_p'' = C(x^2 + 4x + 2)e^x$

Substituting these into the differential equation yields

$$C(x^2+4x+2)e^x - 2[C(x^2+2x)e^x + D] + Cx^2e^x + Dx + E = e^x + x$$

Collecting terms

$$(2C-1)e^x + (D-1)x + E - 2D = 0$$

From which we can see that $C = 1/2$, $D = 1$ and $E = 2$.

$$\therefore y = (A+Bx)e^x + \frac{x^2e^x}{2} + x + 2$$

9.4 Problem sheet

Problem 9.1 (see Worked Solution 9.1)

Solve $\dfrac{d^2y}{dx^2} + 25y = 5x^2 + x$

Problem 9.2 (see Worked Solution 9.2)

Solve $\dfrac{d^2y}{dx^2} - 4y = 12e^{2x}$

Problem 9.3 (see Worked Solution 9.3)

Solve $\dfrac{d^2y}{dx^2} + 6\dfrac{dy}{dx} + 9y = 2e^{-2x}$

given that $y(0) = 1$ and $y'(0) = -2$.

Problem 9.4 (see Worked Solution 9.4)

Solve $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} + y = 4\sinh x$

9.5 Worked solutions

Worked Solution 9.1 (see Problem 9.1)

$$\frac{d^2y}{dx^2} + 25y = 5x^2 + x$$

First find the complementary function (CF), u, where

$$y_h'' + 25y_h = 0$$

$$m^2 + 25 = 0$$

$$\therefore m = \pm i5$$

$$\therefore y_h = A\sin 5x + B\cos 5x$$

Next, find the particular integral (PI), y_p, where

$$y_p'' + 25y_p = 5x^2 + x$$

Try $y_p = Cx^2 + Dx + E$: $\therefore y_p' = 2Cx + D$ and $y_p'' = 2C$

$$\therefore 2C + 25(Cx^2 + Dx + E) = 5x^2 + x$$

Grouping coefficients of x:

$2C+25E+(25D-1)x+(25C-5)x^2=0$

It follows that $C=1/5$, $D=1/25$, $E=-2/125$

$$\therefore y_p=\frac{x^2}{5}+\frac{x}{25}-\frac{2}{125}$$

$$\therefore y=y_h+y_p=A\sin 5x+B\cos 5x+\frac{x^2}{5}+\frac{x}{25}-\frac{2}{125}$$

Worked Solution 9.2 (see Problem 9.2)

$$\frac{d^2y}{dx^2}-4y=12e^{2x}$$

$m^2-4=0$, and so the CF, $y_h=Ae^{2x}+Be^{-2x}$

$y_p''-4y_p=12e^{2x}$

Try $y_p=Cxe^{2x}$, so $y_p'=C(1+2x)e^{2x}$ and $y_p''=C(4+4x)e^{2x}$

$\therefore C(4+4x)e^{2x}-4Cxe^{2x}=12e^{2x}$ so $C=3$

$\therefore y=(A+3x)e^{2x}+Be^{-2x}$

Worked Solution 9.3 (see Problem 9.3)

Solve $\dfrac{d^2y}{dx^2}+6\dfrac{dy}{dx}+9y=2e^{-2x}$

$m^2+6m+9=0 \Rightarrow (m+3)^2=0$

$\therefore y_h=(A+Bx)e^{-3x}$

$y_p = Ce^{-2x}$

$\therefore 4Ce^{-2x} - 12Ce^{-2x} + 9Ce^{-2x} = 2e^{-2x}$ and $C = 2$

$\therefore y = (A + Bx)e^{-3x} + 2e^{-2x}$

Given that $y(0) = 1$ and $y'(0) = -2$.

Now $y' = (B - 3A - 3Bx)e^{-3x} - 4e^{-2x}$

$\therefore 1 = A + 2$ and $-2 = B - 3A - 4$

$\therefore A = -1$ and $B = -1$

$y = -(1 + x)e^{-3x} + 2e^{-2x}$

Worked Solution 9.4 (see Problem 9.4)

Solve $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} + y = 4\sinh x$

$m^2 + 2m + 1 = 0 \Rightarrow (m + 1)^2 = 0$

$\therefore y_h = (A + Bx)e^{-x}$

Now $\sinh x = (e^x - e^{-x})/2$

So try $y_p = Ce^x + Dx^2e^{-x}$

$y_p' = Ce^x + D(2x - x^2)e^{-x}$

$y_p'' = Ce^x + D(2 - 4x + x^2)e^{-x}$

$$\therefore Ce^x + D(2-4x+x^2)e^{-x} + 2Ce^x + 2D(2x-x^2)e^{-x}$$
$$+Ce^x + Dx^2e^{-x} = 2(e^x - e^{-x})$$

Collecting terms

$$(4C-2)e^x + [D(2-4x+x^2) + 2D(2x-x^2) + Dx^2 + 2]e^{-x} = 0$$

which simplifies to get $(4C-2)e^x + (2D+2)x^2e^{-x} = 0$

From which we see that $C = 1/2$ and $D = -1$.

$$\therefore y = (A + Bx - x^2)e^{-x} + \frac{e^x}{2}$$

10

Series and approximations

10.1 Learning outcomes

You should be able to:

- Derive Maclaurin's power series using polynomial fitting.
- Keep track of the order of truncation error using the O notation.
- Derive power series expansions for a range of different functions.
- Determine the radius of convergence of a given function using knowledge of arithmetic, harmonic and geometric series.
- Determine the radius of convergence of a given function using D'Alembert's ratio test.

10.2 Maclaurin's power series

Consider a function, $f(x)$, which can be closely approximated by fitting a polynomial function. This can be achieved as follows:

First $f(x)$ is approximated using a fourth-order polynomial, $a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$, such that it can be stated that

$$f(x) \approx a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \tag{10.1}$$

where the coefficients, a_0, a_1, a_2, a_3 are yet to be defined.

At present there is one equation and four unknowns (a_0, a_1, a_2,a_3). To obtain more equations, consider the derivatives of $f(x)$:

$$f'(x) \equiv f^{(1)}(x) \approx a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3$$

$$f''(x) \equiv f^{(2)}(x) \approx 2a_2 + 6a_3x + 12a_4x^2$$

$$f^{(3)}(x) \approx 6a_3 + 24a_4x$$

$$f^{(4)}(x) \approx 24a_4$$

By setting $x = 0$ it can be seen that:

$$a_0 \approx f^{(0)}(0)$$

$$a_1 \approx f^{(1)}(0)$$

$$a_2 \approx \frac{f^{(2)}(0)}{2}$$

$$a_3 \approx \frac{f^{(3)}(0)}{6}$$

$$a_4 \approx \frac{f^{(4)}(0)}{24}$$

which on substitution into Eq. (10.1) leads to

$$f(x) \approx f^{(0)}(0) + f^{(1)}(0)x + f^{(2)}(0)\frac{x^2}{2} + f^{(3)}(0)\frac{x^3}{6} + f^{(4)}(0)\frac{x^4}{24}$$

Notice that the above equation can also be written using factorials:

$$\begin{aligned} f(x) \approx \quad & f^{(0)}(0)\frac{x^0}{0!} + f^{(1)}(0)\frac{x^1}{1!} + f^{(2)}(0)\frac{x^2}{2!} \\ & + f^{(3)}(0)\frac{x^3}{3!} + f^{(4)}(0)\frac{x^4}{4!} \end{aligned} \tag{10.2}$$

such that it is also possible to state that

$$f(x) \approx \sum_{n=0}^{4} f^{(n)}(0)\frac{x^n}{n!} \tag{10.3}$$

where $f^{(n)}(x)$ and $f^{(n)}(0)$ are defined by

$$f^{(n)}(x) = \frac{d^n f}{dx^n} \quad \text{and} \quad f^{(n)}(0) = \left.\frac{d^n f}{dx^n}\right|_{x=0}$$

The Scottish mathematician, Colin Maclaurin, showed in the early 18th century that the series in Eq. (10.3) becomes an exact expression for $f(x)$ when one considers an infinite number of terms, i.e.,

$$f(x) = \sum_{n=0}^{\infty} f^{(n)}(0)\frac{x^n}{n!} \tag{10.4}$$

Eq. (10.4) above is now widely referred to as a power series or the Maclaurin series. As will be shown in the subsequent exercise, the Maclaurin series can be used to generate approximate asymptotic expansions for many of the different trigonometric and hyperbolic functions previously discussed.

10.3 The big O notation

Consider the approximation given in Eq. (10.2). Given the existence of the power series in Eq. (10.4), it is possible to exchange the $\approx$ sign in Eq. (10.2) for an $=$ sign by including the truncation error term, $O(x^5)$, i.e.,

$$\begin{aligned} f(x) = \; & f^{(0)}(0)\frac{1}{0!} + f^{(1)}(0)\frac{x}{1!} + f^{(2)}(0)\frac{x^2}{2!} \\ & + f^{(3)}(0)\frac{x^3}{3!} + f^{(4)}(0)\frac{x^4}{4!} + O(x^5) \end{aligned}$$

In a similar way it is possible to state that:

$$f(x) = f^{(0)}(0) + f^{(1)}(0)x + f^{(2)}(0)\frac{x^2}{2} + f^{(3)}(0)\frac{x^3}{6} + O(x^4)$$

$$f(x) = f^{(0)}(0) + f^{(1)}(0)x + f^{(2)}(0)\frac{x^2}{2} + O(x^3)$$

or even

$$f(x) = f^{(0)}(0) + O(x)$$

The O term is often used to denote truncation error. If we write $O(x^m)$, it is implied that the associated equation represents a truncated series expansion and the error associated with truncation is of "order" x^m. Note that the truncated terms in this case includes powers of x of order m or above. Therefore, if $x^m \ll 1$ it can be understood that the truncated terms associated with $O(x^m)$ are likely to be negligible.

Keeping track of the order of truncation error enables a robust methodology for simplifying the mathematics where appropriate for given scientific applications. It is also very helpful for determining limits of functions, as will be discovered later in this module.

10.4 Power series expansions of some common functions

The power series can be used to derive infinite series representations of many common functions.

10.4.1 Power series of e^x

Recall the power series in Eq. (10.4). Now consider $f(x) = e^x$. In this case, $f^n(x) = e^x$ and $f^n(0) = 1$. It therefore follows that:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + O(x^3)$$

10.4.2 Power series of $\sin x$

Now consider $f(x) = \sin x$ where:

$$\begin{array}{lll} f^{(0)}(x) = \sin x, & f^{(1)}(x) = \cos x, & f^{(2)}(x) = -\sin x, \\ f^{(3)}(x) = -\cos x, & f^{(4)}(x) = \sin x, & f^{(5)}(x) = \cos x, \\ f^{(6)}(x) = -\sin x, & f^{(7)}(x) = -\cos x, & \end{array}$$

and

$$\begin{array}{lll} f^{(0)}(0) = 0, & f^{(1)}(0) = 1, & f^{(2)}(0) = 0, \\ f^{(3)}(0) = -1, & f^{(4)}(0) = 0, & f^{(5)}(0) = 1, \\ f^{(6)}(0) = 0, & f^{(7)}(0) = -1, & \end{array}$$

and from Eq. (10.4) it can be said that

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + O(x^9)$$

However, it can be seen there are at least two challenges for specifying an appropriately simplified infinite series for this function. The first is that every other term is zero, the second is that every other non-zero term is negative.

10.4.3 Tricks with series

Use $(-1)^n$ to alternate signs. This works because $(-1)^0 = 1$, $(-1)^1 = -1$, $(-1)^2 = 1$, $(-1)^3 = -1$ and so on. For example,

$$\sum_{n=0}^{\infty} (-1)^n b_n = b_0 - b_1 + b_2 - b_3 + \dots$$

Use $2n$ to only include even terms. Sometimes in a series all the odd terms vanish, that is $b_1 = b_3 = b_5 = 0$ and so on. But note that

$$b_0 + b_2 + b_4 + b_6 + \dots = \sum_{n=0}^{\infty} b_{2n}$$

Use $2n+1$ to only include odd terms. Sometimes in a series all the even terms vanish, that is $b_0 = b_2 = b_4 = 0$ and so on. But note that

$$b_1 + b_3 + b_5 + b_7 + \dots = \sum_{n=0}^{\infty} b_{2n+1}$$

10.4.4 $\sin x$ revisited

Recall that

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + O(x^9) \tag{10.5}$$

Because only odd terms are present, the infinite series will take the form

$$\sum_{n=0}^{\infty} b_{2n+1}$$

and given that the odd values of n are negative we can further specify that the form will be something like

$$\sum_{n=0}^{\infty} (-1)^n b_{2n+1} \tag{10.6}$$

By comparing Eq. (10.6) with Eq. (10.5) it can be seen that $b_n = x^n/n!$ and $b_{2n+1} = x^{2n+1}/(2n+1)!$, and finally that

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

10.4.5 Power series of $(a+x)^m$

Consider the function $f(x) = (a+x)^m$. It can be seen that

$$f^{(0)}(x) = (a+x)^m,$$

$$f^{(1)}(x) = m(a+x)^{m-1},$$

$$f^{(2)}(x) = m(m-1)(a+x)^{m-2},$$

$$f^{(3)}(x) = m(m-1)(m-2)(a+x)^{m-3},$$

$$f^{(m)}(x) = m!,$$

$$f^{(m+1)}(x) = m!(m-m)(a+x)^{-1} = 0,$$

$$f^{(m+2)}(x) = m!(m-m)(m-m-1)(a+x)^{-2} = 0,$$

and

$$f^{(0)}(0) = a^m$$

$$f^{(1)}(0) = ma^{m-1}$$

$$f^{(2)}(0) = m(m-1)a^{m-2}$$

$$f^{(3)}(0) = m(m-1)(m-2)a^{m-3}$$

$$f^{(m)}(0) = m!$$

$$f^{(m+1)}(0) = 0$$

$$f^{(m+1)}(0) = 0$$

with all higher order terms also being zero.

So it can be seen that

$$f^n(0) = \begin{cases} \dfrac{m!a^{m-n}}{(m-n)!}, & n \leq m \\ 0, & n > m \end{cases}$$

from which it follows that

$$(a+x)^m = \sum_{n=0}^{m} \frac{m!a^{m-n}}{(m-n)!} \frac{x^n}{n!} \equiv \sum_{n=0}^{m} \left(\begin{array}{c} m \\ n \end{array} \right) a^{m-n} x^n$$

which is the Binomial theorem.

10.5 Convergence and divergence

Let S_N be the sum of the first N terms, a_n, in a series, i.e.,

$$S_N = \sum_{n=1}^{N} a_n$$

If S_N approaches a definite value as $N \to \infty$, then the series is considered to be convergent. If S_N does not approach a definite number as $N \to \infty$, the series is said to be divergent.

10.5.1 Arithmetic series example

Consider the arithmetic series

$$S_N = \sum_{n=1}^{N} n = 1 + 2 + \ldots + (N-1) + N \qquad (10.7)$$

In the late 1700s, a primary school teacher wanted a rest from the children and so asked them to find the sum of all the numbers from 1 to 100. Unfortunately for the teacher, one of his students was Carl Friedrich Gauss, who solved the problem in less than five minutes. The same approach can be used to determine S_N in Eq. (10.7).

Gauss noticed that if N is an even number, he can split the numbers into two groups like:

$$\begin{array}{ll}
1 & +N \\
+2 & +(N-1) \\
+3 & +(N-2) \\
+4 & +(N-3) \\
+\dots & +\dots \\
+(N/2-2) & +(N/2+3) \\
+(N/2-1) & +(N/2+2) \\
+N/2 & +(N/2+1)
\end{array}$$

he could add them horizontally to get $N+1$:

$$1+N=N+1$$

$$2+(N-1)=N+1$$

$$3+(N-2)=N+1$$

$$(N/2-2)+(N/2+3)=N+1$$

$$(N/2-1)+(N/2+2)=N+1$$

$$N/2+(N/2+1)=N+1$$

Given that there will be $N/2$ pairs of terms, it can then be understood that

$$S_N=\sum_{n=1}^{N} n=\frac{N(N+1)}{2}$$

Alternatively, if N is an odd number there will be $(N-1)/2$ pairs plus an additional solo term in the middle of value (N-1)/2+1. It therefore follows that for odd values of N it can also be said that

$$S_N = \sum_{n=1}^{N} n = \frac{(N-1)(N+1)}{2} + \frac{(N-1)}{2} + 1 = \frac{N(N+1)}{2}$$

It can then be seen that the arithmetics series is an example of a divergent series because

$$\lim_{N\to\infty} S_N = \infty$$

10.5.2 Harmonic series example

Another simple series to consider is the harmonic series

$$S_N = \sum_{n=1}^{N} \frac{1}{n}$$

Each term added is smaller than the last, but the sum increases forever as $N \to \infty$ and never converges to a value. Hence this series is also divergent.

10.5.3 Geometric series example

Now consider the geometric series

$$S_N = \sum_{n=0}^{N} \lambda^n$$

Note that

$$S_N = 1 + \lambda + \lambda^2 + \ldots + \lambda^{N-1} + \lambda^N$$

Multiplying both sides by λ then leads to

$$\lambda S_N = \lambda + \lambda^2 + \ldots + \lambda^{N-1} + \lambda^N + \lambda^{N+1}$$

from which it can be seen that

$$\lambda S_N = S_N - 1 + \lambda^{N+1}$$

and further that

$$S_N = \frac{\lambda^{N+1} - 1}{\lambda - 1}$$

From this we can see that if $|\lambda| < 1$

$$\lim_{N \to \infty} S_N = \frac{1}{1 - \lambda}$$

Therefore it can be said that the geometric series is conditionally convergent, the condition being that $|\lambda| < 1$.

10.5.4 D'Alembert's ratio test

For the above examples, it was possible to determine whether the series was convergent by evaluating closed-form expressions for the series (as was done for the arithmetic and geometric series) or by looking for obvious signs of divergence in the series (as was done for the harmonic series). However, for many series these simple techniques are not possible. Instead, alternative convergence tests are required. Here we will consider D'Alembert's ratio test. Jean-Baptiste le Rond D'Alembert was a famous mathematician, philosopher and musician from 18th century France.

Consider the series

$$S = \sum_{n=0}^{\infty} a_n$$

and the ratio term

$$L = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

The ratio test states:

- if $L < 1$ then the series is convergent;
- if $L > 1$ then the series is divergent;
- if $L = 1$ or the limit fails to exist, then the test is inconclusive, because there exist both covergent and divergent series that satisfy this case.

Consider the case when $a_n = b_n x^n$, then

$$L = \lim_{n\to\infty} \left| \frac{b_{n+1} x^{n+1}}{b_n x^n} \right| = \lim_{n\to\infty} \left| \frac{b_{n+1}}{b_n} \right| |x|$$

The series is therefore convergent if

$$\lim_{n\to\infty} \left| \frac{b_{n+1}}{b_n} \right| |x| < 1$$

so that a necessary condition of convergence is that

$$|x| < \lim_{n\to\infty} \left| \frac{b_n}{b_{n+1}} \right|$$

In this way, $\lim\limits_{n\to\infty} |b_n/b_{n+1}|$ represents the radius of convergence for the series , $\sum\limits_{n=0}^{\infty} b_n x^n$.

10.6 Problem sheet

Problem 10.1 (see Worked Solution 10.1)

Derive the power series expansions for the following functions:

a) $\cos x$ b) $\ln(1+ax)$ c) $\ln(1+x^2)$

Problem 10.2 (see Worked Solution 10.2)

Using results obtained during class and results obtained in Question 1 above, determine the following limits:

a) $\lim_{x\to 0} \dfrac{\sin x}{x}$ b) $\lim_{x\to 0} \dfrac{1-\cos x}{x^2}$ c) $\lim_{x\to 0} \dfrac{\ln(1+ax)}{x}$

d) $\lim_{x\to 0} \dfrac{1-2e^x+e^{2x}}{x^2}$ e*) $\lim_{x\to 1} \dfrac{\ln x}{x^2-1}$

*For part e) make the substitution $x = y+1$ and use the series expansion for $\ln(1+y)$.

Problem 10.3 (see Worked Solution 10.3)

Determine whether the functions listed in Question 1 above are convergent and determine their radii of convergence where appropriate.

Problem 10.4 (see Worked Solution 10.4)

Approximate the integral, $\displaystyle\int_0^1 \frac{1}{1+\varepsilon x^{\pi}} dx$, to order ε^2.

10.7 Worked solutions

Worked Solution 10.1 (see Problem 10.1)

Recall that $f(x) = \sum_{n=0}^{\infty} f^{(n)}(0)\frac{x^n}{n!}$

a) Obtain the series expansion for $\cos x$

$$\begin{array}{ll}
y = \cos x & y(0) = 1 \\
y^{(1)}(x) = -\sin x & y^{(1)}(0) = 0 \\
y^{(2)}(x) = -\cos x & y^{(2)}(0) = -1 \\
y^{(3)}(x) = \sin x & y^{(3)}(0) = 0 \\
y^{(4)}(x) = \cos x & y^{(4)}(0) = 1
\end{array}$$

from which it follows that $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + O(x^6)$

Only the even terms are non-zero, $\therefore \cos x = \sum_{n=0}^{\infty} a_{2n}$

and the odd values of n are negative, $\therefore \cos x = \sum_{n=0}^{\infty} (-1)^n b_{2n}$

It can now be seen that $b_{2n} = \frac{x^{2n}}{(2n)!}$

Hence $\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$

b) Obtain the series expansion for $\ln(1+ax)$

$$\begin{array}{ll}
y = \ln(1+ax) & y(0) = 0 \\
y^{(1)}(x) = a(1+ax)^{-1} & y^{(1)}(0) = a \\
y^{(2)}(x) = -a^2(1+ax)^{-2} & y^{(2)}(0) = -a^2 \\
y^{(3)}(x) = 2a^3(1+ax)^{-3} & y^{(3)}(0) = 2a^3 \\
y^{(4)}(x) = -6a^4(1+ax)^{-4} & y^{(4)}(0) = -3!a^4
\end{array}$$

from which it can be seen that $y^{(n)}(0) = (-1)^{n-1}(n-1)!a^n$

Hence it can be said that $\ln(1+ax) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(ax)^n}{n}$

Note especially that the summation is taken from $n = 1$ as the zeroth term is zero.

c) Obtain the series expansion for $\ln(1+x^2)$

Let $y = \ln(1+x^2)$ and $u = x^2$.

It follows that $y = \ln(1+u) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}u^n}{n}$

Therefore $\ln(1+x^2) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^{2n}}{n}$

Worked Solution 10.2 (see Problem 10.2)

a) Determine $\lim_{x\to 0} \frac{\sin x}{x}$

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

It follows that $\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} + O(x^6)$

Therefore $\lim_{x\to 0} \frac{\sin x}{x} = 1$

b) Determine $\lim_{x\to 0} \frac{1-\cos x}{x^2}$

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

It follows that $\dfrac{1-\cos x}{x^2} = \dfrac{1}{2!} - \dfrac{x^2}{4!} + O(x^4)$

Therefore $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x^2} = \dfrac{1}{2}$

c) Determine $\lim\limits_{x\to 0} \dfrac{\ln(1+ax)}{x}$

Recall that $\ln(1+ax) = ax - \dfrac{a^2x^2}{2} + O(x^3)$

It follows that $\dfrac{\ln(1+ax)}{x} = a - \dfrac{a^2x}{2} + O(x^2)$

Therefore $\lim\limits_{x\to 0} \dfrac{\ln(1+ax)}{x} = a$

d) Determine $\lim\limits_{x\to 0} \dfrac{1-2e^x+e^{2x}}{x^2}$

Recall that $e^x = 1 + x + \dfrac{x^2}{2!} + O(x^3)$

It follows that:

$$\begin{aligned} 1-2e^x+e^{2x} &= 1-2\left(1+x+\frac{x^2}{2!}\right)+1+2x+\frac{4x^2}{2!}+O(x^3) \\ &= 1-2+1-2x+2x-x^2+\frac{4x^2}{2}+O(x^3) \\ &= x^2+O(x^3) \end{aligned}$$

Therefore $\lim\limits_{x\to 0} \dfrac{1-2e^x+e^{2x}}{x^2} = 1$

e) Determine $\lim\limits_{x\to 1} \dfrac{\ln x}{x^2-1}$

Let $x = y + 1$ such that it can be said that $\lim_{x \to 1} \frac{\ln x}{x^2 - 1} = \lim_{y \to 0} \frac{\ln(y+1)}{(y+1)^2 - 1}$

Recall that $\ln(y+1) = y - \frac{y^2}{2} + O(y^3)$

It follows that $\frac{\ln(y+1)}{(y+1)^2 - 1} = \frac{y - y^2/2 + O(y^3)}{y^2 + 2y} = \frac{y + O(y^2)}{2y + O(y^2)}$

from which it can be seen that $\lim_{y \to 0} \frac{\ln(y+1)}{(y+1)^2 - 1} = \frac{1}{2}$

Hence $\lim_{x \to 1} \frac{\ln x}{x^2 - 1} = \frac{1}{2}$

Worked Solution 10.3 (see Problem 10.3)

Here we will apply the ratio test and determine $L = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right|$

a) Recall $\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$

from which it can be seen that $a_n = \frac{(-1)^n x^{2n}}{(2n)!}$

and $a_{n+1} = \frac{(-1)^{n+1} x^{2n+2}}{(2n+2)!}$

Hence $\frac{a_{n+1}}{a_n} = \frac{(-1)^{n+1} x^{2n+2}}{(2n+2)!} \frac{(2n)!}{(-1)^n x^{2n}} = \frac{-x^2}{(2n+2)(2n+1)}$

and $L = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = 0$

which is < 1, from which it can be understood that $\cos x$ is unconditionally convergent.

b) Recall $\ln(1+ax) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(ax)^n}{n}$

from which it can be seen that $a_n = \frac{(-1)^{n+1}(ax)^n}{n}$

and $a_{n+1} = \frac{(-1)^{n+2}(ax)^{n+1}}{n+1}$

Hence $\frac{a_{n+1}}{a_n} = \frac{(-1)^{n+2}(ax)^{n+1}}{n+1} \frac{n}{(-1)^{n+1}(ax)^n} = -\frac{nax}{n+1}$

and $L = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = ax$

The series will converge when $L < 1$, which is true when $|ax| < 1$. So the radius of convergence is $|a^{-1}|$.

c) Recall $\ln(1+x^2) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^{2n}}{n}$

from which it can be seen that $a_n = \frac{(-1)^{n+1}x^{2n}}{n}$

and $a_{n+1} = \frac{(-1)^{n+2}x^{2n+2}}{n+1}$

Hence $\frac{a_{n+1}}{a_n} = \frac{(-1)^{n+2}x^{2n+2}}{n+1} \frac{n}{(-1)^{n+1}x^{2n}} = -\frac{nx^2}{n+1}$

and $L = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = x^2$

The series will converge when $L < 1$, which is true when $|x^2| < 1$. So the radius of convergence is 1.

Worked Solution 10.4 (see Problem 10.4)

To evaluate the integral, $\int_0^1 \frac{1}{1+\varepsilon x^{\pi}} dx$, to order ε^2, it first necessary to find a series expansion for $(1+\varepsilon x^{\pi})^{-1}$.

$$\begin{array}{ll} y(\varepsilon) = (1+\varepsilon x^{\pi})^{-1} & y(0) = 1 \\ y^{(1)}(\varepsilon) = -x^{\pi}(1+\varepsilon x^{\pi})^{-2} & y^{(1)}(0) = -x^{\pi} \\ y^{(2)}(\varepsilon) = 2x^{2\pi}(1+\varepsilon x^{\pi})^{-3} & y^{(2)}(0) = 2x^{2\pi} \end{array}$$

from which it follows that $\dfrac{1}{1+\varepsilon x^{\pi}} = 1-\varepsilon x^{\pi}+\varepsilon^2 x^{2\pi}+O(\varepsilon^3)$

and hence

$$\begin{aligned} \int_0^1 \frac{1}{1+\varepsilon x^{\pi}}dx &= \int_0^1 1-\varepsilon x^{\pi}+\varepsilon^2 x^{2\pi}dx+O(\varepsilon^3) \\ &= \left[x-\frac{\varepsilon x^{\pi+1}}{\pi+1}+\frac{\varepsilon^2 x^{2\pi+1}}{2\pi+1}\right]_0^1+O(\varepsilon^3) \\ &= 1-\frac{\varepsilon}{\pi+1}+\frac{\varepsilon^2}{2\pi+1}+O(\varepsilon^3) \end{aligned}$$

11

Partial differentiation

11.1 Learning outcomes

You should be able to:

- Understand the difference between partial derivatives and complete derivatives.
- Use logarithmic differentiation to derive an expression for the total derivative of a function in terms of the partial derivatives of that function with respect to all its dependent variables.
- Derive the same expression as above using a power series expansion.
- Check solutions of partial differential equations.

11.2 A result from logarithmic differentiation

Challenge 11.1 Given that

$$\frac{d\ln f}{dx} = \frac{1}{f}\frac{df}{dx} \tag{11.1}$$

provide an expression for the derivative of y with respect to x where

$$y = PQ^{-2}\sin R$$

and P, Q and R are all functions of x.

Taking the logs of both sides of the equation for y leads to

$$\ln y = \ln\left(PQ^{-2}\sin R\right) = \ln P - 2\ln Q + \ln\sin R$$

Given Eq. (11.1), differentiating both sides of the above equation with respect to x leads to

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{P}\frac{dP}{dx} - \frac{2}{Q}\frac{dQ}{dx} + \frac{1}{\sin R}\frac{d\sin R}{dx}$$

The derivative of $\sin R$ with respect to x can be found by making the substitution, $u = \sin R$, and applying the chain-rule:

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{P}\frac{dP}{dx} - \frac{2}{Q}\frac{dQ}{dx} + \frac{1}{\sin R}\frac{du}{dR}\frac{dR}{dx}$$

from which it follows that

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{P}\frac{dP}{dx} - \frac{2}{Q}\frac{dQ}{dx} + \frac{\cos R}{\sin R}\frac{dR}{dx}$$

Multiplying both sides by y then leads to

$$\frac{dy}{dx} = Q^{-2}\sin R\frac{dP}{dx} - 2PQ^{-3}\sin R\frac{dQ}{dx} + PQ^{-2}\cos R\frac{dR}{dx} \tag{11.2}$$

Challenge 11.2 Find the derivatives of y with respect to P, Q and R and substitute these into Eq. (11.2).

$$\frac{dy}{dP} = Q^{-2}\sin R, \quad \frac{dy}{dQ} = -2PQ^{-3}\sin R, \quad \frac{dy}{dR} = PQ^{-2}\cos R$$

Substituting these into Eq. (11.2) leads to

$$\frac{dy}{dx} = \frac{dy}{dP}\frac{dP}{dx} + \frac{dy}{dQ}\frac{dQ}{dx} + \frac{dy}{dR}\frac{dR}{dx} \tag{11.3}$$

The concept of partial derivatives

In fact there are two types of derivative in Eq. (11.3). The dy/dx, dP/dx, dQ/dx and dR/dx represent the total derivatives of y, P, Q and R with respect to x, respectively. However, note that y is a function of P, Q and R, which in turn are functions of x. Therefore, the term dy/dP represents the derivative of y with respect to P whilst the other variables Q and R are assumed to be independent of P (i.e., held constant).

This latter type of derivative is referred to as a partial derivative because y has been only partially differentiated with respect to P, whilst the other variables have been held constant. A similar statement can be made for the dy/dQ and dy/dR terms. Consequently a more appropriate way to write Eq. (11.3) is to say

$$\frac{dy}{dx} = \left(\frac{\partial y}{\partial P}\right)_{Q,R}\frac{dP}{dx} + \left(\frac{\partial y}{\partial Q}\right)_{P,R}\frac{dQ}{dx} + \left(\frac{\partial y}{\partial R}\right)_{Q,P}\frac{dR}{dx} \tag{11.4}$$

Note that we should now only use d for total derivatives and ∂ for partial derivatives. Furthermore, because the symbol ∂ implies that $\partial y/\partial P$ is the partial derivative of y where all other variables have been held constant, it is also reasonable to state more simply that

$$\frac{dy}{dx} = \frac{\partial y}{\partial P}\frac{dP}{dx} + \frac{\partial y}{\partial Q}\frac{dQ}{dx} + \frac{\partial y}{\partial R}\frac{dR}{dx} \tag{11.5}$$

It is possible to generalize Eq. (11.5) to consider the total derivative of y with respect to an unspecified independent variable by writing

$$dy = \frac{\partial y}{\partial P}dP + \frac{\partial y}{\partial Q}dQ + \frac{\partial y}{\partial R}dR$$

such that if instead we wanted the total derivative of y with respect to an alternative variable, say t, we can immediately see that (in this case)

$$\frac{dy}{dt} = \frac{\partial y}{\partial P}\frac{dP}{dt} + \frac{\partial y}{\partial Q}\frac{dQ}{dt} + \frac{\partial y}{\partial R}\frac{dR}{dt}$$

11.3 A result from power series

Recall the power series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n = f(0) + f^{(1)}(0)x + O(x^2) \tag{11.6}$$

The above equation shows an expansion of $f(x)$ about a point where $x = 0$. In a similar way, $f(x)$ can be expanded about a point where x is some arbitrary value, a, as follows:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f^{(1)}(a)(x-a) + O((x-a)^2)$$

which is known as the Taylor series.

Now consider the situation when $a = z$ and $x = z + \delta z$. It then follows that

$$f(z+\delta z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z)}{n!}\delta z^n = f(z) + f^{(1)}(z)\delta z + O(\delta z^2)$$

from which it can be seen that the function, f, changes by a value of $f(z+\delta z) - f(z)$, as one moves from z to $z+\delta z$. Defining this change in f as δf, it follows that

$$\delta f = f^{(1)}(z)\delta z + O(\delta z^2)$$

Now consider a function of three variables: P, Q and R, defined by $y = f(P,Q,R)$. If P increases by a small amount, δP, there will be an associated small increase in y, δy, which can be found from

$$\delta y = A\delta P + O(\delta P^2)$$

where A is an, as yet, undefined constant.

There will also be changes in y if Q or R are increased instead:

$$\delta y = B\delta Q + O(\delta Q^2)$$

$$\delta y = C\delta R + O(\delta R^2)$$

where B and C are additional unknown coefficients. Furthermore, if these changes in P, Q and R occur simultaneously, it can be understood that

$$\delta y = A\delta P + B\delta Q + C\delta R + O(\delta P^2, \delta Q^2, \delta R^2)$$

Challenge 11.3 Consider a function $y = f(P, Q, R)$.

From the Taylor series we can say that

$$\delta y = A\delta P + B\delta Q + C\delta R + O(\delta P^2, \delta Q^2, \delta R^2)$$

a) Determine an expression for $\delta y / \delta P$ when $\delta Q = \delta R = 0$.

b) Determine an expression for A in terms of a partial derivative of y.

c) Determine similar expressions for B and C.

d) Determine an expression for δy in terms of partial derivatives of y, with A, B and C all replaced with the expressions from b) and c) above.

a) $\dfrac{\delta y}{\delta P} = A + O(\delta P)$

b) $\lim\limits_{\delta P \to 0} \dfrac{\delta y}{\delta P} = \left(\dfrac{\partial y}{\partial P}\right)_{Q,R}$ therefore $A = \left(\dfrac{\partial y}{\partial P}\right)_{Q,R}$

c) $B = \left(\dfrac{\partial y}{\partial Q}\right)_{P,R}$ and $C = \left(\dfrac{\partial y}{\partial R}\right)_{P,Q}$

d) $\delta y = \left(\dfrac{\partial y}{\partial P}\right)_{Q,R} \delta P + \left(\dfrac{\partial y}{\partial Q}\right)_{P,R} \delta Q + \left(\dfrac{\partial y}{\partial R}\right)_{P,R} \delta R + O(\delta P^2, \delta Q^2, \delta R^2)$

11.4 The concept of a total derivative

It can be seen that for a general function of multiple variables, $y = f(P,Q,R)$, the total derivative, dy, is obtained from

$$dy = \left(\frac{\partial y}{\partial P}\right)_{Q,R} dP + \left(\frac{\partial y}{\partial Q}\right)_{P,R} dQ + \left(\frac{\partial y}{\partial R}\right)_{P,R} dR$$

The above example is for a function of three different variables. However, the same idea holds for any number of variable dependencies.

Consider a function, $g = f(x,y,z,t)$, where x [L], y [L] and z [L] are three orthogonal directions in space and t [T] is time. The total derivative of g with respect to time should therefore be written as

$$\frac{dg}{dt} = \frac{\partial g}{\partial x}\frac{dx}{dt} + \frac{\partial g}{\partial y}\frac{dy}{dt} + \frac{\partial g}{\partial z}\frac{dz}{dt} + \frac{\partial g}{\partial t}$$

From the field of thermodynamics, consider internal energy per unit mass, u [L^2T^{-2}], which is a function of pressure, P [$ML^{-1}T^{-2}$], temperature, T [Θ], and molar volume, v [L^3]. The total derivative of u with respect to time is found from

$$\frac{du}{dt} = \left(\frac{\partial u}{\partial P}\right)_{T,v} \frac{dP}{dt} + \left(\frac{\partial u}{\partial T}\right)_{P,v} \frac{dT}{dt} + \left(\frac{\partial u}{\partial v}\right)_{P,T} \frac{dv}{dt}$$

Furthermore, if u is a function of x, y, z and t, it can also be seen that

$$\frac{\partial u}{\partial x} = \left(\frac{\partial u}{\partial P}\right)_{T,v} \frac{\partial P}{\partial x} + \left(\frac{\partial u}{\partial T}\right)_{P,v} \frac{\partial T}{\partial x} + \left(\frac{\partial u}{\partial v}\right)_{P,T} \frac{\partial v}{\partial x}$$

$$\frac{\partial u}{\partial y} = \left(\frac{\partial u}{\partial P}\right)_{T,v} \frac{\partial P}{\partial y} + \left(\frac{\partial u}{\partial T}\right)_{P,v} \frac{\partial T}{\partial y} + \left(\frac{\partial u}{\partial v}\right)_{P,T} \frac{\partial v}{\partial y}$$

$$\frac{\partial u}{\partial z} = \left(\frac{\partial u}{\partial P}\right)_{T,v} \frac{\partial P}{\partial z} + \left(\frac{\partial u}{\partial T}\right)_{P,v} \frac{\partial T}{\partial z} + \left(\frac{\partial u}{\partial v}\right)_{P,T} \frac{\partial v}{\partial z}$$

$$\frac{\partial u}{\partial t} = \left(\frac{\partial u}{\partial P}\right)_{T,v} \frac{\partial P}{\partial t} + \left(\frac{\partial u}{\partial T}\right)_{P,v} \frac{\partial T}{\partial t} + \left(\frac{\partial u}{\partial v}\right)_{P,T} \frac{\partial v}{\partial t}$$

Dimensional analysis

At this stage it is useful to also to discuss the concept of dimensional analysis. Four physical dimensions of interest are:

[M]	-	mass
[L]	-	length
[T]	-	time
[Θ]	-	temperature

Challenge 11.4 Given that Energy = Force × Distance and Force = Mass × Acceleration, determine the physical dimensions of an energy per unit mass.

First determine the dimensions of a force:

$$\text{Force} = \text{Mass} \times \text{Acceleration} = [\text{M}] \times [\text{LT}^{-2}] = [\text{MLT}^{-2}]$$

Now determine the dimensions of an energy:

$$\text{Energy} = \text{Force} \times \text{Distance} = [\text{MLT}^{-2}] \times [\text{L}] = [\text{ML}^2\text{T}^{-2}]$$

Now determine the dimensions of an energy per unit mass:

$$\text{Energy per unit Mass} = \text{Energy} \div \text{Mass} = [\text{ML}^2\text{T}^{-2}] \div [\text{M}] = [\text{L}^2\text{T}^{-2}]$$

Challenge 11.5 Imagine we want to transform a temperature field from a Cartesian coordinate system, $T = f(x,y)$, to a polar coordinate system, $T = g(r,\theta)$, where $x = r\cos\theta$ and $y = r\sin\theta$. Use partial differentiation to determine expressions for:

$$\frac{\partial T}{\partial x} \quad \text{and} \quad \frac{\partial T}{\partial y} \tag{11.7}$$

purely in terms of T, r and θ.

Note that r [L] is the radial distance from the origin and θ [-] is the anti-clockwise angle of direction about the origin from the x axis

Because $T = g(r,\theta)$, the general expression for the total derivative of T takes the form

$$dT = \frac{\partial T}{\partial r}dr + \frac{\partial T}{\partial \theta}d\theta$$

and therefore:

$$\frac{\partial T}{\partial x} = \frac{\partial T}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial T}{\partial \theta}\frac{\partial \theta}{\partial x} \tag{11.8}$$

$$\frac{\partial T}{\partial y} = \frac{\partial T}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial T}{\partial \theta}\frac{\partial \theta}{\partial y} \tag{11.9}$$

Recall that $x = r\cos\theta$ and $y = r\sin\theta$, therefore:

$$\frac{\partial x}{\partial r} = \cos\theta, \quad \frac{\partial y}{\partial r} = \sin\theta, \quad \frac{\partial x}{\partial \theta} = -r\sin\theta, \quad \frac{\partial y}{\partial \theta} = r\cos\theta$$

which on substitution into Eq. (11.8) and (11.9) lead to:

$$\frac{\partial T}{\partial x} = \sec\theta\frac{\partial T}{\partial r} - \frac{\mathrm{cosec}\theta}{r}\frac{\partial T}{\partial \theta}$$

$$\frac{\partial T}{\partial y} = \mathrm{cosec}\theta\frac{\partial T}{\partial r} + \frac{\sec\theta}{r}\frac{\partial T}{\partial \theta}$$

11.5 Problem sheet

Problem 11.1 (see Worked Solution 11.1)

Determine $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ for the following functions:

a) $f(x,y) = 3x + 4y$ b) $f(x,y) = xy^3 + x^2y^2$

c) $f(x,y) = x^3y + e^x$ d) $f(x,y) = xe^{2x+3y}$

e) $f(x,y) = \dfrac{x-y}{x+y}$ f) $f(x,y) = 2x\sin(x^2y)$

Problem 11.2 (see Worked Solution 11.2)

Determine $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$ and $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$ for the following functions:

a) $f(x,y) = x^2 + xy - y^2$ b) $f(x,y) = x^2 \sin y + y^2 \cos x$

c) $f(x,y) = \dfrac{y}{x}\ln x$ d) $f(x,y) = \dfrac{1}{x^2 + y^2}$

Problem 11.3 (see Worked Solution 11.3)

Consider the volcano Mauna Kea. Assuming the volcano to be comprised of a perfect cone of height, $H(t)$, with a base of radius, $R(t)$, the volume, $V(R,H)$, can be determined from

$$V = \frac{\pi}{3}R(t)^2H(t)$$

where t is time.

a) Determine an expression for the total derivative of V with respect to time in terms of R, H and t.

b) The angle of volcano slope, $\theta(R,H)$, is found from

$$\theta = \arctan\left(\frac{H(t)}{R(t)}\right)$$

Determine an expression for the total derivative of θ with respect to time in terms of R, H and t.

Problem 11.4 (see Worked Solution 11.4)

Consider the partial differential equation (PDE)

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = 0$$

Suppose that

$$f(x,y) = g(s)$$

where $s = y/x$. Obtain expressions for the partial derivatives of $f(x,y)$ with respect to x and y in terms of g, s, x and y, and verify that $f(x,y) = g(s)$ is indeed a solution to the PDE above.

11.6 Worked solutions

Worked Solution 11.1 (see Problem 11.1)

a) $f = 3x + 4y$

therefore $\dfrac{\partial f}{\partial x} = 3$ and $\dfrac{\partial f}{\partial y} = 4$

b) $f = xy^3 + x^2y^2$

therefore $\dfrac{\partial f}{\partial x} = y^3 + 2xy^2$ and $\dfrac{\partial f}{\partial y} = 3xy^2 + 2x^2y$

c) $f = x^3y + e^x$

therefore $\dfrac{\partial f}{\partial x} = 3x^2y + e^x$ and $\dfrac{\partial f}{\partial y} = x^3$

d) $f = xe^{2x+3y}$

therefore $\dfrac{\partial f}{\partial x} = e^{2x+3y} + 2xe^{2x+3y} = (1+2x)e^{2x+3y}$

and $\dfrac{\partial f}{\partial y} = 3xe^{2x+3y}$

e) $f = \dfrac{x-y}{x+y}$

Taking logs of both sides leads to $\ln f = \ln(x-y) - \ln(x+y)$

It then follows that $\dfrac{1}{f}\dfrac{\partial f}{\partial x} = \dfrac{1}{x-y} - \dfrac{1}{x+y}$

therefore $\dfrac{\partial f}{\partial x} = \dfrac{x-y}{x+y}\left(\dfrac{1}{x-y} - \dfrac{1}{x+y}\right) = \dfrac{2y}{(x+y)^2}$

and $\dfrac{1}{f}\dfrac{\partial f}{\partial y} = -\dfrac{1}{x-y} - \dfrac{1}{x+y}$

therefore $\dfrac{\partial f}{\partial y} = -\dfrac{x-y}{x+y}\left(\dfrac{1}{x-y} + \dfrac{1}{x+y}\right) = -\dfrac{2x}{(x+y)^2}$

f) $f(x,y) = 2x\sin(x^2y)$

$$\frac{\partial f}{\partial x} = 2\sin(x^2y) + 2x \times 2xy\cos(x^2y) = 2\sin(x^2y) + 4x^2y\cos(x^2y)$$

$$\frac{\partial f}{\partial y} = 2x^3\cos(x^2y)$$

Worked Solution 11.2 (see Problem 11.2)

a) $f(x,y) = x^2 + xy - y^2$

$$\frac{\partial f}{\partial x} = 2x + y \qquad \frac{\partial f}{\partial y} = x - 2y$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 1 \qquad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = 1$$

b) $f(x,y) = x^2 \sin y + y^2 \cos x$

$$\frac{\partial f}{\partial x} = 2x\sin y - y^2 \sin x \qquad \frac{\partial f}{\partial y} = x^2 \cos y + 2y\cos x$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 2x\cos y - 2y\sin x \quad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = 2x\cos y - 2y\sin x$$

c) $f(x,y) = \dfrac{y}{x}\ln x$

$$\frac{\partial f}{\partial x} = \frac{y}{x^2} - \frac{y}{x^2}\ln x \qquad \frac{\partial f}{\partial y} = \frac{1}{x}\ln x$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{1}{x^2}(1 - \ln x) \quad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{1}{x^2}(1 - \ln x)$$

d) $f(x,y) = \dfrac{1}{x^2 + y^2}$

$$\frac{\partial f}{\partial x} = -\frac{2x}{(x^2+y^2)^2} \qquad \frac{\partial f}{\partial y} = -\frac{2y}{(x^2+y^2)^2}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{8xy}{(x^2+y^2)^3} \quad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{8xy}{(x^2+y^2)^3}$$

Worked Solution 11.3 (see Problem 11.3)

a) $V = \frac{\pi}{3}R(t)^2H(t)$

$$\frac{dV}{dt} = \frac{\partial V}{\partial R}\frac{dR}{dt} + \frac{\partial V}{\partial H}\frac{dH}{dt}$$

$$\frac{\partial V}{\partial R} = \frac{\pi}{3}2R(t)H(t)$$

$$\frac{\partial V}{\partial H} = \frac{\pi}{3}R(t)^2$$

Therefore $\frac{dV}{dt} = \frac{\pi}{3}R(t)\left(2H(t)\frac{dR}{dt} + R(t)\frac{dH}{dt}\right)$

b) $\theta = \arctan\left(\frac{H(t)}{R(t)}\right)$

$$\frac{d\theta}{dt} = \frac{\partial\theta}{\partial R}\frac{dR}{dt} + \frac{\partial\theta}{\partial H}\frac{dH}{dt}$$

Let $u = H(t)/R(t)$ such that $\theta = \arctan(u)$.

Recall that $\frac{d\theta}{du} = \frac{1}{1+u^2}$

Note that a "d" is used above, as opposed to a "∂", because the fact that $\theta = \arctan(u)$ suggests that θ is only a function of u, and therefore its derivative with respect to u has to be a total derivative.

Application of the chain rule then yields:

$$\frac{\partial\theta}{\partial R} = \frac{\partial u}{\partial R}\frac{d\theta}{du} = -\frac{H}{R^2}\frac{1}{1+u^2} = -\frac{H}{R^2}\frac{R^2}{R^2+H^2} = -\frac{H}{R^2+H^2}$$

$$\frac{\partial\theta}{\partial H} = \frac{\partial u}{\partial H}\frac{d\theta}{du} = \frac{1}{R}\frac{1}{1+u^2} = \frac{1}{R}\frac{R^2}{R^2+H^2} = \frac{R}{R^2+H^2}$$

Therefore it can be seen that

$$\frac{d\theta}{dt} = \frac{1}{R^2 + H^2}\left(R\frac{dH}{dt} - H\frac{dR}{dt}\right)$$

Worked Solution 11.4 (see Problem 11.4)

The strategy to be used here is to differentiate $g(s)$ with respect to x and y and then substitute these derivatives into the original PDE.

Using the chain rule, it can be said that:

$$\frac{\partial g}{\partial x} = \frac{\partial s}{\partial x}\frac{dg}{ds} \quad \text{and} \quad \frac{\partial g}{\partial y} = \frac{\partial s}{\partial y}\frac{dg}{ds}$$

Note that g is only a function of s. Therefore the derivative of g with respect to s is total as opposed to partial. Consequently, a "d" is used as opposed to "∂" for the derivative of g.

Then given the identity $s = y/x$:

$$\frac{\partial s}{\partial x} = -\frac{y}{x^2} \quad \text{and} \quad \frac{\partial s}{\partial y} = \frac{1}{x}$$

and hence it can be said that

$$\frac{\partial g}{\partial x} = -\frac{y}{x^2}\frac{dg}{ds} \quad \text{and} \quad \frac{\partial g}{\partial y} = \frac{1}{x}\frac{dg}{ds}$$

which, when substituted into the original PDE yields

$$-\frac{xy}{x^2}\frac{dg}{ds} + \frac{y}{x}\frac{dg}{ds} = 0$$

which is indeed true and $f(x,y) = g(s)$ is confirmed.

12

Diffusion and the error function

12.1 Learning outcomes

You should be able to:

- Derive the diffusion equation using a mass conservation statement and a control-volume.
- Apply a similarity transform to a partial differential equation to reduce it to an ordinary differential equation (ODE).
- Apply a dependent variable transform to a second-order ODE to reduce it to a first-order ODE.
- Discuss what the error function is and how it relates to both the normal distribution function and the diffusion equation.

12.2 Fick's first and second law

An important partial differential equation, used for a multitude of different applications, is the diffusion equation. In this example, we will consider the diffusion of solutes in water contained within a water saturated porous medium of porosity ϕ [-]. However, the same equation can be used to describe heat conduction in solids and fluid flow in porous media. Note that the porosity of a rock is found from the volume of voids in the rock divided by the total volume of the rock.

Let c [ML^{-3}] be the concentration of a solute in a given solution. More specifically, c represents the mass of solute per unit volume of solvent. Solute tends to migrate from areas of high concentration to areas of low concentration. The German physiologist, Adolf Fick, active in the late 19th century and early 20th century, was interested in gas diffusion across membranes. According to Fick's first law, solute flux is linearly proportional to the concentration gradient. The coefficient of proportionality is often referred to as the effective diffusion coefficient, which we will denote as D_E [$\mathrm{L}^2\mathrm{T}^{-1}$].

A solute flux represents the rate of mass movement per unit area of a surface. Fick's first law in Cartesian coordinates takes the form:

$$J_x = -D_E \frac{\partial c}{\partial x}, \quad J_y = -D_E \frac{\partial c}{\partial y}, \quad J_z = -D_E \frac{\partial c}{\partial z} \tag{12.1}$$

where J_x [$\mathrm{ML}^{-2}\mathrm{T}^{-1}$], J_y [$\mathrm{ML}^{-2}\mathrm{T}^{-1}$] and J_z [$\mathrm{ML}^{-2}\mathrm{T}^{-1}$] are solute fluxes in the x, y and z direction, respectively.

The diffusion equation is often referred to as Fick's second law. To obtain the diffusion equation we need to invoke the concept of mass conservation in a control-volume.

12.3 Mass conservation in a control-volume

A control-volume represents an infinitesimal fixed volume of length, breadth and height, δx, δy and δz, respectively.

The fundamental concept of mass conservation is simply that the mass of a substance into a control-volume minus the mass of that substance out of a control-volume equals the change of mass of that substance within the control-volume.

Challenge 12.1 Given Fick's first law, use mass conservation across the control-volume, shown in Fig. 12.1, to derive Fick's second law of diffusion for a rock with porosity, ϕ [-].

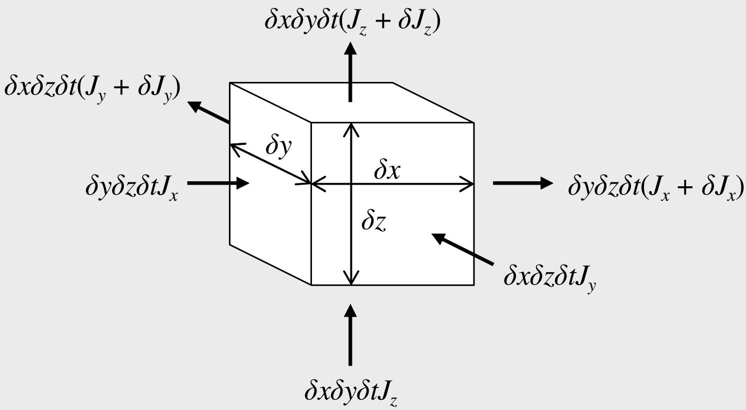


Figure 12.1: Chemical diffusion through a control-volume.

The mass of solute into our control-volume during a finite period of time, δt, can be written as

$$\delta y \delta z \delta t J_x + \delta x \delta z \delta t J_y + \delta x \delta y \delta t J_z$$

The mass of solute out of our control-volume during the same period of time can be written as

$$\delta y \delta z \delta t (J_x + \delta J_x) + \delta x \delta z \delta t (J_y + \delta J_y) + \delta x \delta y \delta t (J_z + \delta J_z)$$

The total mass of solute within the control-volume is found from

$$\delta x \delta y \delta z \phi c$$

The only variable in the above expression that will change as a consequence of solute movement is c. Therefore, the change in mass of solute within the control-volume that takes place during the time period, δt, is found from

$$\delta x \delta y \delta z \phi \delta c$$

Invoking the idea that

$$\text{Mass In} - \text{Mass Out} = \text{Change in Mass}$$

(i.e., mass conservation) then leads to

$$\begin{aligned} &\delta y \delta z \delta t J_x + \delta x \delta z \delta t J_y + \delta x \delta y \delta t J_z \\ &- \delta y \delta z \delta t (J_x + \delta J_x) - \delta x \delta z \delta t (J_y + \delta J_y) - \delta x \delta y \delta t (J_z + \delta J_z) \\ &= \delta x \delta y \delta z \phi \delta c \end{aligned}$$

After some rearranging and simplification, this leads to

$$\phi \frac{\delta c}{\delta t} + \frac{\delta J_x}{\delta x} + \frac{\delta J_y}{\delta y} + \frac{\delta J_z}{\delta z} = 0$$

Now if we consider the limit as $\delta x \to 0$, $\delta y \to 0$, $\delta z \to 0$ and $\delta t \to 0$, it can be said that

$$\phi \frac{\partial c}{\partial t} + \frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z} = 0$$

and substituting Fick's first law (i.e., Eq. (12.1)) then leads to

$$\phi\frac{\partial c}{\partial t} - \frac{\partial}{\partial x}\left(D_E\frac{\partial c}{\partial x}\right) - \frac{\partial}{\partial y}\left(D_E\frac{\partial c}{\partial y}\right) - \frac{\partial}{\partial z}\left(D_E\frac{\partial c}{\partial z}\right) = 0$$

which is Fick's second law of diffusion, otherwise known as the diffusion equation.

Also note that if D_E is constant,

$$\phi\frac{\partial c}{\partial t} - D_E\left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} + \frac{\partial^2 c}{\partial z^2}\right) = 0$$

12.4 Solution by similarity transform

Many one-dimensional partial differential equations (PDE) are self-similar, which means that they admit to a similarity transform such that they can be reduced to an ordinary differential equation (ODE) with respect to the chosen similarity transform. Once transformed to an ODE, the PDE can be solved using methods already developed for solving ODEs.

A PDE is self-similar if it is possible to change the independent variables in the denominator of all the partial derivatives to the generalized similarity transform, $z = x^a t^b$, and choose values of a and b such that x and t are eliminated from the PDE as well as the associated initial and boundary conditions.

Challenge 12.2 Consider the governing equation for one-dimensional diffusion of solute in a semi-infinite homogenous porous medium

$$\phi \frac{\partial c}{\partial t} - D_E \frac{\partial^2 c}{\partial x^2} = 0 \tag{12.2}$$

subjected to the following initial and boundary conditions:

$$\begin{array}{lll} c = c_i, & x \geq 0, & t = 0 \\ c = c_0, & x = 0, & t > 0 \\ c = c_i, & x \to \infty, & t > 0 \end{array} \tag{12.3}$$

where c_i and c_0 are the initial and boundary concentrations, respectively.

Substitute in the generalized similarity transform

$$z = x^a t^b \tag{12.4}$$

and determine values for a and b such that x and t are eliminated and the problem is demonstrated to be self-similar.

Application of the chain rule to Eq. (12.2) leads to

$$\phi \frac{\partial z}{\partial t} \frac{\partial c}{\partial z} - D_E \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} \frac{\partial c}{\partial z} \right) = 0 \tag{12.5}$$

then application of the product rule to the second term in Eq. (12.5) yields

$$\phi \frac{\partial z}{\partial t} \frac{\partial c}{\partial z} - D_E \left[\frac{\partial^2 z}{\partial x^2} \frac{\partial c}{\partial z} + \frac{\partial z}{\partial x} \frac{\partial}{\partial x} \left(\frac{\partial c}{\partial z} \right) \right] = 0$$

and finally application of the chain rule again to the last term leads to

$$\phi \frac{\partial z}{\partial t} \frac{\partial c}{\partial z} - D_E \left[\frac{\partial^2 z}{\partial x^2} \frac{\partial c}{\partial z} + \left(\frac{\partial z}{\partial x} \right)^2 \frac{\partial^2 c}{\partial z^2} \right] = 0 \tag{12.6}$$

From Eq. (12.4), it can be seen that:

$$\frac{\partial z}{\partial t} = bx^a t^{b-1} = \frac{bz}{t}$$

$$\frac{\partial z}{\partial x} = ax^{a-1} t^b = \frac{az}{x}$$

$$\frac{\partial^2 z}{\partial x^2} = a(a-1)x^{a-2} t^b = \frac{a(a-1)z}{x^2}$$

such that it can be said that (from Eq. (12.6))

$$\phi \frac{bz}{t}\frac{\partial c}{\partial z} - D_E \left(a(a-1)\frac{z}{x^2}\frac{\partial c}{\partial z} + \frac{a^2 z^2}{x^2}\frac{\partial^2 c}{\partial z^2} \right) = 0 \tag{12.7}$$

The problem is thought to be self similar if it is possible to choose values of a and b such that all the x and t terms are eliminated.

Multiplying both sides of Eq. (12.7) by x^2 it can be seen that

$$\phi \frac{bzx^2}{t}\frac{\partial c}{\partial z} - D_E \left(a(a-1)z\frac{\partial c}{\partial z} + a^2 z^2 \frac{\partial^2 c}{\partial z^2} \right) = 0$$

Recalling that $z = x^a t^b$, it is now apparent that all the x and t terms are eliminated when $a = 2$ and $b = -1$ such that $z = x^2 t^{-1}$. Setting $a = 2$ and $b = -1$ leads to

$$-\phi z^2 \frac{\partial c}{\partial z} - D_E \left(2z\frac{\partial c}{\partial z} + 4z^2 \frac{\partial^2 c}{\partial z^2} \right) = 0$$

and after some further rearrangement and simplification we have

$$\left(\phi + \frac{2D_E}{z} \right) \frac{\partial c}{\partial z} + 4D_E \frac{\partial^2 c}{\partial z^2} = 0 \tag{12.8}$$

which is an ODE.

Also note that the initial and boundary conditions, given earlier in Eq. (12.3), reduce to

$$\begin{aligned} c &= c_i, \quad z \to \infty \\ c &= c_0, \quad z = 0 \end{aligned} \tag{12.9}$$

12.5 Application of a dependent variable transform

The ODE in Eq. (12.8) remains difficult to solve using methods we have learnt previously. At this stage it is often useful to consider a dependent variable transform, to transform a second-order ODE into a first-order ODE.

Challenge 12.3 Substitute the dependent variable transform

$$u = \frac{\partial c}{\partial z} \tag{12.10}$$

into Eq. (12.8) and solve for u.

Substituting Eq. (12.10) into Eq. (12.8) leads to

$$\left(\phi + \frac{2D_E}{z}\right) u + 4D_E \frac{\partial u}{\partial z} = 0 \tag{12.11}$$

Separating variables we then have

$$\int \left(\phi + \frac{2D_E}{z}\right) dz = -4D_E \int \frac{1}{u} du$$

which leads to

$$\phi z + 2D_E \ln z = -4D_E \ln u + E$$

where E is an integration constant yet to be defined.

After some further rearrangement, it can be shown that

$$u = z^{-1/2} \exp\left(\frac{E-\phi z}{4D_E}\right) = \exp\left(\frac{E}{4D_E}\right) z^{-1/2} \exp\left(-\frac{\phi z}{4D_E}\right) \quad (12.12)$$

Challenge 12.4 Use Eq. (12.12) to obtain a solution for c in terms of an integral function of z.

Equating Eqs. (12.10) and (12.12) it can be understood that

$$\frac{\partial c}{\partial z} = \exp\left(\frac{E}{4D_E}\right) z^{-1/2} \exp\left(-\frac{\phi z}{4D_E}\right)$$

and integrating both sides with respect to z leads to:

$$c = \exp\left(\frac{E}{4D_E}\right) \int z^{-1/2} \exp\left(-\frac{\phi z}{4D_E}\right) dz \quad (12.13)$$

Challenge 12.5 The complementary error function, erfc(ζ), is defined by

$$\mathrm{erfc}(\zeta) = \frac{2}{\sqrt{\pi}} \int_{\zeta}^{\infty} e^{-\zeta^2} d\zeta \quad (12.14)$$

and has the following properties: $\mathrm{erfc}(0) = 1$ and $\mathrm{erfc}(\infty) = 0$.

Substitute the complementary error function into Eq. (12.13) and determine expressions for the two resulting integration constants such that the initial and boundary conditions, given by Eq. (12.9), are satisfied.

Comparing Eqs. (12.13) and (12.14) suggests that we should try to make the substitution

$$\zeta^2 = \frac{\phi z}{4D_E}$$

from which it follows that

$$z = \frac{4D_E\zeta^2}{\phi} \quad \text{and} \quad \frac{\partial z}{\partial \zeta} = \frac{8D_E\zeta}{\phi} = 4\left(\frac{D_E z}{\phi}\right)^{1/2}$$

It can therefore be understood that

$$c = 4\left(\frac{D_E}{\phi}\right)^{1/2} \exp\left(\frac{E}{4D_E}\right) \int e^{-\zeta^2} d\zeta$$

It is now possible to substitute the complementary error function to get

$$c = 2\left(\frac{\pi D_E}{\phi}\right)^{1/2} \exp\left(\frac{E}{4D_E}\right) \operatorname{erfc}(\zeta) + F \tag{12.15}$$

where F is another integration constant.

Recalling that $z = x^2 t^{-1}$ it follows that

$$\zeta^2 = \frac{\phi x^2}{4D_E t}$$

and the initial and boundary conditions, given by Eq. (12.9), reduce to

$$\begin{aligned} c &= c_i, \quad \zeta \to \infty \\ c &= c_0, \quad \zeta = 0 \end{aligned} \tag{12.16}$$

Noting that $\operatorname{erfc}(\infty) = 0$ it follows that

$$F = c_i \tag{12.17}$$

Furthermore, noting that $\operatorname{erfc}(0) = 1$ it follows that

$$2\left(\frac{\pi D_E}{\phi}\right)^{1/2} \exp\left(\frac{E}{4D_E}\right) = c_0 - c_i \tag{12.18}$$

Substituting Eqs. (12.17) and (12.18) into Eq. (12.15) leads to

$$\frac{c - c_i}{c_0 - c_i} = \operatorname{erfc}(\zeta)$$

12.6 The normal distribution function

Data is often described as being normally distributed. What is meant by this is that the histogram of the data looks like the normal distribution. The normal distribution is a continuous probability density function (PDF), which takes the form

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

where f is the probability density of a variable, x, which is normally distributed and has a mean and standard deviation of μ and σ, respectively.

Recall that for a discrete set of data, x_n, the mean and standard deviation can be calculated from:

$$\mu = \frac{1}{N}\sum_{n=1}^{N} x_n, \quad \sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x_n-\mu)^2}$$

where N is the number of available data points.

The meaning of probability density is better understood if ones considers the associated cumulative distribution function, $F(x)$, found from

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

If X is a random variable that is sampled from the PDF, $f(x)$, then $F(x)$ is the probability of X not exceeding x, often denoted as $P(X \leq x)$.

Challenge 12.6 The cumulative distribution function (CDF), $F(x)$, is found from

$$F(x) = \int_{-\infty}^{x} f(x)dx \tag{12.19}$$

where $f(x)$ is the associated probability density function (PDF).

Determine the CDF for the normal distribution PDF:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{12.20}$$

and write your answer in terms of the complementary error function:

$$\operatorname{erfc}(\zeta) = \frac{2}{\sqrt{\pi}} \int_{\zeta}^{\infty} e^{-\zeta^2} d\zeta \tag{12.21}$$

Substituting Eq. (12.20) into Eq. (12.19) leads to

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \tag{12.22}$$

Comparing Eqs. (12.22) and (12.21) suggests we should try to make the substitution

$$\zeta = \frac{x-\mu}{\sigma\sqrt{2}} \tag{12.23}$$

from which we get

$$\frac{d\zeta}{dx} = \frac{1}{\sigma\sqrt{2}} \tag{12.24}$$

and

$$F(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\zeta} e^{-\zeta^2} d\zeta \tag{12.25}$$

It can be seen that this is not an appropriate substitution because the limits in the integral to do not match with those in Eq.

(12.21). An alternative possibility is to try substituting

$$\zeta = \frac{\mu - x}{\sigma\sqrt{2}} \tag{12.26}$$

from which we get

$$\frac{d\zeta}{dx} = -\frac{1}{\sigma\sqrt{2}} \tag{12.27}$$

and

$$F(x) = \frac{1}{\sqrt{\pi}} \int_{\zeta}^{\infty} e^{-\zeta^2} d\zeta \tag{12.28}$$

It therefore follows that

$$F(x) = \frac{1}{2}\operatorname{erfc}\left(\frac{\mu - x}{\sigma\sqrt{2}}\right) \tag{12.29}$$

Relation to the error function

The complementary error function is related to another function of interest, the error function, $\operatorname{erf}(\zeta)$, by the formula $\operatorname{erfc}(\zeta) = 1 - \operatorname{erf}(\zeta)$. The error function gets its name from its use in error analysis where error is assumed to be a random but normally distributed process. These ideas were originally developed by the German mathematician, Carl Friedrich Gauss, during the early 19th century.

It is true that many data appear to be normally distributed. Interestingly, $\operatorname{erfc}(\zeta)$ is a solution to the diffusion equation.

12.7 Problem sheet

Problem 12.1 (see Worked Solution 12.1)

Consider the one-dimensional heat conduction equation

$$\rho c_p \frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial x^2}$$

where ρ [ML^3] is density, c_p [L^{2}T$^{-2}\Theta^{-1}$] is constant-pressure specific heat capacity, T [Θ] is temperature, t [T] is time, k [MLT$^{-3}\Theta^{-1}$] is thermal conductivity and x [L] is distance.

Determine expressions for a and b for when

$$T = C\exp(-ax)\sin(bx - \omega t)$$

is a solution, where C and ω are undefined constants associated with the initial and boundary conditions.

Problem 12.2 (see Worked Solution 12.2)

Consider the one-dimensional groundwater flow equation

$$\phi(c_r + c_w)\frac{\partial p}{\partial t} = \frac{k}{\mu}\frac{\partial^2 p}{\partial x^2}$$

where ϕ [-] is porosity, c_r [M^{-1}LT^2] is rock compressibility c_w [M^{-1}LT^2] is water compressibility, p [ML^{-1}T^{-2}] is pore-pressure, t is time, k [L^2] is permeability, μ [ML^{-1}T^{-1}] is the dynamic viscosity of water and x [L] is distance.

Show that

$$p(x,t) = \sqrt{\frac{\mu}{\phi(c_r + c_w)t}}\exp\left(-\frac{\phi(c_r + c_w)\mu x^2}{4kt}\right)$$

is a solution.

Problem 12.3 (see Worked Solution 12.3)

Consider the one-dimensional chemical diffusion equation

$$\frac{\partial c}{\partial t} = D_A\frac{\partial^2 c}{\partial x^2}$$

where c [ML^{-3}] is solute concentration, D_A [L^{2}T^{-1}] is the apparent diffusion coefficient, t is time and x [L] is distance.

a) Make the change of variables

$$c(x,t) = t^{-1/2} f(s), \quad s = \frac{x}{\sqrt{4D_A t}}$$

to derive a new equation.

b) Show that the total amount of c, found from

$$M = \int_{-\infty}^{\infty} c(x,t)dx$$

is constant.

c) Show that

$$f(s) = \exp(-s^2)$$

is a solution.

12.8 Worked solutions

Worked Solution 12.1 (see Problem 12.1)

The first step is to partially differentiate the proposed solution:

$$
\begin{aligned}
T &= Ce^{-ax}\sin(bx-\omega t) \\
\frac{\partial T}{\partial t} &= -\omega Ce^{-ax}\cos(bx-\omega t) \\
\frac{\partial T}{\partial x} &= Ce^{-ax}\left[b\cos(bx-\omega t)-a\sin(bx-\omega t)\right] \\
\frac{\partial^2 T}{\partial x^2} &= Ce^{-ax}\left[-b^2\sin(bx-\omega t)-ab\cos(bx-\omega t)\right] \\
&\quad -aCe^{-ax}\left[b\cos(bx-\omega t)-a\sin(bx-\omega t)\right] \\
&= Ce^{-ax}\left[(a^2-b^2)\sin(bx-\omega t)-2ab\cos(bx-\omega t)\right]
\end{aligned}
$$

Substituting these back into the original PDE then leads to

$$
\begin{aligned}
&-\rho c_p \omega Ce^{-ax}\cos(bx-\omega t) \\
&= kCe^{-ax}\left[(a^2-b^2)\sin(bx-\omega t)-2ab\cos(bx-\omega t)\right]
\end{aligned}
$$

which reduces further to

$$k(a^2-b^2)\tan(bx-\omega t)-2abk+\rho c_p\omega=0$$

The above equation cannot be a solution to the original PDE with the $\tan(bx-\omega t)$ term. Therefore, it is necessary for $b=\pm a$ to eliminate the tan term.

If we consider $b=a$, then

$$-2a^2k+\rho c_p\omega=0$$

from which it can now be understood that

$$a=\pm\left(\frac{\rho c_p\omega}{2k}\right)^{1/2}$$

Alternatively, if we consider $b = -a$, then

$$2a^2k + \rho c_p \omega = 0$$

from which it can now be understood that

$$a = \pm i \left(\frac{\rho c_p \omega}{2k} \right)^{1/2}$$

Worked Solution 12.2 (see Problem 12.2)

Again, it is necessary to partially differentiate the proposed solution. However, before doing this, it is convenient to make the following substitutions:

$$A = \sqrt{\frac{\mu}{\phi(c_r + c_w)}}, \quad B = -\frac{\phi(c_r + c_w)\mu}{4k}$$

such that

$$\begin{aligned}
p &= At^{-1/2} \exp\left(Bx^2 t^{-1}\right) \\
\frac{\partial p}{\partial t} &= -At^{-1/2} \exp\left(Bx^2 t^{-1}\right) \left(\frac{1}{2t} + \frac{Bx^2}{t^2} \right) \\
\frac{\partial p}{\partial x} &= At^{-1/2} \exp\left(Bx^2 t^{-1}\right) \left(\frac{2Bx}{t} \right) \\
\frac{\partial^2 p}{\partial x^2} &= At^{-1/2} \exp\left(Bx^2 t^{-1}\right) \left(\frac{2B}{t} + \frac{4B^2x^2}{t^2} \right)
\end{aligned}$$

which on substituting back into the original PDE leads to

$$-\phi(c_r + c_w)At^{-1/2} \exp\left(Bx^2 t^{-1}\right) \left(\frac{1}{2t} + \frac{Bx^2}{t^2} \right)$$

$$= \frac{k}{\mu} At^{-1/2} \exp\left(Bx^2 t^{-1}\right) \left(\frac{2B}{t} + \frac{4B^2x^2}{t^2} \right)$$

After some further rearranging, this reduces to

$$-\phi(c_r + c_w) = 4B\frac{k}{\mu}$$

and

$$B = -\frac{\phi(c_r + c_w)\mu}{4k}$$

Hence the proposed function is a solution to the specified PDE.

Worked Solution 12.3 (see Problem 12.3)

a) To make the change in variables to derive a new equation it is necessary first to apply the dependent variable transform, $c = t^{-1/2}f(s)$.

Note that

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial t}(t^{-1/2}f) = \frac{1}{t^{1/2}}\left(\frac{\partial f}{\partial t} - \frac{f}{2t}\right)$$

and

$$\frac{\partial^2 c}{\partial x^2} = \frac{1}{t^{1/2}}\frac{\partial^2 f}{\partial x^2}$$

which when substituted back into the original PDE leads to

$$\frac{\partial f}{\partial t} - \frac{f}{2t} = D_A\frac{\partial^2 f}{\partial x^2} \tag{12.30}$$

To eliminate the original independent variables and apply the proposed similarity transform, $s = x(4D_At)^{-1/2}$, it is necessary to apply the chain-rule as follows:

Note that

$$\frac{\partial s}{\partial t} = -\frac{s}{2t}, \quad \frac{\partial s}{\partial x} = \frac{s}{x}, \quad \frac{\partial^2 s}{\partial x^2} = 0$$

and therefore:

$$
\begin{aligned}
\frac{\partial f}{\partial t} &= \frac{\partial s}{\partial t}\frac{\partial f}{\partial s} = -\frac{s}{2t}\frac{\partial f}{\partial s} \\
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial s}{\partial x}\frac{\partial f}{\partial s}\right) = \frac{\partial^2 s}{\partial x^2}\frac{\partial f}{\partial s} + \frac{\partial s}{\partial x}\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial s}\right) \\
&= \frac{\partial^2 s}{\partial x^2}\frac{\partial f}{\partial s} + \left(\frac{\partial s}{\partial x}\right)^2 \frac{\partial^2 f}{\partial s^2} = \frac{s^2}{x^2}\frac{\partial^2 f}{\partial s^2}
\end{aligned}
$$

which on substitution back into Eq. (12.30) leads to

$$-\frac{s}{2t}\frac{\partial f}{\partial s} - \frac{f}{2t} = D_A \frac{s^2}{x^2}\frac{\partial^2 f}{\partial s^2}$$

which reduces to

$$0 = \frac{\partial^2 f}{\partial s^2} + 2s\frac{\partial f}{\partial s} + 2f \tag{12.31}$$

b) Determine the integral

$$M = \int_{-\infty}^{\infty} c(x,t)dx$$

is a constant.

First apply the dependent variable transform

$$M = \int_{-\infty}^{\infty} t^{-1/2} f(s)dx$$

Now apply the similarity transform using the chain rule

$$M = \int_{-\infty}^{\infty} t^{-1/2} f(s)\frac{\partial x}{\partial s}ds$$

Recall that

$$\frac{\partial s}{\partial x} = \frac{s}{x}$$

therefore

$$\begin{aligned} M &= \int_{-\infty}^{\infty} t^{-1/2} f(s) \frac{x}{s} ds = \int_{-\infty}^{\infty} t^{-1/2} f(s) \frac{x\sqrt{4D_A t}}{x} ds \\ &= \sqrt{4D_A} \int_{-\infty}^{\infty} f(s) ds \end{aligned}$$

Hence M is indeed a constant with time.

c) Consider

$$f(s) = \exp(-s^2) \tag{12.32}$$

from which it can be seen that:

$$\frac{\partial f}{\partial s} = -2s\exp(-s^2)$$

$$\frac{\partial^2 f}{\partial s^2} = (4s^2 - 2)\exp(-s^2)$$

which when substituted into Eq. (12.31) yield

$$0 = (4s^2 - 2)\exp(-s^2) - 4s^2\exp(-s^2) + 2\exp(-s^2)$$

Hence Eq. (12.32) is a solution to Eq. (12.31).

13

Fourier's law, series and transform

13.1 Learning outcomes

You should be able to:

- Discuss the link between Fourier's law of heat conduction, the Fourier series and the Fourier transform.
- Transform a partial differential equation (PDE) to a set of ordinary differential equations (ODE) using separation of variables.
- Constrain the boundary condition of a PDE to an arbitrary spatial distribution of a state variable using an integral transform.
- Simplify trigonometric functions evaluated at integer values.
- Integrate products of piecewise periodic functions with trigonometric functions and other polynomials.

13.2 Fourier's law of heat conduction

Joseph Fourier was a 18th/19th century French mathematician. There are many mathematical subjects associated with Fourier. Three of note are Fourier's law, the Fourier series and the Fourier transform.

Fourier's law is the idea that heat conduction can be modeled using a diffusion equation of the form

$$\rho c_p \frac{\partial T}{\partial t} - \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) - \frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right) - \frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right) = 0$$

where ρ [ML^{-3}] is density, c_p [L^{2}T$^{-2}\Theta^{-1}$] is constant-pressure specific heat capacity, T [Θ] is temperature, t [T] is time and k [MLT$^{-3}\Theta^{-1}$] is thermal conductivity.

Now consider two dimensional steady state heat conduction in a homogenous medium:

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0 \tag{13.1}$$

where the following boundary conditions apply:

$$\begin{array}{lll} T = 0, & x = 0, & 0 \leq y \leq b \\ T = 0, & x = a, & 0 \leq y \leq b \\ T = 0, & 0 \leq x \leq a, & y = 0 \\ T = f(x), & 0 \leq x \leq a, & y = b \end{array}$$

and a [L] and b [L] are the breadth and length of the domain, respectively and $f(x)$ [Θ] is an arbitrary function of x.

In what follows, we will transform the PDE in Eq. (13.1) into two ODEs using separation of variables. The three zero temperature boundaries are easily enforced by determining values for the integration constants. The $f(x)$ boundary is then imposed using an innovative integration method developed by Fourier, which gave rise to the development of the Fourier series and Fourier transform.

Although the original motivation of the Fourier series was to solve PDEs, the series has become invaluable in the field of signal processing and inverse modeling.

Challenge 13.1 Consider the partial differential equation for two-dimensional steady-state heat conduction

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0$$

Substitute $T(x,y) = F(x)G(y)$ and separate the variables to derive two associated ordinary differential equations.

Making the specified substitution leads to

$$\frac{\partial^2 (FG)}{\partial x^2} + \frac{\partial^2 (FG)}{\partial y^2} = 0$$

Because G is constant in x it can come outside of the partial derivative with respect to x. Similarly, because F is constant in y it can come outside of the partial derivative with respect to y. It follows that

$$G\frac{\partial^2 F}{\partial x^2} + F\frac{\partial^2 G}{\partial y^2} = 0$$

To separate the variables, we need to get all the x factors on one side and the y factors on the other side:

$$\frac{1}{G}\frac{\partial^2 G}{\partial y^2} = -\frac{1}{F}\frac{\partial^2 F}{\partial x^2}$$

Note that the left-hand-side of the above equation is purely a function of x whereas the right-hand-side is purely a function of y. It follows that both sides are therefore actually constant in both x and y such that it can be said that

$$\frac{1}{G}\frac{\partial^2 G}{\partial y^2} = -\frac{1}{F}\frac{\partial^2 F}{\partial x^2} = \alpha$$

where α is a constant yet to be defined.

In this way, two ODEs are derived:

$$\frac{\partial^2 F}{\partial x^2} + \alpha F = 0 \tag{13.2}$$

$$\frac{\partial^2 G}{\partial y^2} - \alpha G = 0 \tag{13.3}$$

Challenge 13.2 Derive general solutions for F and G.

Eq. (13.2) represents a linear and homogenous second-order differential equation. The associated auxiliary equation takes the form

$$\lambda^2 + \alpha = 0$$

which has imaginary roots $\lambda = \pm i\sqrt{\alpha}$. It follows that the general solution for Eq. (13.2) is

$$F = A\cos(\sqrt{\alpha}x) + B\sin(\sqrt{\alpha}x)$$

where A and B are integration constants yet to be defined.

Eq. (13.3) also represents a linear and homogenous second-order differential equation. The associated auxiliary equation takes the form

$$\lambda^2 - \alpha = 0$$

which has the roots $\lambda = \pm\sqrt{\alpha}$. It follows that the general solution for Eq. (13.2) is

$$G = C\cosh(\sqrt{\alpha}y) + D\sinh(\sqrt{\alpha}y)$$

where C and D are another set of integration constants yet to be defined.

Challenge 13.3 Given the general solutions for F and G apply the following boundary conditions:

$$\begin{array}{lll} T=0, & x=0, & 0 \le y \le b \\ T=0, & x=a, & 0 \le y \le b \\ T=0, & 0 \le x \le a, & y=0 \end{array}$$

Recall that $T(x,y) = F(x)G(y)$ where

$$F = A\cos(\sqrt{\alpha}x) + B\sin(\sqrt{\alpha}x)$$

$$G = C\cosh(\sqrt{\alpha}y) + D\sinh(\sqrt{\alpha}y)$$

First consider the boundary at $x = 0$. We need to choose a value of A or B such that $F = 0$ when $x = 0$. Noting that $\sin 0 = 0$ it follows that we should choose A to be zero. Therefore

$$F = B\sin(\sqrt{\alpha}x)$$

Now we will consider the boundary at $y = 0$. We need to choose a value for C or D such that $G = 0$ when $y = 0$. Noting that $\sinh 0 = 0$ it follows that we should choose C to be zero. Therefore

$$G = D\sinh(\sqrt{\alpha}y)$$

Given that $T(x,y) = F(x)G(y)$ we can now say that

$$T(x,y) = E\sin(\sqrt{\alpha}x)\sinh(\sqrt{\alpha}y)$$

where $E = BD$.

Now we will consider the boundary at $x = a$. We need to choose a value for E or α such that $T = 0$ when $x = a$. It is clear that there is no non-zero value of E that can satisfy this condition. Therefore we need to choose a value of α such that $\sin(\sqrt{\alpha}a) = 0$. Noting that $\sin(n\pi) = 0$ where n is an integer, we should choose

$$\sqrt{\alpha} = \frac{n\pi}{a}$$

such that we have

$$T(x,y) = E \sin\left(\frac{n\pi x}{a}\right) \sinh\left(\frac{n\pi y}{a}\right) \tag{13.4}$$

13.3 Non-uniform boundary conditions

Unfortunately the boundary condition $T(x, y = b) = f(x)$ is less straightforward to apply. The problem is that it is not possible to choose a constant value of E such that $T = f(x)$ for $0 \leq x \leq a$. However, Fourier recognized that $n = 1, 2, 3, \ldots, \infty$. Therefore, Eq. (13.4) represents a set of solutions:

$$T_n(x,y) = E_n \sin\left(\frac{n\pi x}{a}\right) \sinh\left(\frac{n\pi y}{a}\right)$$

where E_n represents a set of integration constants yet to be defined.

It is therefore possible to write a solution as a summation of these alternative solutions, i.e.,

$$T(x,y) = \sum_{n=1}^{\infty} E_n \sin\left(\frac{n\pi x}{a}\right) \sinh\left(\frac{n\pi y}{a}\right)$$

Challenge 13.4 Consider the solution

$$T(x,y) = \sum_{n=1}^{\infty} E_n \sin\left(\frac{n\pi x}{a}\right) \sinh\left(\frac{n\pi y}{a}\right) \tag{13.5}$$

Apply the boundary condition $T(x,b) = f(x)$ and derive an expression for $f(x)$ in terms of E_n.

The boundary condition implies that $T = f(x)$ when $y = b$. It follows that

$$f(x) = \sum_{n=1}^{\infty} E_n \sin\left(\frac{n\pi x}{a}\right) \sinh\left(\frac{n\pi b}{a}\right) \tag{13.6}$$

Challenge 13.5 Multiply both sides of Eq. (13.6) by $\sin(m\pi x/a)$ (where m is another set of positive integers) and integrate both sides with respect to x from 0 to a.

$$\begin{aligned} &\int_0^a f(x) \sin\left(\frac{m\pi x}{a}\right) dx \\ &= \sum_{n=1}^{\infty} E_n \sinh\left(\frac{n\pi b}{a}\right) \int_0^a \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi x}{a}\right) dx \end{aligned} \tag{13.7}$$

To deal with the integral on the right-hand-side, first invoke the relationship

$$\cos(A+B) = \cos A \cos B - \sin A \sin B$$

from which it can be said that

$$\cos(A-B) = \cos A \cos B + \sin A \sin B$$

and consequently

$$\sin A \sin B = \frac{\cos(A-B) - \cos(A+B)}{2}$$

It follows that

$$
\begin{aligned}
&\int_0^a \sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi x}{a}\right)dx \\
&= \int_0^a \frac{1}{2}\left[\cos\left(\frac{(n-m)\pi x}{a}\right) - \cos\left(\frac{(n+m)\pi x}{a}\right)\right]dx \\
&= \frac{a}{2\pi}\left[\frac{\sin\left(\frac{(n-m)\pi x}{a}\right)}{n-m} - \frac{\sin\left(\frac{(n+m)\pi x}{a}\right)}{n+m}\right]_0^a \\
&= \frac{a}{2\pi}\left[\frac{\sin\left((n-m)\pi\right)}{n-m} - \frac{\sin\left((n+m)\pi\right)}{n+m}\right]
\end{aligned}
\tag{13.8}
$$

Substituting Eq. (13.8) into Eq. (13.7) then yields

$$
\begin{aligned}
&\int_0^a f(x)\sin\left(\frac{m\pi x}{a}\right)dx \\
&= \sum_{n=1}^{\infty} E_n \sinh\left(\frac{n\pi b}{a}\right)\frac{a}{2\pi}\left[\frac{\sin\left((n-m)\pi\right)}{n-m} - \frac{\sin\left((n+m)\pi\right)}{n+m}\right]
\end{aligned}
$$

But note that because n and m are all positive integers, all the terms on the right-hand-side are zero except for the ones where $n = m$. Recall that

$$\lim_{\varepsilon \to 0} \frac{\sin(\varepsilon x)}{\varepsilon} = x$$

It follows that

$$\int_0^a f(x)\sin\left(\frac{m\pi x}{a}\right)dx = E_m \sinh\left(\frac{m\pi b}{a}\right)\frac{a}{2} \tag{13.9}$$

Challenge 13.6 Use Eq. (13.9) to determine E_n and write a solution for $T(x,y)$ that satisfies the boundary condition, $T(x,b) = f(x)$.

Solving Eq. (13.9) for E_m leads to

$$E_m = \frac{2}{a\sinh(m\pi b/a)} \int_0^a f(x)\sin(m\pi x/a)dx$$

If we set $m = n$ we will therefore have

$$E_n = \frac{2}{a\sinh(n\pi b/a)} \int_0^a f(x)\sin(n\pi x/a)dx \tag{13.10}$$

which on substitution back into Eq. (13.5) leads to

$$T(x,y) = \sum_{n=1}^{\infty} \frac{2\sin(n\pi x/a)\sinh(n\pi y/a)}{a\sinh(n\pi b/a)} \int_0^a f(x)\sin(n\pi x/a)dx \tag{13.11}$$

Challenge 13.7 Use Eq. (13.11) to write a solution for $T(x,y)$ when

$$f(x) = \begin{cases} T_0\left(\frac{x}{c}\right), & 0 \le x \le c \\ T_0\left(\frac{a-x}{a-c}\right), & c < x \le a \end{cases} \tag{13.12}$$

Let

$$F_n = \int_0^a f(x)\sin(n\pi x/a)dx \tag{13.13}$$

such that

$$T(x,y) = \sum_{n=1}^{\infty} \frac{2F_n\sin(n\pi x/a)\sinh(n\pi y/a)}{a\sinh(n\pi b/a)}$$

Substituting Eq. (13.12) into Eq. (13.13) leads to

$$F_n = \int_0^c T_0\left(\frac{x}{c}\right)\sin(n\pi x/a)dx + \int_c^a T_0\left(\frac{a-x}{a-c}\right)\sin(n\pi x/a)dx \tag{13.14}$$

It can be shown that:

$$\int_{x_1}^{x_2} \sin(n\pi x/a)dx = \left[-\frac{\cos(n\pi x/a)}{n\pi/a}\right]_{x_1}^{x_2}$$

and

$$\int_{x_1}^{x_2} x\sin(n\pi x/a)dx = \left[\frac{\sin(n\pi x/a)}{(n\pi/a)^2} - \frac{x\cos(n\pi x/a)}{n\pi/a}\right]_{x_1}^{x_2}$$

and therefore

$$\int_{x_1}^{x_2} (a-x)\sin(n\pi x/a)dx = \left[\frac{(x-a)\cos(n\pi x/a)}{n\pi/a} - \frac{\sin(n\pi x/a)}{(n\pi/a)^2}\right]_{x_1}^{x_2}$$

It follows that

$$\begin{aligned}\int_0^c \left(\frac{x}{c}\right)\sin(n\pi x/a)dx &= \frac{1}{c}\left[\frac{\sin(n\pi x/a)}{(n\pi/a)^2} - \frac{x\cos(n\pi x/a)}{n\pi/a}\right]_0^c \\ &= \frac{\sin(n\pi c/a)}{c(n\pi/a)^2} - \frac{\cos(n\pi c/a)}{n\pi/a}\end{aligned}$$

and

$$\begin{aligned}\int_c^a \left(\frac{a-x}{a-c}\right)\sin(n\pi x/a)dx &= \left[\frac{(x-a)\cos(n\pi x/a)}{(a-c)n\pi/a} - \frac{\sin(n\pi x/a)}{(a-c)(n\pi/a)^2}\right]_c^a \\ &= \frac{\sin(n\pi c/a)}{(a-c)(n\pi/a)^2} + \frac{\cos(n\pi c/a)}{n\pi/a}\end{aligned}$$

Substituting the above two expressions into Eq. (13.14) then leads to

$$\begin{aligned} F_n &= T_0\left[\frac{\sin(n\pi c/a)}{c(n\pi/a)^2} - \frac{\cos(n\pi c/a)}{n\pi/a} + \frac{\sin(n\pi c/a)}{(a-c)(n\pi/a)^2} + \frac{\cos(n\pi c/a)}{n\pi/a}\right] \\ &= \frac{T_0 a^3 \sin(n\pi c/a)}{c(a-c)(n\pi)^2} \end{aligned}$$

13.4 The Fourier sine series and sine transform

But consider again Eqs. (13.6) and (13.10):

$$f(x) = \sum_{n=1}^{\infty} E_n \sin\left(\frac{n\pi x}{a}\right) \sinh\left(\frac{n\pi b}{a}\right)$$

$$E_n = \frac{2}{a \sinh(n\pi b/a)} \int_0^a f(x) \sin(n\pi x/a) dx$$

Letting $b_n = E_n \sinh(n\pi b/a)$ suggests that a continuous function, $f(x)$, can be approximated by a summation of sine waves multiplied by a set of weighting coefficients, b_n, i.e.,

$$f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{a}\right) \tag{13.15}$$

where the weighting coefficients are determined from

$$b_n = \frac{2}{a} \int_0^a f(x) \sin(n\pi x/a) dx \tag{13.16}$$

The expansion in Eq. (13.15) is known as the Fourier sine series and Eq. (13.16) is known as the Fourier sine transform.

13.5 Problem sheet

Problem 13.1 (see Worked Solution 13.1)

Simplify the following expressions as much as possible assuming that n is an integer. Take care to consider if more simplifications are possible if n is odd or even.

a) $\sin(n\pi)$ b) $\cos(n\pi)$ c) $\cos^2(n\pi)$

d) $1+\cos(n\pi)$ e) $1-\cos(n\pi)$

Problem 13.2 (see Worked Solution 13.2)

Integrate the following functions with respect to x.

a) $\sin(nx)$ b) $\cos(nx)$ c) $x\sin(nx)$

d) $x\cos(nx)$ e) $x^2\sin(nx)$ f) $x^2\cos(nx)$

Problem 13.3 (see Worked Solution 13.3)

Sketch the following piecewise period functions for x between -2π and 4π assuming that each function has a period of 2π.

a) $f(x) = x - \pi, \quad 0 \leq x \leq 2\pi$

b) $f(x) = x(2\pi - x), \quad 0 \leq x \leq 2\pi$

c) $f(x) = \begin{cases} x, & 0 \leq x < \pi \\ 2\pi - x, & \pi \leq x \leq 2\pi \end{cases}$

d) $f(x) = \begin{cases} -1, & 0 \leq x < \pi \\ 1, & \pi \leq x \leq 2\pi \end{cases}$

Problem 13.4 (see Worked Solution 13.4)

Consider the following notation

$$\langle \cdot \rangle = \frac{1}{\pi}\int_0^{2\pi} \cdot dx$$

such that, for example,

$$\langle f(x)\cos x\rangle = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos x dx$$

For each of the functions in Problem 13.3, determine the following definite integrals, where n (if present) is always an integer value greater than zero.

i) $\langle f(x)\rangle$ ii) $\langle f(x)x\rangle$

iii) $\langle f(x)\sin(nx)\rangle$ iv) $\langle f(x)\cos(nx)\rangle$

Problem 13.5 (see Worked Solution 13.5)

Consider again the $\langle\cdot\rangle$ notation defined in Problem 13.4. Assuming that m and n are integers, determine the following definite integrals with special consideration of the cases $m = n$, $m \neq n$, $m = 0$ and $n = 0$.

a) $\langle \sin(nx)\sin(mx)\rangle$ b) $\langle \sin(nx)\cos(mx)\rangle$

c) $\langle \cos(nx)\cos(mx)\rangle$

13.6 Worked solutions

Worked Solution 13.1 (see Problem 13.1)

a) $\sin(n\pi) = 0$

b) For even values of n, $\cos(n\pi) = 1$, but for odd values of n, $\cos(n\pi) = -1$.

This can be compactly stated by saying $\cos(n\pi) = (-1)^n$

c) From the above discussion it is then clear that $\cos^2(n\pi) = 1$

d) From part b) above $1 + \cos(n\pi) = 1 + (-1)^n$

e) From part b) above $1 - \cos(n\pi) = 1 - (-1)^n = 1 + (-1)^{n+1}$

Worked Solution 13.2 (see Problem 13.2)

a) $$\int \sin(nx)dx = -\frac{\cos(nx)}{n} + C$$

b) $$\int \cos(nx)dx = \frac{\sin(nx)}{n} + C$$

c) $$\begin{aligned}\int x\sin(nx)dx &= -\frac{x\cos(nx)}{n} + \int \frac{\cos(nx)}{n}dx \\ &= -\frac{x\cos(nx)}{n} + \frac{\sin(nx)}{n^2} + C\end{aligned}$$

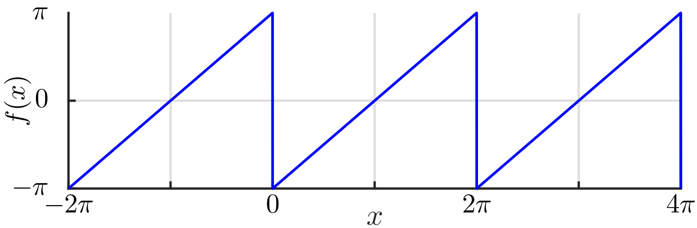
d) $$\begin{aligned}\int x\cos(nx)dx &= \frac{x\sin(nx)}{n} - \int \frac{\sin(nx)}{n}dx \\ &= \frac{x\sin(nx)}{n} + \frac{\cos(nx)}{n^2} + C\end{aligned}$$

e)
$$\int x^2 \sin(nx)dx = -\frac{x^2\cos(nx)}{n} + \frac{2}{n}\int x\cos(nx)dx$$
$$= -\frac{x^2\cos(nx)}{n} + \frac{2x\sin(nx)}{n^2} + \frac{2\cos(nx)}{n^3} + C$$

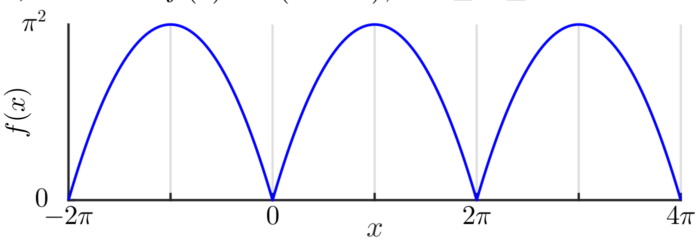
f)
$$\int x^2 \cos(nx)dx = \frac{x^2\sin(nx)}{n} - \frac{2}{n}\int x\sin(nx)dx$$
$$= \frac{x^2\sin(nx)}{n} + \frac{2x\cos(nx)}{n^2} - \frac{2\sin(nx)}{n^3} + C$$

Worked Solution 13.3 (see Problem 13.3)

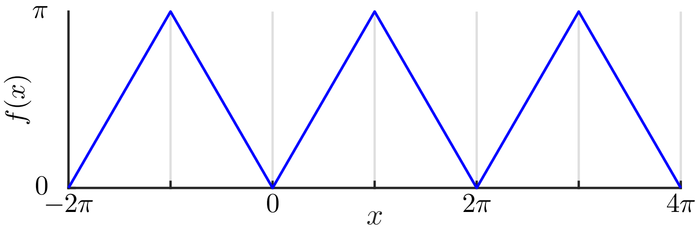
a) Sketch of $f(x) = x - \pi, \quad 0 \le x \le 2\pi$



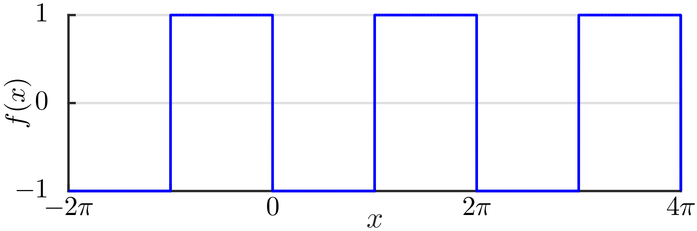
b) Sketch of $f(x) = x(2\pi - x), \quad 0 \le x \le 2\pi$



c) Sketch of $f(x) = \begin{cases} x, & 0 \le x < \pi \\ 2\pi - x, & \pi \le x \le 2\pi \end{cases}$



d) Sketch of $f(x) = \begin{cases} -1, & 0 \leq x < \pi \\ 1, & \pi \leq x \leq 2\pi \end{cases}$



Worked Solution 13.4 (see Problem 13.4)

a) $f(x) = x - \pi$

(i) $\dfrac{1}{\pi}\displaystyle\int_0^{2\pi}(x-\pi)dx$

$$= \left[\frac{x^2}{2\pi} - x\right]_0^{2\pi} = 2\pi - 2\pi = 0$$

(ii) $\dfrac{1}{\pi}\displaystyle\int_0^{2\pi}(x-\pi)xdx$

$$= \left[\frac{x^3}{3\pi} - \frac{x^2}{2}\right]_0^{2\pi} = \frac{8\pi^2}{3} - \frac{4\pi^2}{2} = \frac{2\pi^2}{3}$$

(iii) $\dfrac{1}{\pi}\displaystyle\int_0^{2\pi}(x-\pi)\sin(nx)dx$

$$= \frac{1}{\pi}\left[-\frac{x\cos(nx)}{n} + \frac{\sin(nx)}{n^2} + \frac{\pi\cos(nx)}{n}\right]_0^{2\pi}$$

$$= -\frac{\cos(2n\pi)}{n} + \frac{\cos(2n\pi)}{n} - \frac{1}{n} = -\frac{2}{n}$$

(iv) $\dfrac{1}{\pi}\displaystyle\int_0^{2\pi}(x-\pi)\cos(nx)dx$

$$= \frac{1}{\pi}\left[\frac{x\sin(nx)}{n} + \frac{\cos(nx)}{n^2} - \frac{\pi\sin(nx)}{n}\right]_0^{2\pi}$$

$$= \frac{\cos(2n\pi)}{n^2\pi} - \frac{1}{n^2\pi} = 0$$

b) $f(x) = x(2\pi - x)$

(i) $$\frac{1}{\pi}\int_0^{2\pi} x(2\pi - x)dx = \left[x^2 - \frac{x^3}{3\pi}\right]_0^{2\pi} = 4\pi^2 - \frac{8}{3}\pi^2 = \frac{4\pi^2}{3}$$

(ii) $$\frac{1}{\pi}\int_0^{2\pi} x(2\pi - x)xdx = \left[\frac{2x^3}{3\pi} - \frac{x^4}{4\pi}\right]_0^{2\pi} = \frac{16\pi^3}{3} - \frac{16\pi^3}{4} = \frac{4\pi^3}{3}$$

(iii) $$\frac{1}{\pi}\int_0^{2\pi} x(2\pi - x)\sin(nx)dx = \frac{1}{\pi}\left[2\pi\left(\frac{\sin(nx)}{n^2} - \frac{x\cos(nx)}{n}\right)\right.$$
$$\left. + \left(\frac{x^2\cos(nx)}{n} - \frac{2x\sin(nx)}{n^2} - \frac{2\cos(nx)}{n^3}\right)\right]_0^{2\pi}$$

$$= \left[\left(2 - \frac{2x}{\pi}\right)\frac{\sin(nx)}{n^2} - \left(2x - \frac{x^2}{\pi} + \frac{2}{n^2\pi}\right)\frac{\cos(nx)}{n}\right]_0^{2\pi}$$

$$= -\left(4\pi - \frac{4\pi^2}{\pi} + \frac{2}{n^2}\pi\right)\frac{1}{n} + \left(\frac{2}{n^2}\pi\right)\frac{1}{n} = 0$$

(iv) $$\frac{1}{\pi}\int_0^{2\pi} x(2\pi - x)\cos(nx)dx = \frac{1}{\pi}\left[2\pi\left(\frac{x\sin(nx)}{n} + \frac{\cos(nx)}{n^2}\right)\right.$$
$$\left. - \left(\frac{x^2\sin(nx)}{n} + \frac{2x\cos(nx)}{n^2} - \frac{2\sin(nx)}{n^3}\right)\right]_0^{2\pi}$$

$$= \left[\left(2x - \frac{x^2}{\pi} + \frac{2}{n^2\pi}\right)\frac{\sin(nx)}{n} + \left(2 - \frac{2x}{\pi}\right)\frac{\cos(nx)}{n^2}\right]_0^{2\pi}$$

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

c) $$f(x)=\begin{cases} x, & 0\le x<\pi \\ 2\pi-x, & \pi\le x\le 2\pi \end{cases}$$

(i) $$\langle f(x)\rangle=\frac{1}{\pi}\int_0^{\pi} x\,dx+\frac{1}{\pi}\int_{\pi}^{2\pi}(2\pi-x)\,dx$$

$$=\left[\frac{x^2}{2\pi}\right]_0^{\pi}+\left[2x-\frac{x^2}{2\pi}\right]_{\pi}^{2\pi}=\frac{\pi}{2}+4\pi-\frac{4\pi}{2}-2\pi+\frac{\pi}{2}=\pi$$

(ii) $$\langle f(x)x\rangle=\frac{1}{\pi}\int_0^{\pi} x^2\,dx+\frac{1}{\pi}\int_{\pi}^{2\pi}(2\pi x-x^2)\,dx$$

$$=\left[\frac{x^3}{3\pi}\right]_0^{\pi}+\left[x^2-\frac{x^3}{3\pi}\right]_{\pi}^{2\pi}=\frac{\pi^2}{3}+4\pi^2-\frac{8\pi^2}{3}-\pi^2+\frac{\pi^2}{3}=\pi^2$$

(iii) $$\langle f(x)\sin(nx)\rangle=\frac{1}{\pi}\int_0^{\pi} x\sin(nx)\,dx+\frac{1}{\pi}\int_{\pi}^{2\pi}(2\pi-x)\sin(nx)\,dx$$

$$=\left[\frac{\sin(nx)}{n^2\pi}-\frac{x\cos(nx)}{n\pi}\right]_0^{\pi}+\left[-\frac{2\cos(nx)}{n}-\frac{\sin(nx)}{n^2\pi}+\frac{x\cos(nx)}{n\pi}\right]_{\pi}^{2\pi}$$

$$=-\frac{(-1)^n}{n}-\frac{2(-1)^{2n}}{n}+\frac{2(-1)^{2n}}{n}+\frac{2(-1)^n}{n}-\frac{(-1)^n}{n}=0$$

(iv) $$\langle f(x)\cos(nx)\rangle=\frac{1}{\pi}\int_0^{\pi} x\cos(nx)\,dx+\frac{1}{\pi}\int_{\pi}^{2\pi}(2\pi-x)\cos(nx)\,dx$$

$$=\left[\frac{\cos(nx)}{n^2\pi}+\frac{x\sin(nx)}{n\pi}\right]_0^{\pi}+\left[\frac{2\sin(nx)}{n}-\frac{\cos(nx)}{n^2\pi}-\frac{x\sin(nx)}{n\pi}\right]_{\pi}^{2\pi}$$

$$=\frac{(-1)^n}{n^2\pi}-\frac{1}{n^2\pi}-\frac{(-1)^{2n}}{n^2\pi}+\frac{(-1)^n}{n^2\pi}=\frac{2[(-1)^n-1]}{n^2\pi}$$

d) $$f(x) = \begin{cases} -1, & 0 \leq x < \pi \\ 1, & \pi \leq x \leq 2\pi \end{cases}$$

(i) $$\langle f(x) \rangle = -\frac{1}{\pi}\int_0^{\pi} 1dx + \frac{1}{\pi}\int_{\pi}^{2\pi} 1dx$$

$$= -\left[\frac{x}{\pi}\right]_0^{\pi} + \left[\frac{x}{\pi}\right]_{\pi}^{2\pi} = -1+2-1 = 0$$

(ii) $$\langle f(x)x \rangle = -\frac{1}{\pi}\int_0^{\pi} xdx + \frac{1}{\pi}\int_{\pi}^{2\pi} xdx$$

$$= -\left[\frac{x^2}{2\pi}\right]_0^{\pi} + \left[\frac{x^2}{2\pi}\right]_{\pi}^{2\pi} = -\frac{\pi}{2} + \frac{4\pi}{2} - \frac{\pi}{2} = \pi$$

(iii) $$\langle f(x)\sin(nx) \rangle = -\frac{1}{\pi}\int_0^{\pi} \sin(nx)dx + \frac{1}{\pi}\int_{\pi}^{2\pi} \sin(nx)dx$$

$$= \left[\frac{\cos(nx)}{n\pi}\right]_0^{\pi} + \left[-\frac{\cos(nx)}{n\pi}\right]_{\pi}^{2\pi}$$

$$= \frac{(-1)^n - 1}{n\pi} - \frac{[(-1)^{2n} - (-1)^n)}{n\pi} = \frac{2[(-1)^n - 1]}{n\pi}$$

(iv) $$\langle f(x)\cos(nx) \rangle = -\frac{1}{\pi}\int_0^{\pi} \cos(nx)dx + \frac{1}{\pi}\int_{\pi}^{2\pi} \cos(nx)dx$$

$$= -\left[\frac{\sin(nx)}{n\pi}\right]_0^{\pi} + \left[\frac{\sin(nx)}{n\pi}\right]_{\pi}^{2\pi} = 0$$

Worked Solution 13.5 (see Problem 13.5)

a) $\langle \sin(nx)\sin(mx)\rangle = \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} \sin(nx)\sin(mx)dx$

Recall that $\cos(A+B) = \cos A\cos B - \sin A\sin B$

$\therefore \cos(A-B) = \cos A\cos B + \sin A\sin B$

$\therefore \cos(A-B) - \cos(A+B) = 2\sin A\sin B$

from which it follows that

$$\langle \sin(nx)\sin(nx)\rangle = \frac{1}{2\pi}\int_0^{2\pi}[\cos[(n-m)x] - \cos[(n+m)x]]dx$$

$$= \frac{1}{2\pi}\left[\frac{\sin[(n-m)x]}{n-m} - \frac{\sin[(n+m)x]}{n+m}\right]_0^{2\pi}$$

At a first glance the above expression simply equals zero.

But now recall from the power series expansion for $\sin x$ that

$\sin x = x + O(x^2)$

from which it follows that:

$$\lim_{n\to m}\frac{\sin[(n-m)x]}{n-m} = \frac{(n-m)x}{n-m} = x$$

$$\lim_{n\to -m}\frac{\sin[(n+m)x]}{n+m} = \frac{(n+m)x}{n+m} = x$$

$$\therefore \langle \sin(nx)\sin(mx)\rangle = \begin{cases} 0, & m=0 \text{ or } n=0 \text{ or } m \neq |n| \\ 1, & m=n \text{ and } n\neq 0 \text{ and } m \neq 0 \\ -1, & m=-n \text{ and } n \neq 0 \text{ and } m \neq 0 \end{cases}$$

b) $\langle \sin(nx)\cos(mx)\rangle = \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} \sin(nx)\cos(mx)dx$

Recall that $\sin(A+B) = \sin A\cos B + \cos A \sin B$

$\therefore \sin(A-B) = \sin A \cos B - \cos A \sin B$

$\therefore \sin(A+B) + \sin(A-B) = 2\sin A\cos B$

from which it follows that

$$\langle \sin(nx)\cos(nx)\rangle = \frac{1}{2\pi}\int_0^{2\pi}[\sin[(n+m)x] + \sin[(n-m)x]]dx$$

$$= -\frac{1}{2\pi}\left[\frac{\cos[(n+m)x]}{(n+m)} + \frac{\cos[(n-m)x]}{n-m}\right]_0^{2\pi} = 0$$

c) $\langle \cos(nx)\cos(mx)\rangle = \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} \cos(nx)\cos(mx)dx$

Recall that $\cos(A+B) = \cos A\cos B - \sin A\sin B$

$\therefore \cos(A-B) = \cos A\cos B + \sin A\sin B$

$\therefore \cos(A-B) + \cos(A+B) = 2\cos A\cos B$

from which it follows that

$$\langle \cos(nx)\cos(nx)\rangle = \frac{1}{2\pi}\int_0^{2\pi} [\cos[(n-m)x] + \cos[(n+m)x]]dx$$

$$= \frac{1}{2\pi}\left[\frac{\sin[(n-m)x]}{n-m} + \frac{\sin[(n+m)x]}{n+m}\right]_0^{2\pi}$$

Recalling that

$$\lim_{n\to m} \frac{\sin[(n-m)x]}{n-m} = \frac{(n-m)x}{n-m} = x$$

$$\lim_{n\to -m} \frac{\sin[(n+m)x]}{n+m} = \frac{(n+m)x}{n+m} = x$$

It follows that

$$\langle \cos(nx)\cos(mx)\rangle = \begin{cases} 0, & m \neq |n| \\ 1, & m = |n| \text{ and } n \neq 0 \text{ and } m \neq 0 \\ 2, & m = |n| \text{ and } n = 0 \text{ and } m = 0 \end{cases}$$

14

Fourier series

14.1 Learning outcomes

You should be able to:

- Derive integral expressions for the Fourier coefficients.
- Take advantage of Kronecker delta notation where appropriate.
- Derive Parseval's theorem in Fourier analysis.
- Determine the Fourier series for simple functions.
- Apply Fourier series and Parseval's theorem to determine the sum of given infinite series.

14.2 A general Fourier series

In the previous session, it was shown that some functions can be approximated using series of sine waves multiplied by a set of weighted coefficients. An integral equation was then provided to determine these coefficients. Such a series is known as a sine series. The more general Fourier series expansion of a function takes the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)$$

Introducing the Kronecker delta

As part of the weekly exercise from the last session, some particularly useful expressions were derived, which will help to determine integral expressions for a_0, a_n and b_n, notably:

$$\begin{aligned}
\langle \sin(nx)\sin(mx)\rangle &= \begin{cases} 0, & m=0 \text{ or } n=0 \text{ or } m\neq |n| \\ 1, & m=n \text{ and } n\neq 0 \text{ and } m\neq 0 \\ -1, & m=-n \text{ and } n\neq 0 \text{ and } m\neq 0 \end{cases} \\
\langle \sin(nx)\cos(mx)\rangle &= 0 \\
\langle \cos(nx)\cos(mx)\rangle &= \begin{cases} 0, & m\neq |n| \\ 1, & m=|n| \text{ and } n\neq 0 \text{ and } m\neq 0 \\ 2, & m=|n| \text{ and } n=0 \text{ and } m=0 \end{cases}
\end{aligned}$$

where

$$\langle\cdot\rangle = \frac{1}{\pi}\int_0^{2\pi} \cdot dx$$

Now consider a new notation, the Kronecker delta, $\delta_{i,j}$, defined by

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

from which it can be seen that providing $n > 0$ and $m > 0$

$$\langle \cos(nx)\cos(mx)\rangle = \langle \sin(nx)\sin(mx)\rangle = \delta_{n,m}$$

Challenge 14.1 Given that

$$\langle \cos(nx)\cos(mx)\rangle = \langle \sin(nx)\sin(mx)\rangle = \delta_{n,m} \tag{14.1}$$

and

$$\langle \sin(nx)\cos(mx)\rangle = 0 \tag{14.2}$$

where

$$\langle\cdot\rangle = \frac{1}{\pi}\int_0^{2\pi} \cdot dx$$

determine an expression for the b_n coefficients in the Fourier series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx) \tag{14.3}$$

Considering our experience from the last session, to determine the b_n coefficients, let us consider the following integral expression in relation to Eq. (14.3):

$$\langle f(x)\sin(mx)\rangle = \left\langle \left(\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)\right)\sin(mx)\right\rangle$$

which can be broken down to say

$$\begin{aligned}\langle f(x)\sin(mx)\rangle = \quad & \frac{a_0}{2}\langle \sin(mx)\rangle \\ & + \sum_{n=1}^{\infty} a_n \langle \cos(nx)\sin(mx)\rangle \\ & + \sum_{n=1}^{\infty} b_n \langle \sin(nx)\sin(mx)\rangle\end{aligned}$$

because

$$\int \sum_{n=1}^{\infty} f_n(x)dx = \sum_{n=1}^{\infty} \int f_n(x)dx \tag{14.4}$$

Noting that

$$\langle \sin(mx) \rangle = \frac{1}{\pi} \int_0^{2\pi} \sin(mx)dx = \frac{1}{\pi} \left[-\frac{\cos(mx)}{m} \right]_0^{2\pi} = 0 \tag{14.5}$$

and recalling $\langle \cos(nx) \sin(mx) \rangle = 0$ along with Eq. (14.1) (note that the summations above are both from $n = 1$), it follows that

$$\langle f(x) \sin(mx) \rangle = \sum_{n=1}^{\infty} b_n \delta_{n,m} = b_m \tag{14.6}$$

and therefore

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx)dx$$

Challenge 14.2 Following on from the previous challenge, determine an expression for the a_n coefficients.

To obtain the a_n coefficients let us consider another integral expression in relation to Eq. (14.3):

$$\langle f(x) \cos(mx) \rangle = \left\langle \left(\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx) \right) \cos(mx) \right\rangle$$

which can be broken down to say

$$\begin{aligned} \langle f(x) \cos(mx) \rangle = \ & \frac{a_0}{2} \langle \cos(mx) \rangle \\ & + \sum_{n=1}^{\infty} a_n \langle \cos(nx) \cos(mx) \rangle \\ & + \sum_{n=1}^{\infty} b_n \langle \sin(nx) \cos(mx) \rangle \end{aligned}$$

Noting that

$$\langle \cos(mx) \rangle = \frac{1}{\pi}\int_0^{2\pi} \cos(mx)dx = \frac{1}{\pi}\left[\frac{\sin(mx)}{m}\right]_0^{2\pi} = 0 \quad (14.7)$$

and again recalling $\langle \cos(nx)\sin(mx) \rangle = 0$ along with Eq. (14.1), it follows that

$$\langle f(x)\cos(mx) \rangle = \sum_{n=1}^{\infty} a_n \delta_{n,m} = a_m \qquad (14.8)$$

and therefore

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)dx$$

Challenge 14.3 Determine an associated expression for the a_0 coefficient.

To obtain the a_0 coefficients let us consider another integral expression in relation to Eq. (14.3):

$$\langle f(x) \rangle = \left\langle \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx) \right\rangle$$

which can be broken down to say

$$\langle f(x) \rangle = \left\langle \frac{a_0}{2} \right\rangle + \sum_{n=1}^{\infty} a_n \langle \cos(nx) \rangle + \sum_{n=1}^{\infty} b_n \langle \sin(nx) \rangle$$

Recalling Eqs. (14.5) and (14.7) it is then seen that

$$\langle f(x) \rangle = \frac{a_0(2\pi)}{2\pi} = a_0 \qquad (14.9)$$

and therefore

$$a_0 = \frac{1}{\pi}\int_0^{2\pi} f(x)dx$$

14.3 Parseval's theorem

Another important result to determine is for

$$\langle f^2 \rangle = \frac{1}{\pi}\int_0^{2\pi} f^2 dx \tag{14.10}$$

where $f(x)$ can approximated by the Fourier series, as previously presented in Eq. (14.3).

Note that partial substitution of Eq. (14.3) into Eq. (14.10) leads to

$$\langle f^2 \rangle = \left\langle f(x)\left(\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)\right)\right\rangle$$

which can be broken down to say

$$\langle f^2 \rangle = \frac{a_0}{2}\langle f(x)\rangle + \sum_{n=1}^{\infty} a_n \langle f(x)\cos(nx)\rangle + \sum_{n=1}^{\infty} b_n \langle f(x)\sin(nx)\rangle$$

From Eq. (14.9) it can be seen that

$$\frac{a_0}{2}\langle f(x)\rangle = \frac{a_0^2}{2}$$

From Eq. (14.8) it can be seen that

$$\sum_{n=1}^{\infty} a_n \langle f(x)\cos(nx)\rangle = \sum_{n=1}^{\infty} a_n^2$$

From Eq. (14.6) it can be seen that

$$\sum_{n=1}^{\infty} b_n \langle f(x)\sin(nx)\rangle = \sum_{n=1}^{\infty} b_n^2$$

All of these results lead to the so-called Parseval's theorem, which states that

$$\frac{1}{\pi}\int_0^{2\pi} f^2 dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + \sum_{n=1}^{\infty} b_n^2$$

Antoine Pareseval was another 18th/19th century French mathematician, best known for this theorem in so-called Fourier analysis.

14.4 Summary of key results

Any function, $f(x)$, defined on the interval $[0, 2\pi]$ can be written as a Fourier series of the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx) \tag{14.11}$$

where

$$a_0 = \frac{1}{\pi}\int_0^{2\pi} f(x)dx \tag{14.12}$$

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)dx \tag{14.13}$$

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)dx \tag{14.14}$$

and, according to Parseval's theorem,

$$\frac{1}{\pi}\int_0^{2\pi} f^2 dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + \sum_{n=1}^{\infty} b_n^2 \tag{14.15}$$

14.5 Problem sheet

Problem 14.1 (see Worked Solution 14.1)

Consider the function $f(x) = x - \pi, \quad 0 \le x \le 2\pi$

a) Determine the Fourier series.

b) Use Parseval's theorem to show that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$

Problem 14.2 (see Worked Solution 14.2)

Consider the function $f(x) = x(2\pi - x), \quad 0 \le x \le 2\pi$

a) Determine the Fourier series.

b) By considering $f(\pi)$ show that $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} = -\frac{\pi^2}{12}$

c) Use Parseval's theorem to show that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$

Problem 14.3 (see Worked Solution 14.3)

Consider the function $f(x) = \begin{cases} x, & 0 \le x < \pi \\ 2\pi - x, & \pi \le x \le 2\pi \end{cases}$

a) Determine the Fourier series.

b) By considering $f(\pi)$ show that $\displaystyle\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}$

c) Use Parseval's theorem to show that $\displaystyle\sum_{n=0}^{\infty} \frac{1}{(2n+1)^4} = \frac{\pi^4}{96}$

Problem 14.4 (see Worked Solution 14.4)

Consider the function $f(x) = \begin{cases} -1, & 0 \le x < \pi \\ 1, & \pi \le x \le 2\pi \end{cases}$

a) Determine the Fourier series.

b) By considering $f(\pi/2)$ show that $\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}$

c) Use Parseval's theorem to show that $\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}$

14.6 Worked solutions

Worked Solution 14.1 (see Problem 14.1)

Consider $f(x) = x - \pi, \quad 0 \leq x \leq 2\pi$

a) Considering Eqs. (14.12) to (14.14) in conjunction with the results from Problem 13.4a

$$\begin{aligned} a_0 &= \langle (x-\pi) \rangle &&= 0 \\ a_n &= \langle (x-\pi)\cos(nx) \rangle &&= 0 \\ b_n &= \langle (x-\pi)\sin(nx) \rangle &&= -2/n \end{aligned}$$

and from Eq. (14.11)

$$f(x) = -2 \sum_{n=1}^{\infty} \frac{\sin(nx)}{n}$$

b) From Parseval's theorem, Eq. (14.15)

$$\langle (x-\pi)^2 \rangle = \sum_{n=1}^{\infty} b_n^2 = 4 \sum_{n=1}^{\infty} \frac{1}{n^2}$$

and

$$\begin{aligned}\langle (x-\pi)^2 \rangle &= \frac{1}{\pi}\int_0^{2\pi}(x-\pi)^2 dx = \frac{1}{\pi}\left[\frac{(x-\pi)^3}{3}\right]_0^{2\pi} \\ &= \frac{1}{\pi}\left[\frac{(2\pi-\pi)^3}{3}+\frac{\pi^3}{3}\right] = \frac{2\pi^2}{3}\end{aligned}$$

Equating the two expressions above then leads to

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}$$

Worked Solution 14.2 (see Problem 14.2)

Consider $f(x) = x(2\pi - x), \quad 0 \leq x \leq 2\pi$

a) Considering Eqs. (14.12) to (14.14) in conjunction with the results from Problem 13.4b

$$\begin{aligned} a_0 &= \langle x(2\pi - x)\rangle &&= 4\pi^2/3 \\ a_n &= \langle x(2\pi - x)\cos(nx)\rangle &&= -4/n^2 \\ b_n &= \langle x(2\pi - x)\sin(nx)\rangle &&= 0 \end{aligned}$$

and from Eq. (14.11)

$$f(x) = \frac{2\pi^2}{3} - 4\sum_{n=1}^{\infty}\frac{\cos(nx)}{n^2}$$

b) Considering $f(\pi)$

$$\pi(2\pi - \pi) = \frac{2\pi^2}{3} - 4\sum_{n=1}^{\infty}\frac{\cos(n\pi)}{n^2}$$

from which it follows that

$$\pi^2 = \frac{2\pi^2}{3} - 4\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$$

and therefore

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} = -\frac{\pi^2}{12}$$

c) From Parseval's theorem, Eq. (14.15)

$$\langle x^2(2\pi - x)^2 \rangle = \langle 4\pi^2 x^2 - 4\pi x^3 + x^4 \rangle = \frac{8\pi^4}{9} + 16\sum_{n=1}^{\infty} \frac{1}{n^4}$$

and

$$\begin{aligned}
\langle x^2(x-\pi)^2 \rangle &= \frac{1}{\pi}\int_0^{2\pi} (4\pi^2 x^2 - 4\pi x^3 + x^4)dx \\
&= \frac{1}{\pi}\left[\frac{4\pi^2 x^3}{3} - \pi x^4 + \frac{x^5}{5}\right]_0^{2\pi} \\
&= \frac{1}{\pi}\left[\frac{32\pi^5}{3} - 16\pi^5 + \frac{32\pi^5}{5}\right] = \frac{16\pi^4}{15}
\end{aligned}$$

Equating the two expressions above then leads to

$$\frac{16\pi^4}{15} = \frac{8\pi^4}{9} + 16\sum_{n=1}^{\infty} \frac{1}{n^4}$$

and therefore

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$$

Worked Solution 14.3 (see Problem 14.3)

Consider $f(x) = \begin{cases} x, & 0 \le x < \pi \\ 2\pi - x, & \pi \le x \le 2\pi \end{cases}$

a) Considering Eqs. (14.12) to (14.14) in conjunction with the results from Problem 13.4c from the last session

$$\begin{aligned} a_0 &= \langle f(x) \rangle &&= \pi \\ a_n &= \langle f(x) \rangle &&= \frac{2[(-1)^n - 1]}{n^2\pi} \\ b_n &= \langle f(x) \rangle &&= 0 \end{aligned}$$

and from Eq. (14.11)

$$\begin{aligned} f(x) &= \frac{\pi}{2} + \sum_{n=1}^{\infty} \frac{2[(-1)^n - 1]}{n^2\pi} \cos(nx) \\ &= \frac{\pi}{2} + \sum_{n=0}^{\infty} \frac{2[(-1)^{2n+1} - 1]}{(2n+1)^2\pi} \cos((2n+1)x) \\ &= \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\cos((2n+1)x)}{(2n+1)^2} \end{aligned}$$

b) Considering $f(\pi)$

$$\pi = \frac{\pi}{2} + \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2}$$

from which it follows that

$$\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}$$

c) From Parseval's theorem, Eq. (14.15)

$$\begin{aligned}
\langle f^2\rangle &= \frac{\pi^2}{2}+\sum_{n=1}^{\infty}\left[\frac{2[(-1)^n-1]}{n^2\pi}\right]^2\\
&= \frac{\pi^2}{2}+\frac{4}{\pi^2}\sum_{n=1}^{\infty}\frac{[(-1)^{2n}-2(-1)^n+1]}{n^4}\\
&= \frac{\pi^2}{2}+\frac{8}{\pi^2}\sum_{n=1}^{\infty}\frac{1+(-1)^{n+1}}{n^4}\\
&= \frac{\pi^2}{2}+\frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{1+(-1)^{2n+1+1}}{(2n+1)^4}\\
&= \frac{\pi^2}{2}+\frac{16}{\pi^2}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^4}
\end{aligned}$$

and

$$\begin{aligned}
\langle f^2\rangle &= \frac{1}{\pi}\int_0^{\pi}x^2dx+\frac{1}{\pi}\int_{\pi}^{2\pi}(2\pi-x)^2dx\\
&= \frac{1}{\pi}\left[\frac{x^3}{3}\right]_0^{\pi}+\frac{1}{\pi}\left[\frac{(2\pi-x)^3}{3}\right]_{\pi}^{2\pi}\\
&= \frac{1}{\pi}\left[\frac{\pi^3}{3}+\frac{\pi^3}{3}\right]=\frac{2\pi^2}{3}
\end{aligned}$$

Equating the two expressions above then leads to

$$\frac{2\pi^2}{3}=\frac{\pi^2}{2}+\frac{16}{\pi^2}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^4}$$

and therefore

$$\sum_{n=0}^{\infty}\frac{1}{(2n+1)^4}=\frac{\pi^4}{96}$$

Worked Solution 14.4 (see Problem 14.4)

Consider $f(x) = \begin{cases} -1, & 0 \le x < \pi \\ 1, & \pi \le x \le 2\pi \end{cases}$

a) Considering Eqs. (14.12) to (14.14) in conjunction with the results from Problem 13.4d from the last session

$$\begin{aligned} a_0 &= \langle f(x) \rangle &&= 0 \\ a_n &= \langle f(x) \rangle &&= 0 \\ b_n &= \langle f(x) \rangle &&= \frac{2[(-1)^n - 1]}{n\pi} \end{aligned}$$

and from Eq. (14.11)

$$\begin{aligned} f(x) &= \sum_{n=1}^{\infty} \frac{2[(-1)^n - 1]}{n\pi} \sin(nx) \\ &= -\frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin((2n+1)x)}{2n+1} \end{aligned}$$

b) Considering $f(\pi/2)$

$$-1 = -\frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin((2n+1)\pi/2)}{2n+1} = -\frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$$

from which it follows that

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}$$

c) From Parseval's theorem, Eq. (14.15)

$$\begin{aligned}
\langle f^2 \rangle &= \sum_{n=1}^{\infty} \left[\frac{2[(-1)^n - 1]}{n\pi}\right]^2 \\
&= \frac{4}{\pi^2} \sum_{n=1}^{\infty} \frac{[(-1)^{2n} - 2(-1)^n + 1]}{n^2} \\
&= \frac{8}{\pi^2} \sum_{n=1}^{\infty} \frac{1 + (-1)^{n+1}}{n^2} \\
&= \frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{1 + (-1)^{2n+1+1}}{(2n+1)^2} \\
&= \frac{16}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2}
\end{aligned}$$

and

$$\begin{aligned}
\langle f^2 \rangle &= \frac{1}{\pi} \int_0^{\pi} 1dx + \frac{1}{\pi} \int_{\pi}^{2\pi} 1dx \\
&= \frac{1}{\pi} [x]_0^{\pi} + \frac{1}{\pi} [x]_{\pi}^{2\pi} \\
&= \frac{1}{\pi} [\pi + 2\pi - \pi] = 2
\end{aligned}$$

Equating the two expressions above then leads to

$$2 = \frac{16}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2}$$

and therefore

$$\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}$$

15

Vectors and matrices

15.1 Learning outcomes

You should be able to:

- Understand the difference between scalars, vectors and matrices.
- Perform basic matrix arithmetic including matrix multiplication.
- Understand the meaning of a transpose matrix, identity matrix and the trace of a matrix.
- Determine the dot product of two vectors.
- Derive a rotation matrix, $\mathbf{L}$ for a two-dimensional cartesian system.
- Show that the rotated form of a matrix $\mathbf{A}$ is found from $\mathbf{A}' = \mathbf{LAL}^{\mathrm{T}}$.

15.2 Scalars, vectors and matrices

A scalar is a quantity in which direction is either not applicable (as in temperature) or not specified (as in speed). Scalars are typically observed in typeface as italic letters (e.g. a, b).

A vector is a quantity in which both the magnitude and direction must be specified (such as velocity). Vectors are typically observed in typeface as lower case bold faced letters (e.g. $\mathbf{a}$, $\mathbf{b}$). When handwritten, vectors are commonly indicated by placing a single underline (e.g. $\underline{a}$, $\underline{b}$).

A matrix is a set of quantities contained within a rectangular array. Matrices are typically observed in typeface as upper case bold faced letters (e.g. $\mathbf{A}$, $\mathbf{B}$). When handwritten, matrices are commonly indicated by upper case letters with a single underline (e.g. $\underline{A}$, $\underline{B}$).

The components of a vector, $\mathbf{a}$, are often specified in terms of scalars with a single subscript (e.g. a_i) such that for a vector containing N components, it can be said that

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{N-1} \\ a_N \end{bmatrix}$$

The components of a matrix, $\mathbf{A}$, are often specified in terms of scalars with a double subscripts, $A_{i,j}$, where i denotes the row number and j denotes the column number. An N by M matrix would therefore take the form

$$\mathbf{A} = \begin{bmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,M-1} & A_{1,M} \\ A_{2,1} & A_{2,2} & \dots & A_{2,M-1} & A_{2,M} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{N-2,1} & A_{N-2,2} & \dots & A_{N-2,M-1} & A_{N-2,M} \\ A_{N-1,1} & A_{N-1,2} & \dots & A_{N-1,M-1} & A_{N-1,M} \\ A_{N,1} & A_{N,2} & \dots & A_{N,M-1} & A_{N,M} \end{bmatrix}$$

A vector is a matrix with one column. A scalar is a matrix with one column and one row.

15.3 Matrix arithmetic

Providing matrices are the same size, matrices can be added and subtracted without too much difficulty, for example:

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} A_{1,1} + B_{1,1} & A_{1,2} + B_{1,2} & \dots & A_{1,M} + B_{1,M} \\ A_{2,1} + B_{2,1} & A_{2,2} + B_{2,2} & \dots & A_{2,M} + B_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N,1} + B_{N,1} & A_{N,2} + B_{N,2} & \dots & A_{N,M} + B_{N,M} \end{bmatrix}$$

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} A_{1,1} - B_{1,1} & A_{1,2} - B_{1,2} & \dots & A_{1,M} - B_{1,M} \\ A_{2,1} - B_{2,1} & A_{2,2} - B_{2,2} & \dots & A_{2,M} - B_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N,1} - B_{N,1} & A_{N,2} - B_{N,2} & \dots & A_{N,M} - B_{N,M} \end{bmatrix}$$

It is also straightforward to multiply matrices by scalar quantities:

$$\lambda\mathbf{A} = \begin{bmatrix} \lambda A_{1,1} & \lambda A_{1,2} & \dots & \lambda A_{1,M} \\ \lambda A_{2,1} & \lambda A_{2,2} & \dots & \lambda A_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda A_{N,1} & \lambda A_{N,2} & \dots & \lambda A_{N,M} \end{bmatrix}$$

Multiplying a matrix by another matrix is more complicated. Furthermore, for matrix multiplication to be possible, the first matrix must have the same number of columns as the number of rows in the second matrix. Consider two matrices, **A** and **B**, defined as follows:

$$\mathbf{A} = \begin{bmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,M} \\ A_{2,1} & A_{2,2} & \dots & A_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N,1} & A_{N,2} & \dots & A_{N,M} \end{bmatrix} \tag{15.1}$$

$$\mathbf{B} = \begin{bmatrix} B_{1,1} & B_{1,2} & \dots & B_{1,P} \\ B_{2,1} & B_{2,2} & \dots & B_{2,P} \\ \vdots & \vdots & \ddots & \vdots \\ B_{M,1} & B_{M,2} & \dots & B_{M,P} \end{bmatrix}$$

The product of A and B will take the form

$$\mathbf{AB} = \begin{bmatrix} (\mathbf{AB})_{1,1} & (\mathbf{AB})_{1,2} & \dots & (\mathbf{AB})_{1,P} \\ (\mathbf{AB})_{2,1} & (\mathbf{AB})_{2,2} & \dots & (\mathbf{AB})_{2,P} \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{AB})_{N,1} & (\mathbf{AB})_{N,2} & \dots & (\mathbf{AB})_{N,P} \end{bmatrix}$$

where

$$(\mathbf{AB})_{i,j} = \sum_{k=1}^{M} A_{i,k} B_{k,j}$$

15.4 Transpose, identity and trace

The term $\mathbf{A}^{\mathrm{T}}$ indicates the transpose of the matrix, $\mathbf{A}$. Consider the definition of $\mathbf{A}$ given in Eq. (15.1).

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} A_{1,1} & A_{2,1} & \dots & A_{N,1} \\ A_{1,2} & A_{2,2} & \dots & A_{N,2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1,M} & A_{2,M} & \dots & A_{N,M} \end{bmatrix}$$

The term $\mathbf{I}$ denotes an identity matrix. Identity matrices are square matrices where all elements are zero except for the central diagonal elements which are all one, i.e.,

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Note that, given Eq. (15.1), $\mathbf{IA} = \mathbf{A}$ where $\mathbf{I}$ would be a $N \times N$ square matrix. In contrast, $\mathbf{AI} = \mathbf{A}$ where $\mathbf{I}$ would be a $M \times M$ square matrix.

The trace of a matrix is the sum of the elements on the central diagonal of a square matrix. For example, consider Eq. (15.1) when $M = N$:

$$\mathrm{trace}(\mathbf{A}) = \sum_{i=1}^{N} A_{i,i}$$

15.5 Dot product

From the above discussion about matrix multiplication, it can be understood that the multiplication of two vectors is not possible. However, a useful associated operation is the so-called dot product. Considered two vectors **a** and **b**, both containing N elements. The dot product of **a** and **b** can be written as follows:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^{\mathrm{T}} \mathbf{b} = \begin{bmatrix} a_1 & a_2 & \dots & a_N \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} = \sum_{i=1}^{N} a_i b_i$$

15.6 Rotation matrix

Challenge 15.1 Let u and v be displacements in the x and y direction, respectively (see Fig. 15.1). Now consider an alternative coordinate system, x' and y', obtained by rotating the x and y axes anti-clockwise by an angle, θ, about the origin, (0,0). Now let u' and v' be the associated displacements in the x' and y' direction, respectively. Determine expressions for u' and v' in terms of u and v.

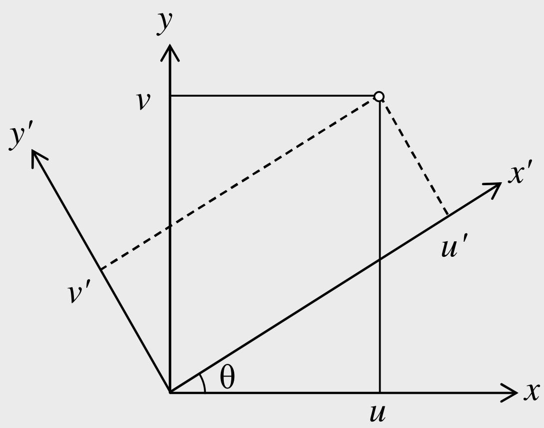


Figure 15.1: Schematic diagram showing a rotated coordinate system.

The first step is to consider a sketch as shown in Fig. 15.2.

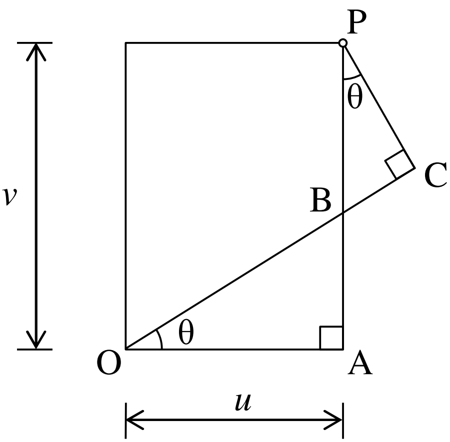


Figure 15.2: Some geometry to help with determining a rotation matrix.

It can be seen that $u' = \mathrm{OB} + \mathrm{BC}$ and $v' = \mathrm{PC}$. It is easily seen that $\mathrm{OB} = u\sec\theta$. Furthermore, $\mathrm{BC} = \mathrm{BP}\sin\theta$ and $\mathrm{PC} = \mathrm{BP}\cos\theta$ where $\mathrm{BP} = v - \mathrm{AB}$. It is also easily seen that $\mathrm{AB} = u\tan\theta$. Therefore $\mathrm{PC} = (v - u\tan\theta)\cos\theta$ and $\mathrm{BC} = (v - u\tan\theta)\sin\theta$. It follows that

$$\begin{aligned}
u' &= \mathrm{OB} + \mathrm{BC} \\
&= u\sec\theta + (v - u\tan\theta)\sin\theta \\
&= u\sec\theta + v\sin\theta - u\sin^2\theta\sec\theta \\
&= u(1 - \sin^2\theta)\sec\theta + v\sin\theta \\
&= u\cos^2\theta\sec\theta + v\sin\theta \\
&= u\cos\theta + v\sin\theta
\end{aligned}$$

and

$$v' = \mathrm{PC} = v\cos\theta - u\sin\theta$$

Challenge 15.2 Given that $u' = u\cos\theta + v\sin\theta$ and $v' = v\cos\theta - u\sin\theta$, determine the rotation matrix, $\mathbf{L}$, that satisfies the equation

$$\mathbf{u}' = \mathbf{L}\mathbf{u}$$

where

$$\mathbf{u}' = \begin{bmatrix} u' \\ v' \end{bmatrix} \quad \text{and} \quad \mathbf{u} = \begin{bmatrix} u \\ v \end{bmatrix}$$

Given that $\mathbf{u}' = \mathbf{L}\mathbf{u}$ it follows that

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} L_{11}u + L_{12}v \\ L_{21}u + L_{22}v \end{bmatrix} \tag{15.2}$$

such that $u' = L_{11}u + L_{12}v$ and $v' = L_{21}u + L_{22}v$.

Given that $u' = u\cos\theta + v\sin\theta$ and $v' = v\cos\theta - u\sin\theta$ it follows that $L_{11} = \cos\theta$, $L_{12} = \sin\theta$, $L_{21} = -\sin\theta$ and $L_{22} = \cos\theta$ such that it can be said that

$$\mathbf{L} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{15.3}$$

Challenge 15.3 Determine $\mathbf{L}\mathbf{L}^{\mathrm{T}}$.

$$\begin{aligned}
\mathbf{L}\mathbf{L}^{\mathrm{T}} &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \cos\theta\sin\theta \\ \cos\theta\sin\theta - \cos\theta\sin\theta & \cos^2\theta + \sin^2\theta \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}
\end{aligned}$$

Orthogonality of rotation matrices

When a matrix $\mathbf{B}$ satisfies the relationship $\mathbf{B}\mathbf{B}^{\mathrm{T}} = \mathbf{I}$, it is said to be orthogonal.

Given an appropriate form of $\mathbf{L}$, Eq. (15.2) can enable the transformation of a vector, $\mathbf{u}$, from any coordinate system (x, y, z) to any other coordinate system (x', y', z'). Furthermore, all rotation matrices are orthogonal. Hence it is a general case that

$$\mathbf{L}\mathbf{L}^{\mathrm{T}} = \mathbf{L}^{\mathrm{T}}\mathbf{L} = \mathbf{I} \tag{15.4}$$

15.7 Rotating a matrix

A rotation matrix can also be used to transform a matrix to a new coordinate system. How this is achieved requires some additional thought concerning matrices and their relationship with vectors through matrix multiplication.

Challenge 15.4 Consider the relationship

$$\mathbf{a} = \mathbf{A}\mathbf{b} \tag{15.5}$$

where $\mathbf{a}$ and $\mathbf{b}$ are vectors and $\mathbf{A}$ is a matrix.

It can be said that $\mathbf{a}'$ and $\mathbf{b}'$ associated with a new coordinate system can be found from

$$\mathbf{a}' = \mathbf{L}\mathbf{a} \quad \text{and} \quad \mathbf{b}' = \mathbf{L}\mathbf{b} \tag{15.6}$$

where $\mathbf{L}$ is a rotation matrix.

Let $\mathbf{A}'$ represent $\mathbf{A}$ in the new coordinate system such that

$$\mathbf{a}' = \mathbf{A}'\mathbf{b}' \tag{15.7}$$

Determine an expression for $\mathbf{A}'$ in terms of $\mathbf{A}$ and $\mathbf{L}$.

Substituting the expressions in Eq. (15.6) into Eq. (15.7) leads to

$$\mathbf{La} = \mathbf{A}'\mathbf{Lb} \tag{15.8}$$

Multiplying both sides of Eq. (15.8) by $\mathbf{L}^{\mathrm{T}}$ leads to

$$\mathbf{L}^{\mathrm{T}}\mathbf{La} = \mathbf{L}^{\mathrm{T}}\mathbf{A}'\mathbf{Lb}$$

and exploiting the identity in Eq. (15.4) yields

$$\mathbf{Ia} = \mathbf{a} = \mathbf{L}^{\mathrm{T}}\mathbf{A}'\mathbf{Lb} \tag{15.9}$$

Comparing Eqs. (15.9) with (15.5) then reveals that

$$\mathbf{A} = \mathbf{L}^{\mathrm{T}}\mathbf{A}'\mathbf{L}$$

and multiplying both sides by $\mathbf{L}$ and $\mathbf{L}^{\mathrm{T}}$ gives

$$\mathbf{LAL}^{\mathrm{T}} = \mathbf{LL}^{\mathrm{T}}\mathbf{A}'\mathbf{LL}^{\mathrm{T}}$$

such that invoking again Eq. (15.4) yields

$$\mathbf{A}' = \mathbf{LAL}^{\mathrm{T}} \tag{15.10}$$

which is an expression that can be used to rotate a matrix , $\mathbf{A}$, to any new coordinate system.

15.8 Problem sheet

Problem 15.1 (see Worked Solution 15.1)

Let

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

and determine the following:

a)	$\mathbf{a}\cdot\mathbf{b}$	b)	$\mathbf{b}\cdot\mathbf{a}$
c)	$(\mathbf{a}+\mathbf{b})\cdot\mathbf{c}$	d)	$\mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c}$
e)	$\mathbf{c}\cdot(\mathbf{a}+\mathbf{b})$	f)	$\mathbf{a}\cdot(\mathbf{b}+\mathbf{c})$
g)	$(\mathbf{a}\cdot\mathbf{b})\mathbf{c}$	h)	$(\mathbf{c}\cdot\mathbf{b})\mathbf{a}$

Problem 15.2 (see Worked Solution 15.2)

Let

$$\mathbf{a}=\begin{bmatrix}3\\1\\4\end{bmatrix}\quad \mathbf{B}=\begin{bmatrix}0&1\\0&-2\\2&3\end{bmatrix}\quad \mathbf{C}=\begin{bmatrix}1&0&-1\\2&3&0\\0&3&4\end{bmatrix}$$

and determine the following:

a)	$\mathbf{CB}$	b)	$\mathbf{B}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}$	c)	$\mathbf{BC}^{\mathrm{T}}$
d)	$\mathbf{C}^2$	e)	$\mathbf{CC}^{\mathrm{T}}$	f)	$\mathbf{C}^{\mathrm{T}}\mathbf{C}$
g)	$\mathbf{B}^2$	h)	$\mathbf{BB}^{\mathrm{T}}$	i)	$\mathbf{B}^{\mathrm{T}}\mathbf{B}$
j)	$\mathbf{a}^{\mathrm{T}}\mathbf{Ca}$				

Problem 15.3 (see Worked Solution 15.3)

Verify that the following matrix is orthogonal

$$\mathbf{G}=\begin{bmatrix}0&-0.80&-0.60\\0.80&-0.36&0.48\\0.60&0.48&-0.64\end{bmatrix}$$

15.9 Worked solutions

Worked Solution 15.1 (see Problem 15.1)

Let

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

a) $\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$

b) $\mathbf{b} \cdot \mathbf{a} = b_1 a_1 + b_2 a_2 + b_3 a_3$
$\therefore \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$

c) $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = (a_1 + b_1)c_1 + (a_2 + b_2)c_2 + (a_3 + b_3)c_3$

d) $\mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c} = a_1 c_1 + a_2 c_2 + a_3 c_3 + b_1 c_1 + b_2 c_2 + b_3 c_3$

e) $\mathbf{c} \cdot (\mathbf{a} + \mathbf{b}) = c_1(a_1 + b_1) + c_2(a_2 + b_2) + c_3(a_3 + b_3)$
$\therefore (\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c} = \mathbf{c} \cdot (\mathbf{a} + \mathbf{b})$

f) $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = a_1(b_1 + c_1) + a_2(b_2 + c_2) + a_3(b_3 + c_3)$
$\therefore \mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) \neq \mathbf{c} \cdot (\mathbf{a} + \mathbf{b})$

g) $(\mathbf{a} \cdot \mathbf{b})\mathbf{c} = \begin{bmatrix} (a_1 b_1 + a_2 b_2 + a_3 b_3)c_1 \\ (a_1 b_1 + a_2 b_2 + a_3 b_3)c_2 \\ (a_1 b_1 + a_2 b_2 + a_3 b_3)c_3 \end{bmatrix}$

h) $(\mathbf{c} \cdot \mathbf{b})\mathbf{a} = \begin{bmatrix} (c_1 b_1 + c_2 b_2 + c_3 b_3)a_1 \\ (c_1 b_1 + c_2 b_2 + c_3 b_3)a_2 \\ (c_1 b_1 + c_2 b_2 + c_3 b_3)a_3 \end{bmatrix}$
$\therefore (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \neq (\mathbf{c} \cdot \mathbf{b})\mathbf{a}$

Worked Solution 15.2 (see Problem 15.2)

Let

$$\mathbf{a} = \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 0 & -2 \\ 2 & 3 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 0 \\ 0 & 3 & 4 \end{bmatrix}$$

a)

$$\mathbf{CB} = \begin{bmatrix} 1\times 0+0\times 0-1\times 2 & 1\times 1+0\times -2-1\times 3 \\ 2\times 0+3\times 0+0\times 2 & 2\times 1+3\times -2+0\times 3 \\ 0\times 0+3\times 0+4\times 2 & 0\times 1+3\times -2+4\times 3 \end{bmatrix}$$

$$= \begin{bmatrix} 0+0-2 & 1+0-3 \\ 0+0+0 & 2-6+0 \\ 0+0+8 & 0-6+12 \end{bmatrix} = \begin{bmatrix} -2 & -2 \\ 0 & -4 \\ 8 & 6 \end{bmatrix}$$

b)

$$\mathbf{B}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}} = \begin{bmatrix} 0 & 0 & 2 \\ 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 3 \\ -1 & 0 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 0+0-2 & 0+0+0 & 0+0+8 \\ 1+0-3 & 2-6+0 & 0-6+12 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 0 & 8 \\ -2 & -4 & 6 \end{bmatrix} = (\mathbf{CB})^{\mathrm{T}}$$

c) $\mathbf{BC}^{\mathrm{T}}$ is not possible because the number of columns in $\mathbf{B}$ is not the same as the number of rows in $\mathbf{C}^{\mathrm{T}}$.

d)

$$\begin{aligned}
\mathbf{C}^2 &= \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 0 \\ 0 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 0 \\ 0 & 3 & 4 \end{bmatrix} \\
&= \begin{bmatrix} 1+0+0 & 0+0-3 & -1+0-4 \\ 2+6+0 & 0+9+0 & -2+0+0 \\ 0+6+0 & 0+9+12 & 0+0+16 \end{bmatrix} \\
&= \begin{bmatrix} 1 & -3 & -5 \\ 8 & 9 & -2 \\ 6 & 21 & 16 \end{bmatrix}
\end{aligned}$$

e)

$$\begin{aligned}
\mathbf{C}\mathbf{C}^{\mathrm{T}} &= \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 0 \\ 0 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 3 \\ -1 & 0 & 4 \end{bmatrix} \\
&= \begin{bmatrix} 1+0+1 & 2+0+0 & 0+0-4 \\ 2+0+0 & 4+9+0 & 0+9+0 \\ 0+0-4 & 0+9+0 & 0+9+16 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 2 & -4 \\ 2 & 13 & 9 \\ -4 & 9 & 25 \end{bmatrix}
\end{aligned}$$

f)

$$\mathbf{C}^{\mathrm{T}}\mathbf{C} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 3 \\ -1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 0 \\ 0 & 3 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 1+4+0 & 0+6+0 & -1+0+0 \\ 0+6+0 & 0+9+9 & 0+0+12 \\ -1+0+0 & 0+0+12 & 1+0+16 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 6 & -1 \\ 6 & 18 & 12 \\ -1 & 12 & 17 \end{bmatrix}$$

g) $\mathbf{B}^2$ is not possible because the number of columns in **B** is not the same as the number of rows in **B**.

h)

$$\mathbf{B}\mathbf{B}^{\mathrm{T}} = \begin{bmatrix} 0 & 1 \\ 0 & -2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 2 \\ 1 & -2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 0+1 & 0-2 & 0+3 \\ 0-2 & 0+4 & 0-6 \\ 0+3 & 0-6 & 4+9 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \\ 3 & -6 & 13 \end{bmatrix}$$

i)

$$\mathbf{B}^{\mathrm{T}}\mathbf{B} = \begin{bmatrix} 0 & 0 & 2 \\ 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & -2 \\ 2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 0+0+4 & 0+0+6 \\ 0+0+6 & 1+4+9 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}$$

j) To calculate $\mathbf{a}^{\mathrm{T}}\mathbf{C}\mathbf{a}$, it is necessary to first calculate $\mathbf{a}^{\mathrm{T}}\mathbf{C}$

$$\begin{aligned}
\mathbf{a}^{\mathrm{T}}\mathbf{C} &= \begin{bmatrix} 3 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 0 \\ 0 & 3 & 4 \end{bmatrix} \\
&= \begin{bmatrix} 3+2+0 & 0+3+12 & -3+0+16 \end{bmatrix} \\
&= \begin{bmatrix} 5 & 15 & 13 \end{bmatrix}
\end{aligned}$$

Now it can be seen that

$$\mathbf{a}^{\mathrm{T}}\mathbf{C}\mathbf{a} = \begin{bmatrix} 5 & 15 & 13 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} = 15+15+52 = 82$$

Worked Solution 15.3 (see Problem 15.3)

To determine if $\mathbf{G}$ is orthogonal it is necessary to check that $\mathbf{G}\mathbf{G}^{\mathrm{T}} = \mathbf{I}$.

$$\begin{aligned}
\mathbf{G}\mathbf{G}^{\mathrm{T}} &= \begin{bmatrix} 0 & -0.80 & -0.60 \\ 0.80 & -0.36 & 0.48 \\ 0.60 & 0.48 & -0.64 \end{bmatrix} \begin{bmatrix} 0 & 0.80 & 0.60 \\ -0.80 & -0.36 & 0.48 \\ -0.60 & 0.48 & -0.64 \end{bmatrix} \\
&= \begin{bmatrix} 0+0.64+0.36 & 0+0.29-0.29 & 0-0.38+0.38 \\ 0+0.29-0.29 & 0.64+0.13+0.23 & 0.48-0.17-0.31 \\ 0-0.38+0.38 & 0.48-0.17-0.31 & 0.36+0.23+0.41 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}
\end{aligned}$$

Therefore $\mathbf{G}$ is orthogonal.

16

Matrix operations with stress and strain

16.1 Learning outcomes

You should be able to:

- Derive the equation for a Mohr circle using a rotation matrix.
- Write out three dimensional Hooke's law in matrix notation.
- Use a rotation matrix to show that Hooke's law applies for a coordinate system that does not necessarily represent the principal stress axes.

16.2 Stress, strain and displacement

Let $\boldsymbol{\tau}$ [$\mathrm{ML}^{-1}\mathrm{T}^{-2}$] and $\boldsymbol{\varepsilon}$ [-] be matrices that contain all the stresses and strains acting on a control-volume of rock, respectively, such that:

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \tau_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \tau_{zz} \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix}$$

Stress equals force divided by area. Strain is a measure of the extent to which a body is deformed when it is subjected to a stress. Quite literally, it is the ratio of a change in length divided by the original length.

The arrangement of the nine stresses in the stress tensor on a control-volume is shown in Fig. 16.1.

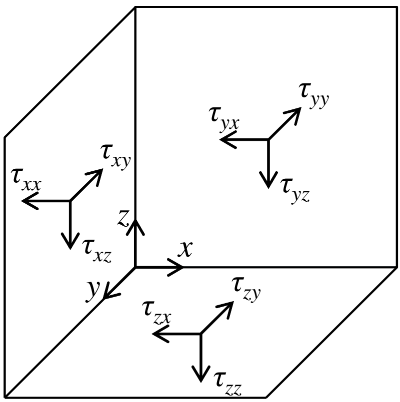


Figure 16.1: Schematic diagram showing the nine stresses in a stress tensor.

There are two types of stress, normal stress and shear stress:

τ_{xx} Normal stress in the x-direction acting on the face normal to the x-direction.
τ_{xy} Shear stress in the y-direction acting on the face normal to the x-direction.
τ_{xz} Shear stress in the z-direction acting on the face normal to the x-direction.
τ_{yx} Shear stress in the x-direction acting on the face normal to the y-direction.
τ_{yy} Normal stress in the y-direction acting on the face normal to the y-direction.
τ_{yz} Shear stress in the z-direction acting on the face normal to the y-direction.
τ_{zx} Shear stress in the x-direction acting on the face normal to the z-direction.
τ_{zy} Shear stress in the y-direction acting on the face normal to the z-direction.
τ_{zz} Normal stress in the z-direction acting on the face normal to the z-direction.

Let u [L], v [L] and w [L] be displacements in the x, y and z direction, respectively. Collectively, these displacements quantify the change in position of a given particle of rock. The strains contained within ε can be written as partial derivatives of u, v and w as follows:

$$
\begin{aligned}
\boldsymbol{\varepsilon} &= \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix} \\
&= \begin{bmatrix} \dfrac{\partial u}{\partial x} & \dfrac{1}{2}\left(\dfrac{\partial u}{\partial y}+\dfrac{\partial v}{\partial x}\right) & \dfrac{1}{2}\left(\dfrac{\partial u}{\partial z}+\dfrac{\partial w}{\partial x}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial u}{\partial y}+\dfrac{\partial v}{\partial x}\right) & \dfrac{\partial v}{\partial y} & \dfrac{1}{2}\left(\dfrac{\partial v}{\partial z}+\dfrac{\partial w}{\partial y}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial u}{\partial z}+\dfrac{\partial w}{\partial x}\right) & \dfrac{1}{2}\left(\dfrac{\partial v}{\partial z}+\dfrac{\partial w}{\partial y}\right) & \dfrac{\partial w}{\partial z} \end{bmatrix}
\end{aligned}
\tag{16.1}
$$

Note that $\varepsilon_{xy} = \varepsilon_{yx}$, $\varepsilon_{xz} = \varepsilon_{zx}$ and $\varepsilon_{yz} = \varepsilon_{zy}$.

16.3 The Mohr circle

Challenge 16.1 Let $\boldsymbol{\tau}$ be a two dimensional stress matrix for a coordinate system defined by an associated x and y axes where

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_{xx} & \tau_{xy} \\ \tau_{yx} & \tau_{yy} \end{bmatrix}$$

Now consider an alternative coordinate system, x' and y', obtained by rotating the x and y axes anti-clockwise by an angle, θ, about the origin, (0,0).

Show that the corresponding stress matrix in this rotated coordinate system takes the form

$$\boldsymbol{\tau}' = \begin{bmatrix} \tau'_{xx} & \tau'_{xy} \\ \tau'_{yx} & \tau'_{yy} \end{bmatrix}$$

where

$$\begin{aligned} \tau'_{xx} &= \tfrac{1}{2}(\tau_{xx}+\tau_{yy}) + \tfrac{1}{2}(\tau_{xx}-\tau_{yy})\cos 2\theta + \tau_{xy}\sin 2\theta \\ \tau'_{xy} &= \tau'_{yx} = \tau_{xy}\cos 2\theta + \tfrac{1}{2}(\tau_{yy}-\tau_{xx})\sin 2\theta \\ \tau'_{yy} &= \tfrac{1}{2}(\tau_{xx}+\tau_{yy}) - \tfrac{1}{2}(\tau_{xx}-\tau_{yy})\cos 2\theta - \tau_{xy}\sin 2\theta \end{aligned}$$

For a two-dimensional system

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_{xx} & \tau_{xy} \\ \tau_{yx} & \tau_{yy} \end{bmatrix}$$

and the rotation matrix takes the form

$$\mathbf{L} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

It follows that the stress matrix for the new coordinate system can be found from $\boldsymbol{\tau}' = \mathbf{L}\boldsymbol{\tau}\mathbf{L}^{\mathrm{T}}$ where

$$\boldsymbol{\tau}' = \begin{bmatrix} \tau'_{xx} & \tau'_{xy} \\ \tau'_{yx} & \tau'_{yy} \end{bmatrix}$$

To determine expressions for the individual components of $\boldsymbol{\tau}'$, first determine the matrix $\mathbf{M}$, defined by $\mathbf{M} = \mathbf{L}\boldsymbol{\tau}$ and

$$\mathbf{M} = \begin{bmatrix} M_{xx} & M_{xy} \\ M_{yx} & M_{yy} \end{bmatrix}$$

It follows that

$$\begin{aligned} M_{xx} &= \tau_{xx}\cos\theta + \tau_{yx}\sin\theta \\ M_{xy} &= \tau_{xy}\cos\theta + \tau_{yy}\sin\theta \\ M_{yx} &= -\tau_{xx}\sin\theta + \tau_{yx}\cos\theta \\ M_{yy} &= -\tau_{xy}\sin\theta + \tau_{yy}\cos\theta \end{aligned}$$

Now consider that

$$\mathbf{L}^{\mathrm{T}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

It follows that:

$$\begin{aligned} \tau'_{xx} &= M_{xx}\cos\theta + M_{xy}\sin\theta \\ \tau'_{xy} &= -M_{xx}\sin\theta + M_{xy}\cos\theta \\ \tau'_{yx} &= M_{yx}\cos\theta + M_{yy}\sin\theta \\ \tau'_{yy} &= -M_{yx}\sin\theta + M_{yy}\cos\theta \end{aligned}$$

and

$$\begin{aligned} \tau'_{xx} &= (\tau_{xx}\cos\theta + \tau_{yx}\sin\theta)\cos\theta + (\tau_{xy}\cos\theta + \tau_{yy}\sin\theta)\sin\theta \\ \tau'_{xy} &= -(\tau_{xx}\cos\theta + \tau_{yx}\sin\theta)\sin\theta + (\tau_{xy}\cos\theta + \tau_{yy}\sin\theta)\cos\theta \\ \tau'_{yx} &= (-\tau_{xx}\sin\theta + \tau_{yx}\cos\theta)\cos\theta + (-\tau_{xy}\sin\theta + \tau_{yy}\cos\theta)\sin\theta \\ \tau'_{yy} &= -(-\tau_{xx}\sin\theta + \tau_{yx}\cos\theta)\sin\theta + (-\tau_{xy}\sin\theta + \tau_{yy}\cos\theta)\cos\theta \end{aligned}$$

which reduces further to

$$\begin{aligned}
\tau'_{xx} &= \tau_{xx}\cos^2\theta + (\tau_{yx}+\tau_{xy})\cos\theta\sin\theta + \tau_{yy}\sin^2\theta \\
\tau'_{xy} &= \tau_{xy}\cos^2\theta + (\tau_{yy}-\tau_{xx})\cos\theta\sin\theta - \tau_{yx}\sin^2\theta \\
\tau'_{yx} &= \tau_{yx}\cos^2\theta + (\tau_{yy}-\tau_{xx})\cos\theta\sin\theta - \tau_{xy}\sin^2\theta \\
\tau'_{yy} &= \tau_{yy}\cos^2\theta - (\tau_{yx}+\tau_{xy})\cos\theta\sin\theta + \tau_{xx}\sin^2\theta
\end{aligned}$$

Then recalling that $\tau_{xy} = \tau_{yx}$:

$$\begin{aligned}
\tau'_{xx} &= \tau_{xx}\cos^2\theta + 2\tau_{xy}\cos\theta\sin\theta + \tau_{yy}\sin^2\theta \\
\tau'_{xy} &= \tau_{xy}(\cos^2\theta - \sin^2\theta) + (\tau_{yy}-\tau_{xx})\cos\theta\sin\theta \\
\tau'_{yx} &= \tau_{xy}(\cos^2\theta - \sin^2\theta) + (\tau_{yy}-\tau_{xx})\cos\theta\sin\theta \\
\tau'_{yy} &= \tau_{yy}\cos^2\theta - 2\tau_{xy}\cos\theta\sin\theta + \tau_{xx}\sin^2\theta
\end{aligned}$$

Now also recall that

$$\begin{aligned}
\cos(A+B) &= \cos A\cos B - \sin A\sin B \\
\sin(A+B) &= \sin A\cos B + \cos A\sin B
\end{aligned}$$

from which it follows that

$$\begin{aligned}
&\cos 2\theta = \cos^2\theta - \sin^2\theta \\
&\sin 2\theta = 2\cos\theta\sin\theta \\
&2\cos^2\theta = \cos^2\theta + 1 - \sin^2\theta = 1 + \cos 2\theta \\
&2\sin^2\theta = \sin^2\theta + 1 - \cos^2\theta = 1 - \cos 2\theta
\end{aligned}$$

and

$$\begin{aligned}
\tau'_{xx} &= \tfrac{1}{2}(\tau_{xx}+\tau_{yy}) + \tfrac{1}{2}(\tau_{xx}-\tau_{yy})\cos 2\theta + \tau_{xy}\sin 2\theta \\
\tau'_{xy} &= \tau'_{yx} = \tau_{xy}\cos 2\theta + \tfrac{1}{2}(\tau_{yy}-\tau_{xx})\sin 2\theta \\
\tau'_{yy} &= \tfrac{1}{2}(\tau_{xx}+\tau_{yy}) - \tfrac{1}{2}(\tau_{xx}-\tau_{yy})\cos 2\theta - \tau_{xy}\sin 2\theta
\end{aligned}$$

Challenge 16.2 Show that for the case when x and y represent the principal stress axes, a plot of τ'_{xy} against τ'_{xx} or τ'_{yy} takes the form of a circle of radius, $|(\tau_{xx}-\tau_{yy})/2|$, and center coordinates, $((\tau_{xx}+\tau_{yy})/2, 0)$, which is the so-called Mohr circle.

Note that in an x-y Cartesian coordinate system, all points (x, y) on a circle with center coordinates, (a, b), and radius, r, satisfy the equation, $(x-a)^2+(y-b)^2=r^2$.

Also note that if $\tau_{xy}=\tau_{yx}=0$, x and y are defined as the principal stress axes.

Consider the case when x and y represent the principal stress axes. In this case $\tau_{xy}=0$ and

$$\begin{aligned}
\tau'_{xx} &= \tfrac{1}{2}(\tau_{xx}+\tau_{yy})+\tfrac{1}{2}(\tau_{xx}-\tau_{yy})\cos 2\theta \\
\tau'_{xy} &= \tau'_{yx} = \tfrac{1}{2}(\tau_{yy}-\tau_{xx})\sin 2\theta \\
\tau'_{yy} &= \tfrac{1}{2}(\tau_{xx}+\tau_{yy})-\tfrac{1}{2}(\tau_{xx}-\tau_{yy})\cos 2\theta
\end{aligned}$$

Now notice that

$$\begin{aligned}
\left[\tau'_{xx}-\tfrac{1}{2}(\tau_{xx}+\tau_{yy})\right]^2 &= \tfrac{1}{4}(\tau_{xx}-\tau_{yy})^2\cos^2 2\theta \\
\left[\tau'_{xy}\right]^2 &= \tfrac{1}{4}(\tau_{xx}-\tau_{yy})^2\sin^2 2\theta \\
\left[\tau'_{yy}-\tfrac{1}{2}(\tau_{xx}+\tau_{yy})\right]^2 &= \tfrac{1}{4}(\tau_{xx}-\tau_{yy})^2\cos^2 2\theta
\end{aligned}$$

Hence it can be seen that:

$$\begin{aligned}
\left[\tau'_{xx}-\left(\frac{\tau_{xx}+\tau_{yy}}{2}\right)\right]^2+\left(\tau'_{xy}\right)^2 &= \left(\frac{\tau_{xx}-\tau_{yy}}{2}\right)^2 \\
\left[\tau'_{yy}-\left(\frac{\tau_{xx}+\tau_{yy}}{2}\right)\right]^2+\left(\tau'_{xy}\right)^2 &= \left(\frac{\tau_{xx}-\tau_{yy}}{2}\right)^2
\end{aligned}$$

which, given the equation for a circle, $(x-a)^2+(y-b)^2=r^2$, show that a plot of τ'_{xy} against τ'_{xx} or τ'_{yy} takes the form of a circle of radius $|(\tau_{xx}-\tau_{yy})/2|$ and origin coordinates, $((\tau_{xx}+\tau_{yy})/2, 0)$. This is the so-called Mohr circle.

Challenge 16.3 Finally, also show that $\mathrm{trace}(\boldsymbol{\tau}') = \mathrm{trace}(\boldsymbol{\tau})$.

Consider $\mathrm{trace}(\boldsymbol{\tau}) = \tau_{xx} + \tau_{yy}$ and $\mathrm{trace}(\boldsymbol{\tau}') = \tau'_{xx} + \tau'_{yy}$

From part a) above:

$$\begin{aligned}\tau'_{xx} &= \tfrac{1}{2}(\tau_{xx}+\tau_{yy}) + \tfrac{1}{2}(\tau_{xx}-\tau_{yy})\cos 2\theta + \tau_{xy}\sin 2\theta \\ \tau'_{yy} &= \tfrac{1}{2}(\tau_{xx}+\tau_{yy}) - \tfrac{1}{2}(\tau_{xx}-\tau_{yy})\cos 2\theta - \tau_{xy}\sin 2\theta\end{aligned}$$

It follows that

$$\mathrm{trace}(\boldsymbol{\tau}') = \frac{1}{2}(\tau_{xx}+\tau_{yy}) + \frac{1}{2}(\tau_{xx}+\tau_{yy}) = \tau_{xx}+\tau_{yy}$$

Hence it can be understood that

$$\mathrm{trace}(\boldsymbol{\tau}) = \mathrm{trace}(\boldsymbol{\tau}')$$

In fact it is generally the case that the trace of a matrix is equal to the trace of its rotated form.

16.4 Hooke's law

For linear elastic systems, Hooke's law can be used to relate stress and strain. Consider the situation when all stresses are zero except for τ_{xx}. Such a situation is often referred to as the uniaxial stress assumption. The resulting strain in the associated direction, ε_{xx}, can be found from

$$\varepsilon_{xx} = \frac{\tau_{xx}}{E}$$

where E [$\mathrm{ML^{-1}T^{-2}}$] is the Young's modulus.

In the above situation it is the case that ε_{yy} and ε_{zz} are also non-zero and found from

$$\varepsilon_{yy} = \varepsilon_{zz} = -\nu\varepsilon_{xx}$$

hence

$$\varepsilon_{xx} = \frac{\tau_{xx}}{E}, \quad \varepsilon_{yy} = \varepsilon_{zz} = -\frac{\nu}{E}\tau_{xx} \tag{16.2}$$

where ν [-] is the so-called Poisson ratio. The other strains, $\varepsilon_{xy} = \varepsilon_{xz} = \varepsilon_{yz} = 0$ in this case.

If instead we consider the case when all stresses are zero except for τ_{yy}, it is found that

$$\varepsilon_{yy} = \frac{\tau_{yy}}{E}, \quad \varepsilon_{xx} = \varepsilon_{zz} = -\frac{\nu}{E}\tau_{yy} \tag{16.3}$$

and again $\varepsilon_{xy} = \varepsilon_{xz} = \varepsilon_{yz} = 0$.

Similarly, if we consider the case when all stresses are zero except for τ_{zz}, it is found that

$$\varepsilon_{zz} = \frac{\tau_{zz}}{E}, \quad \varepsilon_{xx} = \varepsilon_{yy} = -\frac{\nu}{E}\tau_{zz} \tag{16.4}$$

and again $\varepsilon_{xy} = \varepsilon_{xz} = \varepsilon_{yz} = 0$.

3D Hooke's law on principal stress axes

Now it is assumed that only $\tau_{xy} = \tau_{xz} = \tau_{yz} = \tau_{yx} = \tau_{zx} = \tau_{zy} = 0$ and that τ_{xx}, τ_{yy} and τ_{zz} are all non-zero. In this case, the axes x, y and z are said to be the principal stress axes. It is possible to use the principle of superposition to determine the associated strains, ε_{xx}, ε_{yy} and ε_{zz} resulting from the simultaneous application of τ_{xx}, τ_{yy} and τ_{zz}. This is achieved by adding together the individual strain contributions given in Eqs. (16.2) to (16.4) such that it can be said that:

$$\varepsilon_{xx} = \frac{1}{E}\tau_{xx} - \frac{\nu}{E}\tau_{yy} - \frac{\nu}{E}\tau_{zz}$$

$$\varepsilon_{yy} = -\frac{\nu}{E}\tau_{xx} + \frac{1}{E}\tau_{yy} - \frac{\nu}{E}\tau_{zz}$$
$$\varepsilon_{zz} = -\frac{\nu}{E}\tau_{xx} - \frac{\nu}{E}\tau_{yy} + \frac{1}{E}\tau_{zz}$$

Challenge 16.4 When x, y and z are the principal stress axes

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_{xx} & 0 & 0 \\ 0 & \tau_{yy} & 0 \\ 0 & 0 & \tau_{zz} \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{xx} & 0 & 0 \\ 0 & \varepsilon_{yy} & 0 \\ 0 & 0 & \varepsilon_{zz} \end{bmatrix}$$

and

$$\varepsilon_{xx} = \frac{1}{E}\tau_{xx} - \frac{\nu}{E}\tau_{yy} - \frac{\nu}{E}\tau_{zz}$$
$$\varepsilon_{yy} = -\frac{\nu}{E}\tau_{xx} + \frac{1}{E}\tau_{yy} - \frac{\nu}{E}\tau_{zz}$$
$$\varepsilon_{zz} = -\frac{\nu}{E}\tau_{xx} - \frac{\nu}{E}\tau_{yy} + \frac{1}{E}\tau_{zz}$$

Under such conditions, determine a compact expression for $\boldsymbol{\varepsilon}$ as a function of $\boldsymbol{\tau}$.

After some further rearrangement it can be shown that:

$$\varepsilon_{xx} = \frac{(1+\nu)}{E}\tau_{xx} - \frac{\nu}{E}(\tau_{xx} + \tau_{yy} + \tau_{zz})$$

$$\varepsilon_{yy} = \frac{(1+\nu)}{E}\tau_{yy} - \frac{\nu}{E}(\tau_{xx} + \tau_{yy} + \tau_{zz})$$

$$\varepsilon_{zz} = \frac{(1+\nu)}{E}\tau_{zz} - \frac{\nu}{E}(\tau_{xx} + \tau_{yy} + \tau_{zz})$$

Also note that if $\tau_{xy} = \tau_{xz} = \tau_{yz} = \tau_{yx} = \tau_{zx} = \tau_{zy} = 0$, it is said that x, y and z represent the directions of the principal stress axes. Furthermore, in this case, it can be understood that

$\varepsilon_{xy} = \varepsilon_{xz} = \varepsilon_{yz} = 0$. Under these conditions, Hooke's law in three dimensions (3D) can be seen to take the compact form

$$\boldsymbol{\varepsilon} = \frac{(1+\nu)}{E}\boldsymbol{\tau} - \frac{\nu}{E}\text{trace}(\boldsymbol{\tau})\mathbf{I}$$

Challenge 16.5 Given that

$$\boldsymbol{\varepsilon} = \frac{(1+\nu)}{E}\boldsymbol{\tau} - \frac{\nu}{E}\text{trace}(\boldsymbol{\tau})\mathbf{I} \tag{16.5}$$

determine a compact expression for $\boldsymbol{\tau}$ as a function of $\boldsymbol{\varepsilon}$.

First take the trace of both sides of Eq. (16.5) to get

$$\begin{aligned}
\text{trace}(\boldsymbol{\varepsilon}) &= \text{trace}\left[\frac{(1+\nu)}{E}\boldsymbol{\tau} - \frac{\nu}{E}\text{trace}(\boldsymbol{\tau})\mathbf{I}\right] \\
&= \frac{(1+\nu)}{E}\text{trace}(\boldsymbol{\tau}) - \frac{3\nu}{E}\text{trace}(\boldsymbol{\tau}) \\
&= \frac{(1-2\nu)}{E}\text{trace}(\boldsymbol{\tau})
\end{aligned}$$

It follows that

$$\text{trace}(\boldsymbol{\tau}) = \frac{E}{(1-2\nu)}\text{trace}(\boldsymbol{\varepsilon})$$

which on substitution into Eq. (16.5) leads to

$$\boldsymbol{\varepsilon} = \frac{(1+\nu)}{E}\boldsymbol{\tau} - \frac{\nu}{(1-2\nu)}\text{trace}(\boldsymbol{\varepsilon})\mathbf{I}$$

which when rearranged for $\boldsymbol{\tau}$ leads to

$$\boldsymbol{\tau} = \frac{E}{(1+\nu)}\boldsymbol{\varepsilon} + \frac{E\nu}{(1+\nu)(1-2\nu)}\text{trace}(\boldsymbol{\varepsilon})\mathbf{I} \tag{16.6}$$

Shear modulus and the Lamé parameter

At this stage it is worth introducing two related mechanical properties, namely the shear modulus, G [$\mathrm{ML^{-1}T^{-2}}$], and the Lamé parameter, λ [$\mathrm{ML^{-1}T^{-2}}$], which are found from:

$$G = \frac{E}{2(1+\nu)}, \quad \lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}$$

such that Eq. (16.6) can be simplified to

$$\boldsymbol{\tau} = 2G\boldsymbol{\varepsilon} + \lambda \mathrm{trace}(\boldsymbol{\varepsilon})\mathbf{I}$$

16.5 3D Hooke's law on any general axes

Challenge 16.6 Consider a second coordinate system, (x', y', z'), obtained from the principal stress coordinate system (x, y, z) through a rotation matrix, $\mathbf{L}$. The stress and strain matrices in the new coordinate system are given by $\boldsymbol{\tau}' = \mathbf{L}\boldsymbol{\tau}\mathbf{L}^{\mathrm{T}}$ and $\varepsilon' = \mathbf{L}\varepsilon\mathbf{L}^{\mathrm{T}}$.

Given that

$$\boldsymbol{\tau} = 2G\boldsymbol{\varepsilon} + \lambda \mathrm{trace}(\boldsymbol{\varepsilon})\mathbf{I} \tag{16.7}$$

determine an expression for $\boldsymbol{\tau}'$ as a function of ε'.

It follows that

$$\boldsymbol{\tau}' = \mathbf{L}[2G\boldsymbol{\varepsilon} + \lambda \mathrm{trace}(\boldsymbol{\varepsilon})\mathbf{I}]\mathbf{L}^{\mathrm{T}} = 2G\mathbf{L}\boldsymbol{\varepsilon}\mathbf{L}^{\mathrm{T}} + \lambda \mathrm{trace}(\boldsymbol{\varepsilon})\mathbf{L}\mathbf{I}\mathbf{L}^{\mathrm{T}}$$

Recall that because $\mathbf{L}$ is orthogonal, $\mathbf{L}\mathbf{L}^{\mathrm{T}} = \mathbf{I}$. In the same way, it follows that $\mathbf{L}^{\mathrm{T}}\mathbf{L} = \mathbf{I}$. Now consider $\mathbf{L}\mathbf{I}\mathbf{L}^{\mathrm{T}}$. Substituting $\mathbf{I} = \mathbf{L}^{\mathrm{T}}\mathbf{L}$ suggests that $\mathbf{L}\mathbf{I}\mathbf{L}^{\mathrm{T}} = \mathbf{L}\mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{L}^{\mathrm{T}} = \mathbf{I}^2 = \mathbf{I}$. Therefore

$$\boldsymbol{\tau}' = 2G\boldsymbol{\varepsilon}' + \lambda \mathrm{trace}(\boldsymbol{\varepsilon})\mathbf{I}$$

Furthermore, it is known that $\mathrm{trace}(\boldsymbol{\varepsilon}) = \mathrm{trace}(\boldsymbol{\varepsilon}')$ (consider Challenge 1c above). Consequently it can be understood that

$$\boldsymbol{\tau}' = 2G\boldsymbol{\varepsilon}' + \lambda \text{trace}(\boldsymbol{\varepsilon}')\mathbf{I}$$

and hence it can be understood that Eqs. (16.5) and (16.7) are also valid for when x, y and z are not necessarily the principal stress directions.

Challenge 16.7 Given that $\boldsymbol{\tau} = 2G\boldsymbol{\varepsilon} + \lambda \text{trace}(\boldsymbol{\varepsilon})\mathbf{I}$ is valid for non-zero shear stresses, determine expressions for the six shear stresses in the stress tensor.

Recalling that

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

it follows that:

$$\tau_{xy} = 2G\varepsilon_{xy}, \quad \tau_{xz} = 2G\varepsilon_{xz}, \quad \tau_{yz} = 2G\varepsilon_{yz}$$

$$\tau_{yx} = 2G\varepsilon_{yx}, \quad \tau_{zx} = 2G\varepsilon_{zx}, \quad \tau_{zy} = 2G\varepsilon_{zy}$$

and furthermore, given that $\varepsilon_{xy} = \varepsilon_{yx}$, $\varepsilon_{xz} = \varepsilon_{zx}$ and $\varepsilon_{yz} = \varepsilon_{zy}$, it can now also be seen that $\tau_{xy} = \tau_{yx}$, $\tau_{xz} = \tau_{zx}$ and $\tau_{yz} = \tau_{zy}$.

17

Vector calculus

17.1 Learning outcomes

You should be able to:

- Write out vector equations using unit vector notation.
- Find the gradient of scalars and vectors.
- Find the divergence of scalars and vectors.
- Show that the Laplacian of a vector is the divergence of the gradient of a vector.
- Find the Laplacian of scalars and vectors.

17.2 Unit vector notation

When working in three dimensions, there are three axes to consider, x, y and z. Hence, when one considers a velocity or a displacement, for example, there are three components for each of the three directions. In this way, it is necessary to represent these quantities as vectors with three elements.

Let u, v and w be displacements in the x, y and z direction, respectively. These displacements can be contained within a vector called, $\mathbf{u}$. In this way it can be written that

$$\mathbf{u} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

Another popular method for expressing such information is to take advantage of three unit vectors, $\mathbf{i}, \mathbf{j}, \mathbf{k}$, which are defined as follows:

$$\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

such that it can also be stated that

$$\mathbf{u} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$$

Note that

$$\begin{aligned} u\mathbf{i} + v\mathbf{j} + w\mathbf{k} &= u \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + v \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + w \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ v \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix} \\ &= \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \mathbf{u} \end{aligned}$$

17.3 Gradient

Consider a scalar quantity, T. Suppose that we are interested in the gradients of T in the x, y and z direction:

$$\frac{\partial T}{\partial x}, \quad \frac{\partial T}{\partial y}, \quad \frac{\partial T}{\partial z}$$

A convenient shorthand for this is the so-called Del operator, ∇, often referred to as the gradient or grad operator. The gradient of a scalar, T, takes the form:

$$\mathrm{grad}(T) = \nabla T = \begin{bmatrix} \frac{\partial T}{\partial x} \\ \\ \frac{\partial T}{\partial y} \\ \\ \frac{\partial T}{\partial z} \end{bmatrix} = \frac{\partial T}{\partial x}\mathbf{i} + \frac{\partial T}{\partial y}\mathbf{j} + \frac{\partial T}{\partial z}\mathbf{k}$$

Challenge 17.1 Fick's first law in Cartesian coordinates takes the form:

$$J_x = -D_E \frac{\partial c}{\partial x}, \quad J_y = -D_E \frac{\partial c}{\partial y}, \quad J_z = -D_E \frac{\partial c}{\partial z}$$

where D_E [ML^{-2}] is the effective diffusion coefficient, c [ML^{-3}] is solute concentration and J_x [$\mathrm{ML}^{-2}\mathrm{T}^{-1}$], J_y [$\mathrm{ML}^{-2}\mathrm{T}^{-1}$] and J_z [$\mathrm{ML}^{-2}\mathrm{T}^{-1}$] are solute fluxes in the x, y and z direction, respectively. Use vector calculus notation to write Fick's first law in a compact form

The above set of equations can be written in a vector form as follows:

$$\begin{bmatrix} J_x \\ \\ J_y \\ \\ J_z \end{bmatrix} = \begin{bmatrix} -D_E \dfrac{\partial c}{\partial x} \\ \\ -D_E \dfrac{\partial c}{\partial y} \\ \\ -D_E \dfrac{\partial c}{\partial z} \end{bmatrix}$$

Let $\mathbf{J} = J_x\mathbf{i} + J_y\mathbf{j} + J_z\mathbf{k}$. Then

$$\mathbf{J} = -D_E \nabla c$$

Gradient of a vector

Now consider a vector of displacement, **u**, defined by

$$\mathbf{u} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$$

where u, v and w denote displacements in the x, y and z direction, respectively. Convention generally dictates that the gradient of **u** would take the form

$$\mathrm{grad}(\mathbf{u}) = \nabla\mathbf{u} = \begin{bmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x} & \dfrac{\partial w}{\partial x} \\ \\ \dfrac{\partial u}{\partial y} & \dfrac{\partial v}{\partial y} & \dfrac{\partial w}{\partial y} \\ \\ \dfrac{\partial u}{\partial z} & \dfrac{\partial v}{\partial z} & \dfrac{\partial w}{\partial z} \end{bmatrix}$$

which is the same as saying

$$\mathrm{grad}(\mathbf{u}) = \nabla\mathbf{u} = \begin{bmatrix} \nabla u & \nabla v & \nabla w \end{bmatrix} \tag{17.1}$$

In terms of unit vectors, it is possible to state that

$$\mathrm{grad}(\mathbf{u}) = \mathrm{grad}(u)\mathbf{i}^{\mathrm{T}} + \mathrm{grad}(v)\mathbf{j}^{\mathrm{T}} + \mathrm{grad}(w)\mathbf{k}^{\mathrm{T}}$$

Challenge 17.2 Use vector calculus notation to write the following set of relationships between strain and displacement in compact form:

$$\varepsilon_{xx} = \frac{\partial u}{\partial x}, \quad \varepsilon_{xy} = \varepsilon_{yx} = \frac{1}{2}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right),$$

$$\varepsilon_{yy} = \frac{\partial v}{\partial y}, \quad \varepsilon_{xz} = \varepsilon_{zx} = \frac{1}{2}\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right),$$

$$\varepsilon_{zz} = \frac{\partial w}{\partial z}, \quad \varepsilon_{yz} = \varepsilon_{zy} = \frac{1}{2}\left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right).$$

The above set of equations can be written in matrix form as follows:

$$\begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{1}{2}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) & \frac{1}{2}\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right) \\ \frac{1}{2}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) & \frac{\partial v}{\partial y} & \frac{1}{2}\left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right) \\ \frac{1}{2}\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right) & \frac{1}{2}\left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right) & \frac{\partial w}{\partial z} \end{bmatrix}$$

Let $\mathbf{u} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$ and

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix}$$

Recall that

$$\nabla \mathbf{u} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} & \frac{\partial w}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y} & \frac{\partial w}{\partial y} \\ \frac{\partial u}{\partial z} & \frac{\partial v}{\partial z} & \frac{\partial w}{\partial z} \end{bmatrix} \tag{17.2}$$

and

$$(\nabla \mathbf{u})^{\mathrm{T}} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{bmatrix}$$

from which it follows that

$$\boldsymbol{\varepsilon} = \frac{1}{2}[\nabla \mathbf{u} + (\nabla \mathbf{u})^{\mathrm{T}}]$$

17.4 Divergence

A related operator is the divergence operator, div. The divergence of the aforementioned vector, **u**, takes the form

$$\mathrm{div}(\mathbf{u}) = \nabla \cdot \mathbf{u} = \nabla^{\mathrm{T}} \mathbf{u} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}$$

from which it can be better understood that

$$\nabla = \begin{bmatrix} \dfrac{\partial}{\partial x} \\ \dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial z} \end{bmatrix}$$

Challenge 17.3 Solute transport in a porous medium can be described by

$$\phi \frac{\partial c}{\partial t} = -\frac{\partial J_x}{\partial x} - \frac{\partial J_y}{\partial y} - \frac{\partial J_z}{\partial z}$$

where ϕ [-] is porosity, c [ML^{-3}] is solute concentration, t [T] is time and J_x [ML^{-2}T^{-1}], J_y [ML^{-2}T^{-1}] and J_z [ML^{-2}T^{-1}] are solute fluxes in the x, y and z direction, respectively. Use vector calculus notation to write the above equation in a compact form.

$$\phi \frac{\partial c}{\partial t} = -\nabla \cdot \mathbf{J}$$

where $\mathbf{J} = J_x\mathbf{i} + J_y\mathbf{j} + J_z\mathbf{k}$.

Divergence of a matrix

Now consider a matrix, $\mathbf{A}$, defined as

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

Convention dictates that

$$\operatorname{div}(\mathbf{A}) = \begin{bmatrix} \frac{\partial A_{11}}{\partial x} + \frac{\partial A_{21}}{\partial y} + \frac{\partial A_{31}}{\partial z} \\ \\ \frac{\partial A_{12}}{\partial x} + \frac{\partial A_{22}}{\partial y} + \frac{\partial A_{32}}{\partial z} \\ \\ \frac{\partial A_{13}}{\partial x} + \frac{\partial A_{23}}{\partial y} + \frac{\partial A_{33}}{\partial z} \end{bmatrix}$$

Now consider three vectors, $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$, defined by

$$\mathbf{a}_1 = \begin{bmatrix} A_{11} \\ A_{21} \\ A_{31} \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} A_{12} \\ A_{22} \\ A_{32} \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} A_{13} \\ A_{23} \\ A_{33} \end{bmatrix}$$

such that it can be said that

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$$

It can be understood that

$$\operatorname{div}(\mathbf{A}) = \operatorname{div}(\mathbf{a}_1)\mathbf{i} + \operatorname{div}(\mathbf{a}_2)\mathbf{j} + \operatorname{div}(\mathbf{a}_3)\mathbf{k} \qquad (17.3)$$

17.5 Laplacian

Consider the divergence of the gradient of a scalar, T:

$$\operatorname{div}(\operatorname{grad}(T)) = \nabla \cdot \nabla T = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = \nabla^2 T \qquad (17.4)$$

where the ∇^2 is often referred to as the Laplacian.

The name, Laplacian, comes from Laplace's equation for steady state heat conduction:

$$\nabla^2 T = 0 \tag{17.5}$$

Challenge 17.4 Use vector calculus notation to write the following heat conduction equation in a compact form.

$$\rho c_p \frac{\partial T}{\partial t} - \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) - \frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right) - \frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right) = 0$$

where ρ [ML^3] is density, c_p [L^{2}T$^{-2}\Theta^{-1}$] is constant-pressure specific heat capacity, T [Θ] is temperature, t [T] is time and k [MLT$^{-3}\Theta^{-1}$] is thermal conductivity.

$$\rho c_p \frac{\partial T}{\partial t} - \nabla \cdot (k \nabla T) = 0$$

Challenge 17.5 Write a simpler form of your results to the previous challenge that is suitable if thermal conductivity is uniform.

$$\rho c_p \frac{\partial T}{\partial t} - k\nabla^2 T = 0$$

Laplacian of a vector

Now consider the divergence of the gradient of the aforementioned vector, **u**. Substituting Eq. (17.1) into Eq. (17.3) leads to

$$\operatorname{div}(\operatorname{grad}(\mathbf{u})) = \operatorname{div}(\operatorname{grad}(u))\mathbf{i} + \operatorname{div}(\operatorname{grad}(v))\mathbf{j} + \operatorname{div}(\operatorname{grad}(w))\mathbf{k}$$

and from Eq. (17.4)

$$\operatorname{div}(\operatorname{grad}(\mathbf{u})) = (\nabla^2 u)\mathbf{i} + (\nabla^2 v)\mathbf{j} + (\nabla^2 w)\mathbf{k}$$

from which it follows that

$$\operatorname{div}(\operatorname{grad}(\mathbf{u})) = \nabla \cdot \nabla \mathbf{u} = \nabla^2 \mathbf{u} \tag{17.6}$$

such that it can be seen that the divergence of the gradient of a vector or a scalar gives rise to the Laplacian of a vector or a scalar, respectively.

17.6 Problem sheet

Problem 17.1 (see Worked Solution 17.1)

Find the gradient of the following functions:

a) $f = xy$

b) $f = x/y$

c) $f = e^x \sin y$

d) $f = \dfrac{1}{2}\ln(x^2 + y^2)$

e) $f = (x^2 + y^2 + z^2)^{-1/2}$

f) $f = \dfrac{z}{x^2 + y^2}$

Problem 17.2 (see Worked Solution 17.2)

Find the divergence of the following functions:

a) $\mathbf{f} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$

b) $\mathbf{f} = yz\mathbf{i} + zx\mathbf{j} + xy\mathbf{k}$

c) $\mathbf{f} = y^2 e^z\mathbf{i} + x^2 z^2\mathbf{k}$

d) $\mathbf{f} = xyz(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})$

e) $\mathbf{f} = \cos x \cosh y\mathbf{i} + \sin x \sinh y\mathbf{j}$

f) $\mathbf{f} = e^{-xy}\mathbf{i} + e^{-yz}\mathbf{j} + e^{-zx}\mathbf{k}$

Problem 17.3 (see Worked Solution 17.3)

Find the Laplacian of the following functions:

a) $f = x^2 + 4y^2$

b) $f = \dfrac{x - y}{x + y}$

c) $f = e^{xyz}$

d) $f = xz/y$

Problem 17.4 (see Worked Solution 17.4)

Show that

a) $\nabla(fg) = f\nabla g + g\nabla f$

b) $\nabla(f^n) = nf^{n-1}\nabla f$

c) $\nabla(f/g) = \dfrac{1}{g^2}(g\nabla f - f\nabla g)$

d) $\nabla^2(fg) = g\nabla^2 f + 2\nabla f \cdot \nabla g + f\nabla^2 g$

e) $\mathrm{div}(f\mathbf{v}) = f\mathrm{div}(\mathbf{v}) + \mathbf{v} \cdot \nabla f$

f) $\mathrm{div}(f\nabla g) = f\nabla^2 g + \nabla g \cdot \nabla f$

g) $\mathrm{div}(f\nabla g) - \mathrm{div}(g\nabla f) = f\nabla^2 g - g\nabla^2 f$

17.7 Worked solutions

Worked Solution 17.1 (see Problem 17.1)

Recall that $\nabla f = \dfrac{\partial f}{\partial x}\mathbf{i} + \dfrac{\partial f}{\partial y}\mathbf{j} + \dfrac{\partial f}{\partial z}\mathbf{k}$

a) $f = xy$ therefore $\nabla f = y\mathbf{i} + x\mathbf{j} = \begin{bmatrix} y \\ x \\ 0 \end{bmatrix}$

b) $f = x/y$ therefore

$$\nabla f = y^{-1}\mathbf{i} - xy^{-2}\mathbf{j} = y^{-2}(y\mathbf{i} - x\mathbf{j})$$

c) $f = e^x \sin y$ therefore

$$\nabla f = e^x \sin y\mathbf{i} + e^x \cos y\mathbf{j} = e^x(\sin y\mathbf{i} + \cos y\mathbf{j})$$

d) $f = \frac{1}{2}\ln(x^2 + y^2)$ therefore

$$\nabla f = \frac{x}{x^2 + y^2}\mathbf{i} + \frac{y}{x^2 + y^2}\mathbf{j} = \frac{1}{x^2 + y^2}(x\mathbf{i} + y\mathbf{j})$$

e) $f = (x^2 + y^2 + z^2)^{-1/2}$ therefore

$$\begin{aligned} \nabla f &= -\frac{x}{(x^2 + y^2 + z^2)^{3/2}}\mathbf{i} - \frac{y}{(x^2 + y^2 + z^2)^{3/2}}\mathbf{j} \\ &\quad - \frac{z}{(x^2 + y^2 + z^2)^{3/2}}\mathbf{k} = -\frac{(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})}{(x^2 + y^2 + z^2)^{3/2}} \end{aligned}$$

f) $f = \frac{z}{x^2 + y^2}$ therefore

$$\nabla f = -\frac{2xz}{(x^2 + y^2)^2}\mathbf{i} - \frac{2yz}{(x^2 + y^2)^2}\mathbf{j} + \frac{1}{x^2 + y^2}\mathbf{k}$$

Worked Solution 17.2 (see Problem 17.2)

Recall that $\operatorname{div}(\mathbf{f}) \equiv \nabla \cdot \mathbf{f} = \dfrac{\partial f_1}{\partial x} + \dfrac{\partial f_2}{\partial y} + \dfrac{\partial f_3}{\partial z}$

a) $\mathbf{f} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ therefore $\operatorname{div}(\mathbf{f}) = 1 + 1 + 1 = 3$

b) $\mathbf{f} = yz\mathbf{i} + zx\mathbf{j} + xy\mathbf{k}$ therefore $\operatorname{div}(\mathbf{f}) = 0 + 0 + 0 = 0$

c) $\mathbf{f} = y^2 e^z\mathbf{i} + x^2 z^2\mathbf{k}$ therefore $\operatorname{div}(\mathbf{f}) = 0 + 2x^2 z = 2x^2 z$

d) $\mathbf{f} = xyz(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})$ therefore

$\operatorname{div}(\mathbf{f}) = 2xyz + 2xyz + 2xyz = 6xyz$

e) $\mathbf{f} = \cos x \cosh y\mathbf{i} + \sin x \sinh y\mathbf{j}$ therefore

$\operatorname{div}(\mathbf{f}) = -\sin x \cosh y + \sin x \cosh y = 0$

f) $\mathbf{f} = e^{-xy}\mathbf{i} + e^{-yz}\mathbf{j} + e^{-zx}\mathbf{k}$ therefore

$\operatorname{div}(\mathbf{f}) = -ye^{-xy} - ze^{-yz} - xe^{-zx}$

Worked Solution 17.3 (see Problem 17.3)

Recall that

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

a) $f = x^2 + 4y^2$

$\nabla^2 f = 2 + 8 = 10$

b) $f = \dfrac{x-y}{x+y}$

$$\ln f = \ln(x-y) - \ln(x+y)$$

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

$$\frac{\partial^2 f}{\partial x^2} = \frac{-4y}{(x+y)^3}$$

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

$$\frac{\partial^2 f}{\partial y^2} = \frac{4x}{(x+y)^3}$$

$$\nabla^2 f = \frac{4(x-y)}{(x+y)^3}$$

c) $$f = e^{xyz}$$

$$\frac{\partial f}{\partial x} = yze^{xyz}, \quad \frac{\partial^2 f}{\partial x^2} = (yz)^2 e^{xyz}$$

$$\frac{\partial^2 f}{\partial y^2} = (xz)^2 e^{xyz}, \quad \frac{\partial^2 f}{\partial z^2} = (xy)^2 e^{xyz}$$

$$\nabla^2 f = \left[(yz)^2 + (xz)^2 + (xy)^2\right] e^{xyz}$$

d) $$f = xz/y$$

$$\ln f = \ln x + \ln z - \ln y$$

$$\frac{\partial f}{\partial x} = \frac{xz}{y}\frac{1}{x} = \frac{z}{y}, \quad \frac{\partial^2 f}{\partial x^2} = 0$$

$$\frac{\partial f}{\partial y} = -\frac{xz}{y}\frac{1}{y} = \frac{xz}{y^2}, \quad \frac{\partial^2 f}{\partial y^2} = \frac{2xz}{y^3}$$

$$\frac{\partial f}{\partial z} = \frac{xz}{y}\frac{1}{z} = \frac{x}{y}, \quad \frac{\partial^2 f}{\partial z^2} = 0$$

$$\nabla^2 f = \frac{2xz}{y^3}$$

Worked Solution 17.4 (see Problem 17.4)

a) $\nabla(fg)$

$$
\begin{aligned}
\nabla(fg) &= \frac{\partial(fg)}{\partial x}\mathbf{i}+\frac{\partial(fg)}{\partial y}\mathbf{j}+\frac{\partial(fg)}{\partial z}\mathbf{k}\\
&= \left(f\frac{\partial g}{\partial x}+g\frac{\partial f}{\partial x}\right)\mathbf{i}+\left(f\frac{\partial g}{\partial y}+g\frac{\partial f}{\partial y}\right)\mathbf{j}\\
&+\left(f\frac{\partial g}{\partial z}+g\frac{\partial f}{\partial z}\right)\mathbf{k}\\
&= f\left(\frac{\partial g}{\partial x}\mathbf{i}+\frac{\partial g}{\partial y}\mathbf{j}+\frac{\partial g}{\partial z}\mathbf{k}\right)+g\left(\frac{\partial f}{\partial x}\mathbf{i}+\frac{\partial f}{\partial y}\mathbf{j}+\frac{\partial f}{\partial z}\mathbf{k}\right)\\
&= f\nabla g+g\nabla f
\end{aligned}
$$

b) $\nabla(f^n)$

$$
\nabla(f^n) = \frac{\partial(f^n)}{\partial x}\mathbf{i}+\frac{\partial(f^n)}{\partial y}\mathbf{j}+\frac{\partial(f^n)}{\partial z}\mathbf{k}
$$

taking $u=f^n$, $\quad \dfrac{\partial u}{\partial f}=nf^{n-1}$

so it can be seen $\quad du=\dfrac{\partial u}{\partial f}df=nf^{n-1}df$

$$
\therefore \nabla(f^n) = nf^{n-1}\left(\frac{\partial f}{\partial x}\mathbf{i}+\frac{\partial f}{\partial y}\mathbf{j}+\frac{\partial f}{\partial z}\mathbf{k}\right)=nf^{n-1}\nabla f
$$

c) $$\nabla(f/g) = \frac{\partial(f/g)}{\partial x}\mathbf{i} + \frac{\partial(f/g)}{\partial y}\mathbf{j} + \frac{\partial(f/g)}{\partial z}\mathbf{k}$$

$$= \frac{1}{g^2}\left(g\frac{\partial f}{\partial x} - f\frac{\partial g}{\partial x}\right)\mathbf{i} + \frac{1}{g^2}\left(g\frac{\partial f}{\partial y} - f\frac{\partial g}{\partial y}\right)\mathbf{j}$$

$$+\frac{1}{g^2}\left(g\frac{\partial f}{\partial z} - f\frac{\partial g}{\partial z}\right)\mathbf{k}$$

$$= \frac{1}{g}\left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}\right) - \frac{f}{g^2}\left(\frac{\partial g}{\partial x}\mathbf{i} + \frac{\partial g}{\partial y}\mathbf{j} + \frac{\partial g}{\partial z}\mathbf{k}\right)$$

$$= \frac{1}{g^2}\left(g\nabla f - f\nabla g\right)$$

d) $$\nabla^2(fg) = \nabla\cdot\nabla(fg) = \nabla\cdot(f\nabla g + g\nabla f)$$

$$\nabla\cdot(f\nabla g) = \nabla\cdot\left[f\left(\frac{\partial g}{\partial x}\mathbf{i} + \frac{\partial g}{\partial y}\mathbf{j} + \frac{\partial g}{\partial z}\mathbf{k}\right)\right]$$

$$= \frac{\partial f}{\partial x}\frac{\partial g}{\partial x} + f\frac{\partial^2 g}{\partial x^2} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial y} + f\frac{\partial^2 g}{\partial y^2} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial z} + f\frac{\partial^2 g}{\partial z^2}$$

$$= \frac{\partial f}{\partial x}\frac{\partial g}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial z} + f\nabla^2 g$$

$$\nabla\cdot(g\nabla f) = \frac{\partial f}{\partial x}\frac{\partial g}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial z} + g\nabla^2 f$$

$$\nabla^2(fg) = g\nabla^2 f + 2\left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial z}\right) + f\nabla^2 g$$

$$= g\nabla^2 f + 2\nabla f\cdot\nabla g + f\nabla^2 g$$

e) $$\operatorname{div}(f\mathbf{v}) = \frac{\partial(fv_1)}{\partial x} + \frac{\partial(fv_2)}{\partial y} + \frac{\partial(fv_3)}{\partial z}$$

$$= f\frac{\partial v_1}{\partial x} + v_1\frac{\partial f}{\partial x} + f\frac{\partial v_2}{\partial y} + v_2\frac{\partial f}{\partial y} + f\frac{\partial v_3}{\partial z} + v_3\frac{\partial f}{\partial z}$$

$$= f\left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}\right) + v_1\frac{\partial f}{\partial x} + v_2\frac{\partial f}{\partial y} + v_3\frac{\partial f}{\partial z}$$

$$= f\operatorname{div}(\mathbf{v}) + \mathbf{v}\cdot\nabla f$$

f) $$\operatorname{div}(f\nabla g) = f\operatorname{div}(\nabla g) + \nabla g\cdot\nabla f$$

Recall that $\operatorname{div}(\nabla g) = \nabla^2 g$

$$\therefore \operatorname{div}(f\nabla g) = f\nabla^2 g + \nabla g\cdot\nabla f$$

g) Note that it can also be said that

$$\operatorname{div}(g\nabla f) = g\nabla^2 f + \nabla g\cdot\nabla f$$

$$\therefore \operatorname{div}(f\nabla g) - \operatorname{div}(g\nabla f) = f\nabla^2 g - g\nabla^2 f$$

18

Vector calculus with stress and strain

18.1 Learning outcomes

You should be able to:

- Derive the Cauchy momentum equation (CME) by considering the stress equilibrium of a control-volume.
- Derive the P-wave equation by taking the divergence of the CME.
- Derive the S-wave equation by taking the curl of the CME.
- Show that D'Alembert's formula is a solution of the one-dimensional wave equation.

18.2 The Cauchy momentum equation

Challenge 18.1 Consider a control-volume of dimensions, δx [L], δy [L] and δz [L], in the x, y and z directions, respectively. A corner of the control-volume (CV) is located at a reference point located at (x, y, z) and the corner of the CV furthest away from the reference point is located at $(x+\delta x, y+\delta y, z+\delta z)$. Let τ_{xx}, τ_{yx} and τ_{zx} be the stresses acting in the x direction on the CV faces that share the corner at (x, y, z) and are normal to the x, y and z directions, respectively. The stresses acting on the associated opposite faces are $\tau_{xx}+\delta\tau_{xx}$, $\tau_{yx}+\delta\tau_{yx}$ and $\tau_{zx}+\delta\tau_{zx}$, respectively.

In addition to these stresses, the CV is also subject to the gravitational force associated with its own weight, $\delta x\delta y\delta z\rho g_x$, where ρ [ML^{-3}] is the density of the rock and g_x [LT^{-2}] is the acceleration due to gravity in the x direction. This additional force is often referred to as a body force.

Determine an expression for the net force in the x-direction, F_x [MLT^{-2}].

$$\begin{aligned}
F_x &= -\delta y\delta z\tau_{xx} - \delta x\delta z\tau_{yx} - \delta x\delta y\tau_{zx} + \delta y\delta z(\tau_{xx}+\delta\tau_{xx}) \\
&\quad +\delta x\delta z(\tau_{yx}+\delta\tau_{yx}) + \delta x\delta y(\tau_{zx}+\delta\tau_{zx}) + \delta x\delta y\delta z\rho g_x \\
&= \delta y\delta z\delta\tau_{xx} + \delta x\delta z\delta\tau_{yx} + \delta x\delta y\delta\tau_{zx} + \delta x\delta y\delta z\rho g_x \\
&= \delta x\delta y\delta z\left(\frac{\delta\tau_{xx}}{\delta x} + \frac{\delta\tau_{yx}}{\delta y} + \frac{\delta\tau_{zx}}{\delta z} + \rho g_x\right)
\end{aligned}$$

Challenge 18.2 Determine similar expressions for the net force in the y and z directions, F_y and F_z, respectively.

$$F_y = \delta x\delta y\delta z\left(\frac{\delta\tau_{xy}}{\delta x} + \frac{\delta\tau_{yy}}{\delta y} + \frac{\delta\tau_{zy}}{\delta z} + \rho g_y\right)$$

$$F_z = \delta x \delta y \delta z \left(\frac{\delta \tau_{xz}}{\delta x} + \frac{\delta \tau_{yz}}{\delta y} + \frac{\delta \tau_{zz}}{\delta z} + \rho g_z \right)$$

Challenge 18.3 Given Newton's second law it can be said that

$$\mathbf{F} = \delta x \delta y \delta z \rho \mathbf{a}$$

where

$$\mathbf{F} = F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k}$$

and $\mathbf{a}$ [LT^{-2}] is the acceleration vector, defined by

$$\mathbf{a} = a_x \mathbf{i} + a_y \mathbf{j} + a_z \mathbf{k}$$

Using the results from previous two challenges, derive an expression for $\mathbf{a}$ in terms of $\boldsymbol{\tau}$ and $\mathbf{g}$ where

$$\mathbf{g} = g_x \mathbf{i} + g_y \mathbf{j} + g_z \mathbf{k}$$

The first step is to equate the expressions of net force in terms of acceleration with those in terms of stress:

$$\delta x \delta y \delta z \rho a_x = \delta x \delta y \delta z \left(\frac{\delta \tau_{xx}}{\delta x} + \frac{\delta \tau_{yx}}{\delta y} + \frac{\delta \tau_{zx}}{\delta z} + \rho g_x \right)$$

$$\delta x \delta y \delta z \rho a_y = \delta x \delta y \delta z \left(\frac{\delta \tau_{xy}}{\delta x} + \frac{\delta \tau_{yy}}{\delta y} + \frac{\delta \tau_{zy}}{\delta z} + \rho g_y \right)$$

$$\delta x \delta y \delta z \rho a_z = \delta x \delta y \delta z \left(\frac{\delta \tau_{xz}}{\delta x} + \frac{\delta \tau_{yz}}{\delta y} + \frac{\delta \tau_{zz}}{\delta z} + \rho g_z \right)$$

These can be simplified to get:

$$\rho a_x = \frac{\delta \tau_{xx}}{\delta x} + \frac{\delta \tau_{yx}}{\delta y} + \frac{\delta \tau_{zx}}{\delta z} + \rho g_x$$

$$\rho a_y = \frac{\delta \tau_{xy}}{\delta x} + \frac{\delta \tau_{yy}}{\delta y} + \frac{\delta \tau_{zy}}{\delta z} + \rho g_y$$

$$\rho a_z = \frac{\delta\tau_{xz}}{\delta x} + \frac{\delta\tau_{yz}}{\delta y} + \frac{\delta\tau_{zz}}{\delta z} + \rho g_z$$

At this stage it would be useful to exploit vector calculus notation so that we can introduce the stress tensor, $\boldsymbol{\tau}$. To do this we need to make the control-volume infinitesimally small such that $\delta x \to 0$, $\delta y \to 0$ and $\delta z \to 0$. This is useful because

$$\lim_{\delta i \to 0} \frac{\delta\tau_{ij}}{\delta i} = \frac{\partial\tau_{ij}}{\partial i}$$

where $i = x, y, z$ and $j = x, y, z$.

In this way, it can be understood that:

$$\rho a_x = \frac{\partial\tau_{xx}}{\partial x} + \frac{\partial\tau_{yx}}{\partial y} + \frac{\partial\tau_{zx}}{\partial z} + \rho g_x$$

$$\rho a_y = \frac{\partial\tau_{xy}}{\partial x} + \frac{\partial\tau_{yy}}{\partial y} + \frac{\partial\tau_{zy}}{\partial z} + \rho g_y$$

$$\rho a_z = \frac{\partial\tau_{xz}}{\partial x} + \frac{\partial\tau_{yz}}{\partial y} + \frac{\partial\tau_{zz}}{\partial z} + \rho g_z$$

Recall that

$$\nabla \cdot \boldsymbol{\tau} = \begin{bmatrix} \frac{\partial\tau_{xx}}{\partial x} + \frac{\partial\tau_{yx}}{\partial y} + \frac{\partial\tau_{zx}}{\partial z} \\ \\ \frac{\partial\tau_{xy}}{\partial x} + \frac{\partial\tau_{yy}}{\partial y} + \frac{\partial\tau_{zy}}{\partial z} \\ \\ \frac{\partial\tau_{xz}}{\partial x} + \frac{\partial\tau_{yz}}{\partial y} + \frac{\partial\tau_{zz}}{\partial z} \end{bmatrix}$$

from which it follows that

$$\rho \mathbf{a} = \nabla \cdot \boldsymbol{\tau} + \rho \mathbf{g} \tag{18.1}$$

Eq. (18.1) is a form of the Cauchy momentum equation.

Challenge 18.4 From Hooke's law we have

$$\boldsymbol{\tau} = 2G\boldsymbol{\varepsilon} + \lambda \text{trace}(\boldsymbol{\varepsilon})\mathbf{I} \tag{18.2}$$

where G [$\text{ML}^{-1}\text{T}^{-2}$] is the shear modulus, λ [$\text{ML}^{-1}\text{T}^{-2}$] is the Lamé parameter and $\boldsymbol{\varepsilon}$ [-] is a strain tensor found from

$$\boldsymbol{\varepsilon} = \frac{1}{2}\left[\nabla\mathbf{u} + (\nabla\mathbf{u})^{\text{T}}\right] \tag{18.3}$$

and $\mathbf{u} = u\mathbf{i} + v\mathbf{i} + w\mathbf{i}$ [L] is a displacement vector.

Given that

$$\mathbf{a} = \frac{\partial^2\mathbf{u}}{\partial t^2} \tag{18.4}$$

rewrite the result to Challenge 1c with $\mathbf{u}$ as the only dependent variable and simplify as much as possible.

Substituting Eqs. (18.2) and (18.4) into Eq. (18.1) leads to

$$\rho\frac{\partial^2\mathbf{u}}{\partial t^2} = \nabla\cdot[2G\boldsymbol{\varepsilon} + \lambda\text{trace}(\boldsymbol{\varepsilon})\mathbf{I}] + \rho\mathbf{g} \tag{18.5}$$

Substituting Eq. (18.3) into Eq. (18.5) then leads to

$$\rho\frac{\partial^2\mathbf{u}}{\partial t^2} = \nabla\cdot\left[G\left[\nabla\mathbf{u} + (\nabla\mathbf{u})^{\text{T}}\right] + \frac{\lambda}{2}\text{trace}\left[\nabla\mathbf{u} + (\nabla\mathbf{u})^{\text{T}}\right]\mathbf{I}\right] + \rho\mathbf{g}$$

Noting that $\text{trace}(\nabla\mathbf{u}) = \text{trace}[(\nabla\mathbf{u})^{\text{T}}]$ and $\nabla\cdot(\nabla\mathbf{u}) = \nabla^2\mathbf{u}$:

$$\rho\frac{\partial^2\mathbf{u}}{\partial t^2} = G\nabla\cdot(\nabla\mathbf{u}) + G\nabla\cdot\left[(\nabla\mathbf{u})^{\text{T}}\right] + \lambda\nabla\cdot[\text{trace}(\nabla\mathbf{u})\mathbf{I}] + \rho\mathbf{g}$$

The above equation can be simplified further by consideration of the following:

$$\nabla\cdot\left[(\nabla\mathbf{u})^{\mathrm{T}}\right]=\begin{bmatrix}\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\partial v}{\partial x}\right)+\frac{\partial}{\partial z}\left(\frac{\partial w}{\partial x}\right)\\ \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right)+\frac{\partial}{\partial y}\left(\frac{\partial v}{\partial y}\right)+\frac{\partial}{\partial z}\left(\frac{\partial w}{\partial y}\right)\\ \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial z}\right)+\frac{\partial}{\partial y}\left(\frac{\partial v}{\partial z}\right)+\frac{\partial}{\partial z}\left(\frac{\partial w}{\partial z}\right)\end{bmatrix}$$

$$\nabla\cdot[\mathrm{trace}(\nabla\mathbf{u})\mathbf{I}]=\begin{bmatrix}\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)\\ \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)\\ \frac{\partial}{\partial z}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)\end{bmatrix}$$

It therefore follows that

$$\nabla\cdot\left[(\nabla\mathbf{u})^{\mathrm{T}}\right]=\nabla\cdot[\mathrm{trace}(\nabla\mathbf{u})\mathbf{I}]=\nabla(\nabla\cdot\mathbf{u})$$

from which it can be said that

$$\rho\frac{\partial^2\mathbf{u}}{\partial t^2}=G\nabla^2\mathbf{u}+(G+\lambda)\nabla(\nabla\cdot\mathbf{u})+\rho\mathbf{g} \qquad (18.6)$$

18.3 Wave equations

Challenge 18.5 Take the divergence of both sides of Eq. (18.6) and substitute the P-wave potential, Φ [-], defined by

$$\Phi=\nabla\cdot\mathbf{u}$$

Taking the divergence of both sides of Eq. (18.6) leads to

$$\rho\nabla\cdot\left(\frac{\partial^2\mathbf{u}}{\partial t^2}\right) = G\nabla\cdot\left(\nabla^2\mathbf{u}\right) + (G+\lambda)\nabla\cdot\left[\nabla\left(\nabla\cdot\mathbf{u}\right)\right]$$

Substituting the P-wave potential identity then leads to

$$\rho\frac{\partial^2\Phi}{\partial t^2} = G\nabla\cdot\left(\nabla^2\mathbf{u}\right) + (G+\lambda)\nabla\cdot(\nabla\Phi)$$

Recall that $\nabla\cdot(\nabla\Phi) = \nabla^2\Phi$ and therefore

$$\rho\frac{\partial^2\Phi}{\partial t^2} = G\nabla\cdot\left(\nabla^2\mathbf{u}\right) + (G+\lambda)\nabla^2\Phi$$

Now we need to consider the $\nabla\cdot\left(\nabla^2\mathbf{u}\right)$ term:

$$\nabla^2\mathbf{u} = \begin{bmatrix} \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2} \\[2ex] \dfrac{\partial^2 v}{\partial x^2} + \dfrac{\partial^2 v}{\partial y^2} + \dfrac{\partial^2 v}{\partial z^2} \\[2ex] \dfrac{\partial^2 w}{\partial x^2} + \dfrac{\partial^2 w}{\partial y^2} + \dfrac{\partial^2 w}{\partial z^2} \end{bmatrix}$$

$$\begin{aligned}
\nabla\cdot\left(\nabla^2\mathbf{u}\right) &= \frac{\partial}{\partial x}\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\right)\\
&\quad+\frac{\partial}{\partial y}\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial z^2}\right)\\
&\quad+\frac{\partial}{\partial z}\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}+\frac{\partial^2 w}{\partial z^2}\right)\\
&= \frac{\partial^2}{\partial x^2}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)\\
&\quad+\frac{\partial^2}{\partial y^2}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)\\
&\quad+\frac{\partial^2}{\partial z^2}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)\\
&= \nabla^2(\nabla\cdot\mathbf{u}) = \nabla^2\Phi
\end{aligned}$$

and therefore

$$\rho\frac{\partial^2\Phi}{\partial t^2} = (2G+\lambda)\nabla^2\Phi$$

which is a wave equation with a wave velocity of $\sqrt{\rho^{-1}(2G+\lambda)}$.

Challenge 18.6 The curl operator, $\nabla \times \mathbf{u}$, is defined by

$$\nabla \times \mathbf{u} = \begin{bmatrix} \left(\dfrac{\partial w}{\partial y} - \dfrac{\partial v}{\partial z}\right) \\ \\ \left(\dfrac{\partial u}{\partial z} - \dfrac{\partial w}{\partial x}\right) \\ \\ \left(\dfrac{\partial v}{\partial x} - \dfrac{\partial u}{\partial y}\right) \end{bmatrix}$$

Take the curl of both sides of Eq. (18.6) and substitute the S-wave potential, $\boldsymbol{\Psi}$ [-], defined by

$$\boldsymbol{\Psi} = \nabla \times \mathbf{u}$$

Note that:

$$\nabla^2 \mathbf{u} = \begin{bmatrix} \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2} \\ \\ \dfrac{\partial^2 v}{\partial x^2} + \dfrac{\partial^2 v}{\partial y^2} + \dfrac{\partial^2 v}{\partial z^2} \\ \\ \dfrac{\partial^2 w}{\partial x^2} + \dfrac{\partial^2 w}{\partial y^2} + \dfrac{\partial^2 w}{\partial z^2} \end{bmatrix}$$

and

$$\nabla(\nabla \cdot \mathbf{u}) = \begin{bmatrix} \dfrac{\partial}{\partial x}\left(\dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} + \dfrac{\partial w}{\partial z}\right) \\ \\ \dfrac{\partial}{\partial y}\left(\dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} + \dfrac{\partial w}{\partial z}\right) \\ \\ \dfrac{\partial}{\partial z}\left(\dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} + \dfrac{\partial w}{\partial z}\right) \end{bmatrix}$$

from which it can be understood that

$$\nabla \times \nabla^2 \mathbf{u} = \nabla^2 (\nabla \times \mathbf{u})$$

and

$$\nabla \times \nabla (\nabla \cdot \mathbf{u}) = 0$$

Therefore, taking the curl of both sides of Eq. (18.6) leads to

$$\rho \frac{\partial^2 \mathbf{\Psi}}{\partial t^2} = G \nabla^2 \mathbf{\Psi}$$

which is also a wave equation but with a wave velocity of $\sqrt{\rho^{-1} G}$.

18.4 D'Alembert's formula

Challenge 18.7 Consider the Cauchy momentum equation

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = \nabla \cdot \boldsymbol{\tau} + \rho \mathbf{g} \tag{18.7}$$

where ρ [ML^{-3}] is density, $\mathbf{u}$ [L] is a displacement vector, $\boldsymbol{\tau}$ [ML^{-1}T^{-2}] is a stress tensor, and $\mathbf{g}$ [LT^{-2}] is the gravitational acceleration vector. Derive a wave equation for $\partial^2 u / \partial t^2$ when x, y and z are the principal stress directions and strain is zero in the y and z direction.

Helpful auxiliary equations include:

$$\boldsymbol{\tau} = 2G\boldsymbol{\varepsilon} + \lambda \operatorname{trace}(\boldsymbol{\varepsilon}) \mathbf{I} \tag{18.8}$$

and

$$\boldsymbol{\varepsilon} = \frac{1}{2} [\nabla \mathbf{u} + (\nabla \mathbf{u})^{\mathrm{T}}] \tag{18.9}$$

where G [ML^{-1}T^{-2}] is the shear modulus and λ [ML^{-1}T^{-2}] is the Lamé parameter.

Substituting Eq. (18.8) into Eq. (18.7) leads to

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = 2G\text{div}\boldsymbol{\varepsilon} + \lambda\text{div}(\text{trace}(\boldsymbol{\varepsilon})\mathbf{I}) + \rho\mathbf{g}$$

Because x, y and z are the principal stress directions

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{xx} & 0 & 0 \\ 0 & \varepsilon_{yy} & 0 \\ 0 & 0 & \varepsilon_{zz} \end{bmatrix}$$

Because strain is zero in the y and z directions

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{xx} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

It follows that

$$\text{div}\boldsymbol{\varepsilon} = \frac{\partial \varepsilon_{xx}}{\partial x}\mathbf{i}$$

and

$$\text{trace}(\boldsymbol{\varepsilon})\mathbf{I} = \begin{bmatrix} \varepsilon_{xx} & 0 & 0 \\ 0 & \varepsilon_{xx} & 0 \\ 0 & 0 & \varepsilon_{xx} \end{bmatrix}$$

and hence

$$\text{div}(\text{trace}(\boldsymbol{\varepsilon})\mathbf{I}) = \frac{\partial \varepsilon_{xx}}{\partial x}\mathbf{i} + \frac{\partial \varepsilon_{xx}}{\partial y}\mathbf{j} + \frac{\partial \varepsilon_{xx}}{\partial z}\mathbf{k}$$

Furthermore, because $\mathbf{g} = g\mathbf{k}$, it follows that

$$\rho \frac{\partial^2 u}{\partial t^2} = (2G + \lambda)\frac{\partial \varepsilon_{xx}}{\partial x}$$

and because $\varepsilon_{xx} = \partial u/\partial x$:

$$\rho \frac{\partial^2 u}{\partial t^2} = (2G + \lambda)\frac{\partial^2 u}{\partial x^2} \tag{18.10}$$

Challenge 18.8 Show that the wave equation, derived from the previous challenge, is satisfied by D'Alembert's formula:

$$u(x,t) = f(x-ct) + g(x+ct)$$

where $c^2 = (2G+\lambda)/\rho$.

Substituting $c^2 = (2G+\lambda)/\rho$ into Eq. (18.10) leads to

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \tag{18.11}$$

Now consider the proposed solution

$$u(x,t) = f(x-ct) + g(x+ct) \tag{18.12}$$

To verify that this is a solution to the wave equation we need to partially differentiate the proposed solution and substitute it back into the original partial differential equation (PDE).

To this end, let $s_1 = x - ct$ and $s_2 = x + ct$ such that

$$u(x,t) = f(s_1) + g(s_2)$$

Now consider the first derivative with respect to x

$$\frac{\partial u}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x} = \frac{\partial s_1}{\partial x}\frac{\partial f}{\partial s_1} + \frac{\partial s_2}{\partial x}\frac{\partial g}{\partial s_2} = \frac{\partial f}{\partial s_1} + \frac{\partial g}{\partial s_2}$$

and now the second derivative with respect to x

$$\begin{aligned}\frac{\partial^2 u}{\partial x^2} &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 g}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial s_1}\right) + \frac{\partial}{\partial x}\left(\frac{\partial g}{\partial s_2}\right) \\ &= \frac{\partial s_1}{\partial x}\frac{\partial^2 f}{\partial s_1^2} + \frac{\partial s_2}{\partial x}\frac{\partial^2 g}{\partial s_2^2} = \frac{\partial^2 f}{\partial s_1^2} + \frac{\partial^2 g}{\partial s_2^2}\end{aligned} \tag{18.13}$$

In the same way, it can be seen that

$$\frac{\partial u}{\partial t} = \frac{\partial f}{\partial t} + \frac{\partial g}{\partial t} = \frac{\partial s_1}{\partial t}\frac{\partial f}{\partial s_1} + \frac{\partial s_2}{\partial t}\frac{\partial g}{\partial s_2} = -c\frac{\partial f}{\partial s_1} + c\frac{\partial g}{\partial s_2}$$

and

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} &= \frac{\partial^2 f}{\partial t^2} + \frac{\partial^2 g}{\partial t^2} = \frac{\partial}{\partial t}\left(-c\frac{\partial f}{\partial s_1}\right) + \frac{\partial}{\partial t}\left(c\frac{\partial g}{\partial s_2}\right) \\ &= -\frac{\partial s_1}{\partial t}c\frac{\partial^2 f}{\partial s_1^2} + \frac{\partial s_2}{\partial t}c\frac{\partial^2 g}{\partial s_2^2} = c^2\left(\frac{\partial^2 f}{\partial s_1^2} + \frac{\partial^2 g}{\partial s_2^2}\right) \end{aligned} \tag{18.14}$$

Substituting Eqs. (18.13) and (18.14) into Eq. (18.11) leads to

$$c^2\left(\frac{\partial^2 f}{\partial s_1^2} + \frac{\partial^2 g}{\partial s_2^2}\right) - c^2\left(\frac{\partial^2 f}{\partial s_1^2} + \frac{\partial^2 g}{\partial s_2^2}\right) = 0$$

which confirms that Eq. (18.12) is indeed a solution to Eq. (18.10), and therefore the full wave equation when x, y, z represent the principal stress axes and strain only occurs in the x direction.