H2 Mathematics

Ryan Joo

January 28, 2024

Abstract

This book is written with the intention to provide readers with a brief summary of each topic in the Singapore GCE A-Level H2 Mathematics, as well as some topics covered under H2 Further Mathematics.

Some typical examples are included after each section to illustrate how concepts are usually applied in problems. Challenging sample problems are provided at the end of each topic for the reader to try out.

Graphing Calculator

The use of a graphing calculator is crucial in H2 Mathematics (and H2 Further Mathematics), not only because of its ability to solve equations and plot graphs that are rather challenging to do by hand, but also to save time – since completing the paper is a vital stepping stone in guaranteeing high marks and thus better grades.

Here are some combinations of keys that you might find helpful.

General

- Reset: 2 8 ON
- Switch on: ON
- Switch off: OFF
- Return to main page: QUIT
- Display MATH menu: MATH
- Change calculator mode: MODE

Basic input methods

- Input fractions: MATH FRAC
- Change decimal to fraction: MATH 1:▸Frac
- Change fraction to decimal: MATH 2:▸Dec
- Cube root: MATH 4: $\sqrt[3]{}$
- Equality and inequality symbols: TEST CONDITIONS

Graphing

- Input equation of graph: Y=
- View graph of equation: GRAPH
- Restrict domain/range of graph: WINDOW

- Zoom in or out: ZOOM
- Find y value at a specific x value: CALC 1:value
- Find x-intercept: CALC 2:zero
- Find point of intersection between graphs: CALC 5:intersect
- Find minimum or maximum point of graph: CALC 3:minimum or 4:maximum
- Parametric functions: MODE FUNCTION PARAMETRIC then proceed to graph
- Conic sections: APPS 2:Conics
- Piecewise function: MATH B:piecewise

Algebra

- Solve quadratic equation: APPS 4:Plysmlt2 1:POLYNOMIAL ROOT FINDER
- Solve system of linear equations: APPS 4:Plysmlt2 2:SIMULTANEOUS EQN SOLVER
- Find solution(s) to a complicated equation: graph the functions, then find point(s) of intersection

Sequences and Series

- Use graph to determine behaviour of sequence/series: TABLE
- Evaluate summation: MATH 0:summation

Calculus

- Evaluate derivative/gradient at a point (graph): CALC 6:dy/dx
- Evaluate integral given the lower and upper limits (graph): CALC 7:$\int$f(x)dx
- Evaluate derivative at an x value: MATH 8:nDeriv
- Evaluate definite integral: MATH 9:fnInt

Probability

- Permutation: MATH PROB 2:nPr
- Combination: MATH PROB 3:nCr
- Factorial: MATH PROB 4:!

Statistics

- Calculate expectation and variance of discrete random variable:

Contents

Part I

Pure Mathematics

§1 Functions

§1.1 Formulae

§1.1.1 Domain, Range, Rule

Domain: set of possible inputs

Range: set of possible outputs

Vertical line test is used to check for functions: For a function, every vertical line $x = k$, $k \in D_f$ cuts the graph *at most once*.

A **one-one function** is a function where no two distinct elements in the given domain have the same image under f.

$$\forall x_1, x_2 \in D_f, x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$$

This means that every output corresponds to exactly one input.

Horizontal line test is used to check for one-one functions: Every horizontal line $y = k$, $k \in R_f$ cuts the graph *at most once*. Conversely, for a function that is not one-one, provide a *specific counter-example* of a horizontal line that cuts the graph at more than one point.

§1.1.2 Inverse Functions

Let f be a one-one function. Then f has an **inverse function** f^{-1} defined as

$$f^{-1}(y) = x \iff f(x) = y, \forall x \in D_f$$

$D_{f^{-1}} = R_f$, $R_{f^{-1}} = D_f$

For the inverse function f^{-1} to exist, f is a one-one function (check using horizontal line test).

To find the inverse of a given function, make x the subject.

Graphically, the graphs of f and f^{-1} are **reflections** of each other in the line $y = x$.[1] Hence $y = f(x)$, $y = f^{-1}(x)$ and $y = x$ **intersect at one same point**.[2]

[1]Proof: Let (a, b) be a point on the curve $y = f(x)$. Then (b, a) is a point on the curve $y = f^{-1}(x)$ since $f(a) = b \iff a = f^{-1}(b)$.

[2]However, this is not the case for functions where $y = f(x)$ and $y = f^{-1}(x)$ do not intersect in the given domain, or for functions where $f(x) = f^{-1}(x)$.

§1.1.3 Composite Functions

A **composite function** gf is where *all* elements in the domain of function f are directly mapped to the elements in the range of function g.

The composite function gf exists iff $R_f \subseteq D_g$.[3]

$D_{gf} = D_f$, $R_{gf} = R_g$

To determine the range of a composite function, use **two-stage mapping**.

An **identity function** is a function which returns the same value, which was used as its input.

$$f^{-1}f(x) = ff^{-1}(x) = x$$

Remark. Even though the composite functions $f^{-1}f$ and ff^{-1} have the same rule, they may have different domains. $D_{f^{-1}f} = D_f$ whereas $D_{ff^{-1}} = D_{f^{-1}}$.

[3]Proof: The domain of function g must include values of the range of f, so that the function g is well-defined as every element in its domain, as well as in R_f, is mapped to something.

§1.2 Common Mistakes

§1.3 Common Approaches

Problems

§2 Graphs

§2.1 Formulae

§2.1.1 Graph Sketching

Features to include in graph sketch:

1. **Stationary points**
   - Maximum point
   - Minimum point
   - Point of inflexion
2. **Intercepts**
3. **Asymptotes**
   - Horizontal asymptote: line $y = a$ where $x \to \pm\infty$, $y \to a$
   - Vertical asymptote: line $x = a$ where $x \to a$, $y \to \pm\infty$
   - Oblique asymptote: line $y = ax + b$ where $x \to \pm\infty$, $y - (ax + b) \to 0$

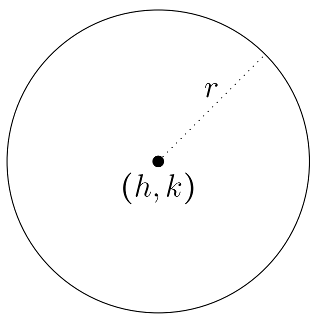
To determine the restriction on possible values of x or y, use *discriminant*.

§2.1.2 Conic Sections

Simplify Tikz code

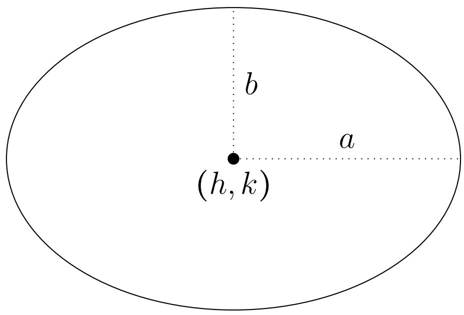
- Circle

$$(x - h)^2 + (y - k)^2 = r^2 \tag{1}$$



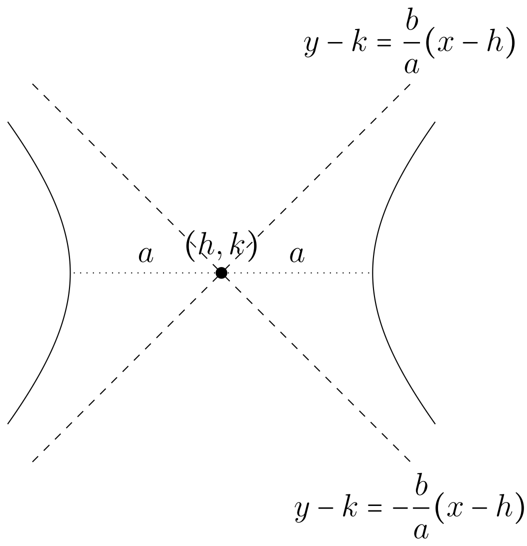
- Ellipse

$$\frac{(x - h)^2}{a^2} + \frac{(y - k)^2}{b^2} = 1 \tag{2}$$



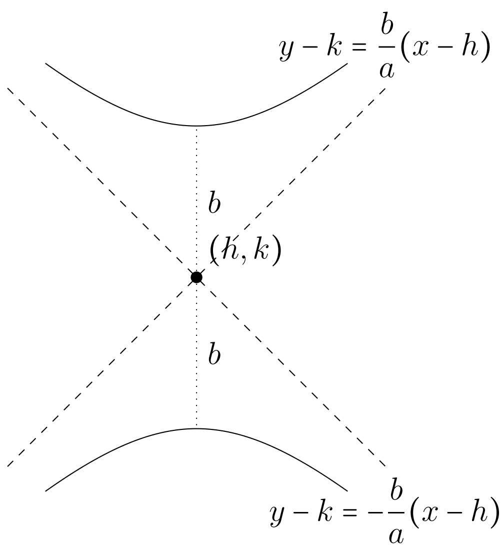
- Horizontal hyperbola

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \tag{3}$$



Vertical hyperbola

$$\frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1 \tag{4}$$



- Regular parabola

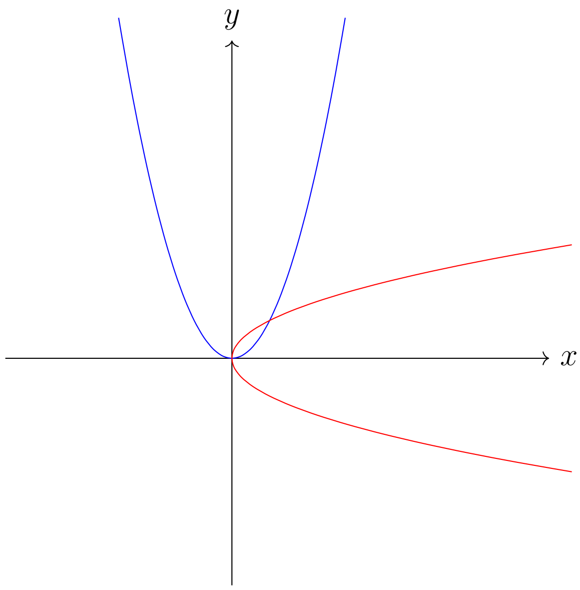
  $$y = a(x-h)^2 + k \tag{5}$$

  where (h, k) denotes the vertex.

  Sideways parabola

  $$x = a(y-k)^2 + h \tag{6}$$

  where (h, k) denotes the vertex.



- Rectangular hyperbola

  In the case of horizontal and vertical asymptotes,

  $$y = \frac{ax+b}{cx+d} \tag{7}$$

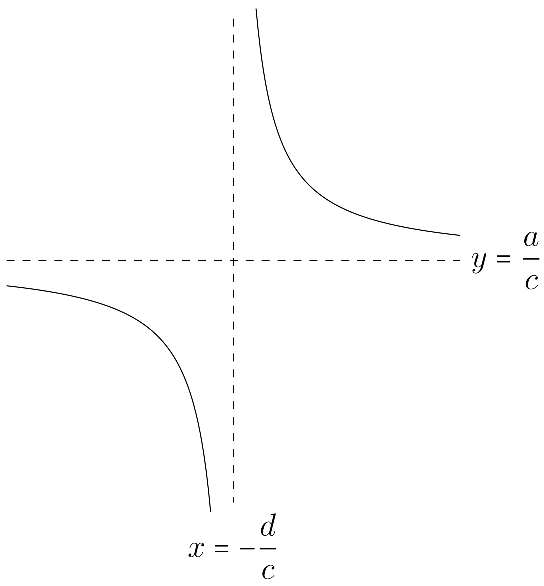
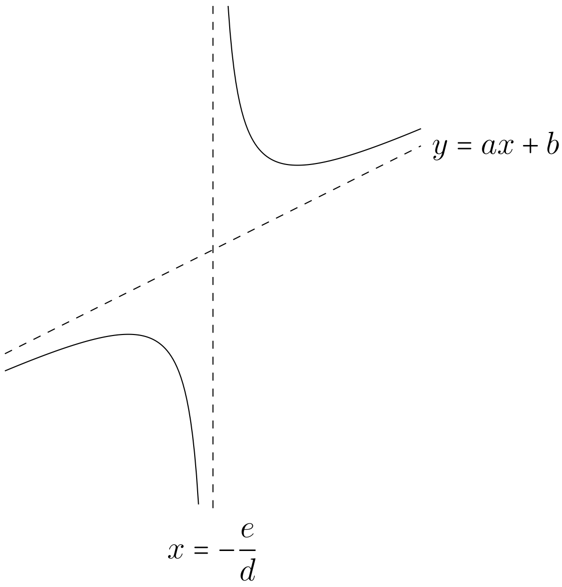


Figure 1: Rectangular hyperbola

In the case of oblique asymptotes,

$$y = ax + b + \frac{c}{dx + e} \tag{8}$$



§2.1.3 Parametric equations

- To sketch the Cartesian graph of parametric equations, using MODE to change the mode from FUNCTION to PARAMETRIC on GC. Remember to take note of the domain of the parameter t.
- To find the Cartesian equation, eliminate the parameter t by solving the equations simultaneously. A Cartesian equation should only contain the variables x and y.
- To convert a Cartesian equation to its parametric form, if involving trigonometric expressions, the Pythagorean trigonometric identity $\sin^2 x + \cos^2 x = 1$ may be handy.

§2.1.4 Transformations

To get a function $cf(bx + a) + d$ from $f(x)$,

1. translate by a units in the negative x direction
2. scale by factor of $\frac{1}{b}$ parallel to x-axis
3. scale by factor of c parallel to the y-axis
4. translate by d units in the positive y direction.

Translation

Equation	Replace	Graph	Point
$y = f(x) + a$	$y \mapsto y - a$	Translate a units in positive y-direction	$(x, y) \to (x, y + a)$
$y = f(x) - a$	$y \mapsto y + a$	Translate a units in negative y-direction	$(x, y) \to (x, y - a)$
$y = f(x - a)$	$x \mapsto x - a$	Translate a units in positive x-direction	$(x, y) \to (x + a, y)$
$y = f(x + a)$	$x \mapsto x + a$	Translate a units in negative x-direction	$(x, y) \to (x - a, y)$

Reflection

Equation	Replace	Graph	Point
$y = -f(x)$	$y \mapsto -y$	Reflected in x-axis	$(x, y) \to (x, -y)$
$y = f(-x)$	$x \mapsto -x$	Reflected in y-axis	$(x, y) \to (-x, y)$

Scaling

Equation	Replace	Graph	Point
$y = af(x)$	$y \mapsto \dfrac{y}{a}$	Scale by factor of a parallel to y-axis	$(x, y) \to (x, ay)$
$y = f\left(\dfrac{x}{a}\right)$	$x \mapsto \dfrac{x}{a}$	Scale by factor of a parallel to x-axis	$(x, y) \to (ax, y)$

Modulus

Equation	Graph	Point
$y = \lvert f(x) \rvert$	Retain the portion of graph above x-axis, reflect the portion below x-axis in the x-axis	$(x, y) \to (x, \lvert y \rvert)$
$y = f(\lvert x \rvert)$	For positive x, retain the portion of graph. For negative x, reflect the positive portion of graph in the y-axis	$(x, y) \to (\lvert x \rvert, y)$

Reciprocal

$$y = \frac{1}{f(x)}$$

Derivative

$y = f'(x)$

§2.2 Common Mistakes

§2.3 Common Approaches

Problems

§3 Equations and Inequalities

§3.1 Formulae

Systems of linear equations can be solved efficiently using PlySmlt2 on GC.

Types of solutions:

1. Unique solution
2. Infinitely many solutions
3. No solutions

Questions often involve practical problems, from which systems of linear equations are set up.

Important points to take note when solving inequalities:

1. Do not cross multiply without knowing whether terms are positive or not.
2. Know the difference between "and" and "or", i.e. intersection and union of sets.
3. Solutions should not be equal to roots of denominator.

Some manipulations before using the methods below:

- Directly deduce after moving all terms to one side.
- Multiply the square of a term in the denominator.
- Either the numerator or denominator is always positive (working is required to show this).

§3.1.1 Test-value method

To use the test-value method:

1. Indicate the critical value(s) on a number line.
2. Choose an x-value within each interval as the *test-value*.
3. Plug in the test-value to evaluate whether the polynomial is positive or negative within that interval.

§3.1.2 Graphical method

For polynomial inequalities in the form of $P(x) > 0$, sketching the graph of $y = P(x)$ and its x-intercepts gives us the solution.

Alternatively, for inequalities in the form of $P(x) > Q(x)$, sketch the graph of $y = P(x)$ and $y = Q(x)$, then identify the region of the graph where the inequality holds.

§3.1.3 Solutions of related inequalities

To deduce solutions of related inequalities, replace x with some expression of x using the solutions of inequalities solved in earlier parts of the question.

§3.2 Common Mistakes

§3.3 Common Approaches

Problems

§4 Sequences and Series

§4.1 Formulae

§4.1.1 Sequences

A **sequence** is a set of numbers $u_1, u_2, \ldots, u_n$ where $n \in \mathbb{Z}^+$. A sequence can be generated by giving a formula u_n for the nth term.

When describing the behaviour of a sequence, describe its

1. **Trend**: (strictly) increasing/decreasing, constant, alternating
   For a strictly increasing sequence, show that $u_{n+1} > u_n$.
   For a strictly decreasing sequence, show that $u_{n+1} < u_n$.

2. **Convergence**: converges, diverges
   For a convergent sequence, show that as $n \to \infty^+$, u_n approaches a *unique value*.
   For a divergent sequence, show that as $n \to \infty^+$, u_n approaches ∞^+ or ∞^-.

§4.1.2 Series

A **series** is the sum of terms of a sequence. The sum to n terms is denoted by S_n.

To find the term for a series when given the sequence,

$$u_n = S_n - S_{n-1} \tag{9}$$

For the sum to infinity S_∞ of a series to exist, the series converges; conversely, the sum to infinity does not exist if the series diverges.

§4.1.3 Arithmetic Progression

An **arithmetic progression** is a sequence in which successive terms differ by a common difference.

The formula for the nth term is

$$u_n = a + (n-1)d \tag{10}$$

The formula for the sum is

$$S_n = \frac{n}{2}[2a + (n-1)d] \tag{11}$$

Test for AP

To show that u_n is an AP,

show that $u_n - u_{n-1}$ is a constant.

§4.1.4 Geometric Progression

An **geometric progression** is a sequence in which successive terms differ by a common ratio.

The formula for the nth term is

$$u_n = ar^{n-1} \tag{12}$$

The formula for the sum is

$$\begin{aligned} S_n &= \frac{a(1-r^n)}{1-r} \quad |r|<1 \\ S_n &= \frac{a(r^n-1)}{r-1} \quad |r|>1 \end{aligned} \tag{13}$$

Test for GP

To show that u_n is a GP,

show that $\dfrac{u_n}{u_{n-1}}$ is a constant.

The formula for sum to infinity is

$$S_\infty = \frac{a}{1-r} \quad |r|<1 \tag{14}$$

Existence of sum to infinity of GP

For sum to infinity of a GP to exist, the GP converges, hence

show that the common ratio $|r|<1$.

§4.1.5 Summation Series

Standard series Note: the lower limit *must* be 1.

$$\sum_{r=1}^{n} r = \frac{n(n+1)}{2} \tag{15}$$

$$\sum_{r=1}^{n} r^2 = \frac{n(n+1)(2n+1)}{6} \tag{16}$$

$$\sum_{r=1}^{n} r^3 = \frac{n^2(n+1)^2}{4} \tag{17}$$

Method of differences The general term $u_r = f(r) - f(r-1)$, then

$$\begin{aligned}\sum_{r=1}^{n} u_r &= \sum_{r=1}^{n}(f(r) - f(r-1)) \\ &= \begin{cases} \cancel{f(1)} - f(0) + \\ \cancel{f(2)} - \cancel{f(1)} + \\ \qquad \vdots \\ \cancel{f(n-1)} - \cancel{f(n-2)} + \\ f(n) - \cancel{f(n-1)} \end{cases} \\ &= f(n) - f(0)\end{aligned}$$

Remark. Must show **diagonal** cancellation of intermediate terms in the working.

§4.2 Common Mistakes

§4.3 Common Approaches

Problems

§5 Differentiation

§5.1 Formulae

§5.1.1 Differentiation rules

Scalar multiplication

$$\frac{\mathrm{d}}{\mathrm{d}x} k f(x) = k \frac{\mathrm{d}}{\mathrm{d}x} f(x) \tag{18}$$

Sum/Difference Rule

$$\frac{\mathrm{d}}{\mathrm{d}x}[f(x) \pm g(x)] = \frac{\mathrm{d}}{\mathrm{d}x} f(x) + \frac{\mathrm{d}}{\mathrm{d}x} g(x) \tag{19}$$

Product Rule

$$\frac{\mathrm{d}}{\mathrm{d}x}[f(x) \cdot g(x)] = f(x) \frac{\mathrm{d}}{\mathrm{d}x} g(x) + g(x) \frac{\mathrm{d}}{\mathrm{d}x} f(x) \tag{20}$$

Quotient Rule

$$\frac{\mathrm{d}}{\mathrm{d}x} \frac{f(x)}{g(x)} = \frac{g(x) f'(x) - f(x) g'(x)}{g^2(x)} \tag{21}$$

Chain Rule

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u} \times \frac{\mathrm{d}u}{\mathrm{d}x} \tag{22}$$

§5.1.2 New functions

Apply Chain Rule wherever necessary.

- Exponential functions

$$\frac{\mathrm{d}}{\mathrm{d}x} e^x = e^x$$

$$\frac{\mathrm{d}}{\mathrm{d}x} a^x = a^x \ln a$$

- Logarithmic functions

$$\frac{\mathrm{d}}{\mathrm{d}x} \ln x = \frac{1}{x}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \log_a x = \frac{1}{x \ln a}$$

- Trigonometric functions

$$\frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \cos x = -\sin x$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \tan x = \sec^2 x$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{cosec} x = -\operatorname{cosec} x \cot x$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\sec x = \sec x \tan x$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cot x = -\operatorname{cosec}^2 x$$

- Inverse trigonometric functions

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin^{-1} x = \frac{1}{\sqrt{1-x^2}}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cos^{-1} x = -\frac{1}{\sqrt{1-x^2}}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan^{-1} x = \frac{1}{1+x^2}$$

§5.2 Common Mistakes

§5.3 Common Approaches

- Increasing and decreasing function

  As long as $f(x)$ is continuous,

  - $f'(x) > 0 \implies$ curve is increasing
  - $f'(x) < 0 \implies$ curve is decreasing
  - $f'(x) = 0 \implies$ curve is stationary

- Concavity

  - $f''(x) > 0 \implies$ curve is concave upwards
  - $f''(x) < 0 \implies$ curve is concave downwards

- Nature of stationary points

$$f'(a) = 0$$

  First derivative test:

  Second derivative test:

  - $f''(a) > 0 \implies$ minimum point
  - $f''(a) < 0 \implies$ maximum point

- Tangents and normals

  Equation of **tangent** to curve $y = f(x)$ at point (a, b):

$$y - b = f'(a)(x - a)$$

  where gradient of tangent $= f'(a)$.

  Equation of **normal** to curve $y = f(x)$ at point (a, b):

$$y - b = -\frac{1}{f'(a)}(x - a)$$

  where gradient of normal $= -\dfrac{1}{f'(a)}$.

- Maximisation and minimisation problems

- Connected rates of change

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u} \times \frac{\mathrm{d}u}{\mathrm{d}x}$$

- Implicit differentiation

  To differentiate an implicit function, make use of the Chain Rule.

- Parametric differentiation

  If x and y are functions of a parameter t, by applying Chain Rule, we have

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t} \cdot \frac{\mathrm{d}t}{\mathrm{d}x} = \frac{\frac{\mathrm{d}y}{\mathrm{d}t}}{\frac{\mathrm{d}x}{\mathrm{d}t}}$$

Problems

§6 Maclaurin Series

§6.1 Formulae

Binomial theorem: $\forall n \in \mathbb{Z}^+$

$$(a+b)^n = a^n + \binom{n}{1} a^{n-1} b + \cdots + b^n \tag{23}$$

Binomial series: $\forall n \in \mathbb{R}, n \neq 0$, which includes **negative** and **fractional** n

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \cdots + \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}x^r + \cdots \tag{24}$$

which has validity range $|x| < 1$.

The **Taylor series** is a power series centered at $x = a$:

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(0)}{3!}(x-a)^3 + \frac{f^{(4)}(0)}{4!}(x-a)^4 + \cdots \tag{25}$$

The **Maclaurin series** is a special case of the Taylor series, centered at $x = 0$:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \cdots \tag{26}$$

§6.1.1 Standard Series

Expansion of standard series and their validity range:

$$\begin{aligned}
(1+x)^n &= 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!} + \cdots && |x| < 1 \\
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots && \text{all } x \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots && \text{all } x \text{ in radians} \\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots && \text{all } x \text{ in radians} \\
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots && -1 < x \leq 1
\end{aligned}$$

§6.1.2 Small angle approximation

Approximation of trigonometric functions for a sufficiently small x such that *higher powers of x can be neglected.*

$$\begin{cases} \sin x & \approx x \\ \cos x & \approx 1 - \frac{x^2}{2} \\ \tan x & \approx x \end{cases} \tag{27}$$

§6.2 Common Mistakes

§6.3 Common Approaches

Problems

§7 Integration Techniques

§7.1 Formulae

- Standard functions

$$\int (ax+b)^n \, dx = \frac{(ax+b)^{n+1}}{a(n+1)}$$

$$\int \frac{1}{ax+b} \, dx = \frac{1}{a} \ln |ax+b|$$

$$\int e^{ax+b} \, dx = \frac{1}{a} e^{ax+b}$$

- Systematic integration

$$\int f'(x)[f(x)]^n \, dx = \frac{[f(x)]^{n+1}}{n+1}$$

$$\int \frac{f'(x)}{f(x)} \, dx = \ln \left| f(x) \right|$$

$$\int f'(x) e^{f(x)} \, dx = e^{f(x)}$$

- Trigonometric functions

$$\int \sin x \, dx = -\cos x$$

$$\int \cos x \, dx = \sin x$$

$$\int \sec^2 x \, dx = \tan x$$

$$\int \operatorname{cosec}^2 x \, dx = -\cot x$$

$$\int \sec x \tan x \, dx = \sec x$$

$$\int \operatorname{cosec} x \cot x \, dx = -\operatorname{cosec} x$$

$$\int \tan x \, dx = \ln |\sec x|$$

$$\int \sec x \, dx = \ln |\tan x + \sec x|$$

$$\int \operatorname{cosec} x \, dx = \ln |\operatorname{cosec} x - \cot x| = \ln \left| \tan \frac{x}{2} \right|$$

$$\int \cot x \, dx = \ln |\sin x|$$

Some transformations:

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

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

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

More importantly,

$$\cos^2 x = \frac{1 + \cos 2x}{2}$$
$$\sin^2 x = \frac{1 - \cos 2x}{2}$$

Sum to product:

$$\sin P + \sin Q = 2 \sin \frac{P+Q}{2} \cos \frac{P-Q}{2}$$
$$\sin P - \sin Q = 2 \cos \frac{P+Q}{2} \sin \frac{P-Q}{2}$$
$$\cos P + \cos Q = 2 \cos \frac{P+Q}{2} \cos \frac{P-Q}{2}$$
$$\cos P - \cos Q = -2 \sin \frac{P+Q}{2} \sin \frac{P-Q}{2}$$

- Algebraic fractions

$$\int \frac{1}{\sqrt{a^2 - x^2}} \, \mathrm{d}x = \sin^{-1} \frac{x}{a}$$
$$\int \frac{1}{a^2 + x^2} \, \mathrm{d}x = \frac{1}{a} \tan^{-1} \frac{x}{a}$$

- Partial fractions

$$\int \frac{1}{x^2 - a^2} \, \mathrm{d}x = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right|$$
$$\int \frac{1}{a^2 - x^2} \, \mathrm{d}x = \frac{1}{2a} \ln \left| \frac{a+x}{a-x} \right|$$

- Substitution

- Integration by parts

$$\int uv' = uv - \int u'v$$

Guideline on choosing "u":

Abbreviation	**Function Type**
L	logarithmic
I	inverse trigonometric
A	algebraic
T	trigonometric
E	exponential

Using **DI method** (also known as tabular method),

Example 7.1

Evaluate

$$\int x^2 e^x \, \mathrm{d}x \, .$$

Solution. We choose x^2 as "D", and e^x as "I".

sign	**D**	**I**
$+$	x^2	e^x
$-$	$2x$	e^x
$+$	2	e^x
$-$	0	e^x

Multiplying terms diagonally and summing them up gives us

$$\int x^2 e^x \, \mathrm{d}x = x^2 e^x - 2xe^x + 2e^x + c.$$

□

§7.2 Common Mistakes

§7.3 Common Approaches

Problems

§8 Applications of Integration

§8.1 Formulae

§8.2 Common Mistakes

§8.3 Common Approaches

§8.4 Integral as a limit of sum

Integration is a process of **summation**.

In general, $\int_a^b f(x)\,\mathrm{d}x$ may be thought of as the value of the area bounded by the curve $y = f(x)$, x-axis, $x = a$, $x = b$ where $a \geq b$.

Since this area can be divided into vertical strips of equal width Δx, where area ΔA of one strip is $y\Delta x$, the area A is given by

$$\lim_{\Delta x \to 0} \sum_{x=a}^{b} y\Delta x = \int_a^b y\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x \tag{28}$$

§8.5 Area under curve

When calculating area of the region required, it should always be **positive**. This means **adding a negative sign** to an integral that is negative.

§8.5.1 Area between curve and axes

For area between curve and x-axis,

1. Sketch the curve
2. Observe where the curve cuts the x-axis.
3. Split the area into one above x-axis (positive), one below x-axis (negative).

For area between curve and y-axis,

1. Sketch the curve
2. Observe where the curve cuts the y-axis.
3. Split the area into one right of y-axis (positive), one left of y-axis (negative).

When 2 or more curves are involved,

1. Evaluate point(s) of intersection.
2. Split the area into different parts.

§8.5.2 Area between curves

For all $x \in [a, b]$, for $f(x) \geq g(x)$, area between curves is

$$\int_a^b [f(x) - g(x)]\,\mathrm{d}x \tag{29}$$

For all $y \in [a, b]$, for $f(y) \geq g(y)$, area between curves is

$$\int_a^b [f(y) - g(y)]\,\mathrm{d}y \tag{30}$$

Remember this as: “upper" minus “lower" curve

§8.5.3 Area under curve parametrically

For $x = h(t)$ and $y = g(t)$, area under curve

$$\int_a^b y\,\mathrm{d}x = \int_{t_1}^{t_2} g(t)\frac{\mathrm{d}x}{\mathrm{d}t}\,\mathrm{d}t \tag{31}$$

where t_1 and t_2 are values of t when $x = a$ and $x = b$ respectively.

Similarly,

$$\int_c^d x\,\mathrm{d}y = \int_{t_3}^{t_4} h(t)\frac{\mathrm{d}y}{\mathrm{d}t}\,\mathrm{d}t \tag{32}$$

where t_3 and t_4 are values of t when $y = c$ and $y = d$ respectively.

Remark. We do not find the Cartesian equation of the curve.

§8.6 Solid of revolution (Volume)

Rotate region bounded by $y = f(x)$, x-axis, $x = a$, $x = b$ 2π around x-axis, volume of revolution is

$$\pi \int_a^b y^2\,\mathrm{d}x = \pi \int_a^b \left[f(x)\right]^2 \mathrm{d}x \tag{33}$$

Rotate region bounded by $x = f(y)$, y-axis, $y = a$, $y = b$ 2π around y-axis, volume of revolution is

$$\pi \int_a^b x^2\,\mathrm{d}y = \pi \int_a^b \left[f(y)\right]^2 \mathrm{d}y \tag{34}$$

Volume of standard shapes

- Cone: $\frac{1}{3}\pi r^2 h$
- Cylinder: $\pi r^2 h$

Rotate region bounded by two curves $y = f(x)$ and $y = g(x)$ 2π around x-axis, volume of revolution is

$$\pi \int_a^b \left[f(x)\right]^2 - \left[g(x)\right]^2 \mathrm{d}x \tag{35}$$

Rotate region bounded by two curves $x = f(y)$ and $x = g(y)$ 2π around y-axis, volume of revolution is

$$\pi \int_a^b \left[f(y)\right]^2 - \left[g(y)\right]^2 \mathrm{d}y \tag{36}$$

Problems

§9 Differential Equations

To remove non-H2 stuff

§9.1 Formulae

§9.2 Common Mistakes

§9.3 Common Approaches

§9.4 1st Order Differential Equations

A **1st order differential equation** has derivatives up to the 1st derivative.

§9.4.1 Separating the variables, families of curves

Example 9.1

Find the general solution of

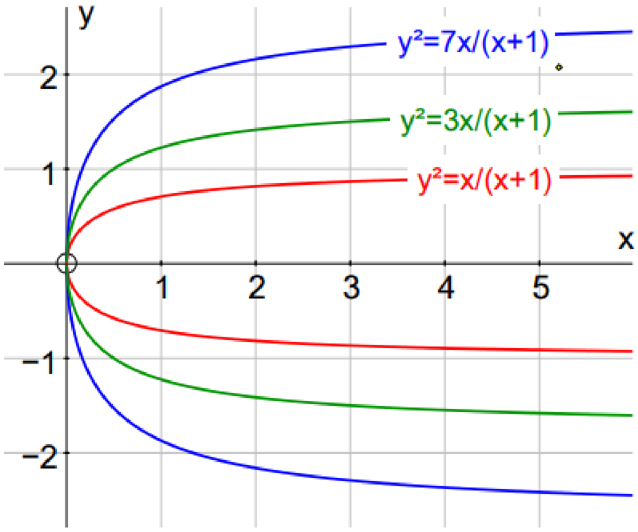
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y}{2x(x+1)} \qquad \text{for } x > 0$$

and sketch some of the family of solution curves.

Solution.

$$\begin{aligned}
&\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y}{2x(x+1)} \\
\implies &\int \frac{2}{y}\,\mathrm{d}y = \int \frac{1}{x(x+1)}\,\mathrm{d}x = \int \left(\frac{1}{x} - \frac{1}{x+1}\right)\mathrm{d}x \\
\implies &2\ln y = \ln x - \ln(x+1) + \ln A \\
\implies &y^2 = \frac{Ax}{x+1}
\end{aligned}$$

Thus for varying values of A and for $x > 0$, we have



□

§9.4.2 Exact Equations*

LHS is an exact derivative.

Example 9.2

Solve

$$\sin x \frac{\mathrm{d}y}{\mathrm{d}x} + y \cos x = 3x^2.$$

Solution. Notice that the LHS is an exact derivative.

$$\begin{aligned}
\sin x \frac{\mathrm{d}y}{\mathrm{d}x} + y \cos x &= \frac{\mathrm{d}}{\mathrm{d}x} y \sin x \\
\frac{\mathrm{d}}{\mathrm{d}x} y \sin x &= 3x^2 \\
y \sin x &= \int 3x^2 \,\mathrm{d}x = x^3 + c \\
y &= \frac{x^3 + c}{\sin x}
\end{aligned}$$

□

§9.4.3 Integrating Factor*

For equations in the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} + Py = Q$$

where P and Q are functions of x only, multiply both sides by an **integrating factor** $R = e^{\int P \mathrm{d}x}$ so that LHS becomes an exact derivative $\frac{\mathrm{d}}{\mathrm{d}x} Ry$.

Derivation. We are looking for an integrating factor R (a function of x), so that LHS becomes an exact derivative.

Multiplying the LHS by R gives

$$R\frac{\mathrm{d}y}{\mathrm{d}x} + RPy$$

If this is to be an exact derivative that we can see, by looking at the first term, we should try

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x} Ry &= R\frac{\mathrm{d}y}{\mathrm{d}x} + y\frac{\mathrm{d}R}{\mathrm{d}x} = R\frac{\mathrm{d}y}{\mathrm{d}x} + RPy \\
y\frac{\mathrm{d}R}{\mathrm{d}x} &= RPy \\
\int \frac{1}{R} \,\mathrm{d}R &= \int P \,\mathrm{d}x \\
\ln R &= \int P \,\mathrm{d}x \\
R &= e^{\int P \mathrm{d}x}
\end{aligned}$$

Hence the required integrating factor is $e^{\int P \mathrm{d}x}$. □

Example 9.3

Solve

$$x\frac{\mathrm{d}y}{\mathrm{d}x} + 2y = 1.$$

Solution. First divide both sides by x

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2}{x}y = \frac{1}{x}$$

Now the equation is in the correct form.

Integrating factor is

$$R = e^{\int P \mathrm{d}x} = e^{\int \frac{2}{x} \mathrm{d}x} = e^{2 \ln x} = x^2$$

Multiplying both sides by this gives us

$$\begin{aligned}
x^2\frac{\mathrm{d}y}{\mathrm{d}x} + 2xy &= x \\
\frac{\mathrm{d}}{\mathrm{d}x} x^y &= x \\
x^2 y &= \int x \,\mathrm{d}x = \frac{x^2}{2} + c \\
y &= \frac{1}{2} + \frac{c}{x^2}
\end{aligned}$$

□

§9.4.4 Using Substitutions

Example 9.4

Use the substitution $z = x + y$ to solve the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \cos(x + y).$$

Solution. From $z = x + y$,

$$\frac{\mathrm{d}z}{\mathrm{d}x} = 1 + \frac{\mathrm{d}y}{\mathrm{d}x} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}x} - 1$$

Substituting this and solving by separable variables,

$$\begin{aligned} \frac{\mathrm{d}z}{\mathrm{d}x} &= 1 + \cos z \\ \int \frac{1}{1 + \cos z}\, \mathrm{d}z &= \int \mathrm{d}x \\ \frac{1}{2}\sec^2 \frac{z}{2}\, \mathrm{d}z &= x + c \quad \text{[double angle formula]} \\ \tan \frac{z}{2} &= x + c \end{aligned}$$

Substituting back gives us

$$\tan \frac{x + y}{2} = x + c$$

□

§9.5 2nd Order Differential Equations

Move this section to University Maths document

A **2nd order differential equation** has derivatives up to the 2nd derivative.

§9.5.1 Linear with constant coefficients

For equations in the form

$$a\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + cy = 0$$

Auxiliary equation:

$$am^2 + bm + c = 0$$

which has solutions

$$m_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Roots of auxiliary equation	General solution
real and distinct	$y = Ae^{m_1x} + Be^{m_2x}$
real and repeated	$y = e^{mx}(A + Bx)$
imaginary: $m = \alpha \pm i\beta$	$y = e^{\alpha x}(A\cos\beta x + B\sin\beta x)$

§9.5.2 Non-linear

Equation

$$a\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + cy = f(x)$$

§9.6 Population Dynamics and Population Growth Models

To look for info from F Maths notes

• exponential growth model • logistic growth model, equilibrium points and their stability, and harvesting

Problems

§10 Complex Numbers

§10.1 Formulae

The **imaginary number** is denoted by $i = \sqrt{-1}$. A **complex number** is of the form $a + bi$ where a and b are real numbers. The set of complex numbers is denoted by $\mathbb{C} = \{z \mid z = a + bi, a, b \in \mathbb{R}\}$.

a is the **real part** of z, denoted by $\operatorname{Re}(z)$; b is the **imaginary part** of z, denoted by $\operatorname{Im}(z)$.

$a + bi$ is the **Cartesian form** of the complex number z.

§10.1.1 Algebra

To add or subtract complex numbers, we just add or subtract their real and imaginary parts:

$$(a+bi)+(c+di)=(a+c)+(b+d)i; \quad (a+bi)-(c+di)=(a-c)+(b-d)i$$

To multiply complex numbers, expand the brackets in the usual fashion and remember that $i^2 = -1$:

$$(a+bi)(c+di) = ac + bci + adi + bdi^2 = (ac-bd)+(ad+bc)i$$

The **conjugate** of $z = a + bi$ is given by $\bar{z} = a - bi$. Multiplying a complex number with its conjugate eliminates the imaginary part:

$$z\bar{z} = a^2 + b^2$$

To divide complex numbers, multiply the numerator and denominator by the conjugate of the denominator:

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

§10.1.2 Complex Roots of Polynomial Equations

With complex numbers, we have the following theorem:

Theorem 10.1: Fundamental Theorem of Algebra

Let $p(z) = a_0 + a_1 z + \cdots + a_n z_n$ be a polynomial of degree $n \geq 1$ with real (or complex) coefficients a_k.
Then the roots of the equation $p(z) = 0$ are complex. That is, there are n (not necessarily distinct) complex numbers $\gamma_1, \ldots, \gamma_n$ such that

$$a_0 + a_1 z + \cdots + a_n z_n = a_n(z - \gamma_1)(z - \gamma_2)\cdots(z - \gamma_n).$$

The theorem shows that a degree n polynomial has n roots in $\mathbb{C}$ (including repetitions).

Remark. The proof of this theorem is far beyond the scope of this text. Note that the theorem only guarantees the existence of the roots of a polynomial somewhere in $\mathbb{C}$ unlike the quadratic formula which determines exactly the roots. The theorem gives no hints as to where in $\mathbb{C}$ these roots are to be found.

Furthermore, if the coefficients of the polynomial equation are real, we have the following result:

Theorem 10.2

Non-real roots of a polynomial equation with *real coefficients* occur in conjugate pairs.

Proof. Consider the equation

$$a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0$$

where for $a_0, \ldots, a_n \in \mathbb{R}$, $a_n \neq 0$, $n \in \mathbb{Z}^+$.

Suppose β is a non-real root of the equation, then

$$a_n \beta^n + a_{n-1} \beta^{n-1} + \cdots + a_1 \beta + a_0 = 0.$$

Taking conjugates on both sides of the equation,

$$a_n \overline{\beta^n} + a_{n-1} \overline{\beta^{n-1}} + \cdots + a_1 \overline{\beta} + a_0 = 0$$

Note that $\overline{\beta^k} = \overline{\beta}^k$. Thus we have

$$a_n \bar{\beta}^n + a_{n-1} \bar{\beta}^{n-1} + \cdots + a_1 \bar{\beta} + a_0 = 0$$

which means that the complex conjugate $\bar{\beta}$ is also a non-real root of the given equation. □

§10.1.3 Cartesian Form

The complex numbers, having two components, their real and imaginary parts, can be represented on a plane known as the **Argand diagram**[4], consisting of a **real axis** and an **imaginary axis**. On this plane, the point (a, b) represents the complex number $a + bi$.

[4]After the Swiss mathematician Jean-Robert Argand (1768-1822)

On the Argand plane, the geometric representation of complex numbers are as follows:

- **Addition and subtraction**

  Addition and subtraction is done similar to that of vectors.
- **Conjugate**

  The conjugate given by the reflection of the corresponding complex number about the real axis.
- **Multiplication and division**

  Multiplying by i, rotate 90° anti-clockwise about the origin.
- **Scalar multiplication**

Remove

Theorem 10.3: Triangle inequality

For complex numbers z_1 and z_2,

$$|z_1| + |z_2| \geq |z_1 + z_2| \tag{37}$$

with equality only if one of them is 0 or $\arg(z_1) = \arg(z_2)$ i.e. z_1 and z_2 are on the same ray from the origin.

§10.1.4 Polar Form

In polar form,

$$z = r(\cos\theta + i\sin\theta) \tag{38}$$

where r is the **modulus**, denoted by $|z|$:

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

and θ is the **argument**, denoted by $\arg(z)$:

$$\arg(z) = \theta = \tan^{-1}\frac{b}{a}$$

where $\theta \in [-\pi, \pi]$.

Theorem 10.4: de Moivre's Theorem

For complex number z with $|z| = r$ and $\arg z = \theta$, then raising z to any real number power n gives us

$$z^n = [r(\cos\theta + i\sin\theta)]^n = r^n(\cos n\theta + i\sin n\theta) \tag{39}$$

Proof. The proof easily follows from mathematical induction.

Alternatively, this is a simple consequence of Euler's formula:

$$(\cos\theta + i\sin\theta)^n = (e^{i\theta})^n = e^{ni\theta} = \cos n\theta + i\sin n\theta$$

□

Exercise 10.1

Find the exact value of $(1+\sqrt{3}i)^6$.

Solution.

$$\begin{aligned}(1+\sqrt{3}i)^6 &= \left[2\left(\cos\frac{\pi}{3}+i\sin\frac{\pi}{3}\right)\right]^6 \\ &= 2^6(\cos 2\pi + i\sin 2\pi) \quad \text{[by de Moivre]} \\ &= \boxed{64}\end{aligned}$$

□

We also have

$$(\cos\theta + i\sin\theta)^{\frac{p}{q}} = \cos\frac{p\theta+2k\pi}{q} + i\sin\frac{p\theta+2k\pi}{q} \tag{40}$$

where $p, q \in \mathbb{Z}$, $q > 0$, $k = 0, 1, 2, \ldots, q-1$.

Applying this theorem, we have

- $z^n + z^{-n} = 2\cos n\theta$
- $z^n - z^{-n} = 2i\sin n\theta$
- $\cos\theta = \frac{1}{2}(z + \frac{1}{z})$
- $\sin\theta = \frac{1}{2i}(z - \frac{1}{z})$

Euler's Formula

Theorem 10.5: Euler's Formula

For a complex number z with modulus r and argument θ, **Euler's Formula** states that z can be expressed as:

$$z = re^{i\theta} = r(\cos\theta + i\sin\theta) \tag{41}$$

Proof. Recall that the Taylor Series for e^x is given by

$$e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$$

Substituting ix for x gives us

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

Separating real and imaginary parts,

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

Observe that the real part is the Taylor series of $\cos x$, while the imaginary part is the Taylor series of $\sin x$. Hence proven. □

Theorem 10.6: Euler's Identity

Euler's Identity is a special case of Euler's Formula, in which $r = 1$ and $\theta = \pi$.

$$e^{i\pi} + 1 = 0 \tag{42}$$

- Magnitude
  $|e^{i\theta}| = 1$

  Proof.
  $$|e^{i\theta}| = |\cos\theta + i\sin\theta| = \sqrt{\cos^2\theta + \sin^2\theta} = 1$$
  □

  This means that $e^{i\theta}$ always lies on the unit circle. Likewise, more generally, if $z = re^{i\theta}$ then $|z| = r$.

- Argument
  If $z = re^{i\theta}$ then $\arg z = \theta$.

  Proof. This is again the definition: the argument is the polar angle θ. □

- Conjugate
  $\overline{re^{i\theta}} = re^{-i\theta}$

  Proof.
  $$\overline{re^{i\theta}} = r(\overline{\cos\theta + i\sin\theta}) = r(\cos\theta - i\sin\theta) = r(\cos(-\theta) + i\sin(-\theta)) = re^{-i\theta}$$
  □

  Thus complex conjugation changes the sign of the argument.

- Multiplication
  If $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$ then
  $$z_1 z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)}$$
  In words, the formula says the for $z_1 z_2$ the magnitudes multiply and the arguments add.

- Division
  Again it's trivial that
  $$\frac{r_1 e^{i\theta_1}}{r_2 e^{i\theta_2}} = \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}$$

Inverse Euler formulae Euler's formula gives a complex exponential in terms of sines and cosines. We can turn this around to get the **inverse Euler formulae**. Euler's formula says:

Move everything from this point onwards

$$e^{it} = \cos t + i\sin t \quad \text{and} \quad e^{-it} = \cos t = i\sin t.$$

By adding and subtracting we get

$$\cos t = \frac{e^{it} + e^{-it}}{2} \quad \text{and} \quad \sin t = \frac{e^{it} - e^{-it}}{2i}.$$

Complex replacement In the next example we will illustrate the technique of complexification or complex replacement. This can be used to simplify a trigonometric integral. It will come in handy when we need to compute certain integrals.

Example 10.1

Use complex replacement to compute

$$I = \int e^x \cos 2x \,\mathrm{d}x$$

Solution. From Euler's formula we have

$$e^{2ix} = cos(2x) + isin(2x)$$

so $\cos 2x = \mathrm{Re}(e^{2ix})$. The complex replacement trick is to replace $\cos 2x$ by e^{2ix}. We get

$$I_c = \int (e^x \cos 2x + ie^x \sin 2x)\,\mathrm{d}x\,, \quad I = \mathrm{Re}(I_c).$$

Computing I_c is straightforward:

$$I_c = \int e^x e^{2ix}\,\mathrm{d}x = \int e^{x(1+2i)}\,\mathrm{d}x = \frac{e^{x(1+2i)}}{1+2i}.$$

Here we will do the computation first in rectangular coordinates

$$\begin{aligned} I_c &= \frac{e^{x(1+2i)}}{1+2i} \cdot \frac{1-2i}{1-2i} \\ &= \frac{e^x(\cos 2x + i\sin 2x)(1-2i)}{5} \\ &= \frac{1}{5}e^x[(\cos 2x + 2\sin 2x) + i(-2\cos 2x + \sin 2x)] \end{aligned}$$

So

$$I = \mathrm{Re}(I_c) = \boxed{\frac{1}{5}e^x(\cos 2x + 2\sin 2x)}$$

□

n-th roots of unity We are going to need to be able to find the n-th **roots of unity**, i.e. solve equations of the form

$$z^n = c$$

where c is a given complex number. This can be done most conveniently by expressing c and z in polar form:

$$c = Re^{i\varphi}, \quad z = re^{i\theta}.$$

Then, upon substituting, we have to solve

$$r^n e^{in\theta} = Re^{i\varphi}$$

For the complex numbers on the left and right to be equal, their absolute values must be same and the arguments can only differ by an integer multiple of 2π, which gives

$$r = R^{\frac{1}{n}} \quad \text{and} \quad n\theta = \varphi + 2k\pi, k = 0, \pm 1, \pm 2, \ldots$$

Solving for θ, we have

$$\theta = \frac{1}{n}(\varphi + 2k\pi) \tag{43}$$

In general $c = Re^{i\varphi}$ has n distinct n-th roots:

$$z_k = r^{1/n} e^{i\varphi/n + i2\pi(k/n)}$$

for $k = 0, 1, 2, \ldots, n-1$.

Geometry of n-th roots Roots are always spaced evenly around a circle centred at the origin. For example, the 5th roots of $1+i$ are spaced at increments of $\frac{2\pi}{5}$ radians around the circle of radius $2^{\frac{1}{5}}$.

Note also that the roots of real numbers always come in conjugate pairs.

§10.2 Common Mistakes

§10.3 Common Approaches

Problems

Problem 10.1. Find all those complex numbers z that satisfy $z^2 = i$.

Solution. Suppose that $z^2 = i$ and $z = a + bi$. Then $i = (a + bi)^2 = (a^2 - b^2) + 2abi$. Comparing real and imaginary parts in the above we obtain two simultaneous equations:

$$\begin{cases} a^2 - b^2 = 0 \\ 2ab = 1 \end{cases}$$

From the first equation, $b = \pm a$. Substituting $b = a$ into the second equation gives $a = b = \frac{1}{\sqrt{2}}$ or $a = b = -\frac{1}{\sqrt{2}}$. Substituting $b = -a$ into the second equation gives $-2a^2 = 1$ which has no real solution a. So the two z which satisfy $z^2 = i$, i.e. the two square roots of i, are $\frac{1+i}{\sqrt{2}}$ and $\frac{-1-i}{\sqrt{2}}$. □

Problem 10.2. What are the complex solutions to the equation

$$x^6 = 1?$$

Solution. We have

$$e^{(2k\pi)i} = 1$$
$$x^6 = e^{(2k\pi)i}$$
$$x = e^{\frac{k\pi}{3}i}$$

Then we simply substitute 6 values of k, from 0 to 5, to get the 6th roots of unity. □

Problem 10.3. Express $\sin 5\theta$ in terms of $\sin\theta$ only.

Solution. From De Moivre's Theorem we know that

$$\begin{aligned}
&\cos 5\theta + i\sin 5\theta \\
&= (\cos\theta + i\sin\theta)^5 \\
&= \cos^5\theta + 5i\cos^4\theta + 10i^2\cos^3\theta\sin^2\theta + 10i^3\cos^2\theta\sin^3\theta + 5i^4\cos\theta\sin^4\theta + i^5\sin^5\theta \\
&= \cos^5\theta + 5i\cos^4\theta - 10\cos^3\theta\sin^2\theta - 10i\cos^2\theta\sin^3\theta + 5\cos\theta\sin^4\theta + i\sin^5\theta
\end{aligned}$$

Equating imaginary parts,

$$\begin{aligned}
\sin 5\theta &= 5\cos^4\theta\sin\theta - 10\cos^2\theta\sin^3\theta + \sin^5\theta \\
&= 5(1-\sin^2\theta)^2\sin\theta - 10(1-\sin^2\theta)\sin^3\theta + \sin^5\theta
\end{aligned}$$
$$\boxed{\sin 5\theta = 16\sin^5\theta - 20\sin^3\theta + 5\sin\theta}$$

□

Problem 10.4. Using de Moivre's Theorem, prove that

$$\sin 4\theta - \sin 8\theta + \sin 12\theta - \cdots + \sin 36\theta = \frac{\sin 20\theta\cos 18\theta}{\cos 2\theta}$$

Solution.

$$
\begin{aligned}
&\sin 4\theta - \sin 8\theta + \sin 12\theta - \cdots + \sin 36\theta \\
&= \operatorname{Im}\left[e^{i(4\theta)} - e^{i(8\theta)} + e^{i(12\theta)} - \cdots + e^{i(36\theta)}\right] \\
&= \operatorname{Im}\left[\frac{e^{i(4\theta)}((-e^{i(4\theta)})^9 - 1)}{(-e^{i(4\theta)}) - 1}\right]
\end{aligned}
$$

□

§11 Vectors

§11.1 Formulae

Magnitude of a vector $\mathbf{a}$ is denoted by $|\mathbf{a}|$. **Unit vector** of vector $\mathbf{a}$ is denoted by $\hat{\mathbf{a}}$, where $|\hat{\mathbf{a}}| = 1$.

$$\frac{\mathbf{a}}{|\mathbf{a}|} = \hat{\mathbf{a}} \tag{44}$$

Convention for unit vectors: $\mathbf{i}$ is unit vector along x-axis, $\mathbf{j}$ is unit vector along y-axis, $\mathbf{k}$ is unit vector along z-axis.

Ratio theorem: if point P divides AB in the ratio of $\lambda : \mu$, then

$$\overrightarrow{OP} = \frac{\mu \cdot \overrightarrow{OA} + \lambda \cdot \overrightarrow{OB}}{\mu + \lambda} \tag{45}$$

Midpoint theorem is a special case where P is the midpoint:

$$\overrightarrow{OX} = \frac{\overrightarrow{OA} + \overrightarrow{OB}}{2}$$

§11.1.1 Dot Product

Dot product (or scalar product) is defined as

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta \tag{46}$$

where θ is the angle in the interval $[0, \pi]$ between vectors $\mathbf{a}$ and $\mathbf{b}$.

Remark. The angle between two vectors which either *both* point outwards or inwards.

For 3D vectors,

$$\mathbf{a} \cdot \mathbf{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3 \tag{47}$$

Remark. Geometrically, dot product measures the **alignment** between the two vectors.

Properties:

- Commutative law applies: $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$
- Distributive law applies: $(\mathbf{a} + \mathbf{b}) \cdot (\mathbf{c} + \mathbf{d}) = \mathbf{a} \cdot \mathbf{c} + \mathbf{a} \cdot \mathbf{d} + \mathbf{b} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{d}$
  Distributive property of scalar multiplication: $(\lambda \mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\lambda \mathbf{b}) = \lambda(\mathbf{a} \cdot \mathbf{b})$
- Associative law applies: $(\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c})$

§11.1.2 Cross Product

Cross product is defined as

$$\mathbf{a}\times\mathbf{b} = |\mathbf{a}||\mathbf{b}|\sin\theta\,\hat{\mathbf{n}} \tag{48}$$

where θ is the angle in the interval $[0,\pi]$ between vectors $\mathbf{a}$ and $\mathbf{b}$, $\hat{\mathbf{n}}$ denotes the unit normal vector (perpendicular to both vectors).

For 3D vectors,

$$\mathbf{a}\times\mathbf{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix} \tag{49}$$

Remark. Geometrically, cross product produces a **new vector** perpendicular to the two vectors.

Properties:

- Commutative law does not apply, in fact $\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}$. Note that $\mathbf{a}\times\mathbf{b}| = |\mathbf{b}\times\mathbf{a}|$.
- Distributive law applies: $\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = (\mathbf{a}\times\mathbf{b}) + (\mathbf{a}\times\mathbf{c})$

  Distributive property of scalar multiplication: $(\lambda\mathbf{a})\times\mathbf{b} = \mathbf{a}\times(\lambda\mathbf{b}) = \lambda(\mathbf{a}\times\mathbf{b})$
- Associative law does not apply: $(\mathbf{a}\times\mathbf{b})\times\mathbf{c} \neq \mathbf{a}\times(\mathbf{b}\times\mathbf{c})$

§11.1.3 Lines

Equation of a line

1. **Vector form**

   $$\ell : \mathbf{r} = \mathbf{a} + \lambda\mathbf{m}, \lambda \in \mathbb{R} \tag{50}$$

   where $\mathbf{r}$ is general position vector of any point on the line; $\mathbf{a}$ is position vector of fixed point; $\mathbf{m}$ is *direction vector*.

2. **Cartesian form**

   $$\frac{x-a_1}{m_1} = \frac{y-a_2}{m_2} = \frac{z-a_3}{m_3} = \lambda \tag{51}$$

   where $\mathbf{r} = x\hat{\imath} + y\hat{\jmath} + z\hat{k}$, $\mathbf{a} = a_1\hat{\imath} + a_2\hat{\jmath} + a_3\hat{k}$ and $\mathbf{m} = m_1\hat{\imath} + m_2\hat{\jmath} + m_3\hat{k}$, $m_1, m_2, m_3 \neq 0$

3. **Parametric form**

   $$x = a_1 + \lambda m_1, \quad y = a_2 + \lambda m_2, \quad z = a_3 + \lambda m_3 \tag{52}$$

Relationships between two lines ($\ell_1 : \mathbf{r} = \mathbf{a}_1 + \lambda\mathbf{m}_1$ and $\ell_2 : \mathbf{r} = \mathbf{a}_2 + \lambda\mathbf{m}_2$)

- **Parallel** (coplanar)

  Direction vectors are parallel to each other: $\mathbf{m}_1 \parallel \mathbf{m}_2$.

- **Intersecting** (coplanar)

  Direction vectors are not parallel to each other: $\mathbf{m}_1 \nparallel \mathbf{m}_2$.

  Solving simultaneously (by equating the two lines) gives one unique solution (λ, μ).

- **Skew**[5] (non-coplanar)

  The lines do not satisfy the conditions in both cases above, i.e. no unique solution (λ, μ).

§11.1.4 Planes

Equation of a plane

1. **Vector equation (parametric form)**

$$\pi : \mathbf{r} = \mathbf{a} + \lambda\mathbf{m}_1 + \mu\mathbf{m}_2, \lambda, \mu \in \mathbb{R} \tag{53}$$

   where $\mathbf{a}$ is position vector of fixed point, $\mathbf{m}_1$ and $\mathbf{m}_2$ are non-zero, non-parallel vectors that are parallel to the plane.

2. **Vector equation (scalar product form)**

$$\pi : \mathbf{r} \cdot \mathbf{n} = D \text{ where } D = \mathbf{a} \cdot \mathbf{n} \tag{54}$$

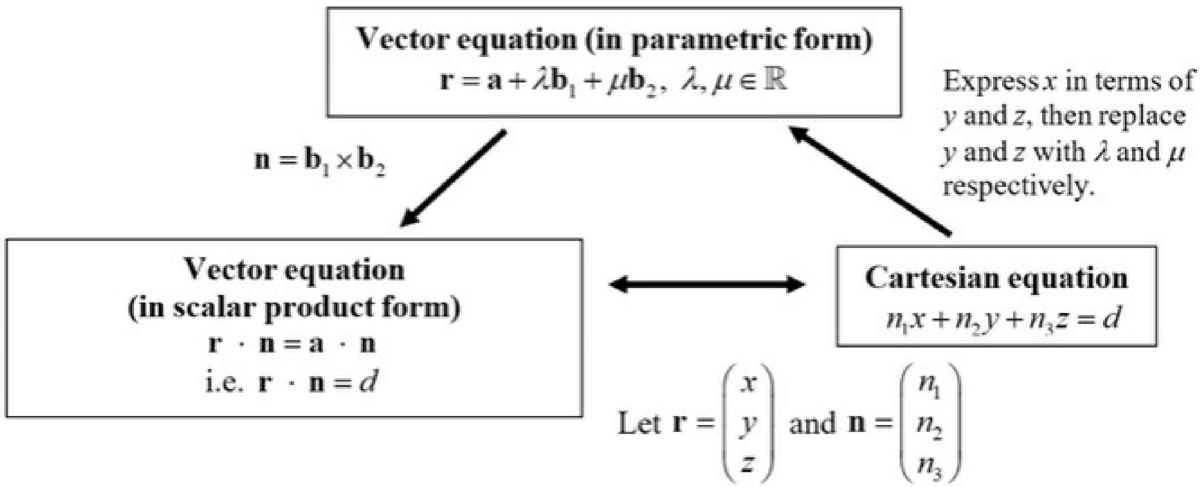
   where $\mathbf{r}$ is position vector of any point, $\mathbf{a}$ is position vector of fixed point, $\mathbf{n}$ is normal vector of plane.[6]

   $D = 0$ if plane passes through origin (since dot product of $\mathbf{0}$ with any vector is 0).

3. **Cartesian equation**

$$\pi : n_1 x + n_2 y + n_3 z = D \tag{55}$$

   where $\mathbf{n} = n_1\hat{i} + n_2\hat{j} + n_3\hat{k}$, $\mathbf{a} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k}$.[7]



---

[5] Do not intersect and are not parallel

[6] If you rearrange it, it is $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0$. Since $\mathbf{r} - \mathbf{a}$ is the vector joining $\mathbf{a}$ to $\mathbf{r}$, this equation says that $\mathbf{n}$ is orthogonal (perpendicular) to the vector joining $\mathbf{a}$ to $\mathbf{r}$ for any $\mathbf{r}$ on the plane. If you draw a diagram, you should be able to intuitively see that this is the case.

[7] This can be derived by expressing vectors in column form, then taking dot product.

Relationships between a line and a plane ($\ell : \mathbf{r} = \mathbf{a} + \lambda\mathbf{m}$ and $\pi : \mathbf{r} \cdot \mathbf{n} = D$)

- ℓ and π **do not intersect**

  $\mathbf{m} \cdot \mathbf{n} = 0$ and $\mathbf{a} \cdot \mathbf{n} \neq D$; that is, ℓ and π are parallel, and line and plane have no common point.

  When solving line and plane simultaneously, no solution.

- ℓ **lies on** π

  $\mathbf{m} \cdot \mathbf{n} = 0$ and $\mathbf{a} \cdot \mathbf{n} = D$; that is, ℓ and π are parallel, and line and plane have infinitely many common points.

  When solving line and plane simultaneously, infinitely many solutions.

- ℓ and π **intersect**

  $\mathbf{m} \cdot \mathbf{n} = 0$; that is, ℓ and π are not parallel, and line and plane have one common point.

  When solving line and plane simultaneously, one solution.

Relationship between two planes

Relationship Among Three Planes

§11.2 Common Mistakes

§11.3 Common Approaches

Insert examples

Basics

- Equal vectors
$$\mathbf{a} = \mathbf{b} \iff \text{same magnitude and direction}$$
- Parallel vectors
$$\mathbf{a} \parallel \mathbf{b} \iff \exists \lambda \in \mathbb{R}, \lambda \neq 0 \text{ s.t. } \mathbf{a} = \lambda \mathbf{b}$$
That is, one vector is a *scalar multiple* of the other.
- Collinear points
$$A, B, C \text{ collinear} \iff \overrightarrow{AB} \parallel \overrightarrow{AC}$$
(and the two vectors share a common point A)
- Coplanar vectors
$$\mathbf{a}, \mathbf{b}, \mathbf{c} \text{ coplanar} \iff \exists \lambda, \mu \in \mathbb{R} \text{ s.t. } \mathbf{a} = \lambda \mathbf{b} + \mu \mathbf{c}$$
where $\mathbf{b}$ and $\mathbf{c}$ are non-parallel, non-zero vectors. That is, one of the vectors can be expressed as a *unique linear combination* of the other two vectors.
- Parallelogram
$$OACB \text{ is a parallogram} \iff \overrightarrow{BC} = \overrightarrow{OA}$$
That is, two opposite sides are equal.

Dot product

- Find length of vector
$$\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$$
- Show perpendicular vectors
$$\mathbf{a} \perp \mathbf{b} \iff \mathbf{a} \cdot \mathbf{b} = 0$$
where $\mathbf{a}$ and $\mathbf{b}$ are non-zero vectors.
- Show parallel vectors
$$\mathbf{a} \cdot \mathbf{b} = \pm |\mathbf{a}||\mathbf{b}|$$
where positive sign implies same direction, negative sign implies opposite directions.
- Angle between two vectors
$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$$
If $\mathbf{a} \cdot \mathbf{b} > 0$, then angle between $\mathbf{a}$ and $\mathbf{b}$ is acute.

  If $\mathbf{a} \cdot \mathbf{b} < 0$, then angle between $\mathbf{a}$ and $\mathbf{b}$ is obtuse.
- Length of projection of $\mathbf{a}$ onto $\mathbf{b}$
$$|\mathbf{a} \cdot \hat{\mathbf{b}}|$$

Remark. Modulus sign, since length must be positive.

- Vector projection of $\mathbf{a}$ onto $\mathbf{b}$
$$(\mathbf{a} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}}$$
Remark. No modulus sign, since vectors can take on both positive and negative values for direction.

Cross product

- Find normal vector[8]
$$\mathbf{c} = \mathbf{a} \times \mathbf{b}$$
Remark. Normal vectors are not unique; any vector parallel to $\mathbf{a} \times \mathbf{b}$ is a normal to the plane containing $\mathbf{a}$ and $\mathbf{b}$.

- Show parallel vectors
$$\mathbf{a} \parallel \mathbf{b} \iff \mathbf{a} \times \mathbf{b} = 0$$

- Show perpendicular vectors
$$\mathbf{a} \perp \mathbf{b} \iff |\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}|$$

- Shortest distance from point to vector or line
$$|\mathbf{a} \times \hat{\mathbf{b}}|$$

- Area of parallelogram
$$|\mathbf{a} \times \mathbf{b}|$$
where $\mathbf{a}$ and $\mathbf{b}$ are two adjacent sides of the parallelogram.

Remark. Modulus sign, since area must be positive.

- Area of triangle
$$\frac{1}{2}|\mathbf{a} \times \mathbf{b}|$$
where $\mathbf{a}$ and $\mathbf{b}$ are two adjacent sides of the triangle.

Remark. Modulus sign, since area must be positive.

Lines

- Determine whether a point lies on a line

To check if point P lies on line ℓ, substitute $\mathbf{p}$ into equation for ℓ. Compare each of the x, y, z components, form three equations and solve for λ. If solution is consistent i.e. λ obtained from each equation is the same, then P lies on the line.

[8]vector perpendicular to two given vectors $\mathbf{a}$ and $\mathbf{b}$

- Acute angle between two lines

  For two lines with direction vectors $\mathbf{m}_1$ and $\mathbf{m}_2$, using dot product,

$$\cos\theta = \frac{|\mathbf{m}_1 \cdot \mathbf{m}_2|}{|\mathbf{m}_1||\mathbf{m}_2|}.$$

- Foot of perpendicular from point to line

  Let F be foot of perpendicular from P to ℓ. Since F lies on ℓ, $\overrightarrow{OF} = \mathbf{a} + \lambda\mathbf{m}$ for some λ (which we want to determine). Find $\overrightarrow{PF}$, in terms of λ. Since $\overrightarrow{PF} \perp \mathbf{m}$, using dot product,

$$\overrightarrow{PF} \cdot \mathbf{m} = 0,$$

  solve for λ. Then substitute the value of λ into the equation of $\overrightarrow{OF}$.

- Perpendicular distance from point to line

  Let h be perpendicular distance from P to ℓ. Using cross product,

$$h = |\overrightarrow{AP} \times \hat{\mathbf{m}}|.$$

- Length of projection of vector onto line

  For line with direction vector $\mathbf{m}$ and vector $\mathbf{v}$,

$$|\mathbf{v} \cdot \hat{\mathbf{m}}|.$$

- Point of reflection of point in line

Planes

- Form vector equation in parametric form

  Find position vector of a fixed point, and two direction vectors.

- Convert from vector equation in parametric form to scalar product form

  To find $\mathbf{n}$, given two vectors parallel to the plane, take cross product.

- Convert from Cartesian equation to vector equation (parametric)

  Express x in terms of y and z (to reduce number of variables to 2). Then replace y and z with λ and μ.

- Point lying on a plane

  A point lies on a plane if its position vector satisfies equation of plane.

  For example, unique (λ, μ) satisfies position vector of a point.

- Perpendicular distance from point to plane

  Let F denote foot of perpendicular from Q to $\pi : \mathbf{r}\cdot\mathbf{n} = D$. Using length of projection,

$$QF = |\overrightarrow{QF}| = |\overrightarrow{QA} \cdot \hat{\mathbf{n}}|$$

  given position vector $\mathbf{q}$, and position vector $\mathbf{a}$ or D.

  Alternative: Find foot of perpendicular, find vector, take magnitude

- Foot of perpendicular from point to plane

  Consider line ℓ_{QF} passing through Q and F. $\mathbf{n}$ can be used as its direction vector. Hence

$$\begin{cases} \pi : \mathbf{r} \cdot \mathbf{n} = D \\ \ell_{QF} : \mathbf{r} = \overrightarrow{OQ} + \lambda \mathbf{n} \end{cases}$$

  Since line and plane intersect at P, solve simultaneously to find λ, substitute back in to find $\overrightarrow{OF}$.

- Acute angle between line and plane
- Acute angle between two planes
- Foot of perpendicular

Problems

Problem 11.1. Given three points P_1, P_2, P_3 in 3-dimensional space (with position vectors $\mathbf{p}_i$ for point P_i). Given two points C_1 and C_2 in the same space (with position vectors $\mathbf{c}_i$ for point C_i), determine if C_1 and C_2 are on the different sides of the plane formed by points P_1, P_2, P_3.

Solution. Firstly find the plane passing through the three points P_1, P_2, P_3. It can be done by taking the normal to the plane to be $(\mathbf{p}_1 - \mathbf{p}_2) \times (\mathbf{p}_1 - \mathbf{p}_3) = \mathbf{n}$.

Notice that the normal $\mathbf{n}$ lies on one of the two sides.

Then find the two vectors connecting the plane to C_1 and C_2, i.e. $\mathbf{p}_1 - \mathbf{c}_1$ and $\mathbf{p}_1 - \mathbf{c}_2$.

Taking the scalar product with the normal indicates how much angle is in between the normal and the two directions. Hence if $\mathbf{n} \cdot (\mathbf{p}_1 - \mathbf{c}_1)$ and $\mathbf{n} \cdot (\mathbf{p}_1 - \mathbf{c}_2)$ have different signs, they lie on different sides of the plane. □

§12 Matrices and Linear Spaces*

• determinant of a square matrix and inverse of a non-singular matrix (2 $\times$ 2 and 3 $\times$ 3 matrices only) • use of matrices to solve a set of linear equations (including row reduction and echelon forms, and geometrical interpretation of the solution)

§12.1 Basics and Notation

A **matrix** $\mathbf{A}$ is an $m \times n$ array of numbers, where m is the number of rows and n is the number of columns.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

The numbers in the array are called the **entries** in the matrix. The entry a_{ij} in the i-th row and j-th column of a matrix is called the (i, j) **entry** of the matrix; the matrix is also denoted by $(\mathbf{A})_{ij}$.

Here are some special matrices:

- **Square matrix** of order n is a matrix with n rows and n columns, i.e. # rows = # columns

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}$$

- **Diagonal matrix**

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{mm} \end{bmatrix}$$

- **Symmetric matrix**

$$\mathbf{A} = \mathbf{A}^T$$

- **Row matrix**: matrix with only one row (sometimes used to represent a vector)

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix}$$

- **Column matrix**: matrix with only one column (sometimes used to represent a vector)

$$\mathbf{A} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$$

§12.2 Matrix Algebra

Matrix operations

- Equality

  $\mathbf{A} = \mathbf{B}$ when $\mathbf{A}$ and $\mathbf{B}$ have same size and same corresponding entries:

  $$a_{ij} = b_{ij} \quad \text{for all } i, j$$

- Addition

  Add like terms and keep the results in place.

  $$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

  Properties:

  - Commutativity

    $$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

  - Associativity

    $$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$$

  - Zero matrix behaves like the number “0" in matrix addition.

- Scalar multiplication

  Multiply each matrix term by the scalar k and keep the results in place.

  $$k \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{bmatrix}$$

- Multiplication

If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is an $n \times p$ matrix,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

the matrix product $\mathbf{C} = \mathbf{AB}$ is defined to be the $m \times p$ matrix:[9]

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1p} + \cdots + a_{2n}b_{np} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mn}b_{n1} & a_{m1}b_{12} + \cdots + a_{mn}b_{n2} & \cdots & a_{m1}b_{1p} + \cdots + a_{mn}b_{np} \end{bmatrix}$$

such that

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, p$. That is, c_{ij} of the product is obtained by multiplying term-by-term the entries of the i-th row of $\mathbf{A}$ and the j-th column of $\mathbf{B}$, and summing these n products. In other words, c_{ij} is the dot product of the i-th row of $\mathbf{A}$ and the j-th column of $\mathbf{B}$.

The product $\mathbf{AB}$ is defined if and only if the number of columns in $\mathbf{A}$ equals the number of rows in $\mathbf{B}$, in this case n.

Properties:

- Non-commutativity

$$\mathbf{AB} \neq \mathbf{BA}$$

  One special case where commutativity does hold is when $\mathbf{D}$ and $\mathbf{E}$ are two (square) diagonal matrices (of the same size); then $\mathbf{DE} = \mathbf{ED}$.

  Note that $\mathbf{AB} = 0$ does not imply $\mathbf{A} = 0$ or $\mathbf{B} = 0$.

- Distributivity

  The matrix product is distributive with respect to matrix addition. That is, if $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ are matrices of respective sizes $m \times n$, $n \times p$, $n \times p$, and $p \times q$, one has left distributivity

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

  and right distributivity

$$(\mathbf{B} + \mathbf{C})\mathbf{D} = \mathbf{BD} + \mathbf{CD}$$

[9] denoted without multiplication signs or dots

  - Scalar multiplication
  - Transpose

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$

  - Complex conjugate
  - Associativity

- Row operations

  There are three types of **row operations**:

  1. Row addition: adding a row to another.
  2. Row multiplication: multiplying all entries of a row by a non-zero constant.
  3. Row switching: interchanging two rows of a matrix.

  These operations are useful in solving linear equations and finding matrix inverses.

- Submatrix

  A **submatrix** of a matrix is obtained by deleting any collection of rows and/or columns.

  For example, from the following 3×4 matrix, we can construct a 2×3 submatrix by removing row 3 and column 2:

$$\mathbf{A} = \begin{bmatrix} 1 & \color{red}{2} & 3 & 4 \\ 5 & \color{red}{6} & 7 & 8 \\ \color{red}{9} & \color{red}{10} & \color{red}{11} & \color{red}{12} \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \end{bmatrix}$$

- Conjugate

§12.3 Identity Matrix, Determinant and Inverse of a Matrix

§12.3.1 Identity Matrix

The **identity matrix**, denoted by I_n, is a square matrix with n rows and columns which as 1s along the leading diagonal (from top left to bottom right) with 0s in all other positions.

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{56}$$

The identity matrix has the property that when multiplied with another matrix it leaves the other matrix unchanged:

$$\mathbf{AI} = \mathbf{A} = \mathbf{IA} \tag{57}$$

§12.3.2 Transpose of Matrix

The **transpose** of an $m \times n$ matrix $\mathbf{A}$ is the $n \times m$ matrix $\mathbf{A}^T$ formed by turning rows into columns and vice versa:

$$(\mathbf{A}^T)_{i,j} = \mathbf{A}_{j,i}$$

For example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -6 & 7 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 2 & -6 \\ 3 & 7 \end{bmatrix}$$

§12.3.3 Determinant of Matrix

The **determinant** of a 2×2 matrix $\mathbf{A}$, denoted by $|\mathbf{A}|$ or $\det \mathbf{A}$, is

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

and the determinant of a 3×3 matrix is

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12} \begin{vmatrix} a_{23} & a_{21} \\ a_{33} & a_{31} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

More generally, for a $n \times n$ matrix $\mathbf{A}$, the formal method is as follows. First we will require some definitions:

- The (i, j)-**minor** of $\mathbf{A}$ is the determinant of the submatrix obtained by deleting the i-th row and j-th column of $\mathbf{A}$. We denote this submatrix as $M_{ij}(\mathbf{A})$.
- The (i, j)-**cofactor** of $\mathbf{A}$ is the matrix $C_{ij}(\mathbf{A}) = (-1)^{i+j} M_{ij}(\mathbf{A})$.

Now, in order to calculate the determinant of an $n \times n$ matrix $\mathbf{A}$, we calculate

$$|\mathbf{A}| = \sum_{i=1}^{n} a_{1n} C_{1n}(\mathbf{A}) = a_{11} C_{11}(\mathbf{A}) + a_{12} C_{12}(\mathbf{A}) + a_{13} C_{13}(\mathbf{A}) + \cdots + a_{1n} C_{1n}(\mathbf{A}) \tag{58}$$

A matrix whose determinant is zero, i.e. $|\mathbf{A}| = 0$, is said to be **singular**; a matrix whose determinant is non-zero, i.e. $|\mathbf{A}| \neq 0$, is said to be **non-singular**.

§12.3.4 Inverse of Matrix

Only non-singular matrices have an inverse matrix. The **inverse** of a matrix $\mathbf{A}$ is denoted $\mathbf{A}^{-1}$ and has the following property:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{59}$$

To find the inverse of a 2×2 matrix $\mathbf{A}$ given by

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

we have the following formula:

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{60}$$

where $|\mathbf{A}| = ad - bc \neq 0$.

Remark. There exists more complicated methods for finding the inverses of 3×3 matrices and square matrices of larger size, which we will not discuss here.

§12.4 Rank of Matrix

§12.5 Orthogonal Matrix

A square matrix A is called **orthogonal** if

$$\mathbf{A}\mathbf{A}^T = \mathbf{I} \text{ and } \mathbf{A}^T\mathbf{A} = \mathbf{I}.$$

Show that if $\mathbf{A}$ and $\mathbf{B}$ are orthogonal matrices, then $\mathbf{AB}$ is an orthogonal matrix.

§12.6 System of Linear Equations

§12.6.1 Linear Systems

Definition 12.1: Linear system of equations

A **linear system of equations** refers to a set of m linear equations in n real variables $x_1, x_2, \ldots, x_n$ which are of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{61}$$

Any vector $(x_1, x_2, \ldots, x_n)$ which satisfies eq. (61) is said to be a **solution**.

Theorem 12.1

Every system of linear equations has either no solution, exactly one solution or infinitely many solutions. (There are no other possibilities)

If a system of equations has no solution, they we say that it is **inconsistent**; if the system has at least one solution, they we say that it is **consistent**.

We can usually solve a linear system by elimination. Note that the method of elimination is to simplify a system of linear equations to another system of linear equations that has exactly the same set of solution(s), but is easier to solve. In the method of elimination, we perform the following three types of operations:

1. Multiply an equation through by a non-zero constant.
2. Interchange two equations.
3. Add a multiple of one equation to another.

§12.6.2 Gaussian and Gauss-Jordan Elimination

Definition 12.2: Augmented matrix

Given a linear system eq. (61) above, the rectangle array of numbers

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

is called the **augmented matrix** of the linear system eq. (61).

Corresponding to the three types of operations in the method of elimination, the following operations on the rows of the augmented matrix are called **elementary row operations**:

1. Multiply a row through by a nonzero constant.
2. Interchange two rows.
3. Add a multiple of one row to another row.

In solving a linear system by the method of elimination, the aim is to *reduce* the linear system by performing these three operations.

Example 12.1

Solve the following linear system by performing elementary row operations:

$$\begin{aligned} x - 3y &= 2 \\ -x + y + 5z &= 2 \\ 2x - 5y + z &= 0 \end{aligned}$$

Solution. The augmented matrix of the linear system is

$$\begin{bmatrix} 1 & -3 & 0 & 2 \\ -1 & 1 & 5 & 2 \\ 2 & -5 & 1 & 0 \end{bmatrix}$$

Hence

$$\begin{pmatrix} 1 & -3 & 0 & 2 \\ -1 & 1 & 5 & 2 \\ 2 & -5 & 1 & 0 \end{pmatrix} \xrightarrow{R2+R1} \begin{pmatrix} 1 & -3 & 0 & 2 \\ 0 & -2 & 5 & 4 \\ 2 & -5 & 1 & 0 \end{pmatrix} \xrightarrow{R3+(-2)R1} \begin{pmatrix} 1 & -3 & 0 & 2 \\ 0 & -2 & 5 & 4 \\ 0 & 1 & 1 & -4 \end{pmatrix} \xrightarrow{R2\leftrightarrow R3} \begin{pmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 1 & -4 \\ 0 & -2 & 5 & 4 \end{pmatrix}$$

$$\xrightarrow{R3+2R2} \begin{pmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 1 & -4 \\ 0 & 0 & 7 & -4 \end{pmatrix} \xrightarrow{\frac{R3}{7}} \begin{pmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 1 & -4 \\ 0 & 0 & 1 & -\frac{4}{7} \end{pmatrix}$$

By backward substitution, we obtain the solution of the linear system:

$$x = -\frac{58}{7}, \quad y = -\frac{24}{7}, \quad z = -\frac{4}{7}.$$

□

Consider the following two linear systems:

$$\begin{aligned} x + 2y - z + 5w &= -1 \\ y + 3z - w &= 2 \\ z + 2w &= 3 \\ w &= 1 \end{aligned} \tag{1}$$

and

$$\begin{aligned} x &= 3 \\ y &= 1 \\ z &= 2 \\ w &= 5 \end{aligned} \tag{2}$$

The solution to (1) can be obtained by backward substitution, while the solution to (2) is immediate.

The augmented matrices of the linear systems (1) and (2) are respectively

$$\begin{bmatrix} 1 & 2 & -1 & 5 & -1 \\ 0 & 1 & 3 & -1 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}$$

The first matrix is an example of a matrix in **row-echelon form**, while the second matrix is an example of a matrix in **reduced row-echelon form**.

Definition 12.3: Row-echelon form

A matrix is said to be in **row-echelon form** if it satisfies all the following properties:

1. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
2. If a row does not consist of entirely of zeros, then the first nonzero number in the row is a 1. We call this a leading 1.
3. In any two successive rows that do not consists entirely of zeros, the leading 1 in the lower row occurs further to the right than the leading 1 in the higher row.

The matrix is said to be in **reduced row-echelon form** if, in addition to the above three properties, the following property is satisfied:

4. Each column that contains a leading 1 has zeros everywhere else in that column.

Example 12.2: Linear system with a unique solution

The augmented matrix of a linear system in (x, y, z) has been reduced to the given row-echelon form:

$$\begin{bmatrix} 1 & 2 & -1 & 2 \\ 0 & 1 & 3 & -1 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

Solve the linear system.

Solution. The corresponding linear system is

$$\begin{aligned} x + 2y - z &= 2 \\ y + 3z &= -1 \\ z &= 4 \end{aligned}$$

By backward substitution, we obtain the solution $x = 32$, $y = -13$ and $z = 4$. □

Example 12.3: Linear system with infinitely many solutions

Write down all the solutions of

$$x + 2y - z = 3.$$

Solution. Let $y = s$ and $z = t$, then $x = 3 - 2s + t$.

Thus all the solutions are $x = 3 - 2s + t$, $y = s$ and $z = t$, where $s, t \in \mathbb{R}$. □

Remark. Note that s and t are called **parameters**, and the set of all solutions expressed in terms of the parameters is called the **general solution** of the linear system.

Example 12.4

The augmented matrix of a linear system in (x, y, z, w) has been reduced to the reduced-row echelon form:

$$\begin{bmatrix} 1 & 0 & 0 & 2 & -7 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Solve the linear system.

Solution. The corresponding linear system is

$$\begin{aligned} x + 2w &= -7 \\ y + w &= 5 \\ z + 3w &= 1 \end{aligned}$$

The variables (unknowns) that corresponding to the leading 1's, namely x, y and z, are called **leading variables**. The non-leading variables (w in this case) are called **free variables**.

Solving for leading variables in terms of variables, we can assign any arbitrary value to the free variable w, say t, which then determines the values of the leading variable. Thus this linear system has *infinitely many solutions* given by

$$x = -7 - 2t, \quad y = 5 - t, \quad z = 1 - 3t, \quad w = t \quad \text{where } t \in \mathbb{R}$$

$\square$

Definition 12.4: Gaussian elimination

The method of solving a linear system by reducing the corresponding augmented matrix to row-echelon form (respectively reduced row-echelon form) is unknown as **Gaussian elimination** (respectively **Gauss-Jordan elimination**).

Example 12.5

Without using a calculator, solve the linear system

$$\begin{aligned} 3x + 4y - 2z + 13w &= 9 \\ x + 2y - 2z + 7w &= 5 \\ 2x + y + 4z + 6w &= -3 \end{aligned}$$

Solution. We write down the augmented matrix of the linear system and then perform elementary row operations to reduce it to row-echelon form or reduced row-echelon form:

$$\begin{pmatrix} 3 & 4 & -2 & 13 & 9 \\ 1 & 2 & -2 & 7 & 5 \\ 2 & 1 & 4 & 6 & -3 \end{pmatrix} \xrightarrow{R1 \leftrightarrow R2} \begin{pmatrix} 1 & 2 & -2 & 7 & 5 \\ 3 & 4 & -2 & 13 & 9 \\ 2 & 1 & 4 & 6 & -3 \end{pmatrix} \xrightarrow[R3-R1\times 2]{R2-R1\times 3} \begin{pmatrix} 1 & 2 & -2 & 7 & 5 \\ 0 & -2 & 4 & -8 & -6 \\ 0 & -3 & 8 & -8 & -13 \end{pmatrix}$$

$$\xrightarrow{R2\times\left(-\frac{1}{2}\right)} \begin{pmatrix} 1 & 2 & -2 & 7 & 5 \\ 0 & 1 & -2 & 4 & 3 \\ 0 & -3 & 8 & -8 & -13 \end{pmatrix} \xrightarrow{R3+R2\times 3} \begin{pmatrix} 1 & 2 & -2 & 7 & 5 \\ 0 & 1 & -2 & 4 & 3 \\ 0 & 0 & 2 & 4 & -4 \end{pmatrix} \xrightarrow{R3\times\frac{1}{2}} \begin{pmatrix} 1 & 2 & -2 & 7 & 5 \\ 0 & 1 & -2 & 4 & 3 \\ 0 & 0 & 1 & 2 & -2 \end{pmatrix}$$

The linear system corresponding to the row-echelon form is

$$\begin{aligned} x + 2y - 2z + 7w &= 5 \\ y - 2z + 4w &= 3 \\ z + 2w &= -2 \end{aligned}$$

which has the same set of solutions as the given linear system. Now x, y and z are the leading variables, and w is the free variable. Let $w = t$ where $t \in \mathbb{R}$ is an arbitrary number. By backward substitution, $z = -2 - 2t, y = -1 - 8t, x = 3 + 5t$. Thus the general solution of the given linear system is

$$x = 3 + 5t, \quad y = -1 - 8t, \quad z = -2 - 2t, \quad w = t \quad \text{where } t \in \mathbb{R}$$

Alternatively, we can further reduce the row-echelon form to reduced row-echelon form, then assign $w = t$ to obtain the same general solution. □

Example 12.6: Geometrical interpretation

The general solution of the system of linear equations

$$\begin{aligned} x + y &= -1 \\ 2x + y + z &= 3 \\ x + z &= 4 \end{aligned}$$

is given by $x = 4 - t, y = -5 + t, z = t$. What is the geometrical interpretation of the solution?

Solution. The three planes $x + y = -1$, $2x + y + z = 3$ and $x + z = 4$ intersect in a common line, with vector equation

$$r = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4-t \\ -5+t \\ t \end{pmatrix} = \begin{pmatrix} 4 \\ -5 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}$$

□

§12.6.3 Homogenous Linear Systems

Definition 12.5: Homogenous linear system

A linear system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

is known as a **homogeneous linear system**.

Every homogeneous linear system is consistent, since $x_1 = x_2 = \cdots = x_n = 0$ is a solution; this solution is called the **trivial solution**; if there are other solutions, then they are called **non-trivial solutions**, i.e. a solution $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ is a non-trivial solution if *at least one* of $s_i \neq 0$.

Theorem 12.2

Every homogeneous system of linear equations with more unknowns than equations has infinity many solutions.

Example 12.7

Determine whether the homogeneous linear system has non-trivial solution.

$$\begin{aligned} x + y + 3z &= 0 \\ -x + 2y + 6z &= 0 \\ 2x - y - 3z &= 0 \end{aligned}$$

Solution. The augmented matrix is

$$\begin{bmatrix} 1 & 1 & 3 & 0 \\ -1 & 2 & 6 & 0 \\ 2 & -1 & -3 & 0 \end{bmatrix}$$

Performing elementary row operations on the augmented matrix gives us:

$$\begin{bmatrix} 1 & 1 & 3 & 0 \\ 0 & 3 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The corresponding homogeneous system

$$\begin{aligned} x + y + 3z &= 0 \\ 3y + 9z &= 0 \end{aligned}$$

has 3 unknowns and 2 equations.

Hence the homogeneous linear system has non-trivial solution. Since it is equivalent to the given homogeneous system, it also has non-trivial solution. □

§12.7 Linear Transformations

• linear spaces and subspaces, and the axioms (restricted to spaces of finite dimension over the field of real numbers only) • linear independence and span • basis and dimension (in simple cases), including use of terms such as 'column space', 'row space', 'range space' and 'null space' • rank of a square matrix and relation between rank, dimension of null space and order of the matrix • linear transformations and matrices from $\mathbb{R}^n$ to $\mathbb{R}^m$ • eigenvalues and eigenvectors of square matrices (2×2 and 3×3 matrices, restricted to cases where the eigenvalues are real and distinct) • diagonalisation of a square matrix M by expressing the matrix in the form QDQ–1, where D is a diagonal matrix of eigenvalues and Q is a matrix whose columns are eigenvectors, and use of this expression such as to find the powers of M

§12.8 Eigenvalues and Eigenvectors

§13 Recurrence Relations*

• sequence generated by a simple recurrence relation including the use of a graphing calculator to generate the sequence defined by the recurrence relation • behaviour of a sequence, such as the limiting behaviour of a sequence • solution of (i) first order linear (homogeneous and nonhomogeneous) recurrence relations with constant coefficients (ii) second order linear homogeneous recurrence relations with constant coefficients • modelling with recurrence relations of the forms above

Part II

Statistics

§14 Permutation and Combination

Factorial notation:

$$n! = \begin{cases} n(n-1)(n-2)\cdots 3\cdot 2\cdot 1 & n \in \mathbb{Z}^+ \\ 1 & n = 0 \end{cases}$$

§14.1 Fundamental Principles of Counting

§14.1.1 Addition Principle

If there are r choices for performing a particular task, and the number of ways to carry out the k-th choice is k_n for $k = 1, 2, 3, \ldots, r$, then the total number of ways of performing the particular task is equal to the sum of the number of ways for all the r different choices, i.e.,

$$n_1 + n_2 + n_3 + \cdots + n_r.$$

Remark. It is important to note that the different choices cannot occur at the same time; otherwise, remember to minus out those that are over-counted more than once.

Example 14.1

How many cards are there in a deck of poker cards that are either a spade or an ace?

Solution. We can consider the problem as counting the total number of ways to perform the task of picking either a spade or an ace from the deck of cards.

We can pick a spade in 13 ways, and pick an ace in 4 ways.

However, note that the two choices can occur at the same time in exactly 1 way, which is picking the ace of spades!

Hence the total number of cards that are either a spade or an ace $= 13 + 4 - 1 = 16$. □

§14.1.2 Multiplication Principle

If one task can be performed in m ways, and following this, a second task can be performed in n ways (regardless of which way the first task was performed), then the number of ways of performing the 2 tasks in succession is $m \times n$.

Note: This can be applied to 2 or more tasks performed independently in succession.

§14.2 Permutation

A **permutation** is an ordered arrangement of objects; in permutations, order matters.

- Permutation of n distinct objects (in a row)

  Given n distinct objects, total number of ways of arranging all these n objects in a row is
  $$n(n-1)(n-2)\cdots 3\cdot 2\cdot 1 = n!$$

- Permutation of r objects taken from n distinct objects (in a row)

  Given n distinct objects, the number of ways arranging r of these objects $(0 \le r \le n)$ is
  $$n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!} =^n \mathrm{P}_r$$

- Permutations involving non-distinct/identical objects

  The number of ways to arrange n objects, of which exactly m objects are identical to one another is
  $$\frac{n!}{m!}$$
  To generalise this, suppose there are n objects, of which there are r different groups of identical / non-distinct / indistinguishable objects. The total number of permutations of all n objects is
  $$\frac{n!}{n_1!\cdots n_r!}$$
  where k_n represents the number of identical objects in the k-th group for $k = 1, 2, 3\ldots, r$.

  Remark. We are essentially dividing away the order created by the identical objects had they been considered to be distinct.

- Permutations of n distinct objects in a circle

  The number of permutations of n distinct objects in a circle (where the seats are indistinguishable) is
  $$\frac{n!}{n} = (n-1)!$$

§14.3 Combination

A **combination** is a selection of objects in which the order of selection does not matter.

- Combination (Selection) of r objects taken from n distinct objects

  The selection of r objects taken from n distinct objects, where $r \le n$, is
  $$\frac{n!}{r!(n-r)!} = \binom{n}{r} =^n \mathrm{C}_r$$

- Select at least 1 object from n distinct objects
  $$2^n - 1$$

Problems

Problem 14.1. Prove that
$$\binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1}$$

Solution. We present a few solutions.

Solution 1: Algebraic Proof

This is fairly straightforward by rewriting binomial expressions in terms of factorials, and then the rest is simply basic algebraic manipulation.

Solution 2: Combinatorial Proof

To choose r objects from $n+1$ distinct objects, we have 2 ways:

- Choose all r objects from the first n objects. This is simply $\binom{n}{r}$.
- Choose the $(n+1)$th object and another $r-1$ objects from the first n objects. This is $1 \times \binom{n}{r-1} = \binom{n}{r-1}$.

Applying addition principle on the two ways gives us the desired answer.

Solution 3: Pascal's Triangle

This is simply adding the $(r-1)$th and rth term in the n row, to give the term below, i.e. rth term in $(n+1)$th row. $\square$

§15 Probability

§15.1 Outcomes, Sample Space and Events

Experiment: a situation involving chance that leads to various possible results.

Outcome (sample point): the result of a single trial of the experiment.

Sample space (or probability space) S: the set of all possible outcomes of the experiment.

Event E: a subset of S containing one or more outcomes of an experiment.

Example 15.1

Consider the experiment of tossing a fair die once. The outcomes are 1, 2, 3, 4, 5, and 6.
The sample space S is given by

$$S = \{1, 2, 3, 4, 5, 6\}.$$

The event of obtaining an even number E is given by

$$E = \{2, 4, 6\}.$$

§15.2 Operations on Events

Let A and B be two events of an experiment.

The **union** of A and B (denoted by $A \cup B$) is the event that either A or B occurs (or both).

The **intersection** of A and B (denoted by $A \cap B$) is the event that both A AND B occurs.

The **complement** of A (denoted by A') is the event that A DOES NOT occur.

Remark. $A \cap A' = \varnothing$ and $A \cup A' = S$.

Two events A and B are **mutually exclusive** or **disjoint** if they cannot occur simultaneously, or equivalently, they do not share any common outcomes i.e. $A \cap B = \varnothing$.

§15.3 Probability

Probability is a measure of how likely an event occurs. Formally, for an experiment where all outcomes in the sample space are equally likely to occur, the probability that event A occurs is given by

$$P(A) = \frac{\text{number of outcomes in } A}{\text{total number of outcomes in sample space } S} = \frac{n(A)}{n(S)}$$

Laws of probability:

- Range of values for probabilities

  Every event A is a subset of sample space S, hence $0 \le n(A) \le n(S)$. Hence

  $$\frac{0}{n(S)} \le \frac{n(A)}{n(S)} \le \frac{n(S)}{n(S)} \implies \boxed{0 \le P(A) \le 1}$$

  If A is impossible (i.e. $A = \varnothing$), then $P(A) = 0$; if A is certain (i.e. $A = S$), then $P(A) = 1$.

- Probability rule for combined events

  For two overlapping events A and B, when you add $n(A)$ and $n(B)$ together you will count the overlap twice. Hence

  $$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

  Dividing all terms by $n(S)$ gives

  $$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

- Probability rule for mutually exclusive events

  If events A and B are mutually exclusive, then

  $$A \cap B = \varnothing \implies n(A \cap B) = 0 \implies \boxed{P(A \cap B) = 0}$$

  Following from the probability rule for combined events, for two mutually exclusive events A and B, since $P(A \cap B) = 0$,

  $$\boxed{P(A \cup B) = P(A) + P(B)}$$

- Probability of the complement

  Since $A \cap A' = \varnothing$ and $A \cup A' = S$, we have

  $$P(A \cup A') = 1 \implies P(A) + P(A') = 1 \implies \boxed{P(A') = 1 - P(A)}$$

§15.4 Conditional Probability

Let A and B be two events of an experiment, where $B \neq \varnothing$, i.e. $P(B) \neq 0$.

The **conditional probability** of A given B, denoted by $P(A|B)$, is the probability of A occurring given that B has occurred:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Rearranging, we obtain the general multiplication rule for probability:

$$P(A \cap B) = P(A|B) \times P(B)$$

and similarly

$$P(A \cap B) = P(B|A) \times P(A)$$

Remark. Note that $P(A|B) \neq P(B|A)$.

Two events A and B are said to be **independent** if the probability that A occurs is not affected by the occurrence of B i.e.

$$P(A|B) = P(A)$$

Following from the general multiplication rule for probability, if A and B are independent, then

$$P(A \cap B) = P(A) \times P(B)$$

§16 Binomial and Poisson Distributions

§16.1 Random Variables

§16.1.1 Random Variable

A **random variable** X for an experiment is one where the event that X takes on a certain value x, corresponds exactly to a possible event E of the experiment. We express the event E as $\{X = x\}$ in this case.

Remark. The convention is to use capital letters, e.g. X and Y, to denote random variables, while corresponding lower case letters, e.g. x and y, are used to represent one of the values it can take.

Remark. A random variable can be either discrete or continuous.

A **discrete random variable** can only take certain numerical values in an interval. For example, the score on a die (1, 2, 3, 4, 5, 6).

§16.1.2 Probability Distribution Function

Let X be a discrete random variable taking values $x_1, x_2, \ldots, x_n$. Then the **probability distribution function** of X is the function f that maps each value x_k to the *probability* that $X = x_k$, i.e.

$$f(x) = P(X = x) \quad \text{for } x = x_1, x_2, \ldots, x_n$$

§16.1.3 Cumulative Distribution Function

If X is a discrete random variable with probability distribution function $P(X = x)$ for $x = x_1, x_2, \ldots$, then the **cumulative distribution function** of X is given by

$$P(X \le x) = \sum_{r \le x} P(X = r)$$

§16.1.4 Expectation of Discrete Random Variable

The **expectation** (mean) of a discrete random variable X indicates its weighted average and is given by

$$\mu = \mathrm{E}(X) = \sum x P(X = x) \tag{62}$$

Example 16.1

Expectation of die toss is given by

$$\mathrm{E}(X) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = \boxed{3.5}$$

For constants a and b,

- $\mathrm{E}(a) = a$ (Average of a constant value is itself)
- $\mathrm{E}(aX) = a\,\mathrm{E}(X)$
- $\mathrm{E}(aX \pm b) = a\,\mathrm{E}(X) \pm b$
- $\mathrm{E}(aX \pm bY) = a\,\mathrm{E}(X) \pm b\,\mathrm{E}(Y)$

§16.1.5 Variance of Discrete Random Variable

Varience $\mathrm{Var}(X)$ is a non-negative number which gives an idea of how widely spread the values of the discrete random variable are; the larger the variance, the more scattered the observations.

$$\sigma^2 = \mathrm{Var}[x] = \sum (x_i - \mu)^2 p_i = \sum ({x_i}^2 p_i - \mu^2) \tag{63}$$

$$\sigma^2 = Var[X] = E\left[(X-\mu)^2\right] = E\left[X^2\right] \smallsmile \mu^2 = E\left[X^2\right] - \left(E[X]\right)^2 \tag{64}$$

• concept of discrete random variables, probability distributions, expectations and variances

§16.2 Binomial distribution

• concept of binomial distribution $B(n,p)$ as an example of a discrete probability distribution and use of B(n, p) as a probability model, including conditions under which the binomial distribution is a suitable model • use of mean and variance of binomial distribution (without proof)

§16.3 Poisson Distribution

§17 Normal Distribution

• concept of a normal distribution as an example of a continuous probability model and its mean and variance; use of $N(\mu, \sigma_2)$ as a probability model • standard normal distribution • finding the value of $P(X < x_1)$ or a related probability, given the values of x_1, μ, σ • symmetry of the normal curve and its properties • finding a relationship between x_1, μ, σ given the value of $P(X < x_1)$, or a related probability • solving problems involving the use of E(aX + b) and Var (aX + b) • solving problems involving the use of E(aX + bY) and Var (aX + bY), where X and Y are independent

§18 Sampling

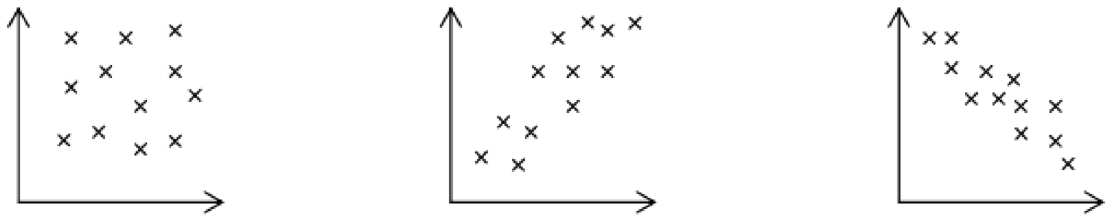
§19 Hypothesis Testing

• concepts of null hypothesis (H0) and alternative hypotheses (H1), test statistic, critical region, critical value, level of significance and p-value • formulation of hypotheses and testing for a population mean based on: – a sample from a normal population of known variance – a large sample from any population • 1-tail and 2-tail tests • interpretation of the results of a hypothesis test in the context of the problem Exclude the use of the term 'Type I' error, concept of Type II error and testing the difference between two population means.

§20 Correlation and Linear Regression

§20.1 Correlation

Scatter diagram to determine if there is a **linear relationship** between two variables



no correlation　　positive correlation　　negative correlation

The pattern of a scatter diagram shows linear correlation in a general manner.

A line of best fit can be draw by eye, but only when the points nearly lie on a straight line.

- correlation coefficient as a measure of the fit of a linear model to the scatter diagram

Product moment correlation coefficient tells us if there is a linear connection between the two variables.

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \tag{65}$$

where

$$S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{1}{N}\left(\sum x_i\right)\left(\sum y_i\right)$$
$$S_{xx} = \sum(x_i - \bar{x})^2 = \sum {x_i}^2 - \frac{1}{N}\left(\sum x_i\right)^2$$
$$S_{yy} = \sum(y_i - \bar{y})^2 = \sum {y_i}^2 - \frac{1}{N}\left(\sum y_i\right)^2$$

It can be shown that $-1 \le r \le 1$.

- $r = 1$ or close to: perfect positive linear correlation
- $r = -1$ or close to: perfect negative linear correlation
- $r = 0$ or close to: no linear correlation

§20.2 Linear regression

• concepts of linear regression and method of least squares to find the equation of the regression line • concepts of interpolation and extrapolation • use of the appropriate regression line to make prediction or estimate a value in practical situations, including explaining how well the situation is modelled by the linear regression model • use of a square, reciprocal or logarithmic transformation to achieve linearity