Triple Math

https://asdia.dev/notes/triple-math.pdf

Notes taken by
Eytan Chong

2024–2025

Mathematics requires a small dose, not of genius,

but of an imaginative freedom,

which in a larger dose would be insanity.

— Angus K. Rodgers

Last updated July 10, 2025.

Preface

About this Book

This book is a collection of notes and exercises based on the mathematics courses offered at Dunman High School[1]. The scope of this book follows that of the 2025 H2 Mathematics (9758), H2 Further Mathematics (9649) and H3 Mathematics (9820) syllabi for the Singapore-Cambridge A-Level examinations.

Notation

All definitions, results, recipes (methods) and examples are colour-coded green, blue, purple and red respectively.

Challenging exercises are marked with a "🌶" symbol.

The area of a polygon $A_1A_2\dots A_n$ is notated $[A_1A_2\dots A_n]$. In particular, the area of a triangle ABC is notated $[\triangle ABC]$.

For formatting reasons, an inline column vector is notated as $(x,\ y,\ z)^{\mathsf{T}}$.

Let n be a positive integer. Then $[n]$ represents the set $\{1, 2, \dots, n\}$.

Contributing

The source code for this book is available on GitHub at `asdia0/TripleMath`. Contributions are more than welcome.

[1] It must be stated that these notes are unofficial and are obviously not endorsed by the school.

Contents

NOTES

Part I

Functions and Graphs

1 Equations and Inequalities

1.1 Quadratic Equations

In this section, we will look at the properties of quadratic equations as well as their roots.

Proposition 1.1.1 (Quadratic Formula)**.** The roots α and β of a quadratic equation $ax^2 + bx + c = 0$, where $a \neq 0$ can be found using the quadratic formula:

$$\alpha,\, \beta = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Proof. Completing the square, we get

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c = 0,$$

which rearranges as

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

Taking roots and simplifying,

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

□

Definition 1.1.2. The expression under the radical, $b^2 - 4ac$, is known as the **discriminant** and is denoted Δ.

Proposition 1.1.3 (Nature of Roots)**.**

- If $\Delta > 0$, the roots are real and distinct.
- If $\Delta = 0$, the roots are equal.
- If $\Delta < 0$, the roots are complex.

Proof. Let the roots to the quadratic equation $ax^2 + bx + c = 0$ be α and β. By the quadratic formula,

$$\alpha,\, \beta = \frac{-b}{2a} \pm \frac{\sqrt{\Delta}}{2a}.$$

Clearly, if $\Delta > 0$, then $\sqrt{D} > 0$, whence the two roots are different. If $\Delta = 0$, then $\sqrt{D} = 0$, whence $\alpha = \beta = -b/2a$. If $\Delta < 0$, then $\sqrt{D}$ is not real, whence α and β are complex. □

Remark. Not only are α and β complex, but they are also *complex conjugates*. We will cover this later in §13.

Proposition 1.1.4 (Vieta's Formula for Quadratics)**.** Let α and β be the roots of the quadratic $ax^2+bx+c=0$, where $a \neq 0$. Then

$$\alpha+\beta=-\frac{b}{a}, \qquad \alpha\beta=\frac{c}{a}.$$

Proof. Since α and β are roots, we can rewrite the quadratic as

$$ax^2+bx+c=a(x-\alpha)(x-\beta)=a\left[x^2-(\alpha+\beta)x+\alpha\beta\right].$$

Comparing coefficients yields

$$\alpha+\beta=-\frac{b}{a}, \qquad \alpha\beta=\frac{c}{a}.$$

□

1.2 System of Linear Equations

Definition 1.2.1. A set of two or more equations to be solved simultaneously is called a **system of equations**. If the system has only equations that contain unknowns of the *first degree*, it is a **system of linear equations**.

Definition 1.2.2. A system of equations is said to be **consistent** if it admits solutions. Conversely, if there are no solutions to the system, it is said to be **inconsistent**.

Example 1.2.3. The system

$$\begin{cases} 3x+6y=3 \\ 3x+8y=9 \end{cases}$$

is consistent, since $x=-5$, $y=3$ is a solution. On the other hand, the system

$$\begin{cases} 3x+6y=3 \\ 6x+12y=7 \end{cases}$$

is inconsistent, as it does not admit any solutions (why?).

Proposition 1.2.4. If a system of linear equations is consistent, it either has a unique solution or infinitely many solutions.

Proof. Geometrically, if a collection of lines has more than one common point, they must all be equivalent. □

1.3 Inequalities

Fact 1.3.1 (Properties of Inequalities)**.** Let $a,b,c,\in\mathbb{R}$.

- (transitivity) If $a>b$ and $b>c$, then $a>c$.
- (addition) If $a>b$, then $a+c>b+c$.
- (multiplication) If $a>b$ and $c>0$, then $ac>bc$; if $c<0$, then $ac<bc$.

1.3.1 Solving Inequalities

In this section, we introduce two main methods of solving inequalities.

Recipe 1.3.2 (Graphical Method). Plot the function and observe which x-values satisfy the inequality.

Recipe 1.3.3 (Test-Value Method).

1. Indicate the root(s) of the function on a number line (i.e. where $f(x) = 0$).
2. Choose an x-value within each interval as your test-value.
3. Using the test-value, evaluate whether the function is positive/negative within that interval.

Note that the test-value method is only useful for inequalities where one side is 0, e.g. $f(x) > 0$.

Sample Problem 1.3.4 (Test-Value Method). Solve the inequality $2x - x^2 \geq -3$.

Solution. In order to apply the test-value method, we must first make one side of the inequality 0:

$$2x - x^2 \geq -3 \implies x^2 - 2x - 3 \leq 0.$$

Since $x^2 - 2x - 3 = (x+1)(x-3)$, the critical values are $x = -1$ and $x = 3$. Picking $x = -2$, $x = 0$ and $x = 4$ as our test-values, we see that $x^2 - 2x - 3$ is only negative on the interval $(-1, 3)$. Hence, the solution is $[-1, 3]$. ☐

In the case where the function is rational, i.e. $f(x)/g(x)$, there is an additional method we can use.

Recipe 1.3.5 (Clearing Denominators). Multiply the square of the denominator, i.e. $[g(x)]^2$, throughout the inequality.

Note that the square ensures that the sign of the inequality is preserved.

1.4 Modulus Function

Definition 1.4.1. The modulus function $|x|$, where $x \in \mathbb{R}$, is defined as

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

The modulus function can be thought of as the "distance" between a number and the origin (the number 0) on the real number line.

Fact 1.4.2 (Properties of Modulus Function). For any $x \in \mathbb{R}$ and $k > 0$,

- $|x| \geq 0$.
- $\left|x^2\right| = |x|^2 = x^2$ and $\sqrt{x^2} = |x|$.
- $|x| < k \iff -k < x < k$.
- $|x| = k \iff x = -k$ or $x = k$.
- $|x| > k \iff x < -k$ or $x > k$.

2 Numerical Methods of Finding Roots

2.1 Bolzano's Theorem

The following theorem forms the basis for finding roots numerically.

Theorem 2.1.1 (Bolzano's Theorem). Let $f(x)$ be a continuous function on the interval $[a, b]$. If $f(a)$ and $f(b)$ have opposite signs, i.e. $f(a)f(b) < 0$, then there exists at least one real root in $[a, b]$.

Additionally, if $f(x)$ is strictly monotonic on $[a, b]$, then there is exactly one real root in $[a, b]$.

2.2 Numerical Methods for Finding Roots

A numerical method for finding roots typically consists of two stages:

1. **Estimate the location of the root**

   Obtain an initial approximate value of this root.

2. **Improve on the estimate (via an iterative process)**

   An iterative process is a repetitive procedure designed to produce a sequence of approximations $\{x_n\}$ so that the sequence converges to a root. The process is continued until the required accuracy is reached.

In this chapter, we will look at three numerical methods for finding roots, namely linear interpolation, fixed point iteration and the Newton-Raphson method.

2.3 Linear Interpolation

Linear interpolation is a numerical method based on approximating the curve $y = f(x)$ to a straight line in the vicinity of the root. The approximate root of the equation $f(x) = 0$ is the intersection of this straight line with the x-axis.

2.3.1 Derivation

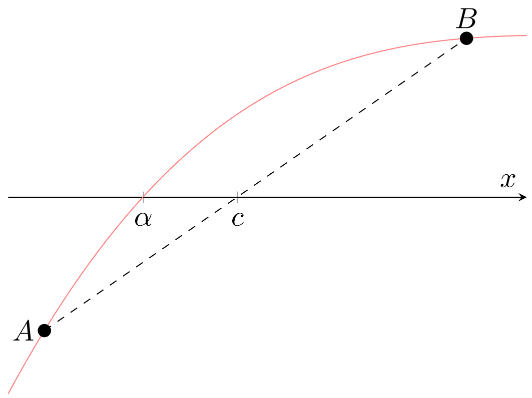


Figure 2.1

Suppose $f(x) = 0$ has exactly one root α in the interval $[a, b]$, where $f(a)$ and $f(b)$ have opposite signs. By the point-slope formula, the line connecting the points $(a, f(a))$ and $(b, f(b))$ is given by

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

At the point $(c, 0)$,

$$0 - f(a) = \frac{f(b) - f(a)}{b - a}(c - a) \implies c = \frac{af(b) - bf(a)}{f(b) - f(a)}.$$

Linear interpolation can be repeatedly applied by replacing either the lower or upper bound of the interval with the previously found approximation.

2.3.2 Convergence

Convergence of the approximations is guaranteed for linear interpolation. However, how good the estimation is depends on how "straight" the graph of $y = f(x)$ is in $[a, b]$, i.e. the rate at which $f'(x)$ is changing in $[a, b]$. This rate also affects the rate of convergence: if $f'(x)$ changes considerably, the rate of convergence is slow; if $f'(x)$ does not change much, the rate of convergence is fast.

2.4 Fixed Point Iteration

Fixed point iteration is used to find a root of an equation $f(x) = 0$ which can be written in the form $x = F(x)$. The roots of the equation are the abscissae of the points of intersection of the line $y = x$ and $y = F(x)$.

2.4.1 Derivation

Let α be a root to $f(x) = 0$. Since $f(x) = 0$ can be written in the form $x = F(x)$, we clearly have $\alpha = F(\alpha)$. Now observe that we can replace the argument α with $F(\alpha)$:

$$\alpha = F(\alpha) = F \circ F(\alpha) = F \circ F \circ F(\alpha) = \dots.$$

Hence,

$$\alpha = F \circ F \circ F \circ \cdots \circ F(x).$$

2.4.2 Geometrical Interpretation

Geometrically, fixed-point iteration can be seen as repeatedly "reflecting" the initial approximation point $(x_1, F(x_1))$ about the line $y = x$, while keeping the resultant point on the curve $y = F(x)$.

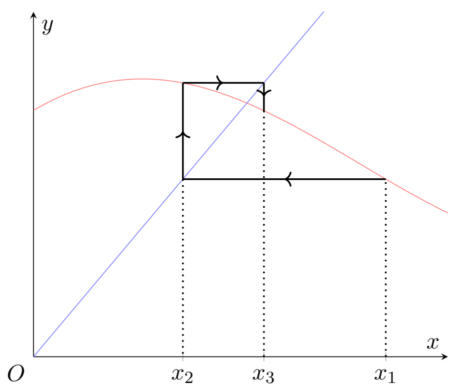


Figure 2.2

2.4.3 Convergence

Convergence is not guaranteed. The rate at which the approximations converge to α depends on the value of $|F'(x)|$ near α. The smaller $|F'(x)|$ is, the faster the convergence. It should be noted that fixed-point iteration fails if $|F'(x)| > 1$ near α.

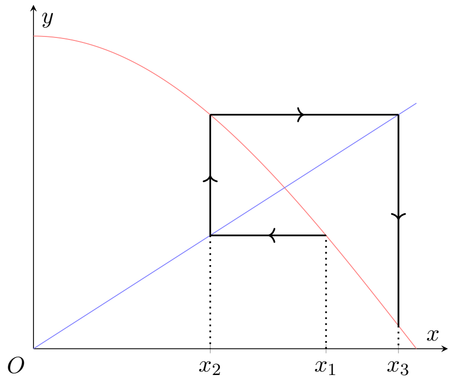


Figure 2.3: Divergence occurs when $|F'(x)| > 1$ near α.

2.5 Newton-Raphson Method

The Newton-Raphson method is a numerical method that improves on linear interpolation by considering the tangent line at the initial approximation to the root.

2.5.1 Derivation

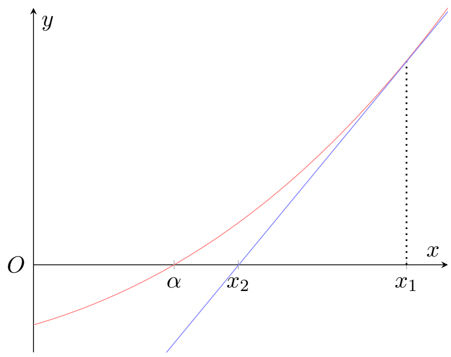


Figure 2.4

Let α be a root to $f(x) = 0$. Consider the tangent to $y = f(x)$ at the point where $x = x_1$. In most circumstances, the point $(x_2, 0)$ where this tangent cuts the x-axis will be nearer to the point $(\alpha, 0)$ than $(x_1, 0)$ was. By the point-slope formula, the equation of the tangent to the curve at $x = x_1$ is

$$y - f(x_1) = f'(x_1)(x - x_1).$$

Since $(x_2, 0)$ lies on the tangent line, we have

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

By repeating the Newton-Raphson process, we are able to get better approximations to α. In general,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

2.5.2 Convergence

The rate of convergence when using the Newton-Raphson method depends on the first approximation used and the shape of the curve in the neighbourhood of the root. In extreme cases, these factors may lead to failure (divergence). The three main cases are:

- $|f'(x_1)|$ is too small (extreme case when $f'(x_1) = 0$),
- $f'(x)$ increases/decreases too rapidly ($|f''(x)|$ is too large),
- x_1 is too far away from α.

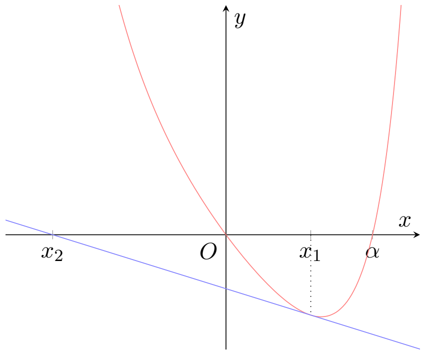


Figure 2.5: Divergence occurs when x_1 is too far away from α.

3 Functions

3.1 Definition and Notation

Definition 3.1.1. A **function** f is a rule or relation that assigns each and every element of $x \in X$ to one and only one element $y \in Y$. We write this as $f : X \to Y$ and read it as "f maps x to Y". X is called the **domain** of f, denoted D_f, while Y is called the **codomain** of f. The elements of y that get mapped to under f is known as the **range** of f, denoted R_f. Mathematically, $R_f = \{f(x) \mid x \in D_f\}$.

To define a function, we must state its rule and specify the domain. There are two ways to represent this:

$$\underbrace{f : x \mapsto x^2 + 1}_{\text{the rule}},\ \underbrace{x \in \mathbb{R}}_{D_f} \qquad \text{or} \qquad \underbrace{f(x) = x^2 + 1}_{\text{the rule}},\ \underbrace{x \in \mathbb{R}}_{D_f}.$$

Note that two functions are equal if and only if they have the same rule and domain. For instance, the function $g : x \mapsto x^2 + 1$, $x \in \mathbb{Z}$ is not equal to f (as defined above) since their domains are not equal ($\mathbb{R} \neq \mathbb{Z}$).

Note that f is not the same as $f(x)$; f is a *map*, while $f(x)$ is the *value* that f maps x to.

3.2 Graph of a Function

Definition 3.2.1. The **graph** of $f(x)$ is the collection of all points (x, y) in the xy-plane such that the values x and y satisfy $y = f(x)$.

Proposition 3.2.2 (Vertical Line Test)**.** A relation f is a function if and only if every vertical line $x = k$, $k \in D_f$ cuts the graph of $y = f(x)$ at one and only one point.

Proof. By definition, a function f is a relation which maps each element in the domain to one and only one image. □

3.3 One-One Functions

Definition 3.3.1. A function is said to be **one-one** if no two distinct elements in the given domain have the same image under f. Mathematically,

$$x_1 \neq x_2 \implies f(x_1) \neq f(x_2).$$

Equivalently, f is one-one if $f(x_1) = f(x_2)$ implies $x_1 = x_2$.

Proposition 3.3.2 (Horizontal Line Test)**.** A function f is one-one if and only if any horizontal line $y = k$, $k \in R_f$ cuts the graph of $y = f(x)$ at one and only one point.

Proof. We only prove the backwards case as the forwards case is trivial. Suppose $y = k$ and $y = f(x)$ intersect more than once. Then there exist two distinct elements x_1 and x_2 in D_f such that $f(x_1) = f(x_2)$, whence f is not one-one. □

Proposition 3.3.3 (Strict Monotonicity Implies One-One)**.** All strictly monotone functions are one-one.

Proof. Seeking a contradiction, assume that there exists a strictly increasing function $f : X \to Y$ which is not one-one. Then there exists $x_1, x_2 \in X$ such that $x_1 \neq x_2 \implies f(x_1) = f(x_2)$. Without loss of generality, assume $x_1 < x_2$, since f is strictly increasing. Then $f(x_1) < f(x_2)$, a contradiction. Therefore, all strictly increasing functions are one-one. Similarly, all strictly decreasing functions are one-one. ☐

To prove that a function is not one-one, it is sufficient to provide a specific counter-example.

3.4 Inverse Functions

Definition 3.4.1. Let $f : X \to Y$ be a function. Its **inverse function**, $f^{-1} : Y \to X$ is a function that undoes the operation of f. Mathematically, for all $x \in D_f$,

$$f^{-1}(y) = x \iff f(x) = y.$$

Fact 3.4.2 (Properties of Inverse Function)**.**

- f^{-1} exists if and only if f is one-one.
- $D_f = R_{f^{-1}}$ and $R_f = D_{f^{-1}}$.
- The graphs of f and f^{-1} are reflections of each other in the line $y = x$.

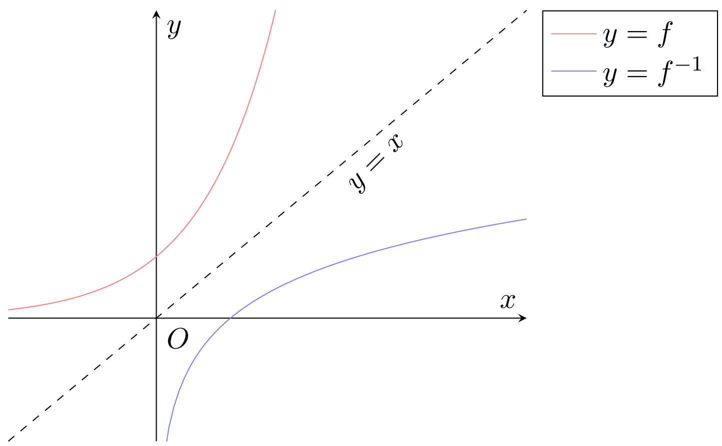


Figure 3.1: The graphs of f and f^{-1} are reflections of each other in the line $y = x$.

3.5 Composite Functions

Definition 3.5.1. Let f and g be functions. Then the **composite function** gf is defined by

$$gf(x) = g(f(x)) = g \circ f(x), \quad x \in D_f.$$

Proposition 3.5.2 (Existence of Composite Function)**.** The composite function gf exists when $R_f \subseteq D_g$.

Proof. Suppose $R_f \not\subseteq D_g$. Then there exists some element y in R_f that is not in D_g. Let the pre-image of y under f be x. Then $gf(x) = g(y)$ is undefined, whence gf is not well-defined and is hence not a function. ☐

Note that in general, composition of functions is not commutative, i.e. $fg \neq gf$.

We write the composition of f with itself n times as $f^n(x)$. For instance, $ff(x) = f(f(x))$ can be written as $f^2(x)$. This should not be confused with $[f(x)]^n$.

3.5.1 Composition of Inverse Function

Suppose $f : x \mapsto y$ has an inverse $f^{-1} : y \mapsto x$. By the definition of an inverse function.

$$f^{-1} \circ f(x) = f \circ f^{-1}(x) = x.$$

Though $f^{-1}f$ and ff^{-1} have the same rule, they may have different domains. This is because $D_{f^{-1}f} = D_f$, while $D_{ff^{-1}} = D_{f^{-1}}$.

4 Graphs and Transformations

4.1 Characteristics of a Graph

When we sketch a graph, we need to take note of the following characteristics and indicate them on the sketch accordingly:

- **Axial intercepts.** x- and y-intercepts.
- **Stationary points.** Maximum, minimum points and stationary points of inflexion.
- **Asymptotes.** Horizontal, vertical and oblique asymptotes.

When sketching a graph, the shape and any symmetry must be clearly seen.

4.2 Asymptotes

Definition 4.2.1. An **asymptote** is a straight line such that the distance between the curve and the line approaches zero at the extreme end(s) of a graph, i.e. the curve approaches the line but never touches it at these ends.

Definition 4.2.2. Let a and b be constants.

- If $x \to \pm\infty$, $y \to a$, then the line $y = a$ is a **horizontal asymptote**.
- If $x \to a$, $y \to \pm\infty$, then the line $x = a$ is a **vertical asymptote**.
- If $x \to \pm\infty$, $y - (ax + b) \to 0$, then the line $y = ax + b$ is an **oblique asymptote**.

4.3 Even and Odd Functions

Definition 4.3.1. A function $f(x)$ is **even** if and only if $f(-x) = f(x)$ for all x in its domain.

Geometrically, a function is even if and only if the graph $y = f(x)$ is symmetrical about the y-axis.

Definition 4.3.2. A function $f(x)$ is **odd** if and only if $f(-x) = -f(x)$ for all x in its domain.

Geometrically, a function is odd if and only if the graph $y = f(x)$ is symmetrical about the origin.

4.4 Graphs of Rational Functions

A rational function f is a ratio of two polynomials $P(x)$ and $Q(x)$, where $Q(x) \not\equiv 0$.

4.4.1 Rectangular Hyperbola

A rectangular hyperbola is a hyperbola with asymptotes that are perpendicular to each other. The general formula for a rectangular hyperbola is $y = \frac{ax+b}{cx+d}$, where a, b, c and d are constants. Note that the curve $y = \frac{ax+b}{cx+d}$ has a vertical asymptote $x = -d/c$ and a horizontal asymptote $y = a/c$. The two possible shapes of a rectangular hyperbola are shown below.

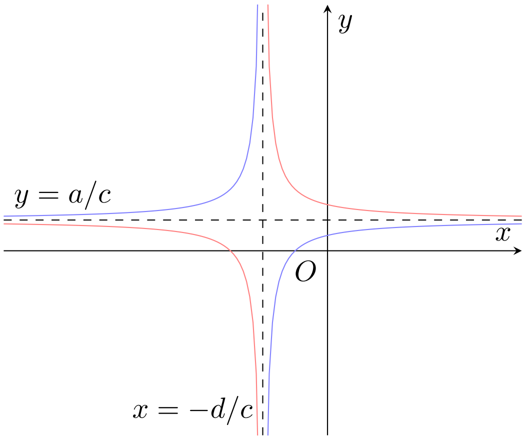


Figure 4.1: Hyperbolas of the form $y = \frac{ax+b}{cx+d}$.

4.4.2 Hyperbolas of the Form $y = \frac{ax^2+bx+c}{dx+e}$

A hyperbola of the form $y = \frac{ax^2+bx+c}{dx+e}$, where a, b, c, d and e are constants, has one vertical and one oblique asymptote. The vertical asymptote has equation $x = -e/d$. To deduce the oblique asymptote, we must first convert the equation to the form $y = px + q + \frac{r}{dx+e}$ (via long division or otherwise). These graphs will generally take one of the two forms below, which can be easily deduced by checking the axial intercepts.

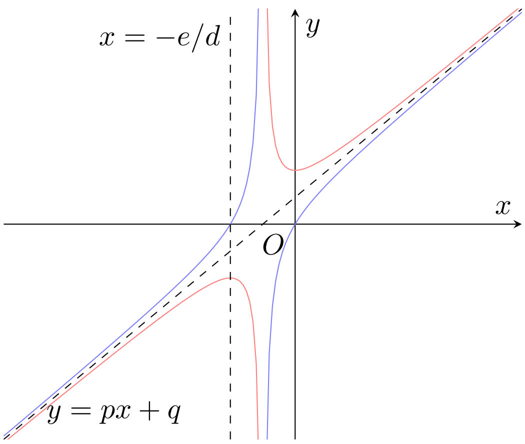


Figure 4.2: Hyperbolas of the form $y = \frac{ax^2+bx+c}{dx+e}$.

4.5 Graphs of Basic Conics

A conic is a curve that can be formed by intersecting a right circular conical surface with a plane. We will examine four types of conics: parabola, circle, ellipse and hyperbola. When sketching graphs of conics, it is important to identify their unique characteristics.

4.5.1 Parabola

Parabolas are curves with equations $y = ax^2$ or $x = by^2$, where a and b are constants.

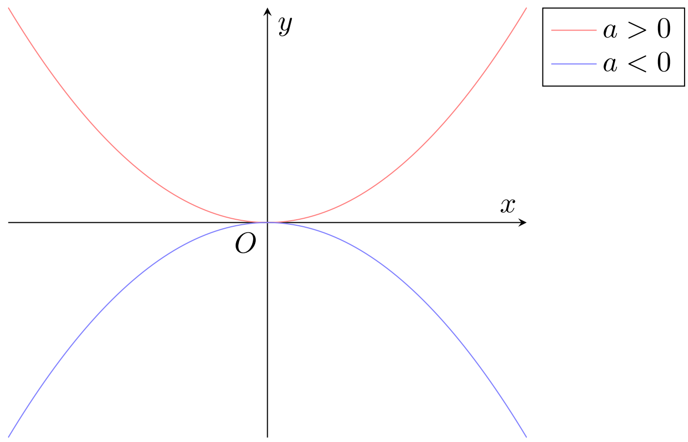


Figure 4.3: Parabolas with equation $y = ax^2$.

Parabolas with equation $y = ax^2$ have a line of symmetry $x = 0$ and a vertex at the origin.

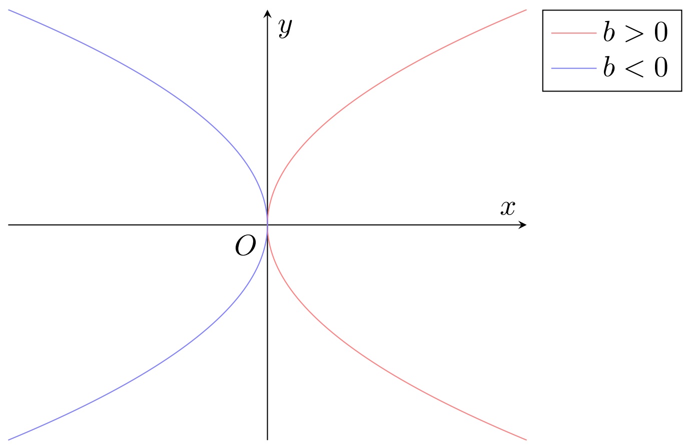


Figure 4.4: Parabolas with equation $x = by^2$.

Parabolas with equation $x = by^2$ have a line of symmetry $y = 0$ and a vertex at the origin.

4.5.2 Circle

A circle is a set of all points in a plane which are the same distance (radius r) from a fixed point (centre). A basic circle with centre at the origin O and radius r is shown below.

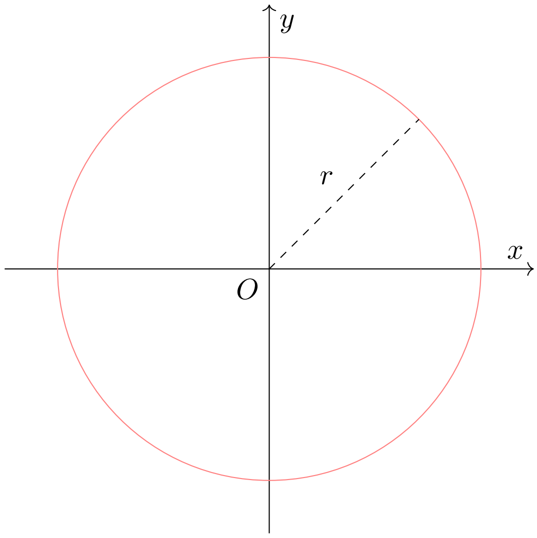


Figure 4.5: Circle with equation $x^2 + y^2 = r^2$.

Any straight line that passes through the centre of the circle is a line of symmetry. The above circle has vertices at $(r, 0)$, $(-r, 0)$, $(0, r)$ and $(0, -r)$.

In general,

- the standard form of the equation of a circle with centre at (h, k) and radius r is $(x-h)^2 + (y-k)^2 = r^2$, where $r > 0$.
- the general form of the equation of a circle is $Ax^2 + Ay^2 + Bx + Cy + D = 0$.

4.5.3 Ellipse

An ellipse is a circle that has been scaled parallel to the x- and/or y-axes. The standard form of the equation of an ellipse centred at $(0, 0)$ is $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where $a, b > 0$. a and b are known as the **horizontal** and **vertical radii** respectively.

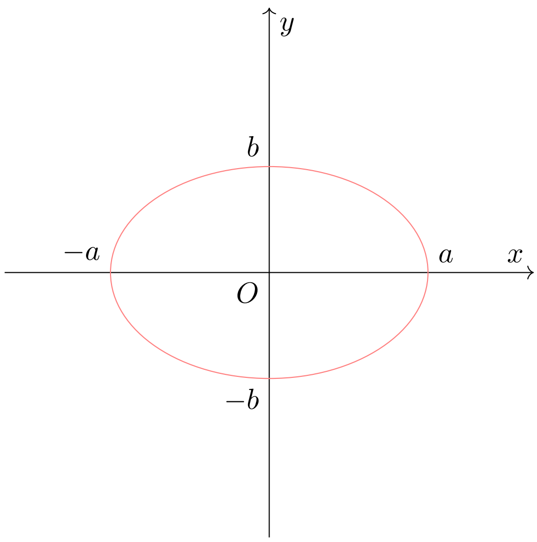


Figure 4.6: Ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

The lines of symmetry for the above ellipse are the x- and y-axes, while its vertices are $(a, 0)$, $(-a, 0)$, $(0, b)$ and $(0, -b)$.

In general,

- the standard form of the equation of an ellipse with centre at (h, k) and radius r is $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$, where $r > 0$.
- the general form of the equation of an ellipse is $Ax^2 + Bx^2 + Cx + Dy + E = 0$.

4.5.4 Hyperbola

The hyperbola is a conic with two oblique asymptotes. The standard form of a hyperbola centred at the origin O is either $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ or $\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1$, where $a, b > 0$, depending on the orientation of the hyperbola.

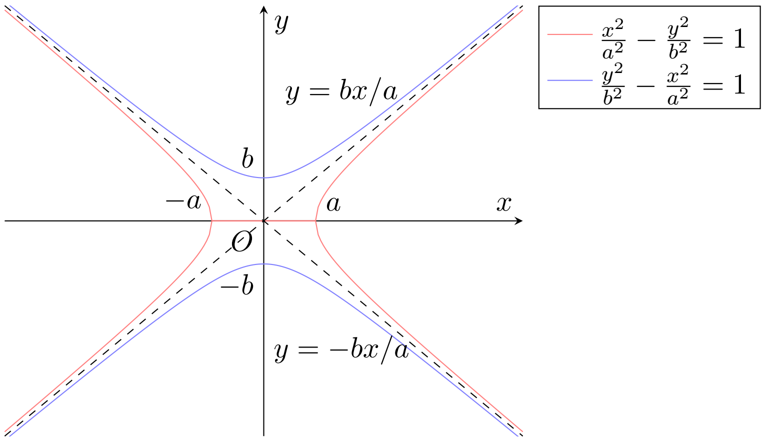


Figure 4.7

Both hyperbolas have the origin as their centres, the x- and y-axes as their lines of symmetry, and their two oblique asymptotes are $y = \pm\frac{b}{a}x$. The hyperbola with equation $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ has vertices $(-a, 0)$ and $(a, 0)$, i.e. a is the horizontal distance from the centre to the vertices. Similarly, the hyperbola with equation $\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1$ has vertices $(0, -b)$ and $(0, b)$, i.e. b is the vertical distance from the centre to the vertices.

In general,

- the standard form of the equation of a hyperbola with centre at (h, k) and radius r is $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$ or $\frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1$ where $a, b > 0$.
- the general form of the equation of a hyperbola is $Ax^2 - Bx^2 + Cx + Dy + E = 0$.

4.6 Parametric Equations

Definition 4.6.1. A set of **parametric equations** define a curve by expressing the coordinates (x, y) in terms of an independent variable t (the **parameter**), i.e. $x = f(t)$ and $y = g(t)$.

Example 4.6.2 (Parametric Equations of a Circle)**.** The parametric equations $x = \cos\theta$, $y = \sin\theta$, $\theta \in [0, 2\pi)$ defines a unit circle.

Note that changing the domain of the parameter may change the shape of the curve, even if the same pair of parametric equations are used. Using the above example, if we instead take $\theta \in [0, \pi)$ the resulting curve is that of a semicircle.

To convert a pair of parametric equations to Cartesian form, the parameter must be eliminated. This can be done by either expressing t in terms of x and/or y.

Example 4.6.3 (Parametric to Cartesian via Substitution)**.** Consider the parametric equations $x = t^2 + 2t$, $y = t^2 - 2t$. Observe that $x - y = 4t$, whence $t = (x - y)/4$. Thus, the Cartesian equation of the resulting curve is

$$y = \left(\frac{x-y}{4}\right)^2 + 2\left(\frac{x-y}{4}\right).$$

A similar process is used to convert implicit Cartesian equations into parametric form. Note that explicit Cartesian equations can be trivially converted: simply take $x = t$.

4.7 Basic Linear Transformations

4.7.1 Translation

For $a > 0$,

How $y = f(x)$ was transformed	**Graphical effect on $y = f(x)$**	**Effect on x or y values**
y replaced with $y - a$	Translated a units in the positive y-direction.	$(x, y) \mapsto (x, y + a)$
y replaced with $y + a$	Translated a units in the negative y-direction.	$(x, y) \mapsto (x, y - a)$
x replaced with $x - a$	Translated a units in the positive x-direction.	$(x, y) \mapsto (x + a, y)$
x replaced with $x + a$	Translated a units in the negative x-direction.	$(x, y) \mapsto (x - a, y)$

4.7.2 Reflection

For $a > 0$,

How $y = f(x)$ was transformed	**Graphical effect on $y = f(x)$**	**Effect on x or y values**
y replaced with $-y$	Reflected in the x-axis.	$(x, y) \mapsto (x, -y)$
x replaced with $-x$	Reflected in the y-axis.	$(x, y) \mapsto (-x, y)$

4.7.3 Scaling

For $a > 0$,

How $y = f(x)$ was transformed	**Graphical effect on $y = f(x)$**	**Effect on x or y values**
y replaced with y/a	Scaled by a factor of a parallel to the y-axis.	$(x, y) \mapsto (x, ay)$
x replaced with x/a	Scaled by a factor of a parallel to the x-axis.	$(x, y) \mapsto (ax, y)$

4.8 Relating Graphs to the Graph of $y = f(x)$

4.8.1 Graph of $y = |f(x)|$

Note that

$$y = |f(x)| = \begin{cases} f(x) & f(x) \geq 0, \\ f(-x) & f(x) < 0. \end{cases}$$

Recipe 4.8.1 (Graph of $y = |f(x)|$). To obtain the graph of $y = |f(x)|$ from the graph of $y = f(x)$,

- Retain the portion of $y = f(x)$ above the x-axis.
- Reflect in the x-axis the portion of $y = f(x)$ below the x-axis.

Example 4.8.2 (Graph of $y = |f(x)|$). Consider the following graph of $y = f(x)$.

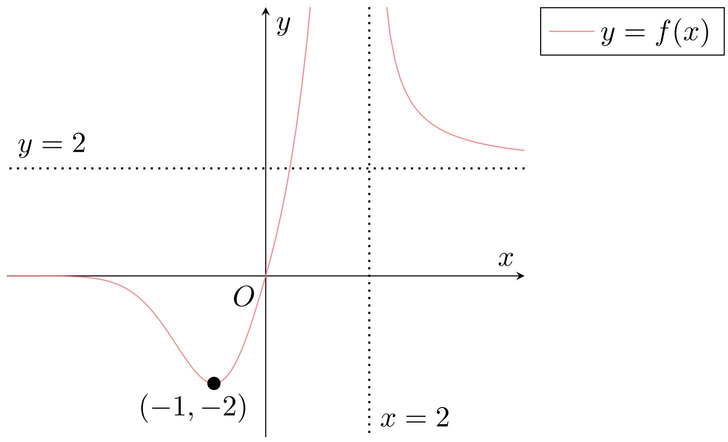


Figure 4.8

Reflecting the portion of the curve below the x-axis, we get the following graph of $y = |f(x)|$.

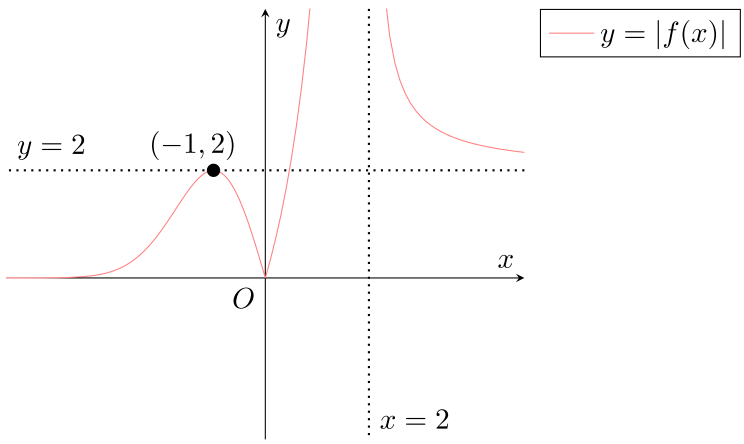


Figure 4.9

4.8.2 Graph of $y = f(|x|)$

Note that

$$y = f(|x|) = \begin{cases} f(x) & x \geq 0, \\ f(-x) & x < 0. \end{cases}$$

Recipe 4.8.3 (Graph of $y = f(|x|)$)**.** To obtain the graph of $y = f(|x|)$ from the graph of $y = f(x)$,

- Retain the portion of $y = f(x)$ where $x \geq 0$.
- Delete the portion of $y = f(x)$ where $x < 0$.
- Copy and reflect in the y-axis the portion of $y = f(x)$ where $x \geq 0$.

Example 4.8.4 (Graph of $y = f(|x|)$)**.** Let the graph of $y = f(x)$ be as in Fig. 4.8. Following the above steps, we see that the graph of $y = f(|x|)$ is

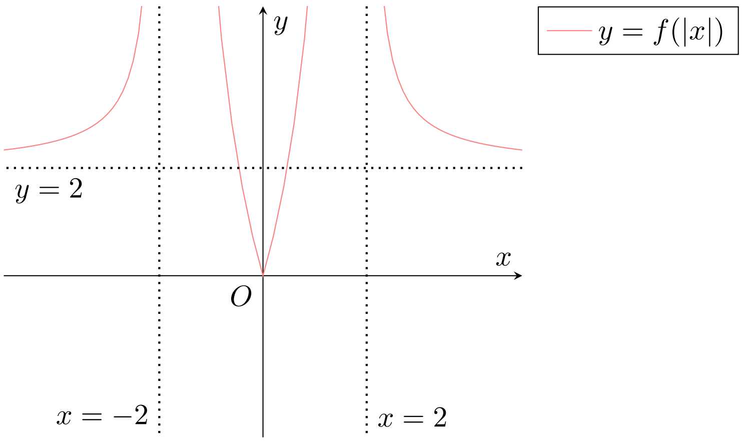


Figure 4.10

4.8.3 Graph of $y = 1/f(x)$

There are several key features and behaviours that we must note when drawing the graph of $y = 1/f(x)$.

- If $y = f(x)$ increases, $1/f(x)$ decreases and vice versa.
- For a minimum point (a, b) where $b \neq 0$ on the graph of $y = f(x)$, it corresponds to a maximum point $(a, 1/b)$ on the graph of $y = 1/f(x)$ and vice versa.
- For an x-intercept $(a, 0)$ on the graph of $y = f(x)$, it corresponds to a vertical asymptote $x = a$ on the graph of $y = 1/f(x)$ and vice versa.
- Oblique asymptotes on the graph of $y = f(x)$ become horizontal asymptotes at $y = 0$ on the graph of $y = 1/f(x)$.

Example 4.8.5 (Graph of $y = 1/f(x)$)**.** Let the graph of $y = f(x)$ be as in Fig. 4.8. Following the above pointers, we see that the graph of $y = 1/f(x)$ is

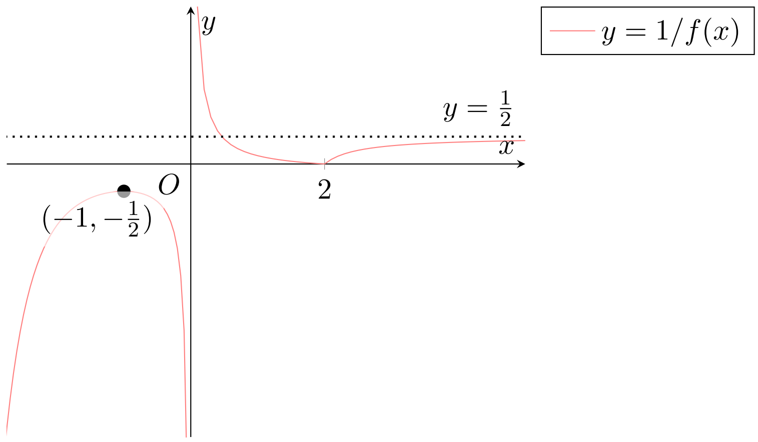


Figure 4.11

5 Polar Coordinates

5.1 Polar Coordinate System

Definition 5.1.1. Let the **pole** (or origin) be a point O in the plane. Let the **initial line** (or polar axis) be a half-line starting at O. Let P be any other point in the plane. Then P has polar coordinates (r, θ), where r is the distance from O to P and θ is the angle between the initial line and the line OP.

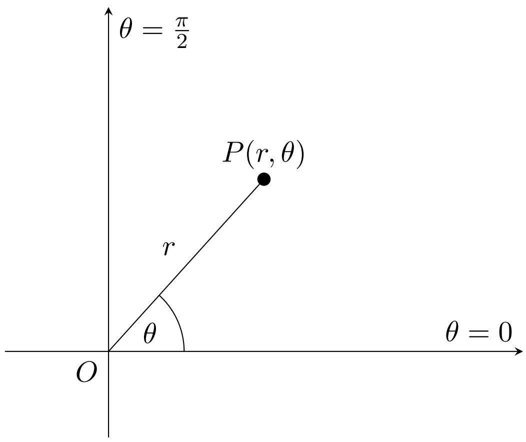


Figure 5.1

There are some conventions regarding the pole and the initial line.

- The initial line is usually drawn horizontally to the right.
- The polar angle θ is positive if measured in the anti-clockwise direction from the initial line and negative in the clockwise direction.
- If $P = 0$, then $r = 0$, and we may use $(0, \theta)$ to represent the pole for any value of θ.

Recall that in the Cartesian coordinate system, each point has a unique representation. This is not the case in the polar coordinate system. For example, the point $(1, \frac{5}{4}\pi)$ could also be written as $(1, \frac{13}{4}\pi)$ or as $(-1, \frac{1}{4}\pi)$. In general, because a complete anti-clockwise rotation is given by the angle 2π, the point (r, θ) can also be represented by $(r, \theta + 2n\pi)$ and $(-r, (2n+1)\pi)$, where n is any integer.

To avoid this ambiguity, it is common to restrict to $0 \leq \theta < 2\pi$ or $-\pi < \theta \leq \pi$ and to take $r \geq 0$.

5.2 Relationship between the Polar and Cartesian Coordinate Systems

Suppose the point P has Cartesian coordinates (x, y) and polar coordinates (r, θ). From the figure above, we have

$$\cos\theta = \frac{x}{r}, \quad \sin\theta = \frac{y}{r}.$$

Thus,

$$x = r\cos\theta, \quad y = r\sin\theta.$$

Note that while the above were deduced from the case where $r > 0$ and $0 < \theta < \frac{\pi}{2}$, these equations are valid for all values of r and θ.

From the figure, we also have

$$r^2 = x^2 + y^2, \quad \tan\theta = \frac{y}{x},$$

which allows us to find r and θ when x and y are known.

5.3 Polar Curves

Definition 5.3.1. The **graph of a polar equation** $r = f(\theta)$ consists of all points $P(r, \theta)$ whose coordinates satisfy the equation.

Fact 5.3.2 (Symmetry of Polar Curves)**.**

- If the equation is invariant under $\theta \mapsto -\theta$, the curve is symmetric about the polar axis.
- If the equation is invariant under $r \mapsto -r$, or when $\theta \mapsto \theta + \pi$, the curve is symmetric about the pole (i.e. the curve remains unchanged when rotated by 180° about the origin).
- If the equation is invariant when $\theta \mapsto \pi - \theta$, the curve is symmetric about the vertical line $\theta = \frac{\pi}{2}$.
- If r is a function of $\cos n\theta$ only, the curve is symmetric about the horizontal half lines $\theta = \frac{k}{n}\pi$, $k \in \mathbb{Z}$.
- If r is a function of $\sin n\theta$ only, the curve is symmetric about the vertical half-lines $\theta = \frac{2k+1}{2n}\pi$, $k \in \mathbb{Z}$.
- If only even powers of r occur in the equation, the curve is symmetric about the pole.

Proposition 5.3.3 (Tangents to Polar Curves)**.** The gradient of the tangent to a polar curve $r = f(\theta)$ at any point is

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{r'\sin\theta + r\cos\theta}{r'\cos\theta - r\sin\theta}.$$

Proof. Recall that

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

Differentiating with respect to θ,

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = r'\cos\theta - r\sin\theta, \qquad \frac{\mathrm{d}y}{\mathrm{d}\theta} = r'\sin\theta + r\cos\theta.$$

Thus,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = \frac{r' \sin\theta + r\cos\theta}{r'\cos\theta - r\sin\theta}.$$

□

Remark. To find horizontal tangents (i.e. $\mathrm{d}y/\mathrm{d}x = 0$), we can solve $\mathrm{d}y/\mathrm{d}\theta = 0$ (provided $\mathrm{d}x/\mathrm{d}\theta \neq 0$). Likewise, to find vertical tangents (i.e. $\mathrm{d}y/\mathrm{d}x$ undefined), we can solve $\mathrm{d}x/\mathrm{d}\theta = 0$ (provided $\mathrm{d}y/\mathrm{d}\theta \neq 0$). Lastly, if we are looking for tangent lines at the pole, where $r = 0$, the equation simplifies to

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \tan\theta,$$

provided $\mathrm{d}r/\mathrm{d}\theta \neq 0$.

Part II

Sequences and Series

6 Sequences and Series

6.1 Sequences

Definition 6.1.1. A **sequence** or **progression** is a set of numbers $u_1, u_2, u_3, \ldots, u_n, \ldots$ arranged in a defined order according to a certain rule. In general, u_n is called the **nth term**.

Remark. A sequence can be thought of as a function with domain $\mathbb{Z}^+$.

Definition 6.1.2. A sequence is said to be **finite** if it terminates; otherwise it is an **infinite sequence**.

Definition 6.1.3. If an infinite sequence u_n approaches a unique value l as $n \to \infty$, then the sequence is said to **converge** to l. We say that l is the **limit** of u_n. A sequence that does not converge is said to **diverge**.

When describing sequences, one should identify

- Trends (increasing/decreasing, constant, alternating)
- Long-run behaviour of an infinite sequence (convergent or divergent)

6.2 Series

Definition 6.2.1. A **series** is the sum of the terms of a sequence u_n. The sum to n terms is denoted by S_n, i.e.

$$S_n = u_1 + u_2 + \cdots + u_{n-1} + u_n.$$

Similar to sequences, a series can be finite or infinite. If a series is infinite, it can further be categorized as convergent or divergent.

6.3 Arithmetic Progression

Definition 6.3.1. An **arithmetic progression** (AP) is a sequence u_n in which each term differs from the preceding term by a constant called the **common difference**. The first term of an AP is usually denoted by a and the common difference by d. Mathematically,

$$u_n = a + (n-1)d.$$

Definition 6.3.2. An **arithmetic series** is obtained by adding the terms of an arithmetic progression.

Proposition 6.3.3. The nth term S_n of an arithmetic series is given by

$$S_n = \frac{n(a+l)}{2},$$

where l is the last term of the AP, i.e.

$$l = u_n = a + (n-1)d.$$

Proof. Note that for all integers $k \in [1, n]$,

$$u_k + u_{n-k+1} = [a + (k-1)d] + [a + (n-k)d] = a + [a + (n-1)d] = a + l.$$

Hence, by pairing the kth term with the $(n-k+1)$th term, we get

$$2S_n = (u_1+u_n)+(u_2+u_{n-1})+\cdots+(u_{n-1}+u_2)+(u_n+u_1) = n\,(a+l) \implies S_n = \frac{n(a+l)}{2}.$$

□

6.4 Geometric Progression

Definition 6.4.1. A **geometric progression** (GP) is a sequence u_n in which each term is obtained form the preceding one by multiplying a non-zero constant, called the **common ratio**. The first term of a GP is usually denoted by a and the common ratio by r. Mathematically,

$$u_n = ar^{n-1}.$$

Remark. In the case where $r = 1$, the geometric progression becomes an arithmetic progression.

Definition 6.4.2. A **geometric series** is the sum of the terms of a geometric progression.

Proposition 6.4.3. The nth term S_n of a geometric series is given by

$$S_n = \frac{a(1-r^n)}{1-r},$$

where $r \neq 1$. If the series is infinite, the sum to infinity S_∞ exists only if $|r| < 1$ and is given by

$$S_\infty = \frac{a}{1-r}.$$

Proof. By the definition of a series, we have

$$S_n = a + ar + \cdots + ar^{n-2} + ar^{n-1}. \tag{1}$$

Multiplying both sides by r yields

$$rS_n = ar + ar^2 + \cdots + ar^{n-1} + ar^n. \tag{2}$$

Subtracting (2) from (1), we have

$$(1-r)S_n = a - ar^n \implies S_n = \frac{a(1-r^n)}{1-r}.$$

Suppose $|r| < 1$. In the limit as $n \to \infty$, we have $r^n \to 0$. Hence,

$$S_\infty = \frac{a(1-0)}{1-r} = \frac{a}{1-r}.$$

□

6.5 Sigma Notation

Definition 6.5.1. The series $u_k + u_{k+1} + \cdots + u_m$ can be denoted using Σ (sigma) notation as

$$u_k + u_{k+1} + \cdots + u_m = \sum_{r=k}^{m} u_r.$$

Here, r is called the **index**, and can be replaced with any letter. k is the **lower limit** of r, while m is the **upper limit** of r. There are a total of $m - k + 1$ terms in the sum.

Fact 6.5.2 (Properties of Sigma Notation)**.**

$$\sum_{r=1}^{n} (u_r \pm v_r) = \sum_{r=1}^{n} u_r \pm \sum_{r=1}^{n} v_r.$$

$$\sum_{r=1}^{n} cu_r = c \sum_{r=1}^{n} u_r.$$

$$\sum_{r=m}^{n} u_r = \sum_{r=1}^{n} u_r - \sum_{r=1}^{m-1} u_r, \quad n > m > 1.$$

Fact 6.5.3 (Standard Series)**.** The sum of the following standard series can be quoted and applied without proof. Note that $m = q - p + 1$ is the number of terms being summed.

- Series of constants
$$\sum_{r=p}^{q} a = ma.$$
- Arithmetic series
$$\sum_{r=p}^{q} r = \frac{m}{2} (p + q).$$
- Geometric series
$$\sum_{r=p}^{q} a^r = \frac{a^p (a^m - 1)}{a - 1}.$$

7 Recurrence Relations

Definition 7.0.1. A **recurrence relation** is an equation that defines a sequence based on a rule that gives the next term as a function of the previous term(s).

7.1 First Order Linear Recurrence Relation with Constant Coefficients

Definition 7.1.1. A **first order linear recurrence relation with constant coefficients** is a recurrence relation of the form

$$u_n = au_{n-1} + b,$$

where a and b are constants. If $b = 0$, the recurrence relation is said to be **homogeneous**.

There are two main ways to solve the above recurrence relation: by converting the recurrence relation into a geometric progression, or solving by procedure.

7.1.1 Converting to Geometrical Progression

Recipe 7.1.2 (Converting to Geometrical Progression)**.** Let k be the constant such that

$$u_n + k = a\left(u_{n-1} + k\right).$$

Then we clearly have $k = \frac{b}{a-1}$. We now define a new sequence $v_n = u_n + k$. This turns our recurrence relation into

$$v_n = av_{n-1},$$

whence v_n is in geometric progression. Thus, $v_n = v_1 a^{n-1}$. Writing this back in terms of u_n, we get

$$u_n + k = (u_1 + k)a^{n-1} \implies u_n = (u_1 + k)a^{n-1} - k.$$

Example 7.1.3 (Solving by GP)**.** Consider the recurrence relation

$$u_1 = 0, \quad u_n = \frac{1}{2}u_{n-1} + 10, \quad n > 1.$$

Let k be the constant such that

$$u_n + k = \frac{1}{2}\left(u_{n-1} + k\right).$$

Then

$$k = \frac{10}{1/2 - 1} = -20.$$

We hence have

$$u_n - 20 = \frac{1}{2}\left(u_{n-1} - 20\right),$$

whence the sequence $\{u_n - 20\}$ is in geometric progression with common ratio $1/2$.

Thus,

$$u_n - 20 = (u_1 - 20)\left(\frac{1}{2}\right)^{n-1}.$$

Rearranging, we obtain the solution

$$u_n = -20\left(\frac{1}{2}\right)^{n-1} + 20 = -40\left(\frac{1}{2}\right)^{n} + 20.$$

7.1.2 Solving by Procedure

Definition 7.1.4. Given a first order linear recurrence relation with constant coefficients $u_n = au_{n-1} + b$,

- $u_n = au_{n-1}$ is the **associated homogeneous recurrence relation**.
- $u_n^{(c)} = Ca^n$ is the general solution of the associated homogeneous recurrence relation and is called the **complementary solution**.
- $u_n^{(p)} = k$ is the **particular solution** to the recurrence relation.

Fact 7.1.5 (Solving by Procedure)**.** The general solution is given by

$$u_n = u_n^{(c)} + u_n^{(p)} = Ca^n + k.$$

Example 7.1.6 (Solving by Procedure)**.** Consider the recurrence relation

$$u_1 = 0, \quad u_n = \frac{1}{2}u_{n-1} + 10, \quad n > 1.$$

Observe that the associated homogeneous recurrence relation is $u_n = \frac{1}{2}u_{n-1}$. Hence, the complementary solution is

$$u_n^{(c)} = C\left(\frac{1}{2}\right)^n$$

for some arbitrary constant C. Let the particular solution be $u_n^{(p)} = k$. Then

$$k = \frac{1}{2}k + 10 \implies k = 20.$$

Hence, the general solution is

$$u_n = u_n^{(c)} + u_n^{(p)} = C\left(\frac{1}{2}\right)^n + 20.$$

Using the initial condition $u_1 = 0$, we have

$$0 = C\left(\frac{1}{2}\right)^1 + 20 \implies C = -40.$$

Thus,

$$u_n = -40\left(\frac{1}{2}\right)^n + 20.$$

7.2 Second Order Linear Homogeneous Recurrence Relation with Constant Coefficients

Definition 7.2.1. A **second order linear homogeneous recurrence relation with constant coefficients** is a recurrence relation of the form

$$u_n = au_{n-1} + bu_{n-2},$$

where a and b are constants.

Recipe 7.2.2 (Solving by Procedure). To solve the recurrence relation

$$u_n = au_{n-1} + bu_{n-2},$$

1. Form the quadratic equation

$$x^2 - ax - b = 0.$$

   This is called the **characteristic equation**.

2. Find the roots α and β of this characteristic equation.

3. Then u_n has the **general solution**
   - $u_n = A\alpha^n + B\beta^n$, if $\alpha \neq \beta$ (distinct roots, may be real or non-real).
   - $u_n = (A + Bn)\alpha^n$, if $\alpha = \beta$ (real and equal roots).
   - $u_n = Ar^n \cos n\theta + Br^n \sin n\theta$, if $\alpha = r\mathrm{e}^{\mathrm{i}\theta}$ and $\beta = r\mathrm{e}^{-\mathrm{i}\theta}$ (non-real roots).

Proof. For $u_{n+1} = pu_n + qu_{n-1}$ with given initial conditions u_1 and u_2, let the constant k be such that

$$u_{n+1} - ku_n = (p-k)(u_n - ku_{n-1}). \tag{1}$$

Note that this is a GP. Comparing coefficients of u_{n-1}, we get

$$(p-k)k = -q \implies k^2 - pk - q = 0.$$

This is the characteristic equation. Let the roots to the characteristic equation be $k = \alpha$ and $k = \beta$. By Vieta's formulas,

$$\alpha + \beta = -\left(\frac{-p}{1}\right) = p.$$

Now, using the fact that (1) is in GP, we get

$$u_{n+1} - ku_n = (p-k)^{n-1}(u_2 - ku_1). \tag{2}$$

Substituting $k = \alpha$ into (2), we obtain

$$u_{n+1} - \alpha u_n = \beta^{n-1}\left(u_2 - \alpha u_1\right). \tag{3a}$$

Substituting $k = \beta$ into (2), we obtain

$$u_{n+1} - \beta u_n = \alpha^{n-1}\left(u_2 - \beta u_1\right). \tag{3b}$$

We now analyse the case where $\alpha = \beta$ and $\alpha \neq \beta$ separately.

Case 1: $\alpha = \beta$. Since the two roots are equal, (3a) and (3b) are equivalent. Taking either,

$$u_{n+1} - \alpha u_n = \alpha^{n-1}(u_2 - \alpha u_1) \implies \frac{u_{n+1}}{\alpha^{n-1}} - \frac{u_n}{\alpha^{n-2}} = u_2 - \alpha u_1.$$

The sequence $\left\{\frac{u_n}{\alpha^{n-2}}\right\}$ is hence in AP with common difference $u_2 - \alpha u_1$. Invoking the closed form for AP, we obtain

$$\frac{u_n}{\alpha^{n-2}} = \frac{u_1}{\alpha^{-1}} + (n-1)(u_2 - \alpha u_1) \implies u_n = \alpha^{n-2}\left(\frac{u_1}{\alpha^{-1}} + (n-1)(u_2 - \alpha u_1)\right).$$

Simplifying,

$$u_n = \left[\left(\frac{2u_1}{\alpha} - \frac{u_2}{\alpha^2}\right) + \left(\frac{u_2}{\alpha^2} - \frac{u_1}{\alpha}\right) n\right] \alpha^n = (A + Bn)\alpha^n.$$

Case 2: $\alpha \neq \beta$. Observe that $\frac{(3b)-(3a)}{\alpha-\beta}$ yields

$$u_n = \frac{\alpha^{n-1}(u_2 - \beta u_1) - \beta^{n-1}(u_2 - \alpha u_1)}{\alpha - \beta}.$$

Simplifying, we have

$$u_n = \left[\frac{u_2 - \beta u_1}{\alpha(\alpha - \beta)}\right] \alpha^n + \left[\frac{u_2 - \alpha u_1}{\beta(\beta - \alpha)}\right] \beta^n = A\alpha^n + B\beta^n.$$

We now consider the case where α and β are non-real. By the conjugate root theorem, we can write $\alpha = r\mathrm{e}^{\mathrm{i}\theta}$ and $\beta = r\mathrm{e}^{-\mathrm{i}\theta}$. Substituting this into the above result, we have

$$u_n = A\left(r\mathrm{e}^{\mathrm{i}\theta}\right)^n + B\left(r\mathrm{e}^{-\mathrm{i}\theta}\right)^n = r^n\left(A\mathrm{e}^{\mathrm{i}n\theta} + B\mathrm{e}^{-\mathrm{i}n\theta}\right).$$

By Euler's identity,

$$u_n = r^n\left[(A+B)\cos n\theta + \mathrm{i}(A-B)\sin n\theta\right] = Cr^n \cos n\theta + Dr^n \sin n\theta.$$

$\square$

Part III

Vector Geometry and Linear Algebra

8 Vectors

8.1 Basic Definitions and Notations

Definition 8.1.1. A **vector** is an object that has both magnitude and direction. Geometrically, we can represent a vector by a **directed** line segment $\overrightarrow{PQ}$, where the length of the line segment represents the magnitude of the vector, and the direction of the line segment represents the direction of the vector. Vectors are typically denoted by bold print (e.g. $\mathbf{a}$) or by $\overrightarrow{PQ}$.

Definition 8.1.2. The **magnitude** of a vector $\mathbf{a}$ is the length of the line representing $\mathbf{a}$, and is denoted by $|\mathbf{a}|$.

Definition 8.1.3. Two vectors $\mathbf{a}$ and $\mathbf{b}$ are said to be **equal vectors** if they both have the same magnitude and direction. $\mathbf{a}$ and $\mathbf{b}$ are said to be **negative vectors** if they have the same magnitude but opposite directions.

Definition 8.1.4 (Multiplication of a Vector by a Scalar)**.** Let λ be a scalar. If $\lambda > 0$, then $\lambda\mathbf{a}$ is a vector of magnitude $\lambda\,|\mathbf{a}|$ and has the same direction as $\mathbf{a}$. If $\lambda < 0$, then $\lambda\mathbf{a}$ is a vector of magnitude $-\lambda\,|\mathbf{a}|$ and is in the opposite direction of $\mathbf{a}$.

Definition 8.1.5. The **zero vector** is the vector with a magnitude of 0 and is denoted $\mathbf{0}$.

Definition 8.1.6. Let $\mathbf{a}$ and $\mathbf{b}$ be non-zero vectors. Then $\mathbf{a}$ and $\mathbf{b}$ are said to be **parallel** if and only if $\mathbf{b}$ can be expressed as a non-zero scalar multiple of $\mathbf{a}$. Mathematically,

$$\mathbf{a} \parallel \mathbf{b} \iff (\exists \lambda \in \mathbb{R} \setminus \{0\}): \quad \mathbf{b} = \lambda\mathbf{a}.$$

Definition 8.1.7. A **unit vector** is a vector with a magnitude of 1. Unit vectors are typically denoted with a hat (e.g. $\hat{\mathbf{a}}$).

Observe that for any non-zero vector $\mathbf{a}$, the unit vector parallel to $\mathbf{a}$ is given by

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|}.$$

Definition 8.1.8. The **Triangle Law of Vector Addition** states that

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}.$$

Geometrically, we add two vectors $\mathbf{a}$ and $\mathbf{b}$ by placing them head to tail, taking the resultant vector as their sum.

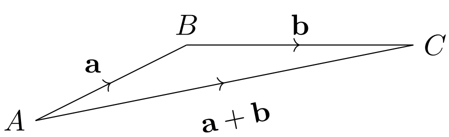


Figure 8.1

We subtract vectors by adding $\mathbf{a} + -(\mathbf{b})$.

Definition 8.1.9. The **angle between two vectors** refers to the angle between their directions when the arrows representing them *both converge* or *both diverge*.

Definition 8.1.10. A **free vector** is a vector that has no specific location in space. The **position vector** of some point A relative to the origin O is unique and is denoted $\overrightarrow{OA}$. A **displacement vector** is a vector that joins its initial position to its final position. For instance, $\overrightarrow{OA}$ is the displacement vector from O to A.

Definition 8.1.11. A set of vectors are said to be **coplanar** if their directions are all parallel to the same plane.

Fact 8.1.12. Any vector $\mathbf{c}$ that is coplanar with $\mathbf{a}$ and $\mathbf{b}$ can be expressed as a **unique linear combination** of $\mathbf{a}$ and $\mathbf{b}$, i.e.

$$(\exists!\, \lambda, \mu \in \mathbb{R}): \quad \mathbf{c} = \lambda\mathbf{a} + \mu\mathbf{b}.$$

Theorem 8.1.13 (Ratio Theorem)**.** If P divides AB in the ratio $\lambda : \mu$, then

$$\overrightarrow{OP} = \frac{\mu\mathbf{a} + \lambda\mathbf{b}}{\lambda + \mu}.$$

Proof. Since P divides AB in the ratio $\lambda : \mu$, we have

$$\overrightarrow{AP} = \frac{\lambda}{\lambda+\mu}\overrightarrow{AB} = \frac{\lambda}{\lambda+\mu}(\mathbf{b} - \mathbf{a}).$$

Thus,

$$\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{AP} = \mathbf{a} + \frac{\lambda}{\lambda+\mu}(\mathbf{b} - \mathbf{a}) = \frac{\mu\mathbf{a} + \lambda\mathbf{b}}{\lambda+\mu}.$$

□

Corollary 8.1.14 (Mid-Point Theorem)**.** If P is the mid-point of AB, then

$$\overrightarrow{OP} = \frac{\mathbf{a} + \mathbf{b}}{2}.$$

8.2 Vector Representation using Cartesian Unit Vectors

8.2.1 2-D Cartesian Unit Vectors

Definition 8.2.1 (2-D Cartesian Unit Vectors)**.** In the 2-D Cartesian plane, $\mathbf{i} = (1,\ 0)^\mathsf{T}$ is defined to be the unit vector in the positive direction of the x-axis, while $\mathbf{j} = (0,\ 1)^\mathsf{T}$ is defined to be the unit vector in the positive direction of the y-axis.

Thus, if P is the point with coordinates (a, b), then we can express $\overrightarrow{OP}$ in terms of the unit vectors $\mathbf{i}$ and $\mathbf{j}$. In particular, $\overrightarrow{OP} = a\mathbf{i} + b\mathbf{j}$.

Proposition 8.2.2 (Magnitude in 2-D)**.**

$$\left|\begin{pmatrix} a \\ b \end{pmatrix}\right| = \sqrt{a^2 + b^2}.$$

Proof. Follows immediately from Pythagoras' theorem. □

8.2.2 3-D Cartesian Unit Vectors

Definition 8.2.3 (3-D Cartesian Unit Vectors)**.** In the 3-D Cartesian plane, $\mathbf{i} = (1,\ 0,\ 0)^\mathsf{T}$, $\mathbf{j} = (0,\ 1,\ 0)^\mathsf{T}$ and $\mathbf{k} = (0,\ 0,\ 1)^\mathsf{T}$ denote the unit vectors in the positive direction of the x, y and z-axes respectively.

Proposition 8.2.4 (Magnitude in 3-D)**.**

$$\left| \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right| = \sqrt{a^2 + b^2 + c^2}.$$

Proof. Use Pythagoras' theorem twice. □

Fact 8.2.5 (Operations on Cartesian Vectors)**.** To add vectors given in Cartesian unit vector form, the coefficients of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are added separately.

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}.$$

Subtraction and scalar multiplication follows immediately.

8.3 Scalar Product

Definition 8.3.1. The **scalar product** (or dot product) of two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined by

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

where θ is the angle between the two vectors (note that $0 \le \theta \le \pi$).

Remark. $\mathbf{a} \cdot \mathbf{b}$ is called the scalar product as the result is a real number (a scalar). It is also called the dot product because of the notation.

Fact 8.3.2 (Algebraic Properties of Scalar Product)**.** Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be vectors and let $\lambda \in \mathbb{R}$. Then

- (commutative) $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$.
- (distributive over addition) $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$.
- $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$.
- $(\lambda\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\lambda\mathbf{b}) = \lambda(\mathbf{a} \cdot \mathbf{b})$.

Proposition 8.3.3 (Geometric Properties of Scalar Product)**.** Let $\mathbf{a}$ and $\mathbf{b}$ be non-zero vectors, and let θ be the angle between them.

- $\mathbf{a} \cdot \mathbf{b} = 0$ if and only if $\theta = \frac{\pi}{2}$, i.e. $\mathbf{a} \perp \mathbf{b}$.
- $\mathbf{a} \cdot \mathbf{b} > 0$ if and only if θ is acute.
- $\mathbf{a} \cdot \mathbf{b} < 0$ if and only if θ is obtuse.

Proof. The sign of $\mathbf{a} \cdot \mathbf{b}$ is determined solely by $\cos\theta$. □

Proposition 8.3.4 (Scalar Product in Cartesian Unit Vector Form).

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = x_1x_2 + y_1y_2 + z_1z_2.$$

Proof. Since $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are pairwise perpendicular, their pairwise scalar products are 0. That is,

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0.$$

Hence, by the distributive property of the scalar product,

$$(x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}) \cdot (x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}) = x_1x_2\mathbf{i} \cdot \mathbf{i} + y_1y_2\mathbf{j} \cdot \mathbf{j} + z_1z_2\mathbf{k} \cdot \mathbf{k}.$$

Lastly, since $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are all unit vectors,

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1.$$

Thus,

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = x_1x_2 + y_1y_2 + z_1z_2.$$

□

8.3.1 Applications of Scalar Product

Proposition 8.3.5 (Angle between Two Vectors). Let θ be the angle between two non-zero vectors $\mathbf{a}$ and $\mathbf{b}$. Then

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|}.$$

Proof. Follows immediately from the definition of the scalar product. □

Definition 8.3.6. Let $\mathbf{a}$ and $\mathbf{b}$ denote the position vectors of A and B respectively, relative to the origin O. Let θ be the angle between $\mathbf{a}$ and $\mathbf{b}$, and let N be the foot of the perpendicular from the point A to the line passing through O and B.

Then, the length ON is defined to be the **length of projection** of the vector $\mathbf{a}$ onto the vector $\mathbf{b}$. Also, $\overrightarrow{ON}$ is the **vector projection** of $\mathbf{a}$ onto $\mathbf{b}$.

Proposition 8.3.7 (Length of Projection). The length of projection of $\mathbf{a}$ onto $\mathbf{b}$ is $\left|\mathbf{a} \cdot \hat{\mathbf{b}}\right|$.

Proof. Consider the case where θ is acute.

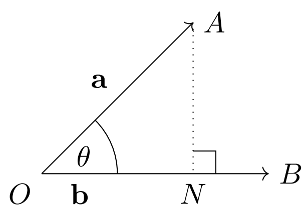


Figure 8.2

From the diagram,

$$ON = OA\cos\theta = |\mathbf{a}|\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|} = \mathbf{a} \cdot \hat{\mathbf{b}}.$$

A similar argument shows that when θ is obtuse, $ON = -\mathbf{a} \cdot \hat{\mathbf{b}}$. Hence, in any case, $ON = \left|\mathbf{a} \cdot \hat{\mathbf{b}}\right|$. ☐

Proposition 8.3.8 (Vector Projection)**.** The vector projection of $\mathbf{a}$ onto $\mathbf{b}$ is $(\mathbf{a} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}}$.

Proof. Case 1: θ is acute. Then $\overrightarrow{ON}$ is in the same direction as $\mathbf{b}$. Hence,

$$\overrightarrow{ON} = |ON|\,\hat{\mathbf{b}} = (\mathbf{a} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}}.$$

Case 2: θ is obtuse. Then $\overrightarrow{ON}$ is in the opposite direction as $\mathbf{b}$. Hence,

$$\overrightarrow{ON} = |ON|\left(-\hat{\mathbf{b}}\right) = -(\mathbf{a} \cdot \hat{\mathbf{b}})(-\hat{\mathbf{b}}) = (\mathbf{a} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}}.$$

☐

8.4 Vector Product

Definition 8.4.1. The **vector product** (or cross product) of two vectors $\mathbf{a}$ and $\mathbf{b}$ is denoted by $\mathbf{a} \times \mathbf{b}$ and is defined by

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

where θ is the angle between $\mathbf{a}$ and $\mathbf{b}$ and $\hat{\mathbf{n}}$ is the unit vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$, in the direction determined by the right-hand grip rule.

Remark. $\mathbf{a} \times \mathbf{b}$ is called the vector product as the result is a vector. It is also called the cross product due to its notation.

Fact 8.4.2 (Algebraic Properties of Vector Product)**.** Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be three vectors, and θ be the angle between $\mathbf{a}$ and $\mathbf{b}$.

- (anti-commutative) $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$.
- (distributive over addition) $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{c})$.
- $|\mathbf{a} \times \mathbf{b}| = |a|\,|b|\sin\theta$.
- $(\lambda\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda\mathbf{b}) = \lambda(\mathbf{a} \times \mathbf{b})$, where $\lambda \in \mathbb{R}$.

Proposition 8.4.3 (Geometric Properties of Vector Product)**.** Let $\mathbf{a}$ and $\mathbf{b}$ be non-zero vectors and θ be the angle between them.

- $|\mathbf{a} \times \mathbf{b}| = 0$ if and only if $\mathbf{a} \parallel \mathbf{b}$.
- $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|$ if and only if $\mathbf{a} \perp \mathbf{b}$.

Proof. Follows from the definition of the vector product (consider $\theta = 0, \frac{\pi}{2}, \pi$). ☐

Proposition 8.4.4 (Vector Product in Cartesian Unit Vector Form)**.**

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} y_1z_2 - z_1y_2 \\ z_1x_2 - x_1z_2 \\ x_1y_2 - y_1x_2 \end{pmatrix}.$$

Proof. From the geometric properties of the vector product, we have

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}.$$

Furthermore, since $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are pairwise perpendicular, by the right-hand grip rule, one has

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}.$$

Hence, by the distributive property of the vector product,

$$\begin{aligned}
&(x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}) \times (x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}) \\
&\quad = x_1y_2\mathbf{k} + x_1z_2(-\mathbf{j}) + y_1x_2(-\mathbf{k}) + y_1z_2\mathbf{i} + z_1x_2\mathbf{j} + z_1y_2(-\mathbf{i}) \\
&\quad = (y_1z_2 - z_1y_2)\mathbf{i} + (z_1x_2 - x_1z_2)\mathbf{j} + (x_1y_2 - y_1x_2)\mathbf{k}.
\end{aligned}$$

□

8.4.1 Applications of Vector Product

Proposition 8.4.5 (Length of Side of Right-Angled Triangle)**.** Let $\mathbf{a}$ and $\mathbf{b}$ denote the position vectors of A and B respectively, relative to the origin O. Let θ be the angle between $\mathbf{a}$ and $\mathbf{b}$, and let N be the foot of the perpendicular from A to OB. Then

$$AN = \left|\mathbf{a} \times \hat{\mathbf{b}}\right|.$$

Proof. With reference to Fig. 8.2, we have

$$AN = OA\sin\theta = |\mathbf{a}|\frac{|\mathbf{a} \times \mathbf{b}|}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{|\mathbf{a} \times \mathbf{b}|}{\mathbf{b}} = \left|\mathbf{a} \times \hat{\mathbf{b}}\right|.$$

□

Proposition 8.4.6 (Area of Triangles and Parallelogram)**.** Let $ABCD$ be a parallelogram, let $\mathbf{a} = \overrightarrow{AB}$ and $\mathbf{b} = \overrightarrow{AC}$, and let θ be the angle between $\mathbf{a}$ and $\mathbf{b}$. Then

$$[\triangle ABC] = \frac{1}{2}|\mathbf{a} \times \mathbf{b}|$$

and

$$[ABCD] = |\mathbf{a} \times \mathbf{b}|.$$

Proof. Recall that the formula for the area of a triangle is

$$[\triangle ABC] = \frac{1}{2}(AB)(AC)\sin\theta = \frac{1}{2}|\mathbf{a}|\,|\mathbf{b}|\sin\theta = \frac{1}{2}|\mathbf{a} \times \mathbf{b}|.$$

Since the area of parallelogram $ABCD$ is twice that of $\triangle ABC$, we immediately have

$$[ABCD] = |\mathbf{a} \times \mathbf{b}|.$$

□

9 Three-Dimensional Vector Geometry

9.1 Lines

9.1.1 Equation of a Line

Definition 9.1.1. The **vector equation** of the line l passing through point A with position vector $\mathbf{a}$ and parallel to $\mathbf{b}$ is given by

$$\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}, \quad \lambda \in \mathbb{R},$$

where $\mathbf{r}$ is the position vector of any point on the line, and λ is a real, scalar parameter. The vector $\mathbf{b}$ is also called the **direction vector** of the line.

Remark. Note that $\mathbf{a}$ can be any position vector on the line and $\mathbf{b}$ can be any vector parallel to the line. Hence, the vector equation of a line is not unique.

Definition 9.1.2. Let $l : \mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, $\lambda \in \mathbb{R}$. By writing $\mathbf{r} = (x,\, y,\, z)^\mathsf{T}$, $\mathbf{a} = (a_1,\, a_2,\, a_3)^\mathsf{T}$ and $\mathbf{b} = (b_1,\, b_2,\, b_3)^\mathsf{T}$, we have

$$\begin{cases} x = a_1 + \lambda b_1 \\ y = a_2 + \lambda b_2, \quad \lambda \in \mathbb{R}. \\ z = a_3 + \lambda b_3 \end{cases}$$

This set of three equations is known as the **parametric equations** of the line l.

Definition 9.1.3. From the parametric form of the line l, by making λ the subject, we have

$$\lambda = \frac{x - a_1}{b_1} = \frac{y - a_2}{b_2} = \frac{z - a_3}{b_3}.$$

This equation is known as the **Cartesian equation** of the line l.

Remark. If $b_1 = 0$, we simply have $x = a_1$. A similar result arises when $b_2 = 0$ or $b_3 = 0$.

9.1.2 Point and Line

Proposition 9.1.4 (Relationship between Point and Line)**.** A point C lies on a line $l : \mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, $\lambda \in \mathbb{R}$, if and only if

$$(\exists\lambda \in \mathbb{R}) : \quad \overrightarrow{OC} = \mathbf{a} + \lambda\mathbf{b}.$$

Proof. Trivial. $\square$

Proposition 9.1.5 (Perpendicular Distance between Point and Line)**.** Let C be a point not on the line $l : \mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, $\lambda \in \mathbb{R}$. Let F be the foot of perpendicular from C to l. Then

$$CF = \left|\overrightarrow{AC} \times \hat{\mathbf{b}}\right|.$$

Proof. Trivial (recall the application of the vector product in finding side lengths of right-angled triangles). $\square$

Recipe 9.1.6 (Finding Foot of Perpendicular from Point to Line). Let F be the foot of perpendicular from C to the line $l : \mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, $\lambda \in \mathbb{R}$. To find $\overrightarrow{OF}$, we use the fact that

- F lies on l, i.e. $\overrightarrow{OF} = \mathbf{a} + \lambda\mathbf{b}$ for some $\lambda \in \mathbb{R}$.
- $\overrightarrow{CF}$ is perpendicular to l, i.e. $\overrightarrow{CF} \cdot \mathbf{b} = 0$.

9.1.3 Two Lines

Definition 9.1.7. The relationship between two lines in 3-D space can be classified as follows:

- **Parallel lines**: The lines are parallel and non-intersecting;
- **Intersecting lines**: The lines are non-parallel and intersecting;
- **Skew lines**: The lines are non-parallel and non-intersecting.

Remark. Note that parallel and intersecting lines are coplanar, while skew lines are non-coplanar.

Recipe 9.1.8 (Relationship between Two Lines). Consider two distinct lines, $l_1 : \mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, $\lambda \in \mathbb{R}$ and $l_2 : \mathbf{r} = \mathbf{c} + \mu\mathbf{d}$, $\mu \in \mathbb{R}$.

- l_1 and l_2 are parallel lines if their direction vectors are parallel.
- l_1 and l_2 are intersecting lines if there are unique values of λ and μ such that $\mathbf{a} + \lambda\mathbf{b} = \mathbf{c} + \mu\mathbf{d}$.
- l_2 and l_2 are skew lines if their direction vectors are not parallel and there are no values of λ and μ such that $\mathbf{a} + \lambda\mathbf{b} = \mathbf{c} + \mu\mathbf{d}$.

Proposition 9.1.9 (Acute Angle between Two Lines). Let the acute angle between two lines with direction vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ be θ. Then

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

Proof. Observe that we are essentially finding the angle between the direction vectors of the two lines, which is given by

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

However, to ensure that θ is acute (i.e. $\cos\theta \geq 0$), we introduce a modulus sign in the numerator. Hence,

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

□

9.2 Planes

9.2.1 Equation of a Plane

Definition 9.2.1. Suppose the plane π passes through a fixed point A with position vector $\mathbf{a}$, and π is parallel to two vectors $\mathbf{b}_1$ and $\mathbf{b}_2$, where $\mathbf{b}_1$ and $\mathbf{b}_2$ are not parallel to each other. Then the vector equation (in **parametric form**) of π is given by

$$\pi : \mathbf{r} = \mathbf{a} + \lambda \mathbf{b}_1 + \mu \mathbf{b}_2,$$

where $\mathbf{r}$ is the position vector of any point P on π, and λ and μ are real parameters.

Definition 9.2.2. Suppose the plane π passes through a fixed point A with position vector $\mathbf{a}$, and π has normal vector $\mathbf{n}$. Let P be an arbitrary point on π. Then $\overrightarrow{AP}$ is perpendicular to the normal vector $\mathbf{n}$, i.e. $\overrightarrow{AP} \cdot \mathbf{n} = 0$. Since $\overrightarrow{AP} = \mathbf{r} - \mathbf{a}$, by the distributivity of the scalar product, one has

$$\mathbf{r} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n}.$$

This is the **scalar product form** of the vector equation of π, which is more commonly written as

$$\mathbf{r} \cdot \mathbf{n} = d.$$

Definition 9.2.3. Let the plane π have scalar product form

$$\pi : \mathbf{r} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n}.$$

Let $\mathbf{r} = (x,\, y,\, z)^\mathsf{T}$, $\mathbf{a} = (a_1,\, a_2,\, a_3)^\mathsf{T}$ and $\mathbf{n} = (n_1,\, n_2,\, n_3)^\mathsf{T}$. Then

$$\pi : n_1 x + n_2 y + n_3 z = a_1 n_1 + a_2 n_2 + a_3 n_3$$

is the **Cartesian equation** of π, which is more commonly written as

$$\pi : n_1 x + n_2 y + n_3 z = d.$$

Recipe 9.2.4 (Converting between Forms). To convert from parametric form to scalar product form, take $\mathbf{n} = \mathbf{b}_1 \times \mathbf{b}_2$. To convert from the Cartesian equation to parametric form, express x in terms of y and z, then replace y and z with λ and μ respectively.

Example 9.2.5 (Parametric to Scalar Product Form). Let the plane π have parametric form $\mathbf{r} = (1,\, 2,\, 3)^\mathsf{T} + \lambda\,(4,\, 5,\, 6)^\mathsf{T} + \mu\,(7,\, 8,\, 9)^\mathsf{T}$. Then the normal vector to π is given by

$$\mathbf{n} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \times \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} -3 \\ 6 \\ -3 \end{pmatrix} \parallel \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Hence,

$$d = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 0,$$

whence π has scalar product form

$$\mathbf{r} \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 0.$$

Example 9.2.6 (Cartesian to Parametric Form). Let the plane π have Cartesian equation

$$x + y + z = 10.$$

Solving for x and replacing y and z with λ and μ respectively, we get

$$x = 10 - \lambda - \mu, \quad y = \lambda, \quad z = \mu.$$

Hence, π has parametric form

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 10 - \lambda - \mu \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 10 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda, \mu \in \mathbb{R}.$$

9.2.2 Point and Plane

Proposition 9.2.7 (Relationship between Point and Plane). A point lies on a plane if and only if its position vector (or its equivalent coordinates) satisfies the equation of the plane.

Proof. Trivial. □

Proposition 9.2.8 (Perpendicular Distance between Point and Plane). Let F be the foot of perpendicular from a point Q to the plane π with vector equation $\pi : \mathbf{r} \cdot \mathbf{n} = d$. Let A be a point on π. Then QF, the perpendicular distance from Q to π, is given by

$$QF = \left|\overrightarrow{QA} \cdot \hat{\mathbf{n}}\right| = \frac{|d - \mathbf{q} \cdot \mathbf{n}|}{|\mathbf{n}|}.$$

Proof. Note that QF is the length of projection of $\overrightarrow{QA}$ onto the normal vector $\mathbf{n}$. Hence,

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

follows directly from the formula for the length of projection. Now, observe that

$$\overrightarrow{QA} \cdot \mathbf{n} = \overrightarrow{OA} \cdot \mathbf{n} - \overrightarrow{OQ} \cdot \mathbf{n} = d - \mathbf{q} \cdot \mathbf{n}.$$

Hence,

$$QF = \frac{\left|\overrightarrow{QA} \cdot \mathbf{n}\right|}{|\mathbf{n}|} = \frac{|d - \mathbf{q} \cdot \mathbf{n}|}{|\mathbf{n}|}.$$

□

Corollary 9.2.9. OF, the perpendicular distance from the plane π to the origin O, is

$$OF = \frac{|d|}{|\mathbf{n}|}.$$

Recipe 9.2.10 (Foot of Perpendicular from Point to Plane). Let F be the foot of perpendicular from a point Q to the plane π with vector equation $\pi : \mathbf{r} \cdot \mathbf{n} = d$. To find the position vector $\overrightarrow{OF}$, we use the fact that

- QF is perpendicular to π, i.e. $\overrightarrow{QF} = \lambda \mathbf{n}$ for some $\lambda \in \mathbb{R}$, and
- F lies on π, i.e. $\overrightarrow{OF} \cdot \mathbf{n} = d$.

Example 9.2.11 (Foot of Perpendicular from Point to Plane). Let the plane π have equation $\pi : \mathbf{r} \cdot (1, 2, 3)^\mathsf{T} = 10$. Let $Q(4, 5, 6)$, and let F be the foot of perpendicular from Q to π. We wish to find $\overrightarrow{OF}$.

Since QF is perpendicular to π, we have

$$\overrightarrow{QF} = \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Hence,

$$\overrightarrow{OF} = \overrightarrow{OQ} + \overrightarrow{QF} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Taking the scalar product on both sides, we get

$$10 = \overrightarrow{OF} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \left[\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right] \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = 32 + 14\lambda.$$

Thus, $\lambda = -11/7$, whence

$$\overrightarrow{OF} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} - \frac{11}{7} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 17 \\ 13 \\ 9 \end{pmatrix}.$$

9.2.3 Line and Plane

Fact 9.2.12 (Relationship between Line and Plane). Given a line $l : \mathbf{r} = \mathbf{a} + \lambda \mathbf{b}$, $\lambda \in \mathbb{R}$, and a plane $\pi : \mathbf{r} \cdot \mathbf{n} = d$, there are three possible cases:

- **l and π do not intersect.** l and π are parallel and have no common point.
- **l lies on π.** l and π are parallel and any point on l is also a point on π.
- **l and π intersect once.** l and π are not parallel.

There are two methods to determine the relationship between a line and a plane.

Recipe 9.2.13 (Using Normal Vector).

- If l and π do not intersect, then $\mathbf{b} \cdot \mathbf{n} = 0$ and $\mathbf{a} \cdot \mathbf{n} \neq d$.
- If l lies on π, then $\mathbf{b} \cdot \mathbf{n} = 0$ and $\mathbf{a} \cdot \mathbf{n} = d$.
- If l and π intersect once, then $\mathbf{b} \cdot \mathbf{n} \neq 0$.

Recipe 9.2.14 (Solving Simultaneous Equations)**.** Solve $l : \mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, $\lambda \in \mathbb{R}$ and $\pi : \mathbf{r} \cdot \mathbf{n} = d$ simultaneously.

- If there are no solutions, then l and π do not intersect.
- If there are infinitely many solutions, then l lies on π.
- If there is a unique solution, then l and π intersect once.

Proposition 9.2.15 (Acute Angle between Line and Plane)**.** Let θ be the acute angle between the line $l : \mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, $\lambda \in \mathbb{R}$ and the plane $\pi : \mathbf{r} \cdot \mathbf{n} = d$. Then

$$\sin\theta = \frac{|\mathbf{b} \cdot \mathbf{n}|}{|\mathbf{b}|\,|\mathbf{n}|}.$$

Proof. We first find ϕ, the acute angle between l and the normal. Recall that

$$\cos\phi = \frac{|\mathbf{b} \cdot \mathbf{n}|}{|\mathbf{b}|\,|\mathbf{n}|}.$$

Since $\phi = \frac{\pi}{2} - \theta$, we have

$$\cos\left(\frac{\pi}{2} - \theta\right) = \sin\theta = \frac{|\mathbf{b} \cdot \mathbf{n}|}{|\mathbf{b}|\,|\mathbf{n}|}.$$

□

9.2.4 Two Planes

Proposition 9.2.16 (Acute Angle between Two Planes)**.** The acute angle θ between two planes $\pi_1 : \mathbf{r} \cdot \mathbf{n}_1 = d_1$ and $\pi_2 : \mathbf{r} \cdot \mathbf{n}_2 = d_2$ is given by

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

Proof. Consider the following diagram.

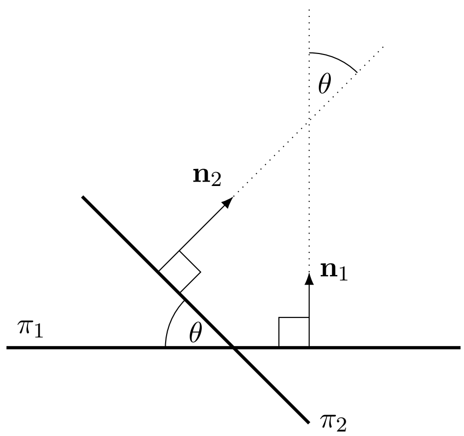


Figure 9.1

It is hence clear that the acute angle between the two planes is equal to the acute angle between the two normal vectors. Thus,

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

□

Fact 9.2.17 (Relationship between Two Planes). Given two distinct planes $\pi_1 : \mathbf{r} \cdot \mathbf{n}_1 = d_1$ and $\pi_2 : \mathbf{r} \cdot \mathbf{n}_2 = d_2$, there are two possible cases:

- **π_1 and π_2 do not intersect.** The two planes are parallel ($\mathbf{n}_1 \parallel \mathbf{n}_2$).
- **π_1 and π_2 intersect at a line.** The two planes are not parallel ($\mathbf{n}_1 \nparallel \mathbf{n}_2$).

Suppose the two planes are not parallel to each other. There are two methods to obtain the equation of the line of intersection.

Recipe 9.2.18 (Via Cartesian Form). Write the equations of the two planes in Cartesian form and solve the two equations simultaneously.

Recipe 9.2.19 (Via Normal Vectors). Observe that as the line of intersection l lies on both planes, l is perpendicular to both the normal vectors $\mathbf{n}_1$ and $\mathbf{n}_2$. Hence, l is parallel to their cross product, $\mathbf{n}_1 \times \mathbf{n}_2$. Thus, if we know a point on the line of intersection l (say point A with position vector $\mathbf{a}$), then the vector equation of l is given by

$$l : \mathbf{r} = \mathbf{a} + \lambda \mathbf{b}, \quad \lambda \in \mathbb{R},$$

where $\mathbf{b}$ is any scalar multiple of $\mathbf{n}_1 \times \mathbf{n}_2$.

10 Matrices

Definition 10.0.1. An $m \times n$ **matrix** $\mathbf{A}$ is an array of numbers with m rows and n columns, with $\mathbf{A} = (a_{ij})$, where a_{ij} is the entry in row i and column j.

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

Example 10.0.2. If

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$

then $\mathbf{A}$ is a 2×3 matrix with $a_{21} = 4$.

Note that row and column vectors are effectively matrices with one row and one column respectively.

10.1 Special Matrices

Definition 10.1.1. A **null matrix** is a matrix with all entries equal to 0. We denote the $m \times n$ null matrix by $\mathbf{0}_{m\times n}$, or simply $\mathbf{0}$.

Example 10.1.2. Examples of null matrices include

$$\begin{pmatrix} 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Definition 10.1.3. A **square matrix** of order n is a matrix with n rows and n columns.

Example 10.1.4. Examples of square matrices include

$$\begin{pmatrix} 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}.$$

Definition 10.1.5. Given a square matrix $\mathbf{A} = (a_{ij})$, the **diagonal** of $\mathbf{A}$ (also called the main, principal or leading diagonal) is the sequence of entries $a_{11}, a_{22}, \ldots, a_{nn}$. The entries a_{ii} are called the **diagonal entries** while a_{ij}, $i \neq j$ are called **non-diagonal entries**.

Definition 10.1.6. A **diagonal matrix** is a square matrix whose non-diagonal entries are zero, i.e. $a_{ij} = 0$ whenever $i \neq j$.

Example 10.1.7. Examples of diagonal matrices include

$$\begin{pmatrix} 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Definition 10.1.8. An **identity matrix** is a diagonal matrix whose diagonal entries are all 1. We denote the identity matrix of order n by $\mathbf{I}_n$, or simply as $\mathbf{I}$.

Example 10.1.9. Examples of identity matrices include

$$\mathbf{I}_1 = \begin{pmatrix} 1 \end{pmatrix}, \quad \mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Definition 10.1.10. A **symmetric matrix** is a square matrix such that $a_{ij} = a_{ji}$ for all i, j.

Example 10.1.11. Examples of symmetric matrices include

$$\begin{pmatrix} 4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 4 \\ 4 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

Definition 10.1.12. A square matrix (a_{ij}) is **upper triangular** if $a_{ij} = 0$ whenever $i > j$; and **lower triangular** if $a_{ij} = 0$ whenever $i < j$.

Example 10.1.13. Examples of triangular matrices include

$$\begin{pmatrix} 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ -2 & -1 & 0 & 1 \end{pmatrix}.$$

The second matrix is an upper triangular matrix, while the third matrix is a lower triangular matrix. The first matrix can be considered both an upper and lower triangular matrix.

Note that a diagonal matrix is both an upper and lower triangular matrix.

10.2 Matrix Operations

10.2.1 Equality

Definition 10.2.1. Two matrices $\mathbf{A}$ and $\mathbf{B}$ are equal if and only if they have the same size and their entries are identical.

10.2.2 Addition

Definition 10.2.2. Let $\mathbf{A}$ and $\mathbf{B}$ be matrices of the same size, and let $\mathbf{C} = \mathbf{A} + \mathbf{B}$ be their sum. Then $(c_{ij}) = (a_{ij} + b_{ij})$. That is, to add two matrices (of the same size), we simply add their corresponding entries.

Example 10.2.3.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 6 & 9 & 12 \\ 10 & 14 & 18 \end{pmatrix}.$$

Fact 10.2.4 (Properties of Matrix Addition). The set of matrices forms an Abelian group under addition.

- Matrix addition is commutative, i.e. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.
- Matrix addition is associative, i.e. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$.
- The null matrix is the additive identity, i.e. $\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}$.
- All matrices have an additive inverse, i.e. $\mathbf{A} - \mathbf{A} = \mathbf{0}$.

10.2.3 Scalar Multiplication

Definition 10.2.5. Let $\mathbf{A}$ be a matrix and let $\lambda \in \mathbb{R}$ be a scalar. Then $\lambda(a_{ij}) = (\lambda a_{ij})$. That is, to multiply a matrix by a scalar λ, we simply multiply each entry by λ.

Example 10.2.6.

$$2\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix}.$$

Fact 10.2.7 (Properties of Scalar Multiplication). Let $\alpha, \beta \in \mathbb{R}$ be scalars, and let $\mathbf{A}$ and $\mathbf{B}$ be matrices of the same size.

- Scalar multiplication is associative, i.e. $\alpha(\beta\mathbf{A}) = (\alpha\beta)\mathbf{A}$.
- Scalar multiplication is distributive over addition, i.e. $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$ and $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$.
- 1 is the multiplicative identity, i.e. $1\mathbf{A} = \mathbf{A}$.
- $0\mathbf{A} = \mathbf{0}$.

10.2.4 Matrix Multiplication

Definition 10.2.8. Let $\mathbf{A}$ be an $m \times p$ matrix, and let $\mathbf{B}$ be a $p \times n$ matrix. Then the matrix product $\mathbf{C} = \mathbf{AB}$ is the $m \times n$ matrix with entries determined by

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Here, c_{ij} can be viewed as the dot product of the ith row of $\mathbf{A}$ with the jth column of $\mathbf{B}$.

Example 10.2.9. Let

$$\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}.$$

Then the matrix product $\mathbf{AB}$ is given by

$$\mathbf{AB} = \begin{pmatrix} (-1)(1) + (0)(3) & (-1)(2) + (0)(0) \\ (2)(1) + (3)(3) & (2)(2) + (3)(0) \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 11 & 4 \end{pmatrix}.$$

Meanwhile, the matrix product $\mathbf{BA}$ is given by

$$\mathbf{BA} = \begin{pmatrix} (1)(-1) + (2)(2) & (1)(0) + (2)(3) \\ (3)(-1) + (0)(2) & (3)(0) + (0)(3) \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ -3 & 0 \end{pmatrix}.$$

Fact 10.2.10 (Properties of Matrix Multiplication)**.**

- Matrix multiplication is *not* commutative, i.e. $\mathbf{AB} \neq \mathbf{BA}$.
- Matrix multiplication is associative, i.e. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.
- Matrix multiplication is distributive over addition, i.e. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ and $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$.
- $\mathbf{AB} = \mathbf{0}$ does not imply that $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$.
- $\mathbf{AB} = \mathbf{AC}$ does not imply that $\mathbf{B} = \mathbf{C}$, i.e. the cancellation law does not apply.

Definition 10.2.11 (Powers of Matrices)**.** If $\mathbf{A}$ is a square matrix, and n is a non-negative integer, we define $\mathbf{A}^n$ as follows:

$$\mathbf{A}^n = \begin{cases} \mathbf{I}, & n = 0, \\ \underbrace{\mathbf{AA}\dots\mathbf{A}}_{n \text{ times}}, & n \geq 1. \end{cases}$$

Here, $\mathbf{I}$ is the identity matrix of the same size as $\mathbf{A}$.

Note that in general, $(\mathbf{AB})^n \neq \mathbf{A}^n\mathbf{B}^n$, where $\mathbf{B}$ is also a square matrix of suitable size.

10.2.5 Transpose

Definition 10.2.12. The **transpose** of a matrix $\mathbf{A} = (a_{ij})$ is denoted $\mathbf{A}^\mathsf{T}$ and is given by (a_{ji}), i.e. the rows and columns are switched.

Example 10.2.13. Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}.$$

Then

$$\mathbf{A}^\mathsf{T} = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

Fact 10.2.14 (Properties of Transpose)**.** Let $\mathbf{A}$ be a matrix and let $c \in \mathbb{R}$ be a scalar.

- The transpose is an involution, i.e. $\left(\mathbf{A}^\mathsf{T}\right)^\mathsf{T} = \mathbf{A}$.
- The transpose is associative, i.e. $(c\mathbf{A})^\mathsf{T} = c\mathbf{A}^\mathsf{T}$.
- The transpose is additive, i.e. $(\mathbf{A}+\mathbf{B})^\mathsf{T} = \mathbf{A}^\mathsf{T} + \mathbf{B}^\mathsf{T}$.
- The transpose reverses the order of matrix multiplication, i.e. $(\mathbf{AB})^\mathsf{T} = \mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T}$.

Note also that $\mathbf{A} = \mathbf{A}^\mathsf{T}$ if and only if $\mathbf{A}$ is a symmetric matrix.

10.3 Solving Systems of Linear Equations

One use of matrix multiplication is to express a system of linear equations. For example,

$$\begin{cases} 3x_1 + 4x_2 + 5x_3 = 6 \\ x_1 + 5x_2 - 6x_3 = 5 \end{cases} \implies \begin{pmatrix} 3 & 4 & 5 \\ 1 & 5 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \end{pmatrix}.$$

The system of equations on the left can be expressed as a matrix equation on the right. What is great about a matrix equation is that we can express a large system of linear equations in a very compact form $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{x}$ and $\mathbf{b}$ are column vectors. In general,

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \implies \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\mathbf{b}}.$$

By translating a system of linear equations into a matrix equation, we can use the power of linear algebra to systematically solve for $\mathbf{x}$, which in turn will yield solutions $(x_1, x_2, \ldots, x_n)$ to our original system of linear equations. We now look at how to systematically solve such matrix equations of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$ using Gaussian elimination.

10.3.1 Elementary Row Operations

Definition 10.3.1. An **elementary row operation** on a matrix refers to one of the following actions performed on it:

- Interchanging row i and row j, denoted $R_i \leftrightarrow R_j$.
- Multiply row i by a non-zero constant k, denoted kR_i.
- Adding k times of row i to row j, denoted $R_j + kR_i$.

Example 10.3.2. The following examples demonstrate the three elementary row operations. Observe how the elementary row operations are written directly to the left of the corresponding rows.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \to_{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 4 & 5 & 6 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \to \begin{array}{c} 10R_1 \\ \\ \\ \end{array} \begin{pmatrix} 10 & 20 & 30 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \to \begin{array}{c} \\ \\ R_3 - 7R_1 \end{array} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & -6 & -12 \end{pmatrix}$$

Multiple elementary row operations can also be combined in a single step:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \to \begin{array}{c} 2R_1 \\ R_2 - R_1 \\ 2R_3 \end{array} \begin{pmatrix} 2 & 4 & 6 \\ 3 & 3 & 3 \\ 14 & 16 & 18 \end{pmatrix}.$$

10.3.2 Gaussian Elimination

Gaussian elimination (also known as Gauss-Jordan elimination, or row reduction) is a systematic algorithm used to convert a system of equations into an *equivalent* system of equations using elementary row operations. That is, the new system of equations has the same solution as the origin system of equations.

Firstly, we rewrite our system of equations as an **augmented matrix** $(\mathbf{A} \mid \mathbf{b})$:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \qquad\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \implies (\mathbf{A} \mid \mathbf{b}) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right).$$

The augmented part is the right-most column, separated by a vertical line to help remind us that these numbers come from the constants in the linear equations ($\mathbf{b}$).

Observe the equivalence between performing elementary row operations on this augmented matrix versus what we might do algebraically to solve the system:

Operations on Equations	Elementary Row Operations on Augmented Matrix
swapping two equations	swapping two rows
multiplying an equation by a non-zero constant	multiplying a row by a non-zero constant
adding a multiple of one equation to another equation	adding a multiple of one row to another row

The objective of Gaussian elimination is thus to repeatedly perform elementary row operations to our augmented matrix until we get a form where we can easily solve for $x_1, \dots, x_n$.

Row-Echelon Form

One such form we aim for is the row-echelon form.

Definition 10.3.3. A matrix is said to be in **row-echelon form** (REF) if

- the first non-zero term in any row (called a **leading term**) is always to the right of the leading term of the previous row, and
- rows consisting of only zeros are at the bottom.

Example 10.3.4. Consider the following matrices:

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 4 & 5 \\ 0 & 4 & 2 & 8 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 3 & 4 & 5 \\ 1 & 4 & 2 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$\mathbf{A}$ is in REF since all leading terms (coloured green) are to the right of the leading term of the previous row. On the other hand, $\mathbf{B}$ is not in REF, since the leading term b_{21} (coloured red) is not to the right of the leading term b_{11}.

Note that a matrix may have multiple row-echelon forms, i.e. REF is not unique.

Once we manipulate our augmented matrix into its REF, we can easily solve for our solutions $x_1, \dots, x_n$ using back-substitution.

Example 10.3.5. Consider the following augmented matrix, which has been manipulated into its REF via elementary row operations:

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 2 \\ 0 & -4 & -4 & 4 \\ 0 & 0 & -15 & 9 \end{array}\right) \implies \begin{cases} x_1 + x_2 + 3x_3 = 2 \\ - 4x_2 - 4x_3 = 4 \\ - 15x_3 = 9 \end{cases}.$$

From the third equation, we easily get $x_3 = -3/5$. Substituting this into the second equation, we get $x_2 = -2/5$. Further substituting this into the first equation, we have $x_1 = 11/5$.

Reduced Row-Echelon Form

Another form we typically aim for when performing Gaussian elimination is the reduced row-echelon form.

Definition 10.3.6. A matrix is said to be in **reduced row-echelon form** (RREF) if it is already in REF, with two further restrictions:

- all leading terms are 1, and
- a column with a leading term has zeroes for all other terms in that column.

Example 10.3.7. Consider the following matrices:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 4 \\ 0 & 4 & 0 \end{pmatrix}.$$

$\mathbf{A}$ is in RREF, since all leading terms (coloured green) are 1 and all other entries in those columns are 0. However, $\mathbf{B}$ is not in RREF. This is because b_{22} is a leading term, but there are non-zero entries in that column (coloured red).

Unlike REF, the RREF of a matrix is unique.

By manipulating our augmented matrix into its RREF, we can easily obtain our solutions $x_1, \ldots, x_n$.

Example 10.3.8. Consider the following augmented, which has been manipulated into RREF using elementary row operations:

$$\left(\begin{array}{ccc|c} 1 & 0 & 3 & 4 \\ 0 & 1 & 4 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right) \implies \begin{cases} x_1 + 3x_3 = 4 \\ x_2 + 4x_3 = 8 \end{cases}.$$

Letting x_3 be a free parameter $\lambda \in \mathbb{R}$, we have

$$x_1 = 4 - 3\lambda, \quad x_2 = 8 - 4\lambda, \quad x_3 = \lambda.$$

10.3.3 Consistent and Inconsistent Systems

Back in §1, we termed a system of linear equations *consistent* if it admits a solution, and *inconsistent* if it does not. We also learnt that a consistent system of linear equations either has a unique solution or infinitely many solutions. Using Gaussian elimination, we can easily determine the number of solutions it admits.

Proposition 10.3.9. Let $\left(\mathbf{A}' \mid \mathbf{b}'\right)$ be the RREF of $\left(\mathbf{A} \mid \mathbf{b}\right)$.

- If $\mathbf{A}' = \mathbf{I}$, the system has a unique solution.
- If the ith row of $\mathbf{A}'$ is all zeroes, and $b_i' = 0$, then the system has infinitely many solutions.
- If the ith row of $\mathbf{A}'$ is all zeroes, and $b_i' = 1$, then the system has no solution.

The first statement is trivially true, since $\mathbf{Ix} = \mathbf{b}' \implies \mathbf{x} = \mathbf{b}'$. To see why the second and third statements are true, consider the following matrices:

$$\mathbf{B} = \left(\begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right), \quad \mathbf{C} = \left(\begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

$\mathbf{B}$ represents the system

$$\begin{cases} x_1 \qquad + 3x_3 = 1 \\ \qquad x_2 + \ \ x_3 = 2 \end{cases}.$$

In this case, we have more unknowns than equations, so we will obtain infinitely many solutions (e.g. by taking $x_3 = \lambda$, where $\lambda \in \mathbb{R}$ is a free parameter). On the other hand, the third row of $\mathbf{C}$ represents the equation

$$0x_1 + 0x_2 + 0x_3 = 1,$$

which is clearly impossible. Thus, there will be no solutions to the system.

10.3.4 Homogeneous Systems of Linear Equations

Recall that a system of linear equations is said to be *homogeneous* if all the constant terms are zero. The corresponding matrix equation is thus $\mathbf{Ax} = \mathbf{0}$. Clearly, every homogeneous system has $\mathbf{x} = \mathbf{0}$ as a solution. This solution is called the **trivial solution**. If there are other solutions, they are called non-trivial solutions.

10.4 Invertible Matrices

While Gaussian elimination remains a good way of solving a system of linear equations, looking at them as a matrix equation can also be useful.

The left side of $\mathbf{Ax} = \mathbf{b}$ may be viewed as a matrix $\mathbf{A}$ acting on a vector $\mathbf{x}$ and sending it to the vector $\mathbf{b}$. Solving the matrix equation hence amounts to finding the pre-image of $\mathbf{b}$ under $\mathbf{A}$. This motivates us to find a multiplicative inverse to $\mathbf{A}$.

Definition 10.4.1. The **multiplicative inverse** of a square matrix $\mathbf{A}$, denoted $\mathbf{A}^{-1}$, has the property

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}.$$

If such a matrix $\mathbf{A}^{-1}$ exists, then $\mathbf{A}$ is said to be **invertible**, or **non-singular**.

If $\mathbf{A}^{-1}$ exists, the solution for the equation $\mathbf{Ax} = \mathbf{b}$ will simply be $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Further, this solution will be unique for each $\mathbf{b}$ (since $\mathbf{A}^{-1}$ will not map $\mathbf{b}$ to multiple vectors).

We now state some properties regarding the inverse of a matrix:

Fact 10.4.2 (Properties of Invertible Matrices)**.** Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices of the same size. Let $a \in \mathbb{R}$ be a scalar and let n be a non-negative integer.

- The inverse of a matrix is unique.
- If $a\mathbf{A}$ is invertible, then $(a\mathbf{A})^{-1} = \frac{1}{a}\mathbf{A}^{-1}$.
- If $\mathbf{A}$ is invertible, then $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$.
- If $\mathbf{A}^\mathsf{T}$ is invertible, then $\left(\mathbf{A}^\mathsf{T}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^\mathsf{T}$.
- If $\mathbf{AB}$ is invertible, then $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
- If $\mathbf{A}^n$ is invertible, then $(\mathbf{A}^n)^{-1} = \mathbf{A}^{-n} = \left(\mathbf{A}^{-1}\right)^n$.

We now discuss how to find the inverses of matrices.

10.4.1 Inverse of a 2×2 Matrix

Proposition 10.4.3 (2×2 Inverse Formula)**.** Let

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

Then its inverse is given by

$$\mathbf{A}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Notice that the 2×2 inverse formula is not valid in the case where $ad - bc = 0$. This quantity, $ad - bc$, is called the **determinant** of the 2×2 matrix, and it plays a special role in determining whether a matrix is invertible. We will discuss more about determinants in the next chapter.

10.4.2 Inverse of an $n \times n$ Matrix

Though there is a general formula for the inverse of an $n \times n$ matrix, it is tedious to compute for $n \geq 3$. Luckily, there is a general procedure that we can employ. This procedure rests on the fact that any elementary row operation can be represented as a left-multiplication by an **elementary matrix**.

Definition 10.4.4. An $n \times n$ matrix is an **elementary matrix** if it can be obtained from the $n \times n$ identity matrix $\mathbf{I}_n$ by performing a single row operation.

Example 10.4.5. As an example, consider

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix}.$$

If we add 3 times the 3rd row to the 1st row, we will obtain

$$\begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} \xrightarrow{R_1 + 3R_3} \begin{pmatrix} 4 & 12 & 14 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix}.$$

Now observe that if we pre-multiply $\mathbf{A}$ by the elementary matrix

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

we get

$$\mathbf{BA} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 12 & 14 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix},$$

which is exactly the same result as doing the row operation.

The correspondence between elementary row operations and elementary matrices allows us to construct the following algorithm to find the inverse of an invertible matrix $\mathbf{A}$.

Recipe 10.4.6 (Finding Matrix Inverse)**.** If $\mathbf{A}$ is invertible, then $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. If we can find a sequence of elementary row operations, corresponding to successive matrix left-multiplications of the elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_k$, such that

$$\mathbf{E}_k \ldots \mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I},$$

then we have $\mathbf{E}_k \ldots \mathbf{E}_2\mathbf{E}_1 = \mathbf{A}^{-1}$.

In practice, however, we will perform the left-multiplications on an augmented matrix of the form $\left(\mathbf{A} \mid \mathbf{I}\right)$:

$$\mathbf{E}_k \ldots \mathbf{E}_2\mathbf{E}_1 \left(\mathbf{A} \mid \mathbf{I}\right) = \left(\mathbf{E}_k \ldots \mathbf{E}_2\mathbf{E}_1\mathbf{A} \mid \mathbf{E}_k \ldots \mathbf{E}_2\mathbf{E}_1\right) = \left(\mathbf{I} \mid \mathbf{A}^{-1}\right).$$

10.5 Determinant of a Matrix

The previous section showed the importance of invertibility and uses elementary row operations to help us determine if a matrix is invertible. Here, we introduce the idea of the determinant of a matrix and how this number tells us if a matrix is invertible.

Definition 10.5.1. The **determinant** of an $n \times n$ matrix $\mathbf{A}$, denoted by

$$|\mathbf{A}| = \det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix},$$

is the minimal polynomial (in the entries of $\mathbf{A}$, i.e. a_{11}, a_{12}, etc.) that is 0 if and only if $\mathbf{A}$ is singular.

10.5.1 The 1×1 and 2×2 Determinant

For 1×1 matrices, $(a)^{-1} = (1/a)$, so the matrix has an inverse if and only if $a \neq 0$. Thus, $|a| = a$.

For $2times2$ matrices, recall that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

The inverse hence does not exist when $ad - bc = 0$. Hence,

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

10.5.2 Cofactor Expansion

Beyond the 2×2 matrix, the closed form of an $n \times n$ determinant becomes much more unwieldy to remember and use. Luckily, there is a general procedure that we can use to calculate the determinant of any $n \times n$ matrix.

Proposition 10.5.2 (Cofactor Expansion)**.** Suppose we have an $n \times n$ matrix $\mathbf{A} = (a_{ij})$. Let $\mathbf{M}_{ij}$ be the $(n-1) \times (n-1)$ matrix obtained from $\mathbf{A}$ by deleting the ith row and the jth column. Then the determinant of $\mathbf{A}$ is given by

$$\det(\mathbf{A}) = \begin{cases} a_{11}, & n = 1, \\ a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}, & n > 1 \end{cases},$$

where $A_{ij} = (-1)^{i+j}\det(\mathbf{M}_{ij})$ is the **cofactor** of entry a_{ij}.

Note that the term $(-1)^{i+j}$ has value 1 when the sum of i and j is even, and -1 when the sum is odd. This may be viewed as a "signed" array as follows:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Example 10.5.3. Using the method of cofactor expansion along the first row, the determinant of a 3×3 matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

is given by

$$\det(\mathbf{A}) = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

The formula given by Proposition 10.5.2 is not unique: we can expand cofactors along any row or column of the matrix to get the determinant. This is particularly useful when a particular row/column contains many zeroes.

Example 10.5.4. Let

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 0 & 4 \\ 3 & 2 & 9 \end{pmatrix}.$$

Expanding along the second column, we see that

$$\det(\mathbf{A}) = -0\begin{vmatrix} 2 & 4 \\ 3 & 9 \end{vmatrix} + 0\begin{vmatrix} 1 & 3 \\ 3 & 9 \end{vmatrix} - 2\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} = 4.$$

10.5.3 Properties

We now look at the properties of determinants.

Fact 10.5.5 (Properties of Determinants)**.** Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices of order n.

- $\det(\mathbf{A}) = \det(\mathbf{A}^\mathsf{T})$.
- $\det(\mathbf{A} + \mathbf{B}) \neq \det(\mathbf{A}) + \det(\mathbf{B})$.
- $\det(c\mathbf{A}) = c^n \det(\mathbf{A})$, where c is a scalar.
- $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$.
- If $\mathbf{A}$ is a triangular matrix, then $\det(\mathbf{A})$ is the product of the diagonal entries of $\mathbf{A}$.
- $\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$.
- If $\mathbf{A}$ is invertible, then $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$.
- If $\mathbf{A}$ has a row or column of zeroes, then $\det(\mathbf{A}) = 0$.

Fact 10.5.6 (Effects of Elementary Row/Column Operations on Determinant)**.**

- If $\mathbf{B}$ is the matrix that results when a row/column of $\mathbf{A}$ is multiplied by a scalar k, then $\det(\mathbf{B}) = k\det(\mathbf{A})$.
- If $\mathbf{B}$ is the matrix that results when two rows/columns of $\mathbf{A}$ are interchanged, then $\det(\mathbf{B}) = -\det(\mathbf{A})$.
- If $\mathbf{B}$ is the matrix that results when a multiple of one row/column of $\mathbf{A}$ is added to another row/column, then $\det(\mathbf{B}) = \det(\mathbf{A})$.

The above results are a result of the fact that $\det(\mathbf{EA}) = \det(\mathbf{E}) \det(\mathbf{A})$, where $\mathbf{E}$ is an elementary matrix.

11 Linear Transformations

Definition 11.0.1. A **linear transformation** is a function $T : \mathbb{R}^n \to \mathbb{R}^m$ is a function that satisfies the following two properties:

- $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$,
- $T(k\mathbf{u}) = kT(\mathbf{u})$ for all scalars $k \in \mathbb{R}$ and vectors $\mathbf{u} \in \mathbb{R}^n$.

Taken together, these two properties mean that linear transformations preserve the structure of linear combinations.

Proposition 11.0.2 (Linear Transformations Preserve Linear Combinations)**.** Let $k_1, \ldots, k_r \in \mathbb{R}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_r \in \mathbb{R}^n$. Then

$$T(k_1\mathbf{v}_1 + \cdots + k_r\mathbf{v}_r) = k_1T(\mathbf{v}_1) + \cdots + k_rT(\mathbf{v}_r).$$

When $k = 0$, the second property of linear transformations also implies that $T(\mathbf{0}) = \mathbf{0}$. That is, a linear transformation must map $\mathbf{0}$ to $\mathbf{0}$.

Recipe 11.0.3 (Determining if a Function is a Linear Transformation)**.** To determine if a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, we go through the following "checklist", arranged in increasing difficulty to see:

- Check if $f(\mathbf{0}) = \mathbf{0}$.
- Check if $f(k\mathbf{v}) = kf(\mathbf{v})$.
- Check if $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$.

If f passes the above checklist, we then proceed to show that $f(k_1\mathbf{v}_1+k_2\mathbf{v}_2) = k_1f(\mathbf{v}_1)+k_2f(\mathbf{v}_2)$. This would immediately imply that f satisfies the two properties and is thus a linear transformation.

Example 11.0.4. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a function defined by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x-y \end{pmatrix}.$$

Clearly, $T(\mathbf{0}) = \mathbf{0}$, so T passes the first check. By inspection, T also satisfies the remaining two checks. We are now confident that T is a linear transformation, so we consider $T(k_1\mathbf{v}_1 + k_2\mathbf{v}_2)$, where $\mathbf{v}_1 = (x_1, y_1)^\mathsf{T}$ and $\mathbf{v}_2 = (x_2, y_2)^\mathsf{T}$. Then

$$\begin{aligned} T(k_1\mathbf{v}_1 + k_2\mathbf{v}_2) = T\begin{pmatrix} k_1x_1 + k_2x_2 \\ k_1y_1 + k_2y_2 \end{pmatrix} &= \begin{pmatrix} k_1x_1 + k_2x_2 \\ (k_1x_1 + k_2x_2) + (k_1y_1 + k_2y_2) \\ (k_1x_1 + k_2x_2) - (k_1y_1 + k_2y_2) \end{pmatrix} \\ &= k_1\begin{pmatrix} x_1 \\ x_1 + y_1 \\ x_1 - y_1 \end{pmatrix} + k_2\begin{pmatrix} x_2 \\ x_2 + y_2 \\ x_2 - y_2 \end{pmatrix} = k_1T(\mathbf{v}_1) + k_2T(\mathbf{v}_2). \end{aligned}$$

Thus, T is indeed a linear transformation.

11.1 Matrix Representation

Observe that the transformation T in the above example may also be written as

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x-y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

This is because matrix multiplication may also be seen as a form of linear transformation.

Proposition 11.1.1 (Matrix Multiplication is a Linear Transformation)**.** Let $\mathbf{A}$ be an $m \times n$ matrix. Then, multiplication by $\mathbf{A}$ will take an n-dimensional vector to an m-dimensional vector, so $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$. Moreover, it is linear, as for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $k \in \mathbb{R}$,

$$T(\mathbf{x}+\mathbf{y}) = \mathbf{A}(\mathbf{x}+\mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y} = T(\mathbf{x}) + T(\mathbf{y})$$

and

$$T(k\mathbf{x}) = \mathbf{A}(k\mathbf{x}) = k\mathbf{A}\mathbf{x} = kT(\mathbf{x}).$$

Surprisingly, there are no other examples of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$; matrix multiplication is the only kind of linear transformation there is for functions between finite-dimensional spaces:

Proposition 11.1.2. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Let $\mathbf{x} \in \mathbb{R}^n$. Then $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ for some $m \times n$ matrix $\mathbf{A}$.

Proof. Let $\mathbf{e}_i$ be the ith standard basis vector. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be an n-dimensional vector. Then

$$T(\mathbf{x}) = T(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1T(\mathbf{e}_1) + \cdots + x_nT(\mathbf{e}_n) = \begin{pmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{pmatrix}\mathbf{x} = \mathbf{A}\mathbf{x}.$$

Since $T(\mathbf{e}_i)$ is an m-dimensional vector (by the definition of T), it follows that $\mathbf{A}$ has m rows and n columns, i.e. $\mathbf{A}$ is an $m \times n$ matrix. □

11.2 Linear Spaces

Definition 11.2.1. A **linear space** (or **vector space**) over $\mathbb{R}$ is a set V equipped with two operations, addition $(+)$ and scalar multiplication $(\cdot)$, such that for any vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and for all $c, d \in \mathbb{R}$, the following ten axioms are satisfied:

1. Closure under addition: $\mathbf{u} + \mathbf{v} \in V$.
2. Addition is commutative: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. Addition is associative: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. Existence of additive identity: There is a zero vector, $\mathbf{0}$, such that $\mathbf{0} + \mathbf{u} = \mathbf{u}$.
5. Existence of additive inverse: There is a vector $-\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. Closure under scalar multiplication: $c\mathbf{u} \in V$.
7. Scalar multiplication is distributive over vector addition: $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. Scalar multiplication is distributive over scalar addition: $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. Scalar multiplication is associative: $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. Existence of scalar multiplicative identity: There exists a scalar, 1, such that $1\mathbf{u} = \mathbf{u}$.

One can think of a linear space as an Abelian group (under addition, Axioms 1-5) with the added structure of "scalar multiplication" (Axioms 6-10).

11.2.1 Examples of Linear Spaces

Definition 11.2.2. The **Euclidean n-space**, denoted by $\mathbb{R}^n$, is the set of all n-vectors (ordered n-tuples) $(u_1, u_2, \ldots, u_n)$ of real numbers.

$$\mathbb{R}^n = \{(u_1, \ldots, u_n) \mid u_1, \ldots, u_n \in \mathbb{R}\}.$$

Proposition 11.2.3. $\mathbb{R}^n$ is a linear space equipped with scalar addition and scalar multiplication.

$\mathbb{R}^n$ is the quintessential example of a linear space, and is the linear space that we will deal with most. We can also generalize the above statements from vectors to matrices:

Proposition 11.2.4. The set of all $m \times n$ matrices with real entries forms a linear space (equipped with matrix addition and scalar multiplication).

There are also more abstract examples of linear spaces:

Proposition 11.2.5. The set of all polynomials with real coefficients of at most degree $n \geq 0$, forms a linear space under the usual addition and multiplication.

Lastly, there is the trivial vector space:

Definition 11.2.6. Let V be a singleton, i.e. $V = \{\mathbf{0}\}$. Define $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $k\mathbf{0} = \mathbf{0}$ for all scalars k. Then V is the **zero vector space**.

11.3 Subspaces

Definition 11.3.1. Suppose V is a linear space under $(+, \cdot)$, and $W \subseteq V$. If W is also a linear space under $(+, \cdot)$, then W is a **subspace** of V.

Example 11.3.2. Consider the set $S = \{(a, b, 0) \mid a, b \in \mathbb{R}\}$. One can clearly show that S is a linear space equipped with the usual addition and scalar multiplication. Since $S \subseteq \mathbb{R}^3$, it follows that S is a subspace of $\mathbb{R}^3$.

Example 11.3.3. If V is a linear space, then V and $\{\mathbf{0}\}$ are both subspaces of V.

Because subspaces inherit addition and multiplication, we do not need to check Axioms 2, 3, 7, 8 and 9. Further, Axiom 5 is guaranteed if Axiom 6 is valid. Thus, we really only need to verify Axioms 1, 4 and 6 when testing for subspaces.

Recipe 11.3.4 (Test for Subspace). Let W be a non-empty subset of a linear space V. Then W is a subspace of V if and only if the following conditions hold

- $\mathbf{0} \in W$.
- (Closure under addition) For all $\mathbf{u}, \mathbf{v} \in W$, we have $\mathbf{u} + \mathbf{v} \in W$.
- (Closure under multiplication) For all $c \in \mathbb{R}$ and $\mathbf{u} \in \mathbf{u}$, we have $c\mathbf{u} \in W$.

Conversely, to show that W is not a subspace, we can try to disprove any of the three conditions. Typically, the first condition ($\mathbf{0} \in W$) is the easiest to disprove. If that fails, we construct a counter-example for closure under addition/multiplication.

Sample Problem 11.3.5. Let W be any plane in $\mathbb{R}^3$ that passes through the origin. Prove that W is a subspace of $\mathbb{R}^3$ under the standard operations.

Solution. Let

$$W = \{\mathbf{r} = \lambda\mathbf{a} + \mu\mathbf{b} \mid \lambda, \mu \in \mathbb{R}\}.$$

- Taking $\lambda = \mu = 0$, we see that $\mathbf{0} \in W$.
- Define $\mathbf{r}_1 = \lambda_1\mathbf{a} + \mu_1\mathbf{b}$ and $\mathbf{r}_2 = \lambda_2\mathbf{a} + \mu_2\mathbf{b}$. Observe that

  $$\mathbf{r}_1 + \mathbf{r}_2 = (\lambda_1\mathbf{a} + \mu_1\mathbf{b}) + (\lambda_2\mathbf{a} + \mu_2\mathbf{b}) = (\lambda_1 + \lambda_2)\,\mathbf{a} + (\mu_1 + \mu_2)\,\mathbf{b}.$$

  Since $\lambda_1 + \lambda_2, \mu_1 + \mu_2 \in \mathbb{R}$, it follows that $\mathbf{r}_1 + \mathbf{r}_2 \in W$, so W is closed under addition.
- Let $k \in \mathbb{R}$. Then

  $$k\mathbf{r} = k\,(\lambda\mathbf{a} + \mu\mathbf{b}) = (k\lambda)\,\mathbf{a} + (k\mu)\,\mathbf{b}.$$

  Since $k\lambda, k\mu \in \mathbb{R}$, it follows that $k\mathbf{r} \in W$, so W is closed under multiplication.

Thus, W is a subspace of $\mathbb{R}^3$. $\square$

Sample Problem 11.3.6. Let W be the set of vectors in $\mathbb{R}^3$ whose length does not exceed 1. Determine whether W is a subspace of $\mathbb{R}^3$.

Solution. Take $\mathbf{u} = (1, 0, 0)^\mathsf{T}$ and $\mathbf{v} = (0, 1, 0)^\mathsf{T}$. Since $|\mathbf{u}| = |\mathbf{v}| = 1 \leq 1$, they are both elements of W. Now consider the length of $\mathbf{u} + \mathbf{v}$:

$$|\mathbf{u} + \mathbf{v}| = \left|\begin{pmatrix}1\\0\\0\end{pmatrix} + \begin{pmatrix}0\\1\\0\end{pmatrix}\right| = \left|\begin{pmatrix}1\\1\\0\end{pmatrix}\right| = \sqrt{2} \geq 1.$$

Thus, $\mathbf{u}+\mathbf{v} \notin W$, so W is not closed under addition. Thus, W is not a linear space, so W is not a subspace of $\mathbb{R}^3$. □

In Sample Problem 11.3.5, we saw how any plane passing through the origin in $\mathbb{R}^3$ is a subspace. We can generalize this further:

Subspaces of $\mathbb{R}^1$	**Subspaces of $\mathbb{R}^2$**	**Subspaces of $\mathbb{R}^3$**
• $\{\mathbf{0}\}$ • $\mathbb{R}^1$	• $\{\mathbf{0}\}$ • Lines through the origin • $\mathbb{R}^2$	• $\{\mathbf{0}\}$ • Lines through the origin • Planes through the origin • $\mathbb{R}^3$

In fact, these are the only subspaces of $\mathbb{R}^1$, $\mathbb{R}^2$ and $\mathbb{R}^3$. Note that this pattern holds for all $\mathbb{R}^n$.

11.4 Span and Linear Independence

11.4.1 Linear Spans

Definition 11.4.1. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ be a non-empty subset of a linear space V. Then the **span** of S, denoted $\operatorname{span} S$ or $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$, is the set containing all linear combinations of vectors of S. That is,

$$\operatorname{span} S = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} = \{a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r \mid a_1, \ldots, a_r \in \mathbb{R}\}.$$

Note that $\operatorname{span} \varnothing = \{\mathbf{0}\}$, since the sum of nothing is $\mathbf{0}$.

Example 11.4.2. Let $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$. Then $S = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \{a\mathbf{v}_1 + b\mathbf{v}_2 \mid a, b \in \mathbb{R}\}$.

- If $\mathbf{v}_1$ and $\mathbf{v}_2$ are non-parallel, then S represents a plane (parallel to $\mathbf{v}_1$ and $\mathbf{v}_2$) that passes through the origin in $\mathbb{R}^n$.
- If $\mathbf{v}_1$ and $\mathbf{v}_2$ are parallel, then S represents a line (parallel to both $\mathbf{v}_1$ and $\mathbf{v}_2$) that passes through the origin in $\mathbb{R}^n$.
- If $\mathbf{v}_1$ and $\mathbf{v}_2$ are both $\mathbf{0}$, then S is simply the origin.

Proposition 11.4.3. Let S be a subset of a linear space V. Then $\operatorname{span} S$ is a subspace of V.

Proof. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$. By definition, we have

$$\operatorname{span} S = \{a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r \mid a_1, \ldots, a_r \in \mathbb{R}\}.$$

- Taking $a_1 = \cdots = a_n = 0$, we see that $\mathbf{0} \in \operatorname{span} S$.
- Let $\mathbf{a}, \mathbf{b} \in \operatorname{span} S$. We can write

  $$\mathbf{a} = a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r \quad \text{and} \quad \mathbf{b} = b_1\mathbf{v}_1 + \cdots + b_r\mathbf{v}_r,$$

  where $a_1, \ldots, a_r, b_1, \ldots, b_r \in \mathbb{R}$. Now consider their sum:

  $$\mathbf{a}+\mathbf{b} = (a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r) + (b_1\mathbf{v}_1 + \cdots + b_r\mathbf{v}_r) = (a_1+b_1)\,\mathbf{v}_1 + \cdots + (a_r+b_r)\,\mathbf{v}_r.$$

  Since $a_1+b_1, \ldots, a_r+b_r \in \mathbb{R}$, it follows that $\mathbf{a}+\mathbf{b}$ is also a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_r$, i.e. $\mathbf{a}+\mathbf{b} \in \operatorname{span} S$. Thus, $\operatorname{span} S$ is closed under addition.

- Let $k \in \mathbb{R}$. Consider $k\mathbf{a}$:
$$k\mathbf{a} = k\left(a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r\right) = ka_1\mathbf{v}_1 + \cdots + ka_r\mathbf{v}_r.$$
Since $ka_1, \ldots, ka_r \in \mathbb{R}$, it follows that $k\mathbf{a} \in \operatorname{span} S$. Thus, $\operatorname{span} S$ is closed under multiplication.

Thus, S is a subspace of V. □

A natural question to ask is "When is a vector in the span of a set of vectors?" For instance, is
$$\begin{pmatrix}1\\2\\3\end{pmatrix} \in \operatorname{span}\left\{\begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}7\\8\\9\end{pmatrix}\right\}?$$
It turns out this is equivalent to finding coefficients $x_1, x_2 \in \mathbb{R}$ such that
$$x_1\begin{pmatrix}4\\5\\6\end{pmatrix} + x_2\begin{pmatrix}7\\8\\9\end{pmatrix} = \begin{pmatrix}1\\2\\3\end{pmatrix}.$$
This, in turn, is equivalent to the matrix equation
$$\begin{pmatrix}4 & 7\\5 & 8\\6 & 9\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}1\\2\\3\end{pmatrix}.$$
Of course, we can use an augmented matrix and calculate its RREF to determine x_1 and x_2:
$$\left(\begin{array}{cc|c}4 & 7 & 1\\5 & 8 & 2\\6 & 9 & 3\end{array}\right) \to \left(\begin{array}{cc|c}1 & 0 & 2\\0 & 1 & -1\\0 & 0 & 0\end{array}\right).$$
This gives $x_1 = 2$ and $x_2 = -1$, so
$$\begin{pmatrix}1\\2\\3\end{pmatrix} \in \operatorname{span}\left\{\begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}7\\8\\9\end{pmatrix}\right\}.$$
This leads us to the following result:

Proposition 11.4.4. The equation $\mathbf{Ax} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of $\mathbf{A}$, i.e. $\mathbf{b}$ is in the span of columns of $\mathbf{A}$.

Sample Problem 11.4.5. Determine if $\mathbb{R}^3$ is spanned by
$$S = \left\{\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}1\\0\\2\end{pmatrix}\right\}.$$

Solution. Let $\mathbf{v} = (a,\, b,\, c)^{\mathsf{T}} \in \mathbb{R}^3$. Consider the equation
$$x_1\begin{pmatrix}1\\2\\1\end{pmatrix} + x_2\begin{pmatrix}1\\0\\2\end{pmatrix} \implies \begin{pmatrix}1 & 1\\2 & 0\\1 & 2\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}a\\b\\c\end{pmatrix}.$$
Using row-operations on the resulting augmented matrix, we obtain
$$\left(\begin{array}{cc|c}1 & 0 & b/2\\0 & 1 & c-a\\0 & 0 & 2a-b/2-c\end{array}\right).$$
The system is only consistent when $2a - b/2 - c = 0$. That is, not all vectors $\mathbf{v} \in \mathbb{R}^3$ can be written as a linear combination of vectors in S. Thus, $\mathbb{R}^3$ is not spanned by S. □

Sample Problem 11.4.6. Determine if $\mathbb{R}^3$ is spanned by

$$S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Solution. Let $\mathbf{v} = (a,\, b,\, c)^{\mathsf{T}} \in \mathbb{R}^3$. Consider the equation

$$x_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a - x_4 \\ b \\ c \end{pmatrix}.$$

Since the matrix on the LHS has non-zero determinant, it is invertible, so there exist $x_1, x_2, x_3, x_4 \in \mathbb{R}$ such that the above equation is satisfied. That is to say, every vector in $\mathbb{R}^3$ can be expressed as a linear combination of vectors in S. Thus, $\mathbb{R}^3$ is spanned by S. □

11.4.2 Linear Independence

Consider the previous sample question. For different choices of x_4, we get different values of x_1, x_2, x_3. That is, for a particular vector $\mathbf{v}$, there is more than one way of expressing $\mathbf{v}$ as a linear combination of the vectors in S. This is because the fourth vector, $(1,\, 0,\, 0)^{\mathsf{T}}$, is redundant as it is a linear combination of the other three vectors, i.e.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = -\frac{2}{3} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \frac{4}{3} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

We say that S is linearly dependent.

Definition 11.4.7. A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is **linear dependent** if there are coefficients $c_1, \dots, c_k$, not all zero, such that

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}.$$

Otherwise, the set of vectors is **linearly independent**.

Equivalently, the set of vectors are linearly dependent if at least one vector is expressible as a linear combination of the other vectors.

Sample Problem 11.4.8. Determine if the following set of vectors is linearly independent:

$$S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Solution. Consider the following equation:

$$c_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Converting to RREF, we obtain

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Thus, the only solutions are $c_1 = c_2 = 0$, so S is linearly independent. □

Sample Problem 11.4.9. Determine if the following set of vectors is linearly independent:

$$S = \left\{ \begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}1\\0\\2\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix} \right\}.$$

Solution. Consider the following equation:

$$c_1 \begin{pmatrix}1\\2\\1\end{pmatrix} + c_2 \begin{pmatrix}1\\0\\2\end{pmatrix} + c_3 \begin{pmatrix}1\\1\\0\end{pmatrix} + c_4 \begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \implies \begin{pmatrix}1 & 1 & 1 & 1\\2 & 0 & 1 & 0\\1 & 2 & 0 & 0\end{pmatrix} \begin{pmatrix}c_1\\c_2\\c_3\\c_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Converting to RREF, we obtain

$$\begin{pmatrix}1 & 0 & 0 & -2/3\\0 & 1 & 0 & 1/3\\0 & 0 & 1 & 4/3\end{pmatrix} \begin{pmatrix}c_1\\c_2\\c_3\\c_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

By backwards substitution, we obtain

$$c_1 = -2\lambda, \quad c_2 = \lambda, \quad c_3 = 4\lambda, \quad c_4 = -\lambda,$$

where $\lambda \in \mathbb{R}$. Thus, there exist non-trivial solutions, so S is linearly dependent. □

We now outline a general strategy to test if a set of vectors is linearly independent.

Recipe 11.4.10 (Test for Linear Independence). We are given r vectors $\mathbf{v}_1, \dots, \mathbf{v}_r \in \mathbb{R}^n$.

Case 1. If $r > n$, then the r vectors must be linearly dependent.

Case 2. If $r \le n$, we find $\mathbf{x} = (x_1, \dots, x_r)^\mathsf{T}$ such that $\mathbf{Ax} = \mathbf{0}$ where $\mathbf{A} = \begin{pmatrix} v_1 & \dots & v_r \end{pmatrix}$ is an $n \times r$ matrix. Whether the r vectors are linearly dependent becomes a question of whether the equation $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$. To answer this question we can

- in general, use row operations to reduce $\mathbf{A}$ to REF. If there are exactly r non-zero rows, then $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.
- (if $r = n$) compute the determinant of $\mathbf{A}$. If $\det \mathbf{A} \neq 0$, then $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

Geometrical Interpretations of Linear Independence

In $\mathbb{R}^2$, two vectors $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent if and only if they lie on the same line (with their initial points at the origin).

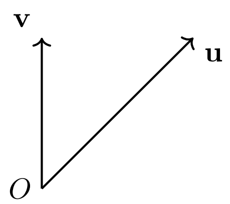


Figure 11.1: Linearly independent

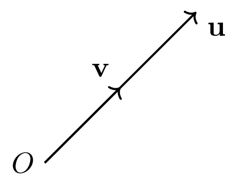


Figure 11.2: Linearly dependent

In $\mathbb{R}^3$, three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent if and only if they lie on the same line or plane (with their initial points at the origin).

11.5 Basis and Dimension

Definition 11.5.1. A **basis** $S = \{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ for a linear space V is a set of vectors such that

- S spans V, and
- S is linearly independent.

Definition 11.5.2. Let $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)$, $\dots$, $\mathbf{e}_n = (0, \dots, 0, 1)$. The set $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is called the **standard basis** of $\mathbb{R}^n$.

Sample Problem 11.5.3. Show that the set

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is a basis for $\mathbb{R}^4$.

Solution. We first show that S spans $\mathbb{R}^4$. Consider $\mathbf{v} = (a,\, b,\, c,\, d)^{\mathsf{T}}$, where $a, b, c, d \in \mathbb{R}$. Consider

$$k_1 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix} + k_3 \begin{pmatrix} 0 \\ 2 \\ 2 \\ 1 \end{pmatrix} + k_4 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \implies \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 \\ 1 & -1 & 2 & 0 \\ 0 & 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

Since the matrix on the LHS has non-zero determinant, every $\mathbf{v}$ can be expressed as a linear combination of the vectors of S. Thus, S spans $\mathbb{R}^4$.

We now show that S is linearly independent. Consider

$$k_1 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix} + k_3 \begin{pmatrix} 0 \\ 2 \\ 2 \\ 1 \end{pmatrix} + k_4 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 \\ 1 & -1 & 2 & 0 \\ 0 & 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since the matrix on the LHS has non-zero determinant, the equation has only the trivial solution. Thus, S is linearly independent. □

One particularly useful property about bases is that there is only one way to build a vector as a linear combination of given basis vector.

Theorem 11.5.4. If $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for a linear space V, then every vector $\mathbf{v} \in V$ can be expressed in the form $\mathbf{v} = k_1 \mathbf{v}_1 + \dots + k_n \mathbf{v}_n$ in exactly one way.

While a linear space can have many bases, the number of basis vectors must be the same. This number is called the dimension of V.

Definition 11.5.5. The **dimension** of a non-zero linear space V is the number of vectors in a basis for V, and is denoted $\dim V$. By convention, we define the dimension of the zero linear space $\{\mathbf{0}\}$ to be 0.

As an example, the linear space $\mathbb{R}^n$ has dimension n (recall that the standard basis consists of n vectors).

We now state several remarks relating spans, linear independence and bases.

Proposition 11.5.6. Let V be a linear space with finite dimension n, and let $S \subseteq V$.

- If $|S| > n$, then S is linearly dependent.
- If $|S| < n$, then S cannot span V.
- If $|S| = n$, then S is a basis of V if and only if S is linearly independent if and only if S spans V.

The last property allows us to easily determine if a set is a basis of a linear space.

Proposition 11.5.7. Let V be a linear space with finite dimension n, and let $S \subseteq V$ be finite.

- If S spans V but is not a basis of V, then it can be reduced to a basis by removing certain vectors from S.
- If S is linearly independent but not a basis of V, then it can be enlarged to a basis by adding in certain vectors from V.

11.6 Vector Spaces Associated with Matrices

11.6.1 Row Space, Column Space and Null Space

Given an $m \times n$ matrix, there are three special subspaces of $\mathbb{R}^m$ and $\mathbb{R}^n$, namely the row space, column space and null space.

Definition 11.6.1. Let $\mathbf{A} = (a)_{ij}$ be an $m \times n$ matrix. Define the **row vectors** of $\mathbf{A}$ to be

$$\mathbf{r}_i = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix}^\mathsf{T}.$$

Then the **row space** of $\mathbf{A}$, denoted $\operatorname{row}\mathbf{A}$, is the span of the row vectors of $\mathbf{A}$.

Because it is the span of vectors in $\mathbb{R}^n$, it is a subspace of $\mathbb{R}^n$.

Definition 11.6.2. Let $\mathbf{A} = (a)_{ij}$ be an $m \times n$ matrix. Define the **column vectors** of $\mathbf{A}$ to be

$$\mathbf{c}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Then the **column space** of $\mathbf{A}$, denoted $\operatorname{col}\mathbf{A}$, is the span of the column vectors of $\mathbf{A}$.

Because it is the span of vectors in $\mathbb{R}^m$, it is a subspace of $\mathbb{R}^m$.

Definition 11.6.3. Let $\mathbf{A}$ be an $m \times n$ matrix. The **null space** of $\mathbf{A}$ is the solution set to the homogeneous system of equations $\mathbf{Ax} = \mathbf{0}$, i.e.

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} = \mathbf{0}\}.$$

The null space is a subspace of $\mathbb{R}^n$.

Proposition 11.6.4. The row space is orthogonal to the null space.

Proof. Let $\mathbf{x}$ be in the null space of $\mathbf{A}$, and let $\mathbf{y}$ be in the row space of $\mathbf{A}$. Let $\mathbf{r}_i$ be the ith row vector of $\mathbf{A}$. Then

$$\mathbf{Ax} = \begin{pmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{x} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

It follows that $\mathbf{r}_i \cdot \mathbf{x} = 0$ for all $1 \leq i \leq m$. Thus,

$$\mathbf{y} \cdot \mathbf{x} = \left(\sum_{i=1}^{m} k_i \mathbf{r}_i \right) \cdot \mathbf{x} = \sum_{i=1}^{m} k_i \left(\mathbf{r}_i \cdot \mathbf{x} \right) = 0,$$

so $\mathbf{y}$ and $\mathbf{x}$ are orthogonal. Thus, the row space is orthogonal to the null space. □

11.6.2 Range Space and Kernel

Let the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ be represented by the $m \times n$ matrix $\mathbf{A}$. In this section, we will introduce two special subspaces related to T, namely the range space and kernel of T. These two subspaces are equal to the column and null spaces of $\mathbf{A}$ respectively.

Definition 11.6.5. The **range space** of T, denoted $\operatorname{range} T$, consists of all vectors $\mathbf{b}$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Proposition 11.6.6. $\operatorname{range} T$ is equal to the column space of $\mathbf{A}$.

Proof. Consider the equation $\mathbf{A}\mathbf{x} = \mathbf{b}$. Let $\mathbf{c}_i$ be the ith column vector of $\mathbf{A}$. Then we have

$$\mathbf{A}\mathbf{x} = \begin{pmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \mathbf{c}_1 + \cdots + x_n \mathbf{c}_n = \mathbf{b}.$$

Any vector $\mathbf{b} \in \operatorname{range} T$ can be expressed as a linear combination of $\mathbf{c}_1, \dots, \mathbf{c}_n$. Thus, $\mathbf{b}$ is in the column space of $\mathbf{A}$. Likewise, any vector $\mathbf{b}$ in the column space of $\mathbf{A}$ is also in the range space of T. Thus, $\operatorname{range} T$ is equal to the column space of $\mathbf{A}$. □

Definition 11.6.7. The **kernel** of T, denoted $\ker T$, is the set of all vectors $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = 0$.

Proposition 11.6.8. $\ker T$ is equal to the null space of $\mathbf{A}$.

Proof. Trivial. □

11.6.3 Basis for Row Space

Definition 11.6.9. Two matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be **row-equivalent** if their row spaces are the same.

Proposition 11.6.10. $\mathbf{A}$ and its REF/RREF are row-equivalent.

Proof. Recall that an elementary row operation produces a new row that is a linear combination of the old rows. Thus, elementary row operations do not change the row space of a matrix. Since the REF/RREF of $\mathbf{A}$ can be obtained solely from elementary row operations, it follows that $\mathbf{A}$ and its REF/RREF are row-equivalent. □

This result allows us to easily find the basis of the row space of $\mathbf{A}$.

Recipe 11.6.11 (Finding Basis of Row Space)**.** Let $\mathbf{B}$ be the REF/RREF of $\mathbf{A}$. Then the non-zero row vectors in $\mathbf{B}$ form a basis for the row space of $\mathbf{A}$.

Example 11.6.12. Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 21 \end{pmatrix}.$$

Its RREF is given by

$$\begin{pmatrix} 1 & 0 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, a row space basis of $\mathbf{A}$ is

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

11.6.4 Basis for Column Space

One way of finding a basis for the column space of $\mathbf{A}$ would be to find a basis for the row space of $\mathbf{A}^\mathsf{T}$. However, there is a much simpler approach, which we now derive.

Proposition 11.6.13. Row operations do not change the linear dependence on columns.

Proof. Suppose we have a matrix $\mathbf{A} = \begin{pmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_n \end{pmatrix}$. The linear independence of the column vectors depends on the solution set $\mathbf{x}$ to the equation

$$x_1\mathbf{c}_1 + \dots x_n\mathbf{c}_n = 0 \implies \mathbf{A}\mathbf{x} = 0.$$

Suppose now that we perform row operations on $\mathbf{A}$ to obtain a new matrix $\mathbf{A}'$. By writing the above equation as an augmented matrix, we see that the row operations do not change the solution set $\mathbf{x}$!

$$\left(\mathbf{A} \mid \mathbf{0}\right) \to \left(\mathbf{A}' \mid \mathbf{0}\right) \implies x_1\mathbf{c}_1' + \cdots + x_n\mathbf{c}_n' = 0.$$

Thus, if $\mathbf{c}_i$ and $\mathbf{c}_j$ were originally linearly independent, the corresponding columns $\mathbf{c}_i'$ and $\mathbf{c}_j'$ will remain linearly independent. Likewise for columns that were originally linearly dependent. Thus, row operations do not change linear dependence on columns. □

Note however, that row operations do not preserve the column space of $\mathbf{A}$. For instance,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

are row-equivalent, but their column spaces are entirely different.

As a consequence of the above result, we obtain the following corollaries:

Corollary 11.6.14. If $\mathbf{A}$ and $\mathbf{B}$ are row-equivalent, a given set of columns of $\mathbf{A}$ forms a basis for $\operatorname{col}(\mathbf{A})$ if and only if the corresponding set of columns of $\mathbf{B}$ forms a basis for $\operatorname{col}(\mathbf{B})$.

With this, we have our standard procedure for finding a basis for the column space of $\mathbf{A}$:

Recipe 11.6.15 (Finding Basis of Column Space)**.** Let $\mathbf{B}$ be the REF/RREF of $\mathbf{A}$. Look at the columns of $\mathbf{B}$ with a leading entry. Then the corresponding columns of $\mathbf{A}$ form a basis of $\operatorname{col}(\mathbf{A})$.

Example 11.6.16. Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 21 \end{pmatrix}.$$

Its RREF is given by

$$\mathbf{B} = \begin{pmatrix} \boxed{1} & 0 & -1 & -2 & 0 \\ 0 & \boxed{1} & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The first, second and fifth columns of $\mathbf{B}$ contain a leading entry. Thus, the first, second and fifth columns of $\mathbf{A}$ form a basis of $\mathrm{col}(\mathbf{A})$:

$$\left\{ \begin{pmatrix} 1 \\ 6 \\ 11 \\ 16 \end{pmatrix}, \begin{pmatrix} 2 \\ 7 \\ 12 \\ 17 \end{pmatrix}, \begin{pmatrix} 5 \\ 10 \\ 15 \\ 21 \end{pmatrix} \right\}.$$

11.6.5 Basis for Null Space

In the proof of Proposition 11.6.13, we saw how row operations do not change the solution set of the equation $\mathbf{Ax} = \mathbf{0}$. Hence, if $\mathbf{B}$ is the REF/RREF of $\mathbf{A}$, then the equations $\mathbf{Ax} = 0$ and $\mathbf{Bx} = \mathbf{0}$ will have the same solution set.

Recipe 11.6.17 (Finding Basis of Null Space). Let $\mathbf{B}$ be the REF/RREF of $\mathbf{A}$. Then the null space of $\mathbf{A}$ is the solution set $\mathbf{x}$ of $\mathbf{Bx} = \mathbf{0}$.

Example 11.6.18. Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 21 \end{pmatrix}.$$

Its RREF is given by

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

We notice that columns 3 and 4 do not have leading entries. The variables corresponding to these columns can thus be set as free variables.

$$\mathbf{Bx} = \begin{pmatrix} 1 & 0 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{cases} x_1 \quad - x_3 - 2x_4 & = 0 \\ x_2 + 2x_3 + 3x_4 & = 0 \\ x_5 & = 0 \end{cases}.$$

Setting $x_3 = s$ and $x_4 = t$, we have

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} s+2t \\ -2s-3t \\ s \\ t \\ 0 \end{pmatrix} = s\begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Thus, the basis of the null space of $\mathbf{A}$ is

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

11.7 Rank and Nullity for Matrices

Definition 11.7.1. The **row rank** of $\mathbf{A}$ is the dimension of the row space of $\mathbf{A}$. The **column rank** of $\mathbf{A}$ is the dimension of the column space of $\mathbf{A}$.

Proposition 11.7.2. Row and column ranks are equal.

Proof. Recall the procedure we took to find the basis for the row and column space of a matrix:

- The column space basis consists of columns in the original matrix corresponding to the leading entries in the REF/RREF.
- The row space basis consists of the rows of the REF/RREF corresponding to the leading entries.

Since each leading entry corresponds to exactly one row and one column, the sizes of the row and column spaces bases must be equal. Hence, the row and column ranks are equal. □

We give this common value a special name:

Definition 11.7.3. The **rank** of $\mathbf{A}$ is the dimension of the row/column space of $\mathbf{A}$. It is denoted by $\operatorname{rank}\mathbf{A}$.

Let $\mathbf{A}$ be an $m \times n$ matrix. Because the row rank is at most m, and the column rank is at most n, we have that $\operatorname{rank}\mathbf{A} \le \min\{m, n\}$. If equality is achieved, we give $\mathbf{A}$ a special name:

Definition 11.7.4. Let $\mathbf{A}$ be an $m \times n$ matrix. If $\operatorname{rank}\mathbf{A} = \min\{m, n\}$, we say $\mathbf{A}$ has **full rank**.

Proposition 11.7.5. $\operatorname{rank}(\mathbf{AB}) \le \min\{\operatorname{rank}\mathbf{A}, \operatorname{rank}\mathbf{B}\}$.

Proof. Every column in $\mathbf{AB}$ can be expressed as a linear combination of the columns of $\mathbf{A}$, so $\operatorname{col}(\mathbf{AB}) \subseteq \operatorname{col}\mathbf{A}$. Taking dimensions, we see that

$$\operatorname{rank}(\mathbf{AB}) = \dim\operatorname{col}(\mathbf{AB}) \le \dim\operatorname{col}\mathbf{A} = \operatorname{rank}\mathbf{A}.$$

Similarly, every row in $\mathbf{AB}$ can be expressed as a linear combination of the rows of $\mathbf{B}$, so $\operatorname{row}(\mathbf{AB}) \subseteq \operatorname{row}\mathbf{B}$. Taking dimensions,

$$\operatorname{rank}(\mathbf{AB}) = \dim\operatorname{row}(\mathbf{AB}) \le \dim\operatorname{row}\mathbf{B} = \operatorname{rank}\mathbf{B}.$$

Combining these two inequalities gives us what we want. □

We can slightly extend the above result:

Proposition 11.7.6. If $\mathbf{B}$ is an invertible $n \times n$ matrix, then $\operatorname{rank}(\mathbf{AB}) = \operatorname{rank}(\mathbf{BA}) = \operatorname{rank}\mathbf{A}$ for all $n \times n$ matrices $\mathbf{A}$.

Proof. Observe that

$$\operatorname{rank}\mathbf{A} = \operatorname{rank}\left(\mathbf{ABB}^{-1}\right) \le \operatorname{rank}(\mathbf{AB}) \le \operatorname{rank}\mathbf{A},$$

so $\operatorname{rank}(\mathbf{AB}) = \operatorname{rank}\mathbf{A}$. Similarly,

$$\operatorname{rank}\mathbf{A} = \operatorname{rank}\left(\mathbf{B}^{-1}\mathbf{BA}\right) \le \operatorname{rank}(\mathbf{BA}) \le \operatorname{rank}\mathbf{A}.$$

so $\operatorname{rank}(\mathbf{BA}) = \operatorname{rank}\mathbf{A}$. □

Definition 11.7.7. The **nullity** of $\mathbf{A}$ is the dimension of the null space of $\mathbf{A}$. It is denoted by $\operatorname{nullity}\mathbf{A}$.

Theorem 11.7.8 (Rank-Nullity Theorem)**.** For an $m \times n$ matrix $\mathbf{A}$,

$$\operatorname{rank}\mathbf{A} + \operatorname{nullity}\mathbf{A} = \text{number of columns of } \mathbf{A},\ n.$$

Proof. $\operatorname{rank}\mathbf{A}$ is equal to the number of columns in the RREF that contains a leading entry, while $\operatorname{nullity}\mathbf{A}$ is equal to the number of columns in the RREF that does not contain a leading entry. Thus, their sum must be the number of columns in the RREF, which is n. □

We can determine the number of solutions to a system of linear equations using the rank of its corresponding matrix:

Recipe 11.7.9 (Finding Number of Solutions)**.** Let $\mathbf{Ax} = \mathbf{b}$ be a system of linear equations in n variables. Then

- if $\operatorname{rank}\mathbf{A} = \operatorname{rank}\left(\mathbf{A} \mid \mathbf{b}\right) = n$, the system if consistent and has a unique solution.
- if $\operatorname{rank}\mathbf{A} = \operatorname{rank}\left(\mathbf{A} \mid \mathbf{b}\right) < n$, then the system is consistent and has an infinite number of solutions.
- if $\operatorname{rank}\mathbf{A} < \operatorname{rank}\left(\mathbf{A} \mid \mathbf{b}\right)$, then the system is inconsistent and thus has no solution.

In the case where the system is consistent, we can apply the following result to find all possible solutions to the system:

Proposition 11.7.10. If $\mathbf{x}_p$ is a particular solution of a consistent non-homogeneous system $\mathbf{Ax} = \mathbf{b}$, then every solution of the system can be written in the form $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h$, where $\mathbf{x}_h$ is a solution to the corresponding homogeneous system $\mathbf{Ax} = \mathbf{0}$.

Proof. Let $\mathbf{x}_p$ be a fixed solution of $\mathbf{Ax} = \mathbf{b}$, and let $\mathbf{x}$ be an arbitrary solution. Then

$$\mathbf{Ax} = \mathbf{b} \quad \text{and} \quad \mathbf{Ax}_p = \mathbf{b}.$$

Subtracting these equations yields

$$\mathbf{A}\left(\mathbf{x} - \mathbf{x}_p\right) = \mathbf{0},$$

so $\mathbf{x} - \mathbf{x}_p$ is a solution of the homogeneous system $\mathbf{Ax} = \mathbf{0}$. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ form a basis for the null space of $\mathbf{A}$. Then there exist $c_1, \dots, c_k \in \mathbb{R}$ such that

$$\mathbf{x} - \mathbf{x}_p = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k.$$

Letting $\mathbf{x}_h = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$, we see that

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h$$

as desired. □

11.8 Rank and Nullity for Linear Transformations

Definition 11.8.1. Let T be a linear transformation. The dimension of the range of T is called the **rank** of T and the dimension of the kernel of T is called the **nullity** of T.

Theorem 11.8.2 (Rank-Nullity Theorem for Linear Transformations)**.** For a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, where $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, we have

$$\operatorname{rank} T + \operatorname{nullity} T = \operatorname{rank} \mathbf{A} + \operatorname{nullity} \mathbf{A} = n.$$

Proof. Recall that the range of T is the column space of $\mathbf{A}$ and the kernel of T is the null space of $\mathbf{A}$. Hence,

$$\operatorname{rank} T = \dim \operatorname{range} T = \dim \operatorname{col} \mathbf{A} = \operatorname{rank} \mathbf{A}$$

and

$$\operatorname{nullity} T = \dim \ker T = \dim(\text{null space of } \mathbf{A}) = \operatorname{nullity} \mathbf{A}.$$

By the Rank-Nullity Theorem for matrices, we have

$$\operatorname{rank} T + \operatorname{nullity} T = \operatorname{rank} \mathbf{A} + \operatorname{nullity} \mathbf{A} = n.$$

□

12 Eigenvalues, Eigenvectors and Diagonal Matrices

12.1 Eigenvalues and Eigenvectors

Definition 12.1.1. Let $\mathbf{A}$ be an $n \times n$ matrix. Let the non-zero vector $\mathbf{x} \in \mathbb{R}^n$ be such that $\mathbf{Ax}$ is a scalar multiple of $\mathbf{x}$. That is, $\mathbf{x}$ satisfies the equation

$$\mathbf{Ax} = \lambda\mathbf{x}$$

for some scalar λ. The scalar λ is an **eigenvector** of $\mathbf{A}$, and $\mathbf{x}$ is the **eigenvector** of $\mathbf{A}$ corresponding to λ.

12.1.1 Geometrical Interpretation

Let $\mathbf{x}$ be an eigenvector of $\mathbf{A}$ with eigenvalue λ. Geometrically, this means $\mathbf{A}$ maps $\mathbf{x}$ along the same line through the origin as $\mathbf{x}$, but scaling it by a factor of λ. If $\lambda < 0$, the direction is reversed.

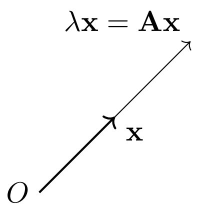


Figure 12.1: $\lambda > 1$

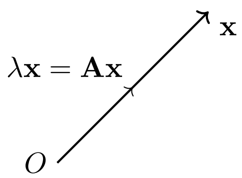


Figure 12.2: $0 \leq \lambda \leq 1$

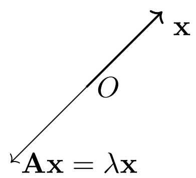


Figure 12.3: $\lambda < 0$

12.1.2 Finding Eigenvalues and Eigenvectors

Definition 12.1.2. The **characteristic polynomial** $\chi(\lambda)$ of an $n \times n$ matrix $\mathbf{A}$ is the n degree polynomial in λ given by

$$\chi(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}).$$

The **characteristic equation** of $\mathbf{A}$ is

$$\chi(\lambda) = 0.$$

Proposition 12.1.3. λ is an eigenvalue of $\mathbf{A}$ if and only if it satisfies the characteristic equation of $\mathbf{A}$.

Proof. To find eigenvalues and eigenvectors, we must solve the equation $\mathbf{Ax} = \lambda\mathbf{x}$. Manipulating this equation, we see that

$$\mathbf{Ax} - \lambda\mathbf{x} = (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = 0.$$

Since $\mathbf{x}$ is non-zero, the null space of $\mathbf{A} - \lambda\mathbf{I}$ must be non-trivial. Thus, $\mathbf{A} - \lambda\mathbf{I}$ must be singular, so

$$\chi(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

Thus, λ satisfies the characteristic equation of $\mathbf{A}$. $\square$

Since the characteristic equation can be easily solved, we now have a straightforward way of finding eigenvalues and eigenvectors.

Recipe 12.1.4 (Finding Eigenvalues and Eigenvectors)**.** We solve the characteristic equation $\chi(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = 0$ to find possible eigenvalues λ. For each λ found, we find its associated eigenvector(s) by finding the basis of the null space of $\mathbf{A} - \lambda\mathbf{I}$.

Sample Problem 12.1.5. Find the eigenvalues and eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}.$$

Sample Problem 12.1.6. The characteristic polynomial is

$$\chi(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \det\begin{pmatrix} 1-\lambda & 2 \\ 5 & 4-\lambda \end{pmatrix} = \lambda^2 - 5\lambda - 6 = (\lambda - 6)(\lambda + 1).$$

Thus, the solutions to the characteristic equation $\chi(\lambda) = 0$ are $\lambda = 6$ and $\lambda = -1$.

Let $\mathbf{x} = (x, y)^{\mathsf{T}}$ be a non-zero vector with $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

Case 1: $\lambda = 6$. We have

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} -5 & 2 \\ 5 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solving, we get $5x - 2y = 0$. Taking $x = 2$ and $y = 5$, the corresponding eigenvector is

$$\mathbf{x} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}.$$

Case 2: $\lambda = -1$. We have

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} 2 & 2 \\ 5 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solving, we get $x + y = 0$. Taking $x = 1$ and $y = -1$, the corresponding eigenvector is

$$\mathbf{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

If $\mathbf{A}$ is a 3×3 matrix, we can use cross products to easily find eigenvectors.

Sample Problem 12.1.7. Let

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

Find the eigenvector of $\mathbf{A}$ corresponding to $\lambda = 1$.

Sample Problem 12.1.8. Let $\mathbf{x}$ be the desired eigenvector. Consider

$$(\mathbf{A} - \mathbf{I})\,\mathbf{x} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

By multiplying out the LHS, we get the following two equations:

$$\mathbf{x} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 0, \quad \mathbf{x} \cdot \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} = 0.$$

These are precisely the equations of two planes, normal to $(1, 0, 1)^{\mathsf{T}}$ and $(-1, 1, 3)^{\mathsf{T}}$ respectively, that also pass through the origin. Thus, $\mathbf{x}$ lies on the line of intersection between the two planes. The direction vector of this line is given by the cross product of the two normal vectors, so

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ -4 \\ 1 \end{pmatrix}.$$

Note that an $n \times n$ matrix may have less than n eigenvalues and eigenvectors. For instance,

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$$

has the sole eigenvalue $\lambda = 3$ with corresponding eigenvector $(1, 0)^{\mathsf{T}}$.

Also, one eigenvalue may have multiple corresponding eigenvectors. For instance,

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

has eigenvalue $\lambda = 2$, which corresponds to two linearly independent eigenvectors:

$$\mathbf{x}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

12.1.3 Useful Results

Proposition 12.1.9. Eigenvectors corresponding to distinct eigenvalues must be linearly independent.

Proof. By way of contradiction, suppose the eigenvectors are linearly dependent. Let j be the maximal j such that $\mathbf{x}_1, \dots, \mathbf{x}_j$ are linearly independent. Then $\mathbf{x}_{j+1}$ can be expressed as a linear combination of $\mathbf{x}_1, \dots, \mathbf{x}_j$:

$$\mathbf{x}_{j+1} = a_1\mathbf{x}_1 + \cdots + a_j\mathbf{x}_j. \tag{1}$$

Applying $\mathbf{A}$ on both sides, we see that

$$\lambda_{j+1}\mathbf{x}_{j+1} = a_1\lambda_1\mathbf{x}_1 + \cdots + a_j\lambda_j\mathbf{x}_j. \tag{2}$$

Since $\mathbf{x}_1, \dots, \mathbf{x}_j$ are linearly independent, we can compare their coefficients in (1) and (2), which gives

$$a_i = a_i \frac{\lambda_i}{\lambda_{j+1}} \implies \lambda_i = \lambda_{j+1}$$

for all $1 \leq i \leq j$. But this clearly contradicts the supposition that the eigenvalues are distinct. Thus, the eigenvectors must be linearly independent. $\square$

Proposition 12.1.10. If $\mathbf{A}$ is a triangular matrix, then the eigenvalues of $\mathbf{A}$ are the entries on the principal diagonal of $\mathbf{A}$.

Proof. Recall that the determinant of a triangular matrix is the product of its principal diagonal entries. Thus,

$$\chi(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = (a_{11} - \lambda)(\alpha_{22} - \lambda) \dots (a_{nn} - \lambda),$$

whence the roots are $\lambda = a_{11}, a_{22}, \dots, a_{nn}$. □

Proposition 12.1.11. Suppose $\mathbf{x}$ is an eigenvector of an $n \times n$ matrix $\mathbf{A}$ with corresponding eigenvalue λ.

(a) For any real number k, $\mathbf{x}$ is an eigenvector of the matrix $k\mathbf{A}$, with corresponding eigenvalue $k\lambda$.

(b) For any positive integer m, $\mathbf{x}$ is an eigenvector of the matrix $\mathbf{A}^m$, with corresponding eigenvalue λ^m.

(c) If $\mathbf{A}$ is invertible, then $\mathbf{x}$ is an eigenvector of $\mathbf{A}^{-1}$ with corresponding eigenvalue λ^{-1} when $\lambda \neq 0$.

(d) If $\mathbf{x}$ is also an eigenvector of an $n \times n$ matrix $\mathbf{B}$ with corresponding eigenvalue μ, then $\mathbf{x}$ is an eigenvector of the sum $\mathbf{A} + \mathbf{B}$, with corresponding eigenvalue $\lambda + \mu$.

Proof of (a). Since $\mathbf{A} = \lambda \mathbf{x}$, we have $(k\mathbf{A})\,\mathbf{x} = (k\lambda)\,\mathbf{x}$. □

Proof of (b). We use induction. Let the statement $P(m)$ be such that

$$P(m) \iff \mathbf{x} \text{ is an eigenvector of the matrix } \mathbf{A}^m \text{ with corresponding eigenvalue } \lambda^m.$$

The base case $m = 1$ is trivial. Suppose $P(k)$ is true for some $k \in \mathbb{N}$. Then

$$\mathbf{A}^{k+1}\mathbf{x} = \mathbf{A}\left(\mathbf{A}^k\mathbf{x}\right) = \mathbf{A}\left(\lambda^k\mathbf{x}\right) = \lambda^k\left(\mathbf{A}\mathbf{x}\right) = \lambda^k\left(\lambda\mathbf{x}\right) = \lambda^{k+1}\mathbf{x}.$$

Thus, $P(k) \implies P(k+1)$. This closes the induction. □

Proof of (c). Since $\mathbf{A} = \lambda \mathbf{x}$, we have

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\lambda\mathbf{x} = \lambda\left(\mathbf{A}^{-1}\mathbf{x}\right) \implies \mathbf{A}^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}.$$

□

Proof of (d). Since $\mathbf{A} = \lambda \mathbf{x}$ and $\mathbf{B} = \mu \mathbf{x}$, we have

$$(\mathbf{A} + \mathbf{B})\,\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x} = (\lambda + \mu)\,\mathbf{x}.$$

□

Corollary 12.1.12. Let $\mathbf{x}$ be an eigenvector of $\mathbf{A}$ with corresponding eigenvalue λ. Define a polynomial $p(X) = a_0 + a_1X + a_2X^2 + \dots + a_nX^n$. Then $p(\lambda)\mathbf{x} = p(\mathbf{A})\mathbf{x}$.

Note that we are taking a_0 to mean $a_0\mathbf{I}$ on the RHS.

Definition 12.1.13. A **submatrix** of $\mathbf{A}$ is a matrix obtained from $\mathbf{A}$ by deleting a collection of rows and/or columns. A **principal minor** of $\mathbf{A}$ is a submatrix whereby the indices of the deleted rows are the same as the indices of the deleted columns.

Example 12.1.14. Given

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$

the following three matrices are submatrices of $\mathbf{A}$:

$$\mathbf{B}_1 = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}, \quad \mathbf{B}_2 = \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix}.$$

To obtain $\mathbf{B}_1$, we deleted the third row and third column. To obtain $\mathbf{B}_2$, we deleted the first row and second column. Note that $\mathbf{B}_1$ is also a principal submatrix.

Proposition 12.1.15. Let $\mathbf{A}$ be an $n \times n$ matrix. Let

$$E_k = \sum_{\mathbf{S} \in S_k} |\mathbf{S}|$$

be the sum of the determinants of all $k \times k$ principal submatrices. We define $E_0 = 1$. Then the characteristic polynomial $\chi(\lambda)$ of $\mathbf{A}$ is given by

$$\chi(\lambda) = \sum_{i=0}^{n} (-1)^i E_{n-i} \lambda^i.$$

Example 12.1.16. Consider

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

Then

$$\begin{aligned} E_1 &= |2| + |2| + |2| = 6, \\ E_2 &= \begin{vmatrix} 2 & 0 \\ -1 & 2 \end{vmatrix} + \begin{vmatrix} 2 & 3 \\ 0 & 2 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 11, \\ E_3 &= \begin{vmatrix} 2 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 0 & 2 \end{vmatrix} = 6. \end{aligned}$$

Invoking the above result, we see that

$$\chi(\lambda) = -\lambda^3 + E_1\lambda^2 - E_2\lambda + E_3 = -\lambda^3 + 6\lambda^2 - 11\lambda + 6.$$

Corollary 12.1.17. If $\mathbf{A}$ is an $n \times n$ matrix,

- The sum of the n eigenvalues of $\mathbf{A}$ (counting multiplicity) is equal to the trace of $\mathbf{A}$.
- The product of the n eigenvalues of $\mathbf{A}$ (counting multiplicity) is equal to the determinant of $\mathbf{A}$.

Proof. Apply Vieta's formula to the above result. □

12.2 Diagonal Matrices

Recall that a diagonal matrix $\mathbf{D}$ is a square matrix where all off-diagonal entries are zero. Diagonal matrices have nice properties that make computations involving them simple and convenient:

- $\det \mathbf{D}$ is the product of its diagonal entries.
- If $\det \mathbf{D} \neq 0$, then $\mathbf{D}^{-1}$ is a diagonal matrix with the corresponding reciprocals in the diagonal.
- $\mathbf{D}^n$ is a diagonal matrix with the corresponding powers in the diagonal.

For instance, if

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

then

$$\mathbf{D}^{100} = \begin{pmatrix} 1^{100} & 0 & 0 \\ 0 & 2^{100} & 0 \\ 0 & 0 & 3^{100} \end{pmatrix} \quad \text{and} \quad \mathbf{D}^{-100} = \begin{pmatrix} 1^{-100} & 0 & 0 \\ 0 & 2^{-100} & 0 \\ 0 & 0 & 3^{-100} \end{pmatrix}.$$

12.2.1 Diagonalization

The useful properties of diagonal matrices motivates us to find a way to write an $n \times n$ matrix in terms of a diagonal matrix, i.e. diagonalize $\mathbf{A}$ in some way.

Definition 12.2.1. A matrix $\mathbf{A}$ is **diagonalizable** if there exists an invertible matrix $\mathbf{Q}$ such that $\mathbf{A} = \mathbf{QDQ}^{-1}$, where $\mathbf{D}$ is a diagonal matrix. We say that $\mathbf{Q}$ **diagonalizes** $\mathbf{A}$.

Proposition 12.2.2. If $\mathbf{A}$ is diagonalizable, then the columns of $\mathbf{Q}$ are the linearly independent eigenvectors of $\mathbf{A}$, and the diagonal matrix $\mathbf{D}$ contains the corresponding eigenvalues.

Proof. Let $\mathbf{A}$ be an $n \times n$ matrix with eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ corresponding to the real eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $\mathbf{Q}$ be the matrix with $\mathbf{x}_1, \ldots, \mathbf{x}_n$ as its columns and let $\mathbf{D}$ be a diagonal matrix with its diagonal entries as $\lambda_1, \ldots, \lambda_n$:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{x}_1 & \ldots & \mathbf{x}_n \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}.$$

Then

$$\mathbf{AQ} = \begin{pmatrix} \mathbf{Ax}_1 & \ldots & \mathbf{Ax}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{x}_1 & \ldots & \lambda_n\mathbf{x}_n \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 & \ldots & \mathbf{x}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix} = \mathbf{QD}.$$

Post-multiplying both sides by $\mathbf{Q}^{-1}$, which exists since the columns of $\mathbf{Q}$ are linearly independent, we have $\mathbf{A} = \mathbf{QDQ}^{-1}$. $\square$

Note that if $\mathbf{A}$ has n real and distinct eigenvalues, it will have n linearly independent eigenvectors, so it will be diagonalizable. However, if it has repeated eigenvalues, it may not be diagonalizable.

Sample Problem 12.2.3. Let

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

Find a matrix $\mathbf{Q}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{QDQ}^{-1}$.

Solution. We previously found the corresponding eigenvectors for eigenvalues 1, 2, 3 to be

$$\begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Thus,

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 & 1 \\ 4 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

□

Note that $\mathbf{Q}$ and $\mathbf{D}$ are not unique. Using the above sample problem, we could have taken

$$\mathbf{Q} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 4 & 1 \\ 1 & -1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

12.2.2 Computing Matrix Powers

One of the more useful purposes of diagonalization is to compute matrix powers.

Proposition 12.2.4. Suppose $\mathbf{A} = \mathbf{QDQ}^{-1}$ is diagonalizable. Then

$$\mathbf{A}^k = \mathbf{QD}^k\mathbf{Q}^{-1}.$$

Proof. Observe that

$$\begin{aligned} \mathbf{A}^k &= \left(\mathbf{QDQ}^{-1}\right)\left(\mathbf{QDQ}^{-1}\right)\dots\left(\mathbf{QDQ}^{-1}\right) = \mathbf{QD}\left(\mathbf{Q}^{-1}\mathbf{Q}\right)\mathbf{D}\left(\mathbf{Q}^{-1}\mathbf{Q}\right)\dots\mathbf{DQ}^{-1} \\ &= \mathbf{QDD}\dots\mathbf{DQ}^{-1} = \mathbf{QD}^k\mathbf{Q}^{-1}. \end{aligned}$$

□

Sample Problem 12.2.5. Let

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

Compute $\mathbf{A}^{10}$.

Solution. We previously found that $\mathbf{A} = \mathbf{QDQ}^{-1}$, where

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 & 1 \\ 4 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Thus,

$$\mathbf{A}^{10} = \mathbf{QD}^{10}\mathbf{Q}^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 4 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1^{10} & 0 & 0 \\ 0 & 2^{10} & 0 \\ 0 & 0 & 3^{10} \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 4 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix}^{-1},$$

which evaluates to

$$\mathbf{A}^{10} = \begin{pmatrix} 29525 & 0 & 29524 \\ 55979 & 1024 & 60071 \\ 29524 & 0 & 29525 \end{pmatrix}.$$

□

Part IV

Complex Numbers

13 Introduction to Complex Numbers

Definition 13.0.1. The **imaginary unit** i is a root to the equation

$$x^2 + 1 = 0.$$

13.1 Cartesian Form

Definition 13.1.1. A **complex number** z has **Cartesian form** $x + \mathrm{i}y$, where x and y are real numbers. We call x the **real part** of z, denoted $\operatorname{Re} z$. Likewise, we call y the **imaginary part** of z, denoted $\operatorname{Im} z$.

Definition 13.1.2. The set of complex numbers is denoted $\mathbb{C}$ and is defined as

$$\mathbb{C} = \{z : z = x + \mathrm{i}y, \quad x, y \in \mathbb{R}\}.$$

Remark. The set of real numbers, $\mathbb{R}$, is a proper subset of the set of complex numbers, $\mathbb{C}$. That is, $\mathbb{R} \subset \mathbb{C}$.

Fact 13.1.3 (Algebraic Operations on Complex Numbers)**.** Let $z_1, z_2, z_3 \in \mathbb{C}$.

- Two complex numbers are equal if and only if their corresponding real and imaginary parts are equal.

$$z_1 = z_2 \iff \operatorname{Re} z_1 = \operatorname{Re} z_2 \text{ and } \operatorname{Im} z_1 = \operatorname{Im} z_2.$$

- Addition of complex numbers is commutative, i.e.

$$z_1 + z_2 = z_2 + z_1$$

  and associative, i.e.

$$(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3).$$

- Multiplication of complex numbers is commutative, i.e.

$$z_1 z_2 = z_2 z_1,$$

  associative, i.e.

$$z_1(z_2 z_3) = (z_1 z_2) z_3$$

  and distributive, i.e.

$$z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3.$$

Proposition 13.1.4. Complex numbers cannot be ordered.

Proof. Seeking a contradiction, suppose $\mathrm{i} > 0$. Multiplying both sides by i, we have $\mathrm{i}^2 = -1 > 0$, a contradiction. Hence, we must have $\mathrm{i} < 0$. However, multiplying both sides by i and changing signs (since $\mathrm{i} < 0$), we have $\mathrm{i}^2 = -1 > 0$, another contradiction. Thus, $\mathbb{C}$ cannot be ordered. $\square$

13.2 Argand Diagram

We can represent complex numbers in the complex plane using an Argand diagram.

Definition 13.2.1. The **Argand diagram** is a modified Cartesian plane where the x-axis represents real numbers and the y-axis represents imaginary numbers. The two axes are called the **real axis** and **imaginary axis** correspondingly.

On the Argand diagram, the complex number $z = x + \mathrm{i}y$, where $x, y \in \mathbb{R}$, can be represented by

- the point $Z(x, y)$ or $Z(z)$; or
- the vector $\overrightarrow{OZ}$.

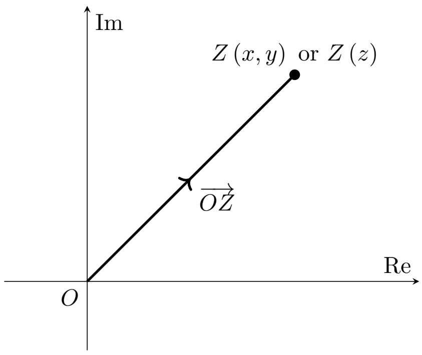


Figure 13.1

In an Argand diagram, let the points Z and W represent the complex numbers z and w respectively. Then $\overrightarrow{OZ}$ and $\overrightarrow{OW}$ are the corresponding vectors representing z and w.

13.2.1 Modulus

Recall in §1, we defined the modulus of a real number x as the "distance" between x and the origin on the real number line. Generalizing this notion to complex numbers, it makes sense to define the modulus of a real number z as the "distance" between z and the origin on the complex plane. This uses Pythagoras' theorem.

Definition 13.2.2. The **modulus** of a complex number z is denoted $|z|$ and is defined as

$$|z| = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}.$$

13.2.2 Complex Conjugate

Definition 13.2.3. The **conjugate** of the complex number $z = x + \mathrm{i}y$ is denoted z^* with definition

$$z^* = x - \mathrm{i}y.$$

We refer to z and z^* as a **conjugate pair** of complex numbers.

On an Argand diagram, the conjugate z^* is the reflection of z about the real axis.

Fact 13.2.4 (Properties of Complex Conjugates).

- (distributive over addition) $(z+w)^* = z^* + w^*$.
- (distributive over multiplication) $(zw)^* = z^* w^*$.
- (involution) $(z^*)^* = z$.
- $z + z^* = 2\,\mathrm{Re}(z)$.
- $z - z^* = 2\,\mathrm{Im}(z)\,\mathrm{i}$.
- $zz^* = \mathrm{Re}(z)^2 + \mathrm{Im}(z)^2 = |z|^2$.

Because conjugation is distributive over addition and multiplication, we also have the following identities:

$$(kz)^* = kz^*, \qquad (z^n)^* = (z^*)^n,$$

where $k \in \mathbb{R}$ and $n \in \mathbb{Z}$.

Using the conjugate of a complex number z, the reciprocal of z can be computed as

$$\frac{1}{z} = \frac{z^*}{zz^*} = \frac{z^*}{|z|^2}.$$

13.2.3 Argument

Definition 13.2.5. The **argument** of a complex number z is the directed angle θ that $Z(z)$ makes with the positive real axis, and is denoted by $\arg(z)$. Note that $\arg(z) > 0$ when measured in an anticlockwise direction from the positive real axis, and $\arg(z) < 0$ when measured in a clockwise direction from the positive real axis.

Note that $\arg(z)$ is not unique; the position of $Z(z)$ is not affected by adding an integer multiple of 2π to θ. Therefore, if $\arg(z) = \phi$, then $\phi + 2k\pi$, where $k \in \mathbb{Z}$, is also an argument of z. We hence introduce the principal argument of z.

Definition 13.2.6. The value of $\arg(z)$ in the interval $(-\pi, \pi]$ is known as the **principal argument** of z.

The modulus $r = |z|$, complex conjugate z^* and argument $\theta = \arg(z)$ of a complex number z can easily be identified on an Argand diagram:

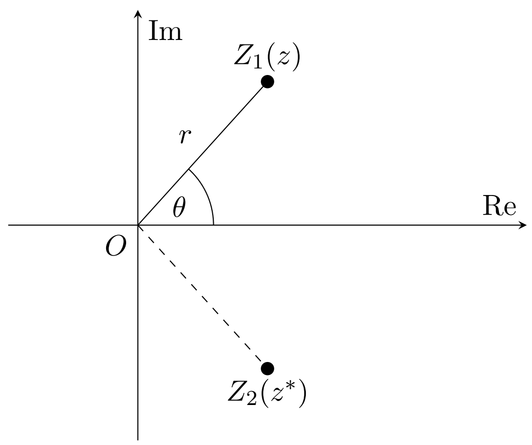


Figure 13.2

13.3 Polar Form

Instead of using Cartesian coordinates on an Argand diagram, we can use polar coordinates, leading to the polar form of a complex number. This polar form can be expressed in two ways: trigonometric form and exponential form.

Definition 13.3.1. The **trigonometric form** of the complex number z is

$$z = r\left(\cos\theta + \mathrm{i}\sin\theta\right),$$

where $r = |z|$ and $\theta = \arg(z)$, $-\pi < \theta \leq \pi$.

Theorem 13.3.2 (Euler's Identity)**.** For all $\theta \in \mathbb{R}$,

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

Proof 1 (Series Expansion). By the standard series expansion of e^x, we have

$$\mathrm{e}^{\mathrm{i}\theta} = 1 + \mathrm{i}\theta + \frac{(\mathrm{i}\theta)^2}{2!} + \frac{(\mathrm{i}\theta)^3}{3!} + \frac{(\mathrm{i}\theta)^4}{4!} + \frac{(\mathrm{i}\theta)^5}{5!} + \dots.$$

Simplifying and grouping real and imaginary parts together,

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

which we recognize to be the standard series expansions of $\cos\theta$ and $\sin\theta$ respectively. Hence,

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

□

Proof 2 (Differentiation). Let $f(\theta) = \mathrm{e}^{-\mathrm{i}\theta}\left(\cos\theta + \mathrm{i}\sin\theta\right)$. Differentiating with respect to θ,

$$f'(\theta) = \mathrm{e}^{-\mathrm{i}\theta}\left(-\sin\theta + \mathrm{i}\cos\theta\right) - \mathrm{i}\mathrm{e}^{-\mathrm{i}\theta}\left(\cos\theta + \mathrm{i}\sin\theta\right) = 0.$$

Hence, $f(\theta)$ is constant. Evaluating $f(\theta)$ at $\theta = 0$, we have $f(\theta) = 1$, whence

$$\mathrm{e}^{-\mathrm{i}\theta}\left(\cos\theta + \mathrm{i}\sin\theta\right) = 1 \implies \mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta.$$

□

Definition 13.3.3. The **exponential form** of the complex number z is

$$z = r\mathrm{e}^{\mathrm{i}\theta},$$

where $r = |z|$ and $\theta = \arg(z)$, $-\pi < \theta \leq \pi$.

Recall z^* is the reflection of z about the real axis. Hence, we clearly have the following:

Proposition 13.3.4 (Conjugation in Polar Form)**.** If $z = r\mathrm{e}^{\mathrm{i}\theta}$, then $z^* = r\mathrm{e}^{-\mathrm{i}\theta}$. Also,

$$\arg(z^*) = -\theta = -\arg(z), \qquad |z| = r = |z^*|.$$

Using the proposition above, we can convert the results $z + z^* = 2\operatorname{Re}(z)$ and $z - z^* = 2\operatorname{Im}(z)\,\mathrm{i}$ into polar form:

Proposition 13.3.5.

$$e^{i\theta} + e^{-i\theta} = 2\cos\theta, \qquad e^{i\theta} - e^{-i\theta} = (2\sin\theta)\,i.$$

Lastly, we observe the effect of multiplication and division on the modulus and argument of complex numbers.

Proposition 13.3.6 (Multiplication in Polar Form)**.** Let $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$. Then

$$|z_1 z_2| = r_1 r_2 = |z_1|\,|z_2|\,, \qquad \arg(z_1 z_2) = \theta_1 + \theta_2 = \arg(z_1) + \arg(z_2)\,.$$

Proof. Observe that

$$z_1 z_2 = \left(r_1 e^{i\theta_1}\right)\left(r_2 e^{i\theta_2}\right) = (r_1 r_2) e^{i(\theta_1+\theta_2)}.$$

The results follow immediately. □

Corollary 13.3.7 (Exponentiation in Polar Form)**.** For $n \in \mathbb{Z}$,

$$|z^n| = r^n = |z|^n\,, \qquad \arg(z^n) = n\theta = n\arg(z)\,.$$

Proof. Repeatedly apply the above proposition. □

Proposition 13.3.8 (Division in Polar Form)**.** Let $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$. Then

$$\left|\frac{z_1}{z_2}\right| = \frac{r_1}{r_2} = \frac{|z_1|}{|z_2|}, \qquad \arg\left(\frac{z_1}{z_2}\right) = \theta_1 - \theta_2 = \arg(z_1) - \arg(z_2)\,.$$

Proof. Observe that

$$\frac{z_1}{z_2} = \frac{r_1 e^{i\theta_1}}{r_2 e^{i\theta_2}} = \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}.$$

The results follow immediately. □

13.4 De Moivre's Theorem

Theorem 13.4.1 (De Moivre's Theorem)**.** For $n \in \mathbb{Q}$, if $z = r(\cos\theta + i\sin\theta) = re^{i\theta}$, then

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

Proof. Write z^n in exponential form before converting it into trigonometric form. □

We now discuss some of the applications of de Moivre's theorem.

Recipe 13.4.2 (Finding nth Roots)**.** Suppose we want to find the nth roots of a complex number $w = re^{i\theta}$. We begin by setting up the equation

$$z^n = w = re^{i(\theta+2k\pi)},$$

where $k \in \mathbb{Z}$. Next, we take nth roots on both sides, which yields

$$z = r^{1/n} e^{i(\theta+2k\pi)/n}.$$

Lastly, we pick values of k such that $\arg z = \frac{\theta+2k\pi}{n}$ lies in the principal interval $(-\pi, \pi]$.

Definition 13.4.3. Let $n \in \mathbb{Z}$. The **nth roots of unity** are the n solutions to the equation

$$z^n - 1 = 0.$$

Proposition 13.4.4 (Roots of Unity in Polar Form)**.** The nth roots of unity are given by

$$z = \cos\frac{2k\pi}{n} + \mathrm{i}\sin\frac{2k\pi}{n} = \mathrm{e}^{\mathrm{i}(2k\pi/n)},$$

where $k \in \mathbb{Z}$.

Proof. Use de Moivre's theorem. □

Fact 13.4.5 (Geometric Properties of Roots of Unity)**.** On an Argand diagram, the nth roots of unity

- all lie on a circle of radius 1.
- are equally spaced apart.
- form a regular n-gon.

De Moivre's theorem can also be used to derive trigonometric identities. The trigonometric identities one will be required to prove typically involve reducing "powers" to "multiple angles" (e.g. expressing $\sin^3\theta$ in terms of $\sin\theta$ and $\sin 3\theta$), or vice versa.

Proposition 13.4.6 (Power to Multiple Angles)**.** Let $z = \cos\theta + \mathrm{i}\sin\theta = \mathrm{e}^{\mathrm{i}\theta}$. Then

$$z^n + z^{-n} = 2\cos n\theta, \qquad z^n - z^{-n} = 2\mathrm{i}\sin n\theta.$$

Proof. Use de Moivre's theorem □

Recipe 13.4.7 (Multiple Angles to Powers)**.** Suppose we want to express $\cos n\theta$ and $\sin n\theta$ in terms of powers of $\sin\theta$ and $\cos\theta$. We begin by invoking de Moivre's theorem:

$$\cos n\theta + \mathrm{i}\sin n\theta = \left(\cos\theta + \mathrm{i}\sin\theta\right)^n.$$

Next, using the binomial theorem,

$$\cos n\theta + \mathrm{i}\sin n\theta = \sum_{k=0}^{n}\binom{n}{k}\cos^k\theta\sin^{n-k}\theta.$$

We then take the real and imaginary parts of both sides to isolate $\cos n\theta$ and $\sin n\theta$:

$$\cos n\theta = \operatorname{Re}\sum_{k=0}^{n}\binom{n}{k}\cos^k\theta\sin^{n-k}\theta, \qquad \sin n\theta = \operatorname{Im}\sum_{k=0}^{n}\binom{n}{k}\cos^k\theta\sin^{n-k}\theta.$$

Example 13.4.8. Suppose we want to write $\sin 2\theta$ in terms of $\sin\theta$ and $\cos\theta$. Using de Moivre's theorem,

$$\cos 2\theta + \mathrm{i}\sin 2\theta = \left(\cos\theta + \mathrm{i}\sin\theta\right)^2 = \cos^2\theta + 2\mathrm{i}\cos\theta\sin\theta - \sin^2\theta.$$

Comparing imaginary parts, we obtain $\sin 2\theta = 2\cos\theta\sin\theta$ as expected.

Another way to derive new trigonometric identities is to differentiate known identities.

Example 13.4.9. Using the "power to multiple angle" formula above, one can show that

$$\cos^6\theta = \frac{1}{32}\left(\cos 6\theta + 6\cos 4\theta + 15\cos 2\theta + 10\right).$$

Differentiating, we obtain a new trigonometric identity:

$$\sin\theta\cos^5\theta = \frac{1}{32}\left(\sin 6\theta + 4\sin 4\theta + 5\sin 2\theta\right).$$

13.5 Solving Polynomial Equations over $\mathbb{C}$

Theorem 13.5.1 (Fundamental Theorem of Algebra)**.** A non-zero, single-variable, degree n polynomial with complex coefficients has n roots in $\mathbb{C}$, counted with multiplicity.

Theorem 13.5.2 (Conjugate Root Theorem)**.** For a polynomial equation with all real coefficients, non-real roots must occur in conjugate pairs.

Proof. Suppose z is a non-real root to the polynomial $P(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, where $a_n, a_{n-1}, \ldots, a_1, a_0 \in \mathbb{R}$. Consider $P(z^*)$.

$$P(z^*) = a_n\left(z^*\right)^n + a_{n-1}\left(z^*\right)^{n-1} + \cdots + a_1\left(z^*\right) + a_0.$$

By conjugation properties, this simplifies to

$$P(z^*) = \left(a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0\right)^*,$$

which clearly evaluates to 0, whence z^* is also a root of $P(z)$. $\square$

14 Geometrical Effects of Complex Numbers

14.1 Geometrical Effect of Addition

The following diagram shows the geometrical effect of addition on complex numbers. Here, the point P represents the complex number $z+w$. Observe that $OWPZ$ is a parallelogram (due to the parallelogram law of vector addition).

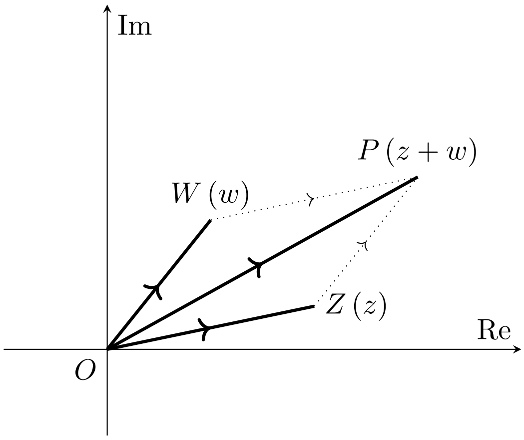


Figure 14.1

14.2 Geometrical Effect of Scalar Multiplication

The following diagram shows the geometrical effect of multiplying a complex number by a real number k. Here, Z_1 represents a point where $k > 1$, Z_2 where $0 < k < 1$, and Z_3 where $k < 0$. Observe that the points lie on the straight line passing through the origin O and the point Z.

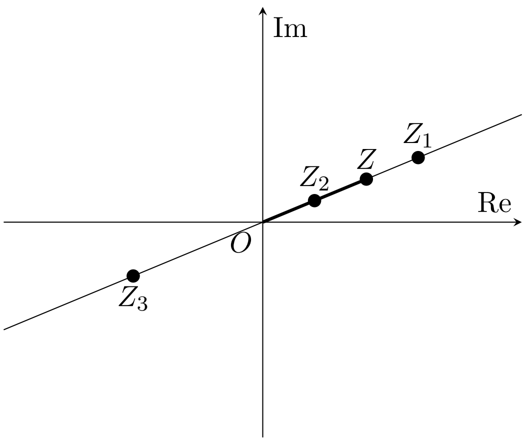


Figure 14.2

14.3 Geometrical Effect of Complex Multiplication

Let points P, Q and R represent the complex numbers z_1, z_2 and z_3 respectively, as illustrated in the Argand diagram below.

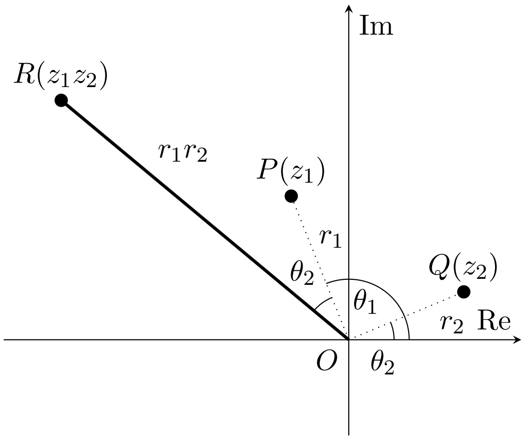


Figure 14.3

Geometrically, the point $R(z_1 z_2)$ is obtained by

1. scaling by a factor of r_2 on $\overrightarrow{OP}$ to obtain a new modulus of $r_1 r_2$, followed by
2. rotating $\overrightarrow{OP}$ through an angle θ_2 about O in an anti-clockwise direction if $\theta_2 > 0$ to obtain a new argument $\theta_1 + \theta_2$ (or in a clockwise direction if $\theta_2 < 0$).

14.4 Loci in Argand Diagram

Definition 14.4.1. The **locus** (plural: loci) of a variable point is the path traced out by the point under certain conditions.

14.4.1 Standard Loci

Fact 14.4.2 (Circle)**.** For $|z - a| = r$, with P representing the complex number z and A representing the fixed complex number a and $r > 0$, the locus of P is a circle with centre A and radius r.

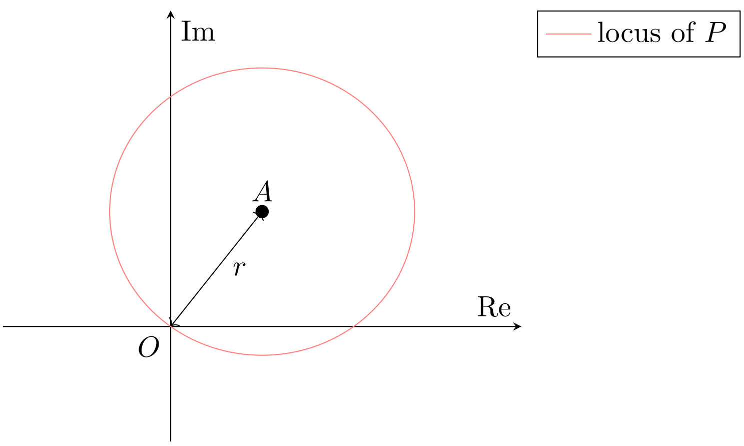


Figure 14.4

Fact 14.4.3 (Perpendicular Bisector). For $|z-a|=|z-b|$, with P representing the complex number z, points A and B representing the fixed complex numbers a and b respectively, the locus of P is the perpendicular bisector of the line segment joining A and B.

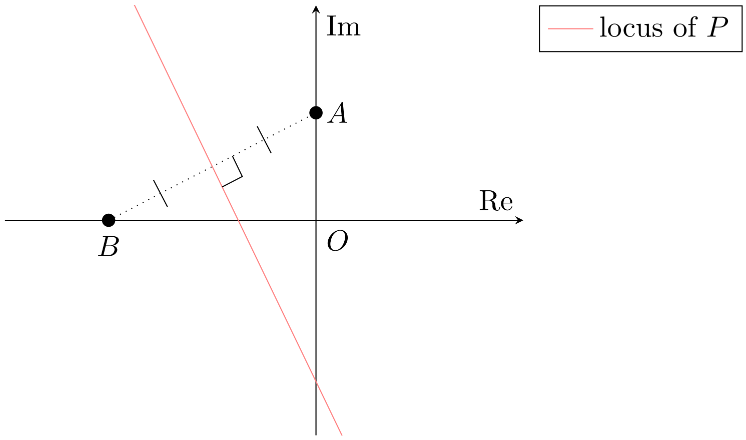


Figure 14.5

Fact 14.4.4 (Half-Line). For $\arg(z-a)=\theta$, with P representing the complex number z and point A representing the fixed complex number a, the locus of P is the half-line starting from A (excluding this point) and inclined at a directed angle θ to the positive real axis.

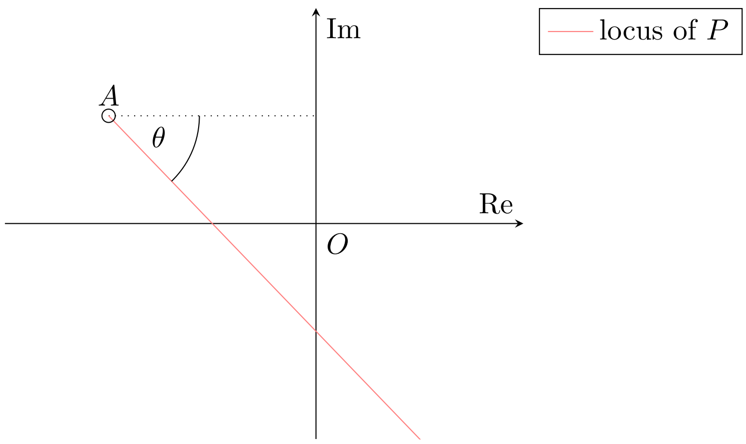


Figure 14.6

14.4.2 Non-Standard Loci

When sketching non-standard loci, one useful technique is to write the equation in Cartesian form, i.e. letting $z=x+\mathrm{i}y$, $x, y \in \mathbb{R}$.

Example 14.4.5. Let P be the point representing the complex number z, where z satisfies the equation $\operatorname{Re} z+2 \operatorname{Im} z=2$. We begin by writing z in Cartesian form, i.e. $z=x+\mathrm{i}y$, $x, y \in \mathbb{R}$. Substituting this into the equation, we have $x+2y=2$. Thus, the locus of P is given by the equation $x+2y=2$.

14.4.3 Loci and Inequalities

We will use the inequality $|z-(3+4\mathrm{i})|<5$ as an example to illustrate the general procedure of finding the locus of an inequality.

We begin by considering the equality case. As we have seen above, $|z-(3+4\mathrm{i})|=5$ corresponds to a circle centred at $(3,4)$ with radius 5. This is the "boundary" of our locus.

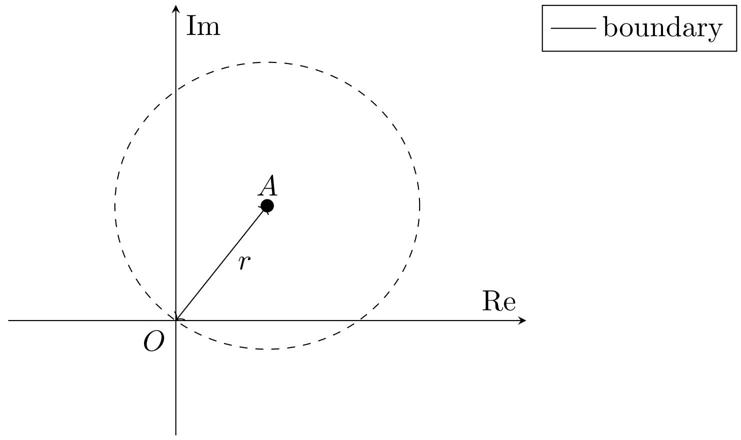


Figure 14.7

Notice that the circle is dashed as the inequality is strict; if the inequality was not strict, i.e. $|z-(3+4\mathrm{i})|\leq 5$, the circle would be drawn with a solid line.

Now, observe that the complex plane has been split into two parts: the interior and exterior of the circle. To determine which region satisfies our inequality, we simply test a complex number in each region.

- Since $3+4\mathrm{i}$ is in the interior of the circle, and $|(3+4\mathrm{i})-(3+4\mathrm{i})|=0<5$, the interior of the circle satisfies the inequality.
- Since $10+4\mathrm{i}$ is in the exterior of the circle, and $|(10+4\mathrm{i})-(3+4\mathrm{i})|=7>5$, the exterior of the circle does not satisfy the inequality.

We thus conclude that the locus of $|z-(3+4\mathrm{i})|<5$ is the interior region of the circle, as shaded below:

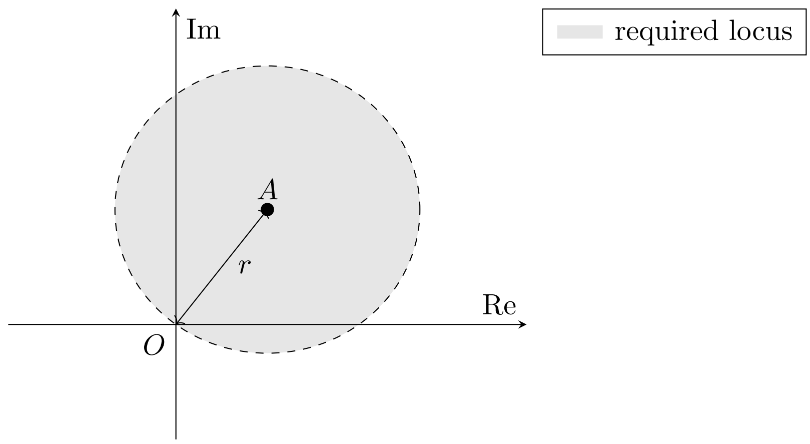


Figure 14.8

14.4.4 Further Use of the Argand Diagram

Many interesting and varied problems involving complex numbers can be solved simply using an Argand diagram. For instance, one may ask what the range of $\arg z$ is, given that z satisfies some other constraint, e.g. $|z-\mathrm{i}|=1$. Given how diverse these problems may be, there is no general approach to solving them. However, there are several tips that one should keep in mind when doing these problems:

- Think geometrically, not algebraically. Draw out the given constraints on an Argand diagram. Most of the time, the given constraints are simply the three standard loci above (circles, perpendicular bisector and half-lines).
- When working with circles and an external point, drawing tangents and diameters may help. This allows one to use properties of circles (e.g. tangents are perpendicular to the radius).
- Keep an eye out for symmetry or similar figures.

Part V

Analysis

15 Differentiation

15.1 Limits

Let a be a constant.

- $x \to a$ means "x approaches the value a",
- $x \to a^-$ means "x approaches the value a from a value slightly more than a",
- $x \to a^+$ means "x approaches the value a from a value slightly more than a",
- $\lim_{x\to a} f(x)$ means "the limit of $f(x)$ as x approaches a".

Definition 15.1.1. The **limit** of $f(x)$ as x approaches a exists if there exists some $l \in \mathbb{R}$ such that

$$\lim_{x\to a^-} f(x) = l = \lim_{x\to a^+} f(x).$$

We write

$$\lim_{x\to a} f(x) = l.$$

15.2 Derivative

Definition 15.2.1. The **gradient of a straight line** is defined as the ratio of the change in the y-coordinate to that of the x-coordinate between any two points on the line. Mathematically, the gradient m is given by

$$m = \frac{y_2 - y_1}{x_2 - x_1},$$

where (x_1, y_1) and (x_2, y_2) are two points on the line.

Definition 15.2.2. The **tangent to the curve at A** is the line *touching* the curve at A.

Definition 15.2.3. The **instantaneous rate of change** or **gradient** of a curve at any point is defined as the gradient of the tangent to the curve at the point.

Definition 15.2.4. The **derivative** of a function $f(x)$, denoted $\frac{\mathrm{d}}{\mathrm{d}x} f(x)$ or $f'(a)$, represents the instantaneous rate of change of $f(x)$ with respect to x.

If $y = f(x)$, we write the derivative as $\frac{\mathrm{d}y}{\mathrm{d}x}$ or y'. Note that the symbol $\frac{\mathrm{d}}{\mathrm{d}x}$ means "the derivative with respect to x of" and should be treated as an operation, not a fraction.

Definition 15.2.5. The **nth derivative** of y with respect to x is

$$\frac{\mathrm{d}^n y}{\mathrm{d}x^n} = f^{(n)}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}}\right),$$

where $n \in \mathbb{Z}^+$.

15.2.1 Differentiation from First Principles

Consider a curve $y = f(x)$. Let $A(x, f(x))$ and $B(x + \Delta x, f(x + \Delta x))$ be two points on the curve, where Δx is a small increment in x.

Observe that the gradient of the tangent to the curve at A can be approximated by the gradient of the chord AB, denoted m_{AB}. The closer B is to A, the better the approximation. Therefore, the gradient of the curve at point A is $\lim_{B\to A} m_{AB}$. Now observe that

$$m_{AB} = \frac{f(x+\Delta x) - f(x)}{(x+\Delta x) - x} = \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

Additionally, as $B \to A$, $\Delta x \to 0$. Hence,

$$\lim_{B\to A} m_{AB} = \lim_{\Delta x\to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x\to 0} \frac{\Delta y}{\Delta x} = \frac{\mathrm{d}y}{\mathrm{d}x}.$$

For convenience, we replace Δx with h. The derivative is hence

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x} f(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}.$$

15.3 Differentiation Rules

Proposition 15.3.1 (Differentiation Rules). Let $k \in \mathbb{R}$ and suppose u and v are functions of x. Then

- (Sum/Difference Rule) If $y = u \pm v$ then $y' = u' \pm v'$.
- (Product Rule) If $y = uv$, then $y' = u'v + uv'$.
- (Quotient Rule) If $y = \frac{u}{v}$, then $y' = \frac{u'v - uv'}{v^2}$.
- (Chain Rule) If $y = f(x)$ and $x = g(t)$, then $\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}$.

The sum, product and quotient rules are easy to prove from first principles. We hence only prove the chain rule. However, we first need to define differentiability of a function:

Definition 15.3.2. A function $f(x)$ is **differentiable** at a if there exists some function $q(x)$ continuous at a such that

$$[q(x) = \frac{f(x) - f(a)}{x - a}.$$

Note that there is at most one such $q(x)$, and if it exists, then $q(x) = f'(x)$.

We now prove the chain rule.

Proof of Chain Rule. Suppose $y = f(x)$ and $x = g(t)$. Suppose also that $f(x)$ is differentiable at $x = g(a)$, and that $g(t)$ is differentiable at a.

Since $f(x)$ is differentiable at $x = g(a)$, by the above definition, there exists a function $q(x)$ such that

$$q(x) = \frac{f(x) - f(g(a))}{x - g(a)}.$$

Replacing x with $g(t)$, we get

$$q(g(t)) = \frac{f(g(t)) - f(g(a))}{g(t) - g(a)} \implies g(t) - g(a) = \frac{f(g(t)) - f(g(a))}{q(g(t))}. \quad (1)$$

Similarly, since $g(t)$ is differentiable at a, by the above definition, there must exist a function $r(t)$ continuous at a such that

$$r(t) = \frac{g(t) - g(a)}{t - a} \implies g(t) - g(a) = r(t)(t - a). \tag{2}$$

Equating (1) and (2), we have

$$\frac{f(g(t)) - f(g(a))}{q(g(t))} = r(t)(t - a).$$

Rearranging,

$$q(g(t))r(t) = \frac{f(g(t)) - f(g(a))}{t - a} = \frac{(f \circ g)(t) - (f \circ g)(a)}{t - a}.$$

By our assumptions, $q(g(t))r(t)$ is continuous at $t = a$. Hence, by the above definition, $q(g(t))r(t)$ is the derivative of $(f \circ g)'(t)$. Since $q(x) = f'(x)$ and $r(t) = g'(t)$, we arrive at

$$(f \circ g)'(t) = f'(g(t))g'(t).$$

In Liebniz notation, this reads as

$$\frac{\mathrm{d}}{\mathrm{d}t} f(g(t)) = \left[\frac{\mathrm{d}}{\mathrm{d}x} f(g(t))\right] \left[\frac{\mathrm{d}}{\mathrm{d}t} g(t)\right].$$

Since $x = g(t)$ and $y = f(x) = f(g(t))$, this can be written more compactly as

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

□

From the chain rule, we can derive the following property:

Proposition 15.3.3. Suppose $\mathrm{d}x/\mathrm{d}y \not\equiv 0$. Then

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

Proof. By the chain rule,

$$1 = \frac{\mathrm{d}y}{\mathrm{d}y} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}y} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\mathrm{d}x/\mathrm{d}y}.$$

□

Note that this property does not generalize to higher derivatives. For instance, $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} \neq \frac{1}{\mathrm{d}^2x/\mathrm{d}y^2}$.

15.4 Derivatives of Standard Functions

Let $n, a \in \mathbb{R}$.

y	y'	y	y'	y	y'
x^n	nx^{n-1}	$\sin x$	$\cos x$	$\cos x$	$-\sin x$
a^x	$a^x \ln a$	$\sec x$	$\sec x \tan x$	$\csc x$	$-\csc x \cot x$
$\log_a x$	$1/(x \ln a)$	$\tan x$	$\sec^2 x$	$\cot x$	$-\csc^2 x$

y	y'
$\arcsin x$	$1/\sqrt{1-x^2}$, $\lvert x\rvert < 1$
$\arccos x$	$-1/\sqrt{1-x^2}$, $\lvert x\rvert < 1$
$\arctan x$	$1/(1+x^2)$

15.5 Implicit Differentiation

Definition 15.5.1. An **explicit function** is one of the form $y = f(x)$, i.e. the dependent variable y is expressed explicitly in terms of the independent variable x, e.g. $y = 2x \sin x + 3$. An **implicit function** is one where the dependent variable y is expressed implicitly in terms of the independent variable x, e.g. $xy + \sin y = 2$.

Recipe 15.5.2 (Implicit Differentiation)**.** y' is found by differentiating every term in the equation with respect to x and with subsequent arrangement, making y' the subject.

Implicit differentiation requires the use of the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}x} g(y) = \frac{\mathrm{d}}{\mathrm{d}y} g(y) \cdot \frac{\mathrm{d}y}{\mathrm{d}x}.$$

Example 15.5.3 (Implicit Differentiation)**.** Consider the implicit function $3y^3 + x^2 y = 2$. Implicitly differentiating each term with respect to x, we obtain

$$9y^2 y' + \left(x^2 y' + 2xy\right) = 0 \implies y' = \frac{-2xy}{9y^2 + x^2}.$$

Proposition 15.5.4 (Derivative of Inverse Functions)**.**

$$\frac{\mathrm{d}}{\mathrm{d}x} f^{-1}(x) = \frac{1}{f'\left(f^{-1}(x)\right)}.$$

Proof. Let $y = f^{-1}(x)$. Then $f(y) = x$. Implicitly differentiating,

$$f'(y)\, y' = 1 \implies y' = \frac{1}{f'(y)} = \frac{1}{f'\left(f^{-1}(x)\right)}.$$

□

We can use the above result to derive the derivatives of the inverse trigonometric functions and the logarithm.

Example 15.5.5 (Derivative of $\arcsin x$)**.** Let $f(x) = \sin x$. Then $f'(x) = \cos x$. Using the above result,

$$\frac{\mathrm{d}}{\mathrm{d}x} \arcsin x = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-x^2}}.$$

Example 15.5.6 (Derivative of $\log_a x$)**.** Let $f(x) = a^x$. Then $f'(x) = a^x \ln a$. Using the above result,

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

15.6 Parametric Differentiation

Sometimes it is difficult to obtain the Cartesian form of a parametric equation, so we are unable to express $\mathrm{d}y/\mathrm{d}x$ in terms of x. However, we are still able to obtain $\mathrm{d}y/\mathrm{d}x$ in terms of the parameter t using the chain rule. If $x = f(t)$ and $g(t)$, then

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

Example 15.6.1 (Parametric Differentiation)**.** Suppose $x = \sin 2\theta$, $y = \cos 4\theta$. Differentiating x and y with respect to θ, we see that

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = 2\cos 2\theta, \qquad \frac{\mathrm{d}y}{\mathrm{d}\theta} = -4\sin 4\theta.$$

Hence, by the chain rule,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}x} = \frac{-2\sin 4\theta}{\cos 2\theta}.$$

16 Applications of Differentiation

16.1 Monotonicity

Definition 16.1.1. Let f be a function, and let $I \subseteq D_f$ be an interval. Let x_1 and x_2 be distinct elements in I.

- f is **strictly increasing** if $x_1 < x_2 \implies f(x_1) < f(x_2)$.
- f is **strictly decreasing** if $x_1 < x_2 \implies f(x_1) > f(x_2)$

Proposition 16.1.2 (Sign of $f'(x)$ Describes Monotonicity)**.** If $f'(x) > 0$ for all $x \in I$, then f is strictly increasing on I. Similarly, if $f'(x) < 0$ for all $x \in I$, then f is strictly decreasing on I.

Proof. Suppose $f'(x) > 0$ for all $x \in I$. By the Mean Value Theorem, there exists some $c \in I$ such that

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Since $f'(c) > 0$ and $x_1 < x_2$, it follows that $f(x_1) < f(x_2)$, whence f is strictly increasing. The proof of the second statement is similar. □

Note that the converse of the above results is not true. Consider the function $f(x) = x^{1/3}$. Clearly, $f(x)$ is increasing on $\mathbb{R}$, yet $f'(x) = x^{-2/3}/3$ is undefined at $x = 0$.

16.2 Concavity

Definition 16.2.1. Let f be a function, and let $I \subseteq D_f$ be an interval.

- f is **concave upwards** on I if the gradient of f increases as x increases.
- f is **concave downwards** on I if the gradient of f decreases as x increases.

Geometrically, f is concave upwards if the graph of $y = f(x)$, $x \in I$ lies above its tangents. Likewise, f is concave downwards if the graph lies below its tangents.

Proposition 16.2.2 (Sign of $f''(x)$ Describes Concavity)**.** If $f''(x) > 0$ for all $x \in I$, then f is concave upwards on I. Similarly, if $f''(x) < 0$ for all $x \in I$, then f is concave downwards on I.

Proof. Suppose $f''(x) > 0$ for all $x \in I$. Then f' is increasing on I. The gradient of f hence increases as x increases, whence f is concave upwards. The proof of the second statement is similar. □

16.3 Stationary Points

Definition 16.3.1. A **stationary point** on a curve $y = f(x)$ is a point where $f'(x) = 0$.

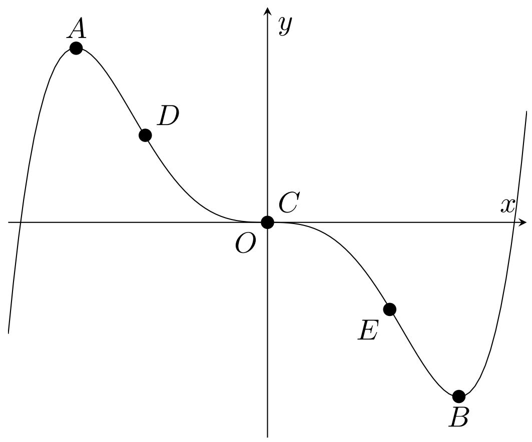


Figure 16.1: Types of stationary points.

There are two types of stationary points:

- turning points: maximum points (A) and minimum points (B)
- stationary points of inflexion: C

Definition 16.3.2. A **point of inflexion** is a point on the curve at which the curve crosses its tangent and the concavity of the curve changes from up to down or vice versa.

Note that a point of inflexion is not necessarily stationary; points D and E in the above figure are **non-stationary points of inflexion**.

16.3.1 Turning Points

In the neighbourhood of turning points, the gradient of the curve, $f'(x)$, changes sign.

Maximum Points

In the neighbourhood of a maximum turning point A, the gradient $f'(x)$ decreases from positive values, through zero at A, to negative values. The y-coordinate of A is known as the **maximum value** of y.

Minimum Points

In the neighbourhood of a minimum turning point B, the gradient $f'(x)$ increases from negative values, through zero at B, to positive values. The y-coordinate of B is known as the **minimum value** of y.

16.3.2 Stationary Points of inflexion

In the neighbourhood of a stationary point of inflexion, the gradient of the curve, $f'(x)$ does not change sign.

16.3.3 Methods to Determine the Nature of Stationary Points

Suppose $y = f(x)$ has stationary point at $x = a$.

Recipe 16.3.3 (First Derivative Test)**.** Check the signs of $f'(x)$ when $x \to a^-$ and $x \to a^+$.

x	a^-	a	a^+	a^-	a	a^+	a^-	a	a^+
$f'(x)$	+ve	0	−ve	−ve	0	+ve	+ve −ve	0 0	+ve −ve
Nature	Maximum point			Minimum point			Stationary point of inflexion		

Example 16.3.4 (First Derivative Test)**.** Let $f(x) = x^2$. Note that $f'(x) = 2x$. Solving for $f'(x) = 0$, we see that $x = 0$ is a stationary point. Checking the signs of y' as $x \to 0^-$ and $x \to 0^+$,

x	0^-	0	0^+
$f'(x)$	−ve	0	+ve

Thus, by the first derivative test, the stationary point at $x = 0$ is a minimum point.

Proposition 16.3.5 (Second Derivative Test)**.** Suppose $f(x)$ has a stationary point at $x = a$.

- If $f''(a) < 0$, then the stationary point is a maximum.
- If $f''(a) > 0$, then the stationary point is a minimum.
- If $f''(a) = 0$, the test is inconclusive.

Proof. At $x = a$, the function $f(x)$ is given by the Taylor series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \cdots.$$

When x is arbitrarily close to a, the terms $(x-a)^3$, $(x-a)^4$, ... become negligibly small, whence $f(x)$ is well-approximated by

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2.$$

Since $x = a$ is a stationary point, $f'(a) = 0$, whence

$$f(x) \approx f(a) + \frac{f''(a)}{2}(x-a)^2.$$

Now observe that $\frac{1}{2}(x-a)^2$ is non-negative. Hence, the sign of $\frac{f''(a)}{2}(x-a)^2$ depends solely on the sign of $f''(a)$: if $f''(a)$ is positive, the entire term is positive and

$$f(x) \approx f(a) + \frac{f''(a)}{2}(x-a)^2 > f(a),$$

whence $f(a)$ is a minimum (since $f(a) < f(x)$ for all x in the neighbourhood of a). Similarly, if $f''(a)$ is negative, the entire term is negative and

$$f(x) \approx f(a) + \frac{f''(a)}{2}(x-a)^2 < f(a),$$

whence $f(a)$ is a maximum. If $f''(a)$ is zero, we cannot say anything about $f(x)$ around $f(a)$ and the test is inconclusive. ☐

Example 16.3.6 (Second Derivative Test). Let $f(x) = x^2$. From the previous example, we know that $x = 0$ is a stationary point. Since $f''(0) = 2 > 0$, by the second derivative test, it must be a minimum point.

16.4 Graph of $y = f'(x)$

The table below shows the relationships between the graphs of $y = f(x)$ and $y = f'(x)$.

	Graph of $y = f(x)$	Graph of $y = f'(x)$
1a	vertical asymptote $x = a$	vertical asymptote $x = a$
1b	horizontal asymptote $y = b$	horizontal asymptote $y = 0$
1c	oblique asymptote $y = mx + c$	horizontal asymptote $y = b$
2	stationary point at $x = a$	$x = a$ is the x-intercept
3a	f is strictly increasing	curve above the x-axis
3b	f is strictly decreasing	curve below the y-axis
4a	f is concave upward	curve is increasing
4b	f is concave downward	curve is decreasing
5	point of inflexion at $x = a$	maximum or minimum point at $x = a$

For most cases, we can deduce the graph of $y = f'(x)$ by using points (1) to (3) only. Points (4) and (5) are usually for checking.

16.5 Tangents and Normals

Let $P(k, f(k))$ be a point on the graph of $y = f(x)$.

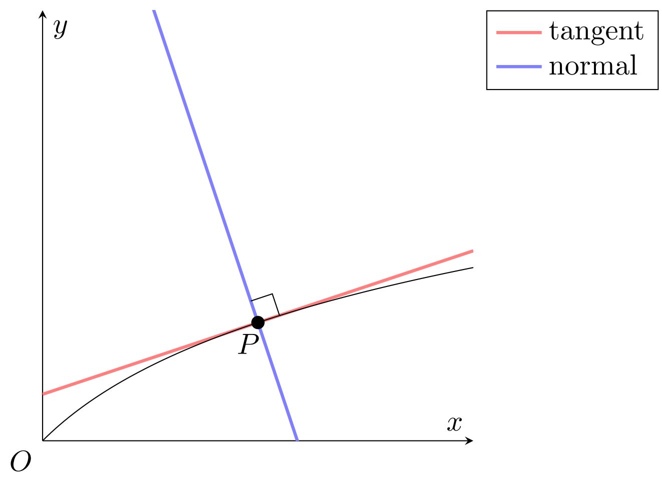


Figure 16.2

The gradient of the tangent to the curve at P is $f'(k)$, while the gradient of the normal to the curve at P is $-1/f'(k)$. This follows from the fact that the tangent and the normal are perpendicular, hence the product of their gradients is -1.

16.6 Optimization Problems

Many real-life situations require that some quantity be minimized (e.g. cost of manufacture) or maximized (e.g. profit on sales). We can use differentiation to solve many of these problems.

Recipe 16.6.1. Suppose we have a dependent variable y that we wish to maximize. We first express y in terms of a single independent variable, say x. We then differentiate y with respect to x and solve for stationary points. Lastly, we determine the nature of the stationary points to obtain the maximum point.

Example 16.6.2. Suppose we wish to enclose the largest rectangular area with only 20 metres of fence. Let x m and y m be the length and width of the rectangular area. The perimeter of the rectangular area is

$$2(x+y)=20 \implies y=10-x.$$

We can hence express the area of the rectangular area A solely in terms of x:

$$A=xy=x(10-x)=-x^2+10x.$$

Differentiating A with respect to x, we see that

$$\frac{\mathrm{d}A}{\mathrm{d}x}=-2x+10.$$

There is hence a stationary point at $x=5$. By the second derivative test, this is a maximum point. Thus, $x=y=5$ gives the largest rectangular area.

16.7 Connected Rates of Change

$\mathrm{d}y/\mathrm{d}x$ measures the instantaneous rate of change of y with respect to x. If t represents time, then $\mathrm{d}y/\mathrm{d}t$ represents the rate of change of the variable y with respect to time t. At the same instant, the rates of change can be connected using the chain rule:

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

Sample Problem 16.7.1. An oil spill spreads on the surface of the ocean, forming a circular shape. The radius of the oil spill r is increasing at a rate of $\mathrm{d}r/\mathrm{d}t=0.5$ m/min. At what rate is the area of the oil spill increasing when the radius is 10 m?

Solution. Let A be the area of the oil spill. Note that $A=\pi r^2$. Differentiating with respect to r, we get $\mathrm{d}A/\mathrm{d}r=2\pi r$. Hence, by the chain rule,

$$\frac{\mathrm{d}A}{\mathrm{d}t}=\frac{\mathrm{d}A}{\mathrm{d}r}\frac{\mathrm{d}r}{\mathrm{d}t}=(2\pi r)(0.5)=\pi r.$$

Thus, when the radius is 10 m, the area of the oil spill is increasing at a rate of 10π m/min.
□

17 Maclaurin Series

Definition 17.0.1. A **power series** is an infinite series of the form

$$\sum_{n=0}^{\infty} a_n(x-c)^n = a_0 + a_1(x-c) + a_2(x-c)^2 + \dots,$$

where a_n is the constant coefficient of the nth term and c is the **centre** of the power series.

Under certain conditions, a function $f(x)$ can be expressed as a power series. This makes certain operations, such as integration, easier to perform. For instance, the integral $\int x e^x \, dx$ is non-elementary. However, we can approximate it by replacing xe^x with its power series and integrating a polynomial instead.

In this chapter, we will learn how to determine the power series of a given function $f(x)$ with centre $c = 0$ by using differentiation. This particular power series is called the Maclaurin series.

17.1 Deriving the Maclaurin Series

Suppose we can express a function $f(x)$ as a power series with centre $c = 0$. That is, we wish to find constant coefficients such that

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \dots. \tag{1}$$

Notice that we can obtain a_0 right away: substituting $x = 0$ into (1) gives

$$f(0) = a_0 + a_1(0) + a_2(0)^2 + \dots = a_0.$$

Now, observe that if we differentiate (1), we get

$$f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \dots. \tag{2}$$

Once again, we can obtain a_1 using the same trick: substituting $x = 0$ into (2) yields

$$f'(0) = a_1 + 2a_2(0) + 3a_3(0)^2 + \dots = a_1.$$

If we continue this process of differentiating and substituting $x = 0$ into the resulting formula, we can obtain any coefficient we so desire. In general,

$$f^{(n)}(0) = \frac{d^n}{dx^n}(a_n x^n). \tag{3}$$

However, by repeatedly applying the power rule, we clearly have

$$\frac{d^n}{dx^n} x^n = \frac{d^{n-1}}{dx^{n-1}} n x^{n-1} = \frac{d^{n-2}}{dx^{n-2}} n(n-1) x^{n-2} = \dots = n(n-1)(n-2)\dots(3)(2)(1) = n!.$$

Thus, a simple rearrangement of (3) gives

$$a_n = \frac{f^{(n)}(0)}{n!}.$$

We thus arrive at the formula for the Maclaurin series of $f(x)$:

Definition 17.1.1. The **Maclaurin series** of $f(x)$ is given by

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \dots.$$

There are a few caveats, though:

- The Maclaurin series of $f(x)$ can only be found if $f^{(n)}(0)$ exists for all values of n. For example, $f(x) = \ln x$ cannot be expressed as a Maclaurin series because $f(0) = \ln 0$ is undefined.
- The Maclaurin series may converge to $f(x)$ for only a specific range of values of x. This range is called the **validity range**.

17.2 Binomial Series

Proposition 17.2.1 (Binomial Series Expansion)**.** Let $n \in \mathbb{Q} \setminus \mathbb{Z}^+$. Then

$$(1+x)^n = \sum_{k=0}^{\infty} \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} x^k,$$

with validity range $|x| < 1$.

Proof. Consider $f(x) = (1+x)^n$, where $n \in \mathbb{Q} \setminus \mathbb{Z}^+$. By repeatedly differentiating $f(x)$, it is not too hard to see that

$$f^{(k)}(x) = n(n-1)(n-2)\dots(n-k+1)\,(1+x)^{n-k}.$$

Hence,

$$f^{(k)}(0) = n(n-1)(n-2)\dots(n-k+1).$$

Substituting this into the formula for the Maclaurin series, we have

$$f(x) = \sum_{k=0}^{\infty} \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} x^k.$$

We now consider the range of validity. If $|x| \geq 1$, then x^k diverges to ∞ as $k \to \infty$. Meanwhile, if $|x| < 1$, then x_k converges to 0 as $k \to \infty$. Hence, the range of validity is $|x| < 1$. $\square$

Note that the binomial theorem is similar to the above result: taking $n \in \mathbb{Z}^+$, we see that

$$\frac{n(n-1)(n-2)\dots(n-k+1)}{k!} = \begin{cases} \binom{n}{k} & k \leq n, \\ 0 & k > n, \end{cases}$$

whence

$$(1+x)^n = \sum_{k=0}^{\infty} \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} x^k = \sum_{k=0}^{n} \binom{n}{k} x^n,$$

which is exactly the binomial theorem. The only difference between the two results is that the range of validity is $\mathbb{R}$ when n is a positive integer. This is because the series is finite (all terms $k > n$ vanish), hence it will always converge.

17.3 Methods to Find Maclaurin Series

17.3.1 Standard Maclaurin Series

Using repeated differentiation, we can derive the following standard Maclaurin series.

$f(x)$	**Standard series**	**Validity range**
$(1+x)^n$	$\displaystyle\sum_{k=0}^{\infty} \frac{n(n-1)(n-2)\dots(n-k+1)}{k!}x^k$	$\lvert x\rvert < 1$
e^x	$\displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!}$	all x
$\sin x$	$\displaystyle\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}$	all x (in radians)
$\cos x$	$\displaystyle\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}$	all x (in radians)
$\ln(1+x)$	$\displaystyle\sum_{k=0}^{\infty} \frac{(-1)^{k+1} x^k}{r}$	$-1 < x \leq 1$

We can use these standard series to find the Maclaurin series of their composite functions.

Example 17.3.1 (Standard Maclaurin Series)**.** Suppose we wish to find the first three terms of the Maclaurin series of $e^x(1+\sin 2x)$. Using the above standard series, we see that

$$e^x = 1 + x + \frac{x^2}{2} + \cdots, \quad \text{and} \quad 1 + \sin 2x = 1 + 2x + \cdots.$$

Hence,

$$\begin{aligned} e^x(1+\sin 2x) &= \left(1 + x + \frac{x^2}{2} + \cdots\right)(1 + 2x + \cdots) \\ &= (1+2x) + \left(x + 2x^2\right) + \left(\frac{x^2}{2}\right) + \cdots = 1 + 3x + \frac{5}{2}x^2 + \cdots. \end{aligned}$$

17.3.2 Repeated Implicit Differentiation

For complicated functions, it is more efficient to repeatedly implicitly differentiate and substitute $x = 0$ to find the values of $y'(0)$, $y''(0)$, etc.

Example 17.3.2 (Repeated Implicit Differentiation)**.** Suppose we wish to find the first three terms of the Maclaurin series of $y = \ln(1+\cos x)$. Rewriting, we get $e^y = 1 + \cos x$. Implicitly differentiating repeatedly with respect to x,

$$\begin{aligned} e^y y' = -\sin x \implies e^y\left[(y')^2 + y''\right] = -\cos x \implies e^y\left[(y')^3 + 3y'y'' + y'''\right] = \sin x \\ \implies e^y\left[(y')^4 + 3(y'')^2 + 6(y')^2 y'' + 4y'y''' + y^{(4)}\right] = \cos x. \end{aligned}$$

Evaluating the above at $x = 0$, we get

$$y(0) = \ln 2, \quad y'(0) = 0, \quad y''(0) = -\frac{1}{2}, \quad y'''(0) = 0, \quad y^{(4)}(0) = -\frac{1}{4}.$$

Thus,

$$\ln(1+\cos x) = \ln 2 + \frac{-1/2}{2!}x^2 + \frac{-1/4}{4!}x^4 + \cdots = \ln 2 - \frac{1}{4}x^2 - \frac{1}{96}x^4 + \cdots.$$

17.4 Approximations using Maclaurin series

Maclaurin series can be used to approximate a function $f(x)$ near $x = 0$.

Example 17.4.1 (Approximating Integrals). Suppose we wish to approximate

$$\int_0^{0.5} \ln(1+\cos x)\ \mathrm{d}x.$$

Doing so analytically is very hard, so we can approximate it using the Maclaurin series of $\ln(1+\cos x)$, which we previously found to be $\ln 2 - \frac{1}{4}x^2 - \frac{1}{96}x^4 + \cdots$. Integrating this expression over the interval $[0, 0.5]$, we get

$$\int_0^{0.5} \ln(1+\cos x)\ \mathrm{d}x \approx \int_0^{0.5} \left(\ln 2 - \frac{1}{4}x^2 - \frac{1}{96}x^4\right) \mathrm{d}x = 0.336092,$$

which is close to the actual value of 0.336091.

Example 17.4.2 (Approximating Constants). For small x,

$$\sin x \approx x - \frac{x^3}{3!}.$$

Since $\sin(\pi/4) = 1/\sqrt{2}$, the numerical value of $1/\sqrt{2}$ can be approximated by substituting $x = \pi/4$ into the above equation:

$$\frac{1}{\sqrt{2}} = \sin\frac{\pi}{4} \approx \frac{\pi}{4} - \frac{(\pi/4)^3}{3} = 0.70465.$$

This is close to the actual value of $1/\sqrt{2} \approx 0.70711$.

To improve the approximation, we can

- choose an x-value closer to 0;
- use more terms of the series.

Example 17.4.3 (Improving Approximations). Continuing on from the previous example, we note that $\sin(3\pi/4)$ is also equal to $1/\sqrt{2}$. If we substitute $x = 3\pi/4$ into $\sin x \approx x - x^3/3!$, we get

$$\frac{1}{\sqrt{2}} = \sin\frac{3\pi}{4} \approx \frac{3\pi}{4} - \frac{(3\pi/4)^3}{3} = 0.17607,$$

which is a worse approximation than if we had used $x = \pi/4$. This is because $|\pi/4| < |3\pi/4|$.

17.5 Small Angle Approximation

For x near zero, we can approximate trigonometric functions with just the first few terms of their respective Maclaurin series:

$$\sin x \approx x, \qquad \cos x \approx 1 - \frac{x^2}{2}, \qquad \tan x \approx x.$$

18 Integration

18.1 Indefinite Integration

In the previous chapters, we learnt about differentiation, which can be thought as finding the derivative $f'(x)$ from a function $f(x)$. Reversing this, we define integration as the process of finding the function $f(x)$ from its derivative $f'(x)$. Simply put, integration "undoes" differentiation and vice versa.

18.1.1 Notation and Terminology

Definition 18.1.1. We write the **indefinite integral** with respect to x of a function $f(x)$ as

$$\int f(x)\,\mathrm{d}x.$$

Here, $f(x)$ is called the **integrand**.

Let the derivative of $F(x)$ be $f(x)$, and let c be an arbitrary constant. Since the derivative of a constant is zero, the function $F(x)+C$ will always have the same derivative: $f(x)$. Thus, when we integrate $f(x)$, we don't get back a single function $F(x)$. Instead, we get back a *class* of functions of the form $F(x)+C$. We call $F(x)$ the **primitive** of $f(x)$, and c the **constant of integration**.

With our notation, we can write down the notion of integration "undoing" differentiation mathematically:

$$\int \frac{\mathrm{d}}{\mathrm{d}x}[f(x)]\,\mathrm{d}x = f(x)+C, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\left[\int f(x)\,\mathrm{d}x\right] = f(x).$$

18.1.2 Basic Rules

Fact 18.1.2 (Properties of Indefinite Integrals)**.** Let $f(x)$ and $g(x)$ be any two functions, and let k be a constant.

- (linearity) $\int [f(x)+g(x)]\,\mathrm{d}x = \int f(x)\,\mathrm{d}x + \int g(x)\,\mathrm{d}x$.
- $\int kf(x)\,\mathrm{d}x = k\int f(x)\,\mathrm{d}x$.

18.2 Definite Integration

Definition 18.2.1. Suppose f is a continuous function defined on the interval $[a,b]$ and $\int f(x)\,\mathrm{d}x = F(x)+C$. Then, the **definite integral** of $f(x)$ from a to b with respect to x is denoted by

$$\int_a^b f(x)\,\mathrm{d}x = [F(x)]_a^b = F(b)-F(a).$$

We call a the **lower limit** and b the **upper limit** of the integral.

Note that the indefinite integral $\int f(x)\,\mathrm{d}x$ is a function in x, while the definite integral $\int_a^b f(x)\,\mathrm{d}x$ is a numerical value. Also note that x is a **dummy variable** as it does not appear in the final expression of the definite integral; it can be replaced by any symbol.

Fact 18.2.2 (Properties of Definite Integrals)**.** Let $f(x)$ and $g(x)$ be any two functions. Let k and c be constants.

- (linearity) $\int_a^b [f(x) + g(x)] \, \mathrm{d}x = \int_a^b f(x) \, \mathrm{d}x + \int g(x) \, \mathrm{d}x$.
- $\int_a^b k f(x) \, \mathrm{d}x = k \int_a^b f(x) \, \mathrm{d}x$.
- $\int_a^b f(x) \, \mathrm{d}x = \int_a^c f(x) \, \mathrm{d}x + \int_c^b f(x) \, \mathrm{d}x$.

Note that from the last property, we can deduce the following properties:

$$\int_a^a f(x) \, \mathrm{d}x = 0, \qquad \text{and} \qquad \int_a^b f(x) \, \mathrm{d}x = -\int_b^a f(x) \, \mathrm{d}x.$$

18.3 Integration Techniques

18.3.1 Systematic Integration

Proposition 18.3.1 (Integrals of Standard Functions)**.**

$$\int x^n \, \mathrm{d}x = \frac{x^{n+1}}{n+1} + C, \qquad (n \neq -1)$$
$$\int \frac{1}{x} \, \mathrm{d}x = \ln|x| + C,$$
$$\int \mathrm{e}^x \, \mathrm{d}x = \mathrm{e}^x + C.$$

Proposition 18.3.2 (Integrals of Trigonometric Functions)**.**

$$\int \sin x \, \mathrm{d}x = -\cos x + C, \qquad \int \cos x \, \mathrm{d}x = \sin x + C,$$
$$\int \sec x \, \mathrm{d}x = -\ln|\sec x - \tan x| + C, \qquad \int \csc x \, \mathrm{d}x = \ln|\csc x - \cot x| + C,$$
$$\int \tan x \, \mathrm{d}x = -\ln|\cos x| + C, \qquad \int \cot x \, \mathrm{d}x = \ln|\sin x| + C.$$

Equivalently,

$$\int \sec x \, \mathrm{d}x = \ln|\sec x + \tan x| \qquad \text{and} \qquad \int \csc x \, \mathrm{d}x = -\ln|\csc x + \cot x|.$$

Products of trigonometric functions can be easily integrated using the following identities:

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

Powers of trigonometric functions can also be integrated using the following identities:

$$\sin^2 x = \frac{1 - \cos 2x}{2}, \qquad \cos^2 x = \frac{1 + \cos 2x}{2},$$
$$\sin^3 x = \frac{3 \sin x - \sin 3x}{4}, \qquad \cos^3 x = \frac{3 \cos x + \cos 3x}{4}.$$

Proposition 18.3.3 (Algebraic Fractions).

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

18.3.2 Integration by Substitution

If the given integrand is not in one of the standard forms, it may be possible to reduce it to a standard form by a change of variable. This method is called **integration by substitution**, and it "undoes the chain rule".

Proposition 18.3.4 (Integration by Substitution). Let $F' = f$. Then

$$\int f(g(x)) g'(x) \, \mathrm{d}x = F(g(x)) + C.$$

Proof. Recall that by the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}x} \left[F(g(x)) \right] = F'(g(x)) g'(x) = f(g(x)) g'(x).$$

Integrating both sides with respect to x,

$$\int f(g(x)) g'(x) \, \mathrm{d}x = F(g(x)) + C.$$

□

A simpler way to interpret the above formula is as follows:

Recipe 18.3.5 (Integration by Substitution). Given an integral $\int f(x) \, \mathrm{d}x$ and a substitution $x = g(u)$, convert all instances of x in terms of u. This includes replacing $\mathrm{d}x$ with $\mathrm{d}u$, which can be found by "splitting" $\mathrm{d}x/\mathrm{d}u$:

$$\frac{\mathrm{d}x}{\mathrm{d}u} = g'(u) \implies \mathrm{d}x = g'(u) \, \mathrm{d}u.$$

If the integral is definite, the bounds should also be converted to their corresponding u values. Once the integral has been evaluated, all instances of u should be converted back to x.

Example 18.3.6 (Definite Integration by Substitution). Consider the definite integral

$$\int_{2/\sqrt{3}}^{2} \frac{1}{x\sqrt{x^2 - 1}} \, \mathrm{d}x.$$

Under the substitution $x = 1/u$, we have

$$\frac{\mathrm{d}x}{\mathrm{d}u} = -\frac{1}{u^2} \implies \mathrm{d}x = -\frac{1}{u^2} \, \mathrm{d}u.$$

When $x = 2/\sqrt{3}$, $u = \sqrt{3}/2$. When $x = 2$, $u = 1/2$. Thus, the integral becomes

$$\int_{\sqrt{3}/2}^{1/2} \frac{u}{\sqrt{u^{-2}-1}} \frac{1}{u^2} \, \mathrm{d}u = \int_{1/2}^{\sqrt{3}/2} \frac{1}{\sqrt{1-u^2}} \, \mathrm{d}u = [\arcsin u]_{1/2}^{\sqrt{3}/2} = \frac{\pi}{6}.$$

Example 18.3.7 (Indefinite Integration by Substitution)**.** Consider the indefinite integral

$$\int \frac{1}{x\sqrt{x^2-1}} \, \mathrm{d}x.$$

Following the same substitution as above ($x = 1/u$), we get

$$\int \frac{1}{x\sqrt{x^2-1}} \, \mathrm{d}x = \int \frac{1}{\sqrt{1-u^2}} \, \mathrm{d}u = \arcsin u + C = \arcsin \frac{1}{x} + C.$$

18.3.3 Integration by Parts

Just like integration by substitution "undoes" the chain rule, **integration by parts** "undoes" the product rule.

Proposition 18.3.8 (Integration by Parts)**.** Let u and v be functions of x. Then

$$\int uv' \, \mathrm{d}x = uv - \int vu' \, \mathrm{d}x.$$

For definite integrals,

$$\int_a^b uv' \, \mathrm{d}x = [uv]_a^b - \int_a^b vu' \, \mathrm{d}x.$$

Proof. By the product rule,

$$(uv)' = uv' + u'v.$$

Integrating both sides and rearranging yields the desired result. □

The statement is also sometimes written as

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

As we just learnt in the previous section, the two forms are perfectly equivalent under substitution (simply substitute x for u and v in the integrands).

Care must be exercised in the choice of the factor u. The aim is to ensure that $u'v$ on the RHS is easier to integrate than uv'. To choose u, we can use the following guideline:

Recipe 18.3.9 (LIATE)**.** In decreasing order of suitability, u should be

- **L**ogarithmic
- **I**nverse trigonometric
- **A**lgebraic
- **T**rigonometric
- **E**xponential

Example 18.3.10 (Integration by Parts). Consider the integral $\int \ln x \, \mathrm{d}x$. Picking $u = \ln x$ and $v' = 1$, we get

$$\int \ln x \, \mathrm{d}x = uv - \int u'v \, \mathrm{d}x = (\ln x)(x) - \int \left(\frac{1}{x}\right)(x) \, \mathrm{d}x = x \ln x - x + C.$$

The astute reader would have noticed that we actually dropped an arbitrary constant when integrating v in the above example. We picked $v' = 1$ but only got $v = x$, instead of the expected $v = x + C$. However, including the arbitrary constant does not matter: if we replace v with $v + C$ into the integration by parts formula, we get

$$\int u \, \mathrm{d}v = u(v + C) - \int (v + C) \, \mathrm{d}u = uv + Cu - \left(\int v \, \mathrm{d}u + Cu\right) = uv - \int v \, \mathrm{d}u,$$

which is what we would have got had we not included the arbitrary constant C.

However, this is not to say that we should always drop the arbitrary constant. In certain situations, including it might actually prove more useful, as demonstrated in the following example.

Example 18.3.11 (Including Arbitrary Constant). Consider the integral $\int \ln(x+1) \, \mathrm{d}x$. Picking $u = \ln(x+1)$ and $v' = 1$ (which implies $v = x + C$), we get

$$\int \ln(x+1) \, \mathrm{d}x = uv - \int u'v \, \mathrm{d}x = (x + C)\ln(x+1) - \int \frac{x + C}{x + 1} \, \mathrm{d}x.$$

Here, a convenient choice for C would be 1, as the integral on the RHS would simplify to $\int 1 \, \mathrm{d}x$, which we can easily integrate. Thus,

$$\int \ln(x+1) \, \mathrm{d}x = (x+1)\ln(x+1) - x + C.$$

If evaluating an integral requires doing multiple integration by parts in succession, the DI method is more convenient.

Recipe 18.3.12 (DI Method). Given the integral $\int uv \, \mathrm{d}x$, construct the following table:

	D	I
$+$	u	v
$-$	u'	$v^{(-1)}$
$+$	u''	$v^{(-2)}$
$\vdots$	$\vdots$	$\vdots$
$\pm$	$u^{(n)}$	$v^{(-n)}$

In other words, keep differentiating the middle column (u) and keep integrating the right column (v), while alternating the sign in the left column. This sign is "attached" to the u terms.

Next, draw diagonal arrows from the middle column to the right column one row below. For instance, u is arrowed to $v^{(-1)}$, while u' is arrowed to $v^{(-2)}$ and so on. Multiply the terms connected by an arrow, keeping in mind the sign of the u terms. Add these terms up, and add the integral of the product of the last row (i.e. $\int u^{(n)} v^{(-n)} \, \mathrm{d}x$).

Essentially, the DI method allows us to easily compute the extended integration by parts formula, which states that

$$\int uv \, \mathrm{d}x = uv^{(-1)} - u'v^{(-2)} + u''v^{(-3)} - u^{(3)}v^{(-4)} + \cdots \pm \int u^{(n)} v^{(-n)} \, \mathrm{d}x,$$

where the sign of the integral depends on the parity of n.

Example 18.3.13 (DI Method). Consider the integral $\int x^3 \sin x \, \mathrm{d}x$. Taking $u = x^3$ and $v = \sin x$, we construct the DI table:

	D	I
$+$	x^3	$\sin x$
$-$	$3x^2$	$-\cos x$
$+$	$6x$	$-\sin x$
$-$	6	$\cos x$

Thus,

$$\begin{aligned}\int x^3 \sin x \, \mathrm{d}x &= x^3(-\cos x) - 3x^2(-\sin x) + 6x(\cos x) - 6\int \cos x \, \mathrm{d}x \\ &= -x^3 \cos x + 3x^2 \sin x + 6x \cos x - 6 \sin x + C.\end{aligned}$$

19 Applications of Integration

19.1 Area

19.1.1 The Riemann Sum and Integral

Suppose we wish to find exact area bounded by the graph of $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$, where $a \leq b$ and $f(x) \geq 0$ for $a \leq x \leq b$.

We can approximate this area by drawing n rectangles of equal width, as shown in the diagram below:

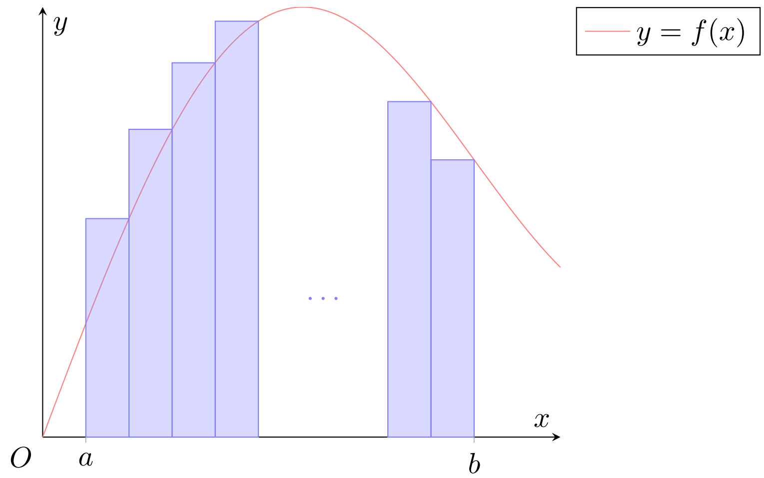


Figure 19.1

Observe that the kth rectangle has width $\Delta x = (b - a)/n$ and height $f(a + k\Delta x)$. The total area of the rectangles is hence

$$\sum_{k=1}^{n} f(a + k\Delta x)\Delta x.$$

This is known as the **Riemann sum** of f over $[a, b]$.

As the number of rectangles approaches ∞, the width Δx of the rectangles approaches 0, and the total area of rectangles approaches the actual area under the curve. In other words,

$$\text{Area} = \lim_{\Delta x \to 0} \sum_{k=1}^{n} f(a + k\Delta x)\Delta x.$$

In the limit, the Riemann sum becomes the **Riemann integral**, which is conventionally written as the definite integral

$$\int_a^b f(x)\,\mathrm{d}x.$$

Note that this is where the integral and differential sign comes from: in the limit, $\sum \to \int$ and $\Delta x \to \mathrm{d}x$.

19.1.2 Definite Integral as the Area under a Curve

Proposition 19.1.1 (Area between a Curve and the x-axis). Let A denote the area bounded by the curve of $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$. Then

$$\text{Area } A = \int_a^b |y| \, \mathrm{d}x = \int_a^b |f(x)| \, \mathrm{d}x.$$

Proposition 19.1.2 (Area between Two Curves). The area A between two curves $y = f(x)$ and $y = g(x)$ is given by

$$\text{Area } A = \int_a^b |f(x) - g(x)| \, \mathrm{d}x.$$

Similar results hold when integrating with respect to the y-axis instead.

Proposition 19.1.3 (Area between a Parametric Curve and the x-axis). Let C be the curve with parametric equations $x = f(t)$ and $y = g(t)$. Then the area A bounded between C and the x-axis is

$$\text{Area } A = \int_a^b |y| \, \mathrm{d}x = \int_{t_1}^{t_2} |g(t)| \, \frac{\mathrm{d}x}{\mathrm{d}t} \, \mathrm{d}t,$$

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

The formula can be applied similarly when we wish to find the area bounded between C and the y-axis.

Proposition 19.1.4 (Area Enclosed by Polar Curve). Let $r = f(\theta)$ be a polar curve, and let A be the area of the region bounded by a segment of the curve and two half-lines $\theta = \alpha$ and $\theta = \beta$. Then

$$\text{Area } A = \frac{1}{2} \int_\alpha^\beta r^2 \, \mathrm{d}\theta.$$

Proof. Divide the enclosed region A into n sectors with the same interior angle $\Delta\theta$. Consider that a typical sector of A can be approximated by a sector of a circle. Thus, the area of that sector is approximately

$$\Delta A \approx \frac{1}{2} r^2 \Delta\theta.$$

Summing up these approximations, we see that

$$A \approx \sum_{\theta=\alpha}^{\theta=\beta} \frac{1}{2} r^2 \Delta\theta.$$

This approximation will improve as the number of sectors increases, i.e. $\Delta\theta \to 0$. Hence,

$$\text{Area } A = \lim_{\Delta\theta \to 0} \sum_{\theta=\alpha}^{\theta=\beta} \frac{1}{2} r^2 \Delta\theta = \frac{1}{2} \int_\alpha^\beta r^2 \, \mathrm{d}\theta.$$

□

19.2 Volume

Definition 19.2.1. If an enclosed region is rotated about a straight line, the three-dimensional object formed is called a **solid of revolution**, and its volume is a **volume of revolution**.

The line about which rotation takes place is always an axis of symmetry for the solid of revolution, and any cross-section of the solid which is perpendicular to the axis of rotation is circular.

19.2.1 Disc Method

Consider the solid of revolution formed when the region bounded between $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$ is rotated about the x-axis.

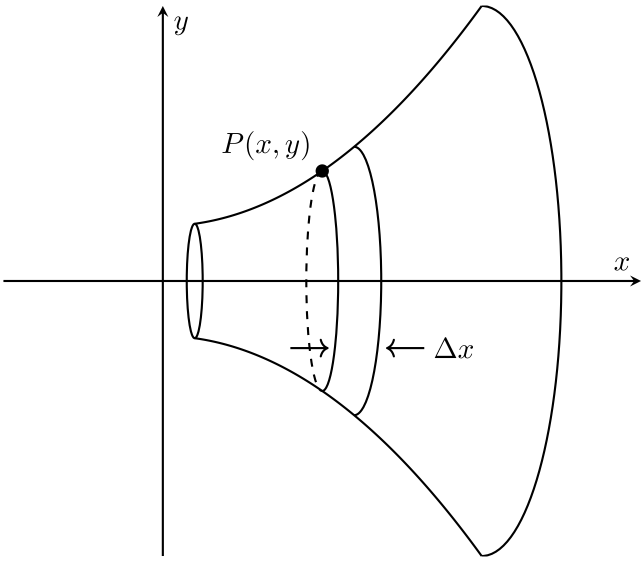


Figure 19.2

To calculate the volume of this solid, we can cut it into thin slices (or discs) of thickness Δx. Each disc is approximately a cylinder and the approximate volume of the solid can be found by summing the volumes of these cylinders. The smaller Δx is, the better the approximation.

Consider a typical disc formed by a one cut through the point $P(x, y)$ and the other cut distant Δx from the first. The volume of this disc is approximately

$$\Delta V \approx \pi y^2 \Delta x.$$

Summing over all discs,

$$V \approx \sum_{x=a}^{b} \pi y^2 \Delta x.$$

As more cuts are made, $\Delta x \to 0$, whence

$$V = \lim_{\Delta x \to 0} \sum_{x=a}^{b} \pi y^2 \Delta x = \pi \int_a^b y^2 \, \mathrm{d}x.$$

Proposition 19.2.2 (Disc Method)**.** When the region bound by the curve $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$ is rotated 2π radians about the x-axis, the volume of the solid of revolution generated is given by

$$V = \pi \int_a^b y^2 \, \mathrm{d}x = \pi \int_a^b [f(x)]^2 \, \mathrm{d}x.$$

Proposition 19.2.3 (Disc Method: Volume Enclosed by Two Curves)**.** When the region enclosed by two curves $y = f(x)$ and $y = g(x)$ is rotated 2π radians about the x-axis, the volume of the solid of revolution generated is given by

$$V = \pi \int_a^b [f(x)]^2 \, \mathrm{d}x - \pi \int_a^b [g(x)]^2 \, \mathrm{d}x = \pi \int_a^b \left([f(x)]^2 - [g(x)]^2 \right) \mathrm{d}x.$$

Similar results hold when the axis of rotation is the y-axis.

19.2.2 Shell Method

Suppose a region R is rotated about the y-axis. Consider a typical vertical strip in the region R with height y and thickness Δx. It will form a cylindrical shell with inner radius x, outer radius $x + \Delta x$ and height y when rotated about the y-axis. Hence, it has volume

$$\Delta V = \pi(x + \Delta x)^2 y - \pi x^2 y = 2\pi xy\Delta x + \pi \Delta x^2 y \approx 2\pi xy\Delta x.$$

Hence, the volume of revolution is approximately

$$V \approx \sum_{x=a}^{b} 2\pi xy\Delta x.$$

As more strips are considered, $\Delta x \to 0$, whence

$$V = \lim_{\Delta x \to 0} = 2\pi \int_a^b xy \, \mathrm{d}x.$$

Proposition 19.2.4 (Shell Method)**.** When the region bound by the curve $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$ is rotated 2π radians about the y-axis, the volume of the solid of revolution is given by

$$V = 2\pi \int_a^b xy \, \mathrm{d}x.$$

A similar result holds when the axis of rotation is the x-axis.

19.3 Arc Length

19.3.1 Parametric Form

Proposition 19.3.1 (Arc Length of Parametric Curve)**.** Let $A(t_1)$ and $B(t_2)$ be points the parametric curve with equations $x = f(t)$, $y = g(t)$, $t \in [t_1, t_2]$. Then

$$\overset{\frown}{AB} = \int_{t_1}^{t_2} \sqrt{[f'(t)]^2 + [g'(t)]^2} \, \mathrm{d}t = \int_{t_1}^{t_2} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2} \, \mathrm{d}t.$$

Proof. Let $s = \wideparen{AB}$ be the arc length of AB. Let P and Q be points on AB with parameters t and $t + \Delta t$ respectively. By the Pythagorean theorem, the straight line PQ is given by

$$PQ^2 = [f(t+\Delta t) - f(t)]^2 + [g(t+\Delta t) - g(t)]\,.$$

Dividing both sides by $(\Delta t)^2$,

$$\left(\frac{PQ}{\Delta t}\right)^2 = \left[\frac{f(t+\Delta t) - f(t)}{\Delta t}\right]^2 + \left[\frac{g(t+\Delta t) - g(t)}{\Delta t}\right]^2.$$

As $\Delta t \to 0$, we can write the RHS in terms of $f'(t)$ and $g'(t)$:

$$\lim_{\Delta t \to 0} \left(\frac{PQ}{\Delta t}\right)^2 = \left[f'(t)\right]^2 + \left[g'(t)\right]^2.$$

Rearranging,

$$\lim_{\Delta t \to 0} PQ = \sqrt{[f'(t)]^2 + [g'(t)]^2}\,\Delta t.$$

However, observe that as $\Delta t \to 0$, the straight line PQ approximates the arc length PQ (i.e. Δs) better and better. Hence,

$$\Delta s = \wideparen{PQ} = \sqrt{[f'(t)]^2 + [g'(t)]^2}\,\Delta t.$$

Integrating from A to B, we thus obtain

$$s = \wideparen{AB} = \int_{t_1}^{t_2} \sqrt{[f'(t)]^2 + [g'(t)]^2}\,\mathrm{d}t = \int_{t_1}^{t_2} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t.$$

□

19.3.2 Cartesian Form

Taking $t = x$ or $t = y$, we get the following formulas involving $\mathrm{d}y/\mathrm{d}x$ and $\mathrm{d}x/\mathrm{d}y$, which is suitable for Cartesian curves.

Proposition 19.3.2 (Arc Length of Cartesian Curve)**.** Let $A(x_1, y_1)$ and $B(x_2, y_2)$ be points on the curve $y = f(x)$. The arc length AB is given by

$$\wideparen{AB} = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = \int_{y_1}^{y_2} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2 + 1}\,\mathrm{d}y.$$

19.3.3 Polar Form

Proposition 19.3.3. Let $A(r_1, \theta_1)$ and $B(r_2, \theta_2)$ be points on the polar curve $r = f(\theta)$. Then the arc length AB is given by

$$\wideparen{AB} = \int_{\theta_1}^{\theta_2} \sqrt{r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta.$$

Proof. Recall that $x = r\cos\theta$ and $y = r\sin\theta$. Hence,

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = \cos\theta\frac{\mathrm{d}r}{\mathrm{d}\theta} - r\sin\theta, \qquad \frac{\mathrm{d}y}{\mathrm{d}\theta} = \sin\theta\frac{\mathrm{d}r}{\mathrm{d}t} + r\cos\theta.$$

It follows that

$$\left(\frac{\mathrm{d}(r\cos\theta)}{\mathrm{d}\theta}\right)^2+\left(\frac{\mathrm{d}(r\sin\theta)}{\mathrm{d}\theta}\right)^2=\left(\cos^2\theta+\sin^2\theta\right)\left[r^2+\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2\right]=r^2+\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2.$$

Taking $t=\theta$,

$$\widehat{AB}=\int_{\theta_1}^{\theta_2}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta=\int_{\theta_1}^{\theta_2}\sqrt{r^2+\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta.$$

□

19.4 Surface Area of Revolution

Definition 19.4.1. The surface area of a solid of revolution is called the **surface area of revolution**.

Proposition 19.4.2 (Surface Area of Revolution of Parametric Curve)**.** Let $A(t_1)$ and $B(t_2)$ be points the parametric curve with equations $x=f(t)$, $y=g(t)$, $t\in[t_1,t_2]$. Then the surface area of revolution about the x-axis of arc AB is given by

$$A=2\pi\int_{t_1}^{t_2}y\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t.$$

Similarly, the surface area of revolution about the y-axis is given by

$$A=2\pi\int_{t_1}^{t_2}x\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t.$$

Proof. Let $s=\widehat{AB}$ be the arc length of AB. Let P and Q be points on AB with parameters t and $t+\Delta t$ respectively. Recall that

$$\Delta s=\widehat{PQ}=\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\Delta t.$$

Now consider the surface area of revolution about the x-axis of arc PQ. For small Δs, the solid of revolution is approximately a disc wish radius y and width Δs. The surface area of this disc can be calculated as

$$\Delta A=2\pi y\Delta s=2\pi y\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\Delta t.$$

Integrating from A to B, we see that

$$A=2\pi\int_{t_1}^{t_2}y\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t.$$

A similar argument is used when the axis of rotation is the y-axis. □

19.5 Approximating Definite Integrals

In §19.1, we saw how Riemann sums could approximate definite integrals using rectangles. This is a blunt tool which utilizes very little information from the curve and thus will often not give a good estimate of the definite integral for a fixed number of rectangles.

In this chapter, we will be exploring two other methods: the trapezium rule and Simpson's rule, for finding the approximate value of an area under a curve. These methods often give better approximations to the actual area as compared to using Riemann sums. Similar to Riemann sums, these methods can be extended to estimate the value of a definite integral.

19.5.1 Trapezium Rule

Consider the curve $y = f(x)$ which is non-negative over the interval $[a, b]$.

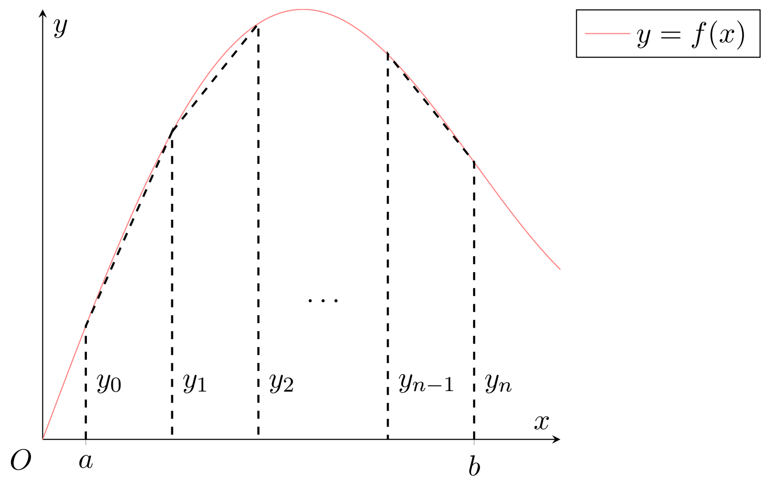


Figure 19.3

Divide the interval $[a, b]$ into n equal intervals (strips) with each having width $h = (b - a)/n$. Then the area of the n trapeziums is given by

$$\text{Area} = \sum_{k=0}^{n} \frac{h}{2}(y_k + y_{k+1}) = \frac{h}{2}\left[y_0 + 2(y_1 + y_2 + \cdots + y_{n-1}) + y_n\right].$$

Recipe 19.5.1 (Trapezium Rule). The trapezium rule with $(n+1)$ ordinates (or n intervals) gives the approximation

$$\int_a^b f(x)\,\mathrm{d}x \approx \sum_{k=0}^{n} \frac{h}{2}(y_k + y_{k+1}) = \frac{h}{2}\left[y_0 + 2(y_1 + y_2 + \cdots + y_{n-1}) + y_n\right],$$

where $h = (b - a)/n$.

Sample Problem 19.5.2. Use the trapezium rule with 4 strips to find an approximation for

$$\int_0^2 \ln(x + 2)\,\mathrm{d}x.$$

Find the percentage error of the approximation.

Solution. Let $f(x) = \ln(x+2)$. By the trapezium rule,

$$\begin{aligned}\int_0^2 \ln(x+2)\ \mathrm{d}x &\approx \frac{1}{2}\cdot\frac{2-0}{4}\Big(f(0)+2\left[f(0.5)+f(1)+f(1.5)\right]+f(2)\Big)\\ &= 2.15369 \text{ (5 d.p.).}\end{aligned}$$

One can easily verify that the integral evaluates to 2.15888 (5 d.p.). Hence, the percentage error is

$$\left|\frac{2.15888-2.15369}{2.15888}\right| = 0.240\%.$$

□

Error in Trapezium Rule Approximation

If the curve is concave upward, the secant lines lie above the curve. Hence, the trapezium rule will give an overestimate.

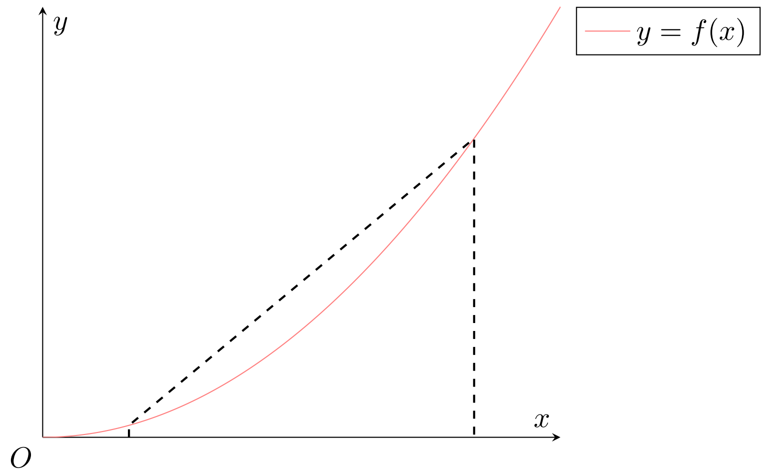


Figure 19.4

If the curve is concave downward, the secant lines lie below the curve. Hence, the trapezium rule will give an underestimate.

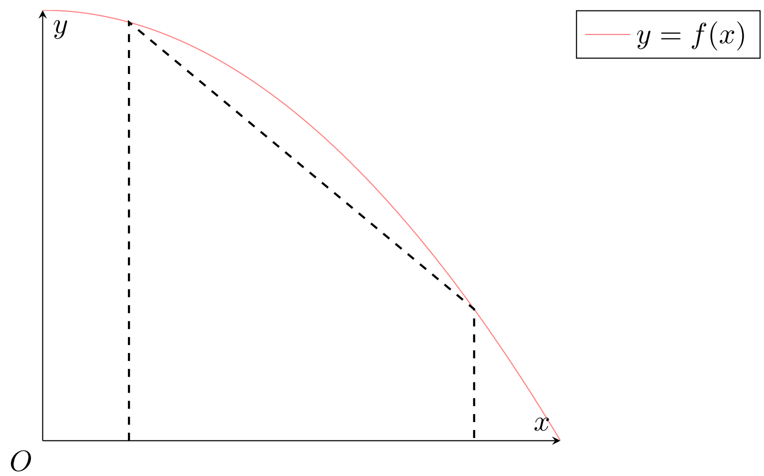


Figure 19.5

19.5.2 Simpson's Rule

Previously, we explored how Riemann sums approximate definite integrals using horizontal lines (i.e. degree 0 polynomials). We also saw how the trapezium rule improves this approximation by using sloped lines (i.e. degree 1 polynomials). Now, we introduce Simpson's rule, which takes this a step further by using quadratics (i.e. degree 2 polynomials) to achieve even greater accuracy in approximating definite integrals.

Consider the curve $y = f(x)$, which is non-negative over the interval $[a, b]$. Suppose the area represented by $\int_a^b f(x)\,\mathrm{d}x$ is divided by the ordinates y_0, y_1, y_2 into two strips each of width h as shown below. A particular parabola can be found passing through the three points on the curve with ordinates y_0, y_1, y_2. Simpson's rule uses the area under the parabola to approximate the area represented by $\int_a^b f(x)\,\mathrm{d}x$.

To deduce the area under the parabola, we consider the case where $y = f(x)$ is translated x_1 units to the left, i.e. the line $x = x_1$ is now the y-axis.

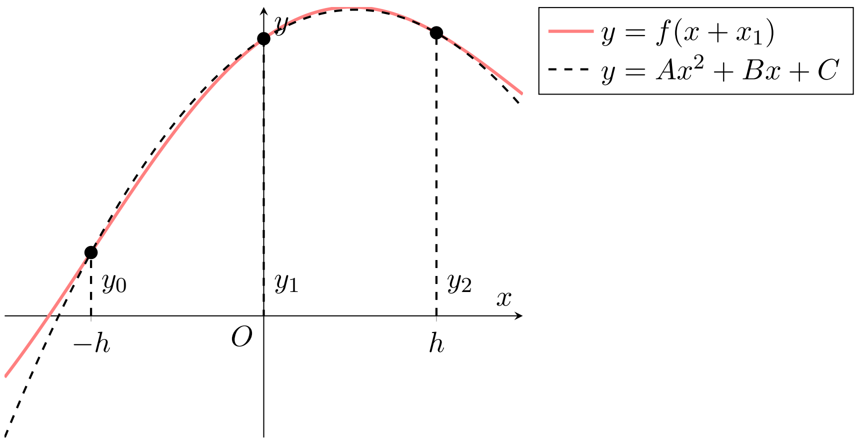


Figure 19.6

Under this translation,

$$\int_a^b f(x)\,\mathrm{d}x = \int_{-h}^{h} f(x + x_1)\,\mathrm{d}x.$$

This area will now be approximated by a parabola $y = g(x) = Ax^2 + Bx + C$, where A, B and C are constants. The area under the parabola is given by

$$\int_{-h}^{h} \left(Ax^2 + Bx + C\right)\,\mathrm{d}x = \left[\frac{A}{3}x^3 + \frac{B}{2}x^2 + Cx\right]_{-h}^{h} = \frac{h}{3}\left(2Ah^2 + 6C\right).$$

Now, observe that the parabola $y = g(x)$ intersects the curve at $(-h, y_1)$, $(0, y_2)$ and (h, y_3). Hence,

$$g(-h) = Ah^2 - Bh + C = y_0, \quad g(0) = C = y_1, \quad g(h) = Ah^2 + Bh + C = y_2.$$

Thus,

$$\frac{h}{3}\left(2Ah^2 + 6C\right) = \frac{h}{3}\left[\left(Ah^2 - Bh + C\right) + 4C + \left(Ah^2 + Bh + C\right)\right] = \frac{h}{3}(y_0 + 4y_1 + y_2).$$

We hence arrive at Simpson's rule with 2 strips:

$$\int_a^b f(x)\,\mathrm{d}x \approx \frac{h}{3}(y_0 + 4y_1 + y_2).$$

We can extend Simpson's rule to cover any even number of strips. In general,

Recipe 19.5.3 (Simpson's Rule). Simpson's rule with $2n$ strips (or $2n+1$ ordinates) gives the approximation

$$\begin{aligned}\int_a^b f(x)\,\mathrm{d}x &\approx \sum_{k=0}^{n} \frac{h}{3}(y_{2k} + 4y_{2k+1} + y_{2k+2}) \\ &= \frac{h}{3}\left[y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{2n-2} + 4y_{2n-1} + y_{2n}\right].\end{aligned}$$

Sample Problem 19.5.4. Use Simpson's rule with 4 strips to find an approximation for

$$\int_0^2 \ln(x+2)\ \mathrm{d}x.$$

Find the percentage error of the approximation.

Solution. Let $f(x) = \ln(x+2)$. By the trapezium rule,

$$\begin{aligned}\int_0^2 \ln(x+2)\ \mathrm{d}x &\approx \frac{1}{3}\cdot\frac{2-0}{4}\Big[f(0) + 4f(0.5) + 2f(1) + 4f(1.5) + f(2)\Big] \\ &= 2.15881 \text{ (5 d.p.)}.\end{aligned}$$

As previously mentioned in Sample Problem 19.5.2 the actual value of the integral is 2.15888 (5 d.p.). Hence, the percentage error is

$$\left|\frac{2.15888 - 2.15881}{2.15888}\right| = 0.00324\%.$$

□

In the previous example, the trapezium rule gave an estimate of 2.15369 (5 d.p.), which has an error of 0.240%. In the case of Simpson's rule, the error is 0.00324%, vastly better than that of the trapezium rule's.

In general, Simpson's rule gives a better approximation than the trapezium rule as the quadratics used account for the concavity of the curve.

20 Functions of Two Variables

In Chapter §3, we learnt that functions can be described as a machine that takes in an input and produces an output according to a rule. Some examples of functions that we have encountered thus fare are $f(x) = x^2$, $g(x) = \cos x$, etc. These are functions of one variable, also called **univariate functions**.

However, in real life, there are functions that depend on more than one variable (i.e. the domain is not a subset of the real numbers). For instance, the cost (output) of a taxi ride may depend on variables (input) like time, distance travelled, traffic conditions, demand, etc. In this case, the function is called a **multivariate function**. The input with many variables can be expressed as a vector.

Similarly, the codomain of a function does not necessarily need to be a subset of the real numbers. Consider the following function $f(s,t)$:

$$f(s,t) = \begin{pmatrix} s+t \\ t \\ 2s-1 \end{pmatrix}.$$

Here, $f(s,t)$ takes in two inputs (s and t), and spits out three outputs ($s+t$, t and $2s-1$).

For the rest of this chapter, we will only study scalar-valued functions of two variables, of the form

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

which we can visualize in 3D space. We will see how the ideas from univariate functions can be extended to two variable functions and how concepts of vectors can be useful in studying these functions.

20.1 Functions of Two Variables and Surfaces

20.1.1 Functions of Two Variables

Definition 20.1.1. A (scalar) **function of two variables**, f, is a rule that assigns each ordered pair of real numbers (x,y) in its domain to a unique real number.

Recall that the domain of a function $g(x)$ is a subset of the real number line, i.e. $D_g \subseteq \mathbb{R}$. Generalizing this to scalar functions of two variables, the domain of f is a subset of the xy-plane, denoted $\mathbb{R} \times \mathbb{R}$ or $\mathbb{R}^2$. Mathematically,

$$D_f \subseteq \mathbb{R}^2.$$

If the domain of $f(x,y)$ is not well specified, then we will take its domain to be the set of all pairs $(x,y) \in \mathbb{R}^2$ for which the given expression is a well-defined real number.

Example 20.1.2 (Domain of $f(x,y)$)**.** Let $f(x,y) = \ln\left(y^2 - x\right)$. For $f(x,y)$ to be well-defined, the argument of the natural logarithm must be positive. That is, we require $y^2 - x > 0$. The domain of f is hence

$$D_f = \left\{(x,y) \in \mathbb{R}^2 \mid y^2 - x > 0\right\}.$$

20.1.2 Surfaces

Recall that we defined the graph of a function $g(x)$ to be the collection of all points (x, y) in the xy-plane such that the values x and y satisfy $y = g(x)$. We can extend this notion to functions of two variables:

Definition 20.1.3. The **graph** of $z = f(x, y)$, or **surface** with equation $z = f(x, y)$, is the collection of all points (x, y, z) in 3D Cartesian space such that the values x, y and z satisfy $z = f(x, y)$.

Visualizing and illustrating a 3D surface can be challenging, especially as surfaces become complicated. We can study the surface by fixing or changing the variables one at a time. This is the idea behind traces, or level curves.

Definition 20.1.4. Horizontal traces (or **level curves**) are the resulting curves when we intersect the surface $z = f(x, y)$ with horizontal planes.

This is like fixing the value of z, giving the 2D graph of the equation $f(x, y) = c$ for some constant c.

Definition 20.1.5. Vertical traces are the resulting curves when we intersect the surface $z = f(x, y)$ with vertical planes.

This is like fixing the value of x or y (or a combination of both, e.g. $y = x$).

Definition 20.1.6. A **contour plot** of $z = f(x, y)$ is a graph of numerous horizontal traces $f(x, y) = c$ for representative values of c (usually spaced-out values).

We may identify a surface by examining these traces to visualize graphs of two variables.

Example 20.1.7 (Graph of $z = f(x, y)$)**.** Let $f(x, y) = \ln(x^2 + y^2)$. Consider the horizontal traces of $z = f(x, y)$. Setting $z = c$, we get

$$\ln(x^2 + y^2) = c \implies x^2 + y^2 = \mathrm{e}^c.$$

Hence, the horizontal trace of $z = f(x, y)$ at $z = c$ corresponds to a circle centred at the origin with radius e^c. Thus, the graph of $z = \ln(x^2 + y^2)$ looks like

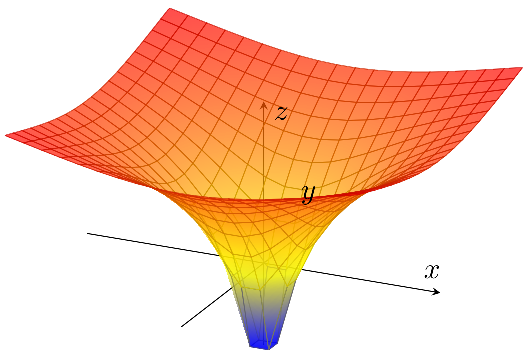


Figure 20.1

20.1.3 Cylinders and Quadric Surfaces

Exploring the traces of a surface allows us to visualize the shape of the surface. We can now look at some of the common surfaces, such as cylinders and quadric surfaces.

Definition 20.1.8. A surface is a **cylinder** if there is a plane P such that all planes parallel to P intersect the surface in the same curve (when viewed in 2D).

Examples of cylinders include the graphs of $x^2 + z^2 = 1$ and $z = y^2$, as shown below:

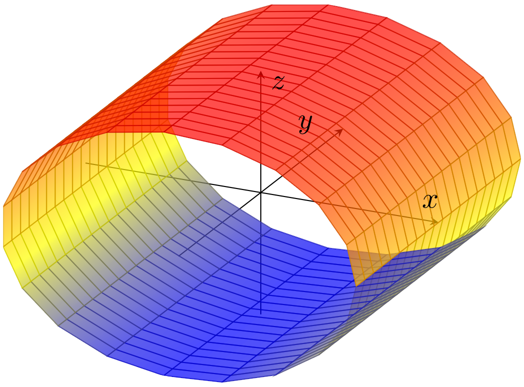


Figure 20.2: Graph of $x^2 + z^2 = 1$.

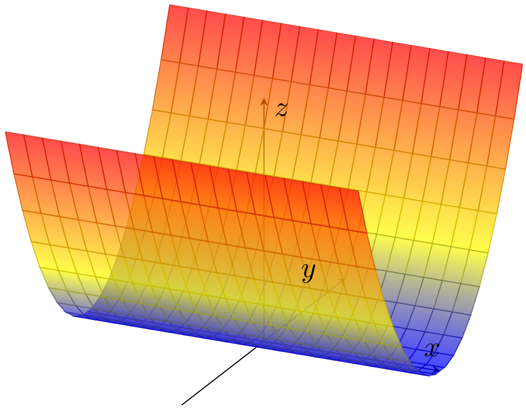


Figure 20.3: Graph of $z = y^2$.

Observe that $x^2 + z^2 = 1$ is a special case of a function of two variables $z = f(x, y)$ that can be reduced to $z = f(x)$ since z is independent of y. Similarly, $z = y^2$ can be reduced to $z = f(y)$ since z is independent of x. Indeed, if a function $z = f(x, y)$ can be reduced to a univariate function, then its surface must be cylindrical.

Another common surface is a quadric surface, which is a 3D generalization of 2D conic sections. Recall that a conic section in 2D has the general form

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$

We can generalize this into 3D to get a quadric surface.

Definition 20.1.9. A **quadric surface** has the form

$$Ax^2 + By^2 + Cz^2 + Dxy + Eyz + Fzx + Gx + Hy + Iz + J,$$

where $A, B, \ldots, J \in \mathbb{R}$ and at least one of A, B and C is non-zero.

An example of a quadric surface is the ellipsoid, which is a generalization of an ellipse and has equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

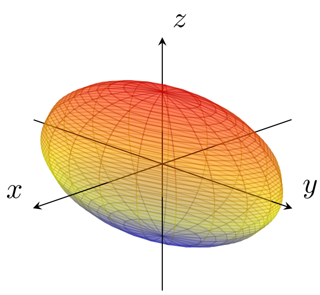


Figure 20.4: An ellipsoid.

When $a = b = c = r$, we get the equation

$$x^2 + y^2 + z^2 = r^2.$$

This represents a sphere centred at the origin with radius r. Observe the similarity between the equation of a circle ($x^2 + y^2 = r^2$) and the equation of a sphere.

20.2 Partial Derivatives

Recall that for a function f of one variable x, we defined the derivative function as

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

The usual notations are $\frac{\mathrm{d}y}{\mathrm{d}x}$ or $\frac{\mathrm{d}f}{\mathrm{d}x}$ if $y = f(x)$.

The notation $\frac{\mathrm{d}y}{\mathrm{d}x}$ gives some insight into how derivatives are derived. We can view

- "$\mathrm{d}x$" as a small change in x, and
- "$\mathrm{d}y$" as the change in y as a result of the small change in x.

Hence, the notation $\frac{\mathrm{d}y}{\mathrm{d}x}$ actually represents the "rise over run", which is a measure of gradient at the point (x, y) on the graph.

We can extend this notion to functions of two variables $z = f(x, y)$. There are now two variables that will affect the change in the value of f. We can choose to vary x slightly (Δx) or vary y slightly Δy and see how f changes (Δf). This gives us some notion of a derivative. However, because we are only varying one independent variable at a time, we are only differentiating the function $f(x, y)$ "partially". We hence call these derivatives the partial derivatives of f.

Definition 20.2.1. The (first-order) **partial derivatives** of $f(x, y)$ are the functions f_x and f_y defined by

$$f_x(x, y) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x},$$
$$f_y(x, y) = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}.$$

In Liebniz notation,

$$f_x(x, y) = \frac{\partial f}{\partial x}, \quad f_y(x, y) = \frac{\partial f}{\partial y}.$$

Recipe 20.2.2 (Partial Differentiation)**.** To partially differentiate a function $f(x, y)$ with respect to x, we differentiate $f(x, y)$ as we normally would, treating y as a constant. Similarly, if we are partially differentiating with respect to y, we treat x as a constant.

Sample Problem 20.2.3. Given $f(x, y) = \cos\left(xy + y^2\right)$, find $f_x(x, y)$.

Solution. To partially differentiate it with respect to x, we treat y as a constant. Using the chain rule,

$$f_x(x, y) = -\sin\left(xy + y^2\right) \frac{\partial}{\partial x} \left[xy + y^2\right].$$

Since y is a constant,

$$\frac{\partial}{\partial x}(xy) = y, \quad \frac{\partial}{\partial x} y^2 = 0.$$

Hence,

$$f_x(x, y) = -y \sin\left(xy + y^2\right).$$

$\square$

20.2.1 Geometric Interpretation

Consider a surface S given by the equation $z = f(x, y)$. Let $P(a, b, c)$ be a point on S.

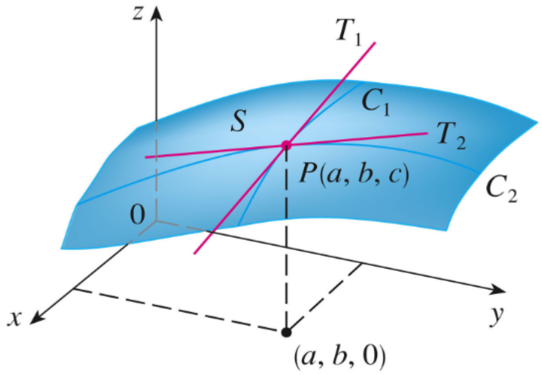


Figure 20.5: Partial derivatives as slopes of tangent lines.[1]

The curve C_1 is the graph of the function $g(x) = f(x, b)$, which is the intersection curve of the surface and the vertical plane $y = b$. The slope of its tangent T_1 at P is $g'(x) = f_x(a, b)$.

Similarly, the curve C_2 is the graph of the function $h(y) = f(a, y)$, which is the intersection curve of the surface and the vertical plane $x = a$. The slope of its tangent T_2 at P is $h'(y) = f_y(a, b)$.

We can hence visualize partial derivatives at the point P on S as slopes to the tangent lines T_1 and T_2 at that point.

20.2.2 Gradient

To represent the "full" derivative of a function, we simply collect its partial derivatives.

Definition 20.2.4. The **gradient** of a function $f(x, y)$, denoted as ∇f, is the collection of all its partial derivatives into a vector.

$$\nabla f = \begin{pmatrix} f_x \\ f_y \end{pmatrix}.$$

Example 20.2.5 (Gradient)**.** Let $f(x, y) = xy^2 + x^3$. Then its gradient is

$$\nabla f = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} y^2 + 3x^2 \\ 2xy \end{pmatrix}.$$

20.2.3 Second Partial Derivatives

Similar to second-order derivatives for univariate functions, we can also consider the partial derivatives of partial derivatives:

$$(f_x)_x, \quad (f_x)_y, \quad (f_y)_x, \quad (f_y)_y.$$

[1]Source: https://www2.victoriacollege.edu/~myosko/m2415sec143notes(7).pdf

If $z = f(x, y)$, we use the following notation for the second partial derivatives:

$$(f_x)_x = f_{xx} = \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 z}{\partial x^2},$$
$$(f_x)_y = f_{xy} = \frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial^2 z}{\partial y\,\partial x},$$
$$(f_y)_x = f_{yx} = \frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 z}{\partial x\,\partial y},$$
$$(f_y)_y = f_{yy} = \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 z}{\partial y^2}.$$

Thus, the notation f_{xy} means that we first partially differentiate with respect to x and then with respect to y. Notice that the order the variables appear in the denominator is reversed when using Liebniz notation, similar to the idea of composite functions:

$$(f_x)_y = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y\,\partial x}.$$

Example 20.2.6 (Second Partial Derivatives)**.** Consider the function $f(x, y) = xy^2 + x^3 + \ln y$. Its partial derivatives are

$$f_x = y^2 + 3x^2, \quad f_y = 2xy + \frac{1}{y},$$

and its second partial derivatives are

$$f_{xx} = 6x, \quad f_{xy} = 2y, \quad f_{yx} = 2y, \quad f_{yy} = 2x - \frac{1}{y^2}.$$

Notice in the above example that $f_{xy} = f_{yx}$. This symmetry of second partial derivatives is known as Clairaut's theorem.

Theorem 20.2.7 (Clairaut's Theorem)**.** If f_{xy} and f_{yx} are both continuous, then $f_{xy} = f_{yx}$

20.2.4 Multivariate Chain Rule

Recall that for a univariate function $y = f(x)$, where the variable x is a function of t, i.e. $x = g(t)$, the chain rule states

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

We can generalize this result to multivariate functions using partial derivatives:

Proposition 20.2.8 (Multivariate Chain Rule)**.** Consider the function $f(x, y)$, where x and y are functions of t. Then

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}.$$

To see why this is morally true, we return to the definition of a partial derivative:

$$f_x(x, y) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x},$$
$$f_y(x, y) = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}.$$

Rewriting these equations, we get

$$f(x+\Delta x, y) = f(x,y) + \Delta x f_x(x,y), \tag{1}$$
$$f(x, y+\Delta y) = f(x,y) + \Delta y f_y(x,y), \tag{2}$$

where Δx and Δy should be thought of as infinitesimally small changes in x and y.

We now consider the quantity $f(x+\Delta x, y+\Delta y)$. Applying (1) and (2) sequentially, we get

$$\begin{aligned} f(x+\Delta x, y+\Delta y) &= f(x, y+\Delta y) + \Delta x f_x(x, y+\Delta y) \\ &= f(x,y) + \Delta y f_y(x,y) + \Delta x f_x(x, y+\Delta y). \end{aligned} \tag{3}$$

Observe that if we partially differentiate (2) with respect to x, we get

$$f_x(x, y+\Delta y) = f_x(x,y) + \Delta y f_{yx}(x,y).$$

Substituting this into (3) yields

$$\begin{aligned} f(x+\Delta x, y+\Delta y) &= f(x,y) + \Delta y f_y(x,y) + \Delta x \left[f_x(x,y) + \Delta y f_{yx}(x,y)\right] \\ &= f(x,y) + \Delta y f_y(x,y) + \Delta x f_x(x,y) + \Delta x \Delta y f_{yx}(x,y). \end{aligned} \tag{4}$$

Since Δx and Δy are both infinitesimally small, the quantity $\Delta x \Delta y$ is negligible and can be disregarded. We thus have

$$\Delta f = f(x+\Delta x, y+\Delta y) - f(x,y) = \Delta x f_x(x,y) + \Delta y f_y(x,y).$$

Dividing throughout by Δt and writing f_x, f_y in Liebniz notation, we have

$$\frac{\Delta f}{\Delta t} = \frac{\partial f}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y}\frac{\Delta y}{\Delta t}.$$

In the limit as $\Delta t \to 0$, we have

$$\frac{\Delta f}{\Delta t} \to \frac{\mathrm{d}x}{\mathrm{d}t}, \quad \frac{\Delta x}{\Delta t} \to \frac{\mathrm{d}x}{\mathrm{d}t}, \quad \frac{\Delta y}{\Delta t} \to \frac{\mathrm{d}y}{\mathrm{d}t}.$$

Thus,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}.$$

□

Observe that if we had applied (2) before (1) on $f(x+\Delta x, y+\Delta y)$, we would have got

$$f(x+\Delta x, y+\Delta y) = f(x,y) + \Delta y f_y(x,y) + \Delta x f_x(x,y) + \Delta x \Delta y f_{xy}(x,y).$$

However, by Clairaut's theorem, we know $f_{xy} = f_{yx}$, so we would still have ended up with (4).

20.2.5 Directional Derivative

So far, we only know how to find the instantaneous rate of change of $f(x,y)$ in two special cases:

- The first case is when we vary x and hold y constant, in which the partial derivative $f_x(x,y)$ represents the instantaneous rate of change of $f(x,y)$.
- The second case is when we vary y and hold x constant, in which the partial derivative $f_y(x,y)$ represents the instantaneous rate of change of $f(x,y)$.

We wish to construct a more general "derivative" which represents the instantaneous rate of change of $f(x, y)$ where x and y are both allowed to vary.

To simplify matters, we assume that x and y are changing at a constant rate. That is, every time x increases by u_x, y will increase by u_y. We can represent this change with a unit vector $\mathbf{u}$ along the xy-plane:

$$\mathbf{u} = \begin{pmatrix} u_x \\ u_y \end{pmatrix}.$$

Because we are measuring the instantaneous rate of change of $f(x, y)$ along a direction, we call this quantity the "directional derivative".

Definition 20.2.9. The **directional derivative** of $f(x, y)$ in the direction of the unit vector $\mathbf{u} = (u_x,\, u_y)^\mathsf{T}$ is denoted $D_\mathbf{u} f(x, y)$ and is defined as

$$D_\mathbf{u} f(x, y) = \lim_{h \to 0} \frac{f(x + hu_x, y + hu_y) - f(x, y)}{h}.$$

We now relate the directional derivative with the gradient of f.

Proposition 20.2.10.

$$D_\mathbf{u} f(x, y) = \nabla f \cdot \mathbf{u} = u_x f_x(x, y) + u_y f_y(x, y).$$

Proof. In §20.2.4, we derived the equation

$$f(x + \Delta x, y + \Delta y) - f(x, y) = \Delta x f_x(x, y) + \Delta y f_y(x, y),$$

where Δx and Δy are infinitesimally small. If we take $(\Delta x,\, \Delta y)^\mathsf{T}$ to be in the same direction as $(u_x,\, u_y)^\mathsf{T}$, i.e.

$$\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \lim_{h \to 0} h \begin{pmatrix} u_x \\ u_y \end{pmatrix},$$

then we have

$$f(x + hu_x, y + hu_y) - f(x, y) = hu_x f_x(x, y) + hu_y f_y(x, y),$$

keeping in mind that we are taking the limit $h \to 0$ on both sides. Dividing both sides throughout by h,

$$\lim_{h \to 0} \frac{f(x + hu_x, y + hu_y) - f(x, y)}{h} = u_x f_x(x, y) + u_y f_y(x, y),$$

which was what we wanted. □

With this relation, we can prove several neat results.

Proposition 20.2.11. Suppose f is differentiable at (x_0, y_0), and $\nabla f(x_0, y_0) \neq \mathbf{0}$. Then $\nabla f(x_0, y_0)$ is perpendicular to the level curve of f through (x_0, y_0).

Proof. Let $f(x, y) = (x(t), y(t))$. Note that the tangent to the level curve at (x_0, y_0) has direction vector $\mathbf{u} = (\mathrm{d}x/\mathrm{d}t,\, \mathrm{d}y/\mathrm{d}t)^\mathsf{T}$.

Let the level curve at (x_0, y_0) have equation $f(x, y) = c$. Implicitly differentiating this with respect to t, we get

$$\frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = \begin{pmatrix} f_x \\ f_y \end{pmatrix} \cdot \begin{pmatrix} \mathrm{d}x/\mathrm{d}t \\ \mathrm{d}y/\mathrm{d}t \end{pmatrix} = \nabla f \cdot \mathbf{u} = 0.$$

Since both ∇f and $\mathbf{u}$ are non-zero vectors, they must be perpendicular to each other. □

Proposition 20.2.12. The greatest rate of change of f occurs in the direction of ∇f, while the smallest rate of change occurs in the direction of $-\nabla f$

Proof. Since $\mathbf{u}$ is a unit vector,

$$D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u} = |\nabla f| \, |\mathbf{u}| \cos\theta = |\nabla f| \cos\theta,$$

where θ is the angle between ∇f and $\mathbf{u}$. Clearly, $D_{\mathbf{u}} f$ is maximal when $\theta = 0$, in which case $\mathbf{u}$ is in the same direction as ∇f. Similarly, $D_{\mathbf{u}} f$ is minimal when $\theta = \pi$, in which case $\mathbf{u}$ is in the opposite direction as ∇f. $\square$

We say that $\nabla f(a,b)$ is the **direction of steepest ascent** at (a,b), while $-\nabla f(a,b)$ is the **direction of steepest descent**.

20.2.6 Implicit Differentiation

Consider the unit circle, which has equation

$$x^2 + y^2 = r^2.$$

Previously, we learnt that to find $\mathrm{d}y/\mathrm{d}x$, we can simply differentiate term by term, treating y as a function of x and using the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}x} g(y) = \frac{\mathrm{d}}{\mathrm{d}y} g(y) \cdot \frac{\mathrm{d}y}{\mathrm{d}x}.$$

Using our example of the unit circle, we get

$$2x + 2y\frac{\mathrm{d}y}{\mathrm{d}x} = 0 \implies \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{y}{x}.$$

While morally true, this approach to implicit differentiate is not entirely rigorous. For a more formal justification, we turn to partial derivatives.

Going back to our example of the unit circle, if we move all terms to one side of the equation, we get

$$x^2 + y^2 - r^2 = 0.$$

Now, observe that the LHS is simply a function of x and y, i.e.

$$f(x, y) = x^2 + y^2 - r^2.$$

Hence, we can define y implicitly as a function of x that satisfies

$$f(x, y) = 0.$$

If we differentiate the above equation with respect to x, by the multivariate chain rule, we get

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}x} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

Clearly, $\mathrm{d}x/\mathrm{d}x = 1$. Rearranging, we get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{f_x(x, y)}{f_y(x, y)}.$$

Since

$$f_x(x, y) = 2x, \quad \text{and} \quad f_y(x, y) = 2y,$$

we get

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

as expected.

More generally,

Proposition 20.2.13 (Implicit Differentiation for Univariate Functions)**.** If the equation

$$f(x, y) = 0$$

implicitly defines y as a function of x, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{f_x(x, y)}{f_y(x, y)},$$

given that $f_y(x, y) \neq 0$.

We can extend this result to functions of two variables.

Proposition 20.2.14 (Implicit Differentiation for Functions of Two Variables)**.** If the equation

$$f(x, y, z) = 0$$

implicitly defines z as a function of x and y, then

$$\frac{\partial z}{\partial x} = -\frac{f_x(x, y, z)}{f_z(x, y, z)} \quad \text{and} \quad \frac{\partial z}{\partial y} = -\frac{f_y(x, y, z)}{f_z(x, y, z),}$$

given that $f_z(x, y, z) \neq 0$.

To see this in action, consider the following sample problem:

Sample Problem 20.2.15. Find the value of $\partial^2 z/\partial x^2$ at $(0, 0, c)$ of the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

Solution. Let

$$f(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1.$$

Applying the above result, we have

$$\frac{\partial z}{\partial x} = -\frac{f_x(x, y, z)}{f_z(x, y, z)} = -\frac{2x/a^2}{2z/c^2} = -\frac{c^2}{a^2}\frac{x}{z}.$$

Partially differentiating with respect to x once more,

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial z}{\partial x}\left(-\frac{c^2}{a^2}\frac{x}{z}\right) = -\frac{c^2}{a^2 z}.$$

Hence,

$$\left.\frac{\partial^2 z}{\partial x^2}\right|_{(0,0,c)} = -\frac{c}{a^2}.$$

□

20.3 Approximations

In §17, we learnt that

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \dots.$$

If we want to approximate $f(x)$ for x near 0, we can truncate the Maclaurin series of $f(x)$. For instance, the linear approximation to x is

$$f(x) \approx f(0) + f'(0),$$

which is the tangent line at $x = 0$. If we want better approximations, we can simply take more terms. For instance, if we take one more term, then we get the quadratic approximation

$$f(x) \approx f(0) + f'(0)x + \frac{f''(0)}{2!}x^2.$$

In some sense, we can get a good approximation to $f(x)$ around $x = 0$ if we can find a simpler function which

- has the same value as f at $x = 0$, and
- has the same derivatives as f at $x = 0$ (up to the order of derivatives we prefer).

The same idea is extended to functions of two variables (or any multivariate functions) at a general point. The idea of approximation $f(x, y)$ at a point (a, b) is to find a simpler function which

- has the same value as f at (a, b), and
- has the same nth-order partial derivatives as f at (a, b) (where n is the highest order we prefer).

In this subsection, we look at the case where $n = 1$ (linear approximation) and $n = 2$ (quadratic approximation).

20.3.1 Tangent Plane

To find a linear approximation of $f(x, y)$ at (a, b) is to find a simpler function which

- has the same value as f at (a, b), and
- has the same partial derivatives as f at (a, b).

Let this approximation be $T(x, y)$. As the name suggests, $T(x, y)$ is linear and is hence of the form

$$T(x, y) = C_1 + C_2(x - a) + C_3(y - b),$$

where C_1, C_2 and C_3 are constants to be determined.

From the first condition, we require $f(a, b) = T(a, b)$. Hence,

$$f(a, b) = T(a, b) = C_1.$$

From the second condition, we require $f_x(a, b) = T_x(a, b)$ and $f_y(a, b) = T_y(a, b)$. This gives

$$f_x(a, b) = T_x(a, b) = C_2$$

and

$$f_y(a, b) = T_y(a, b) = C_3.$$

We hence have:

Proposition 20.3.1 (Linear Approximation)**.** The linear approximation at (a, b) is given by

$$T(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Recall that the linear approximation to a univariate function at $x = a$ is the tangent line at that point. Generalizing this up a dimension, the linear approximation $T(x, y)$ is the **tangent plane** to $f(x, y)$ at (a, b).

Using 3D vector geometry, we can find the normal vector to $z = f(x, y)$ at (a, b):

$$\mathbf{n} = \begin{pmatrix} f_x(a, b) \\ f_y(a, b) \\ -1 \end{pmatrix}.$$

20.3.2 Quadratic Approximation

To find a quadratic approximation of $f(x, y)$ at (a, b) is to find a simpler function which

- has the same value as f at (a, b), and
- has the same first and second partial derivatives as f at (a, b).

Remark. In univariate functions, the word "quadratic" refers to functions with terms of order 2, such as x^2. Similarly with multivariables, "quadratic" refers to terms with order 2, but it could be x^2, y^2 or xy; all variables contribute to the total order of the term. For instance, x^2y^3 is a term of order $2 + 3 = 5$.

To get the quadratic approximation $Q(x, y)$, we simply add terms of order 2 to the linear approximation $T(x, y)$:

$$Q(x, y) = T(x, y) + C_1(x - a)^2 + C_2(x - a)(y - b) + C_3(y - b)^2,$$

where C_1, C_2 and C_3 are constants. We can determine them by equating the second partial derivatives of $Q(x, y)$ with that of $f(x, y)$'s:

$$\begin{aligned} f_{xx}(a, b) &= Q_{xx}(a, b) = 2C_1, \\ f_{xy}(a, b) &= Q_{xx}(a, b) = \ \ C_2, \\ f_{yy}(a, b) &= Q_{xx}(a, b) = 2C_3. \end{aligned}$$

We hence have:

Proposition 20.3.2 (Quadratic Approximation)**.** The quadratic approximation at (a, b) is given by

$$\begin{aligned} Q(x, y) = f(a, b) &+ f_x(a, b)(x - a) + f_y(a, b)(y - b) \\ &+ \frac{1}{2} f_{xx}(a, b)(x - a)^2 + f_{xy}(a, b)(x - a)(y - b) + \frac{1}{2} f_{yy}(a, b)(y - b)^2. \end{aligned}$$

Note that by Clairaut's theorem, we can interchange f_{xy} and f_{yx} in the formula above, so long as they are continuous.

20.4 Maxima, Minima and Saddle Points

One important application of calculus is the optimization of functions which have many dependent variables. For example, one may maximize the amount of profit based on parameters such as the cost of raw materials, workers' salaries, time needed for production, etc.

To find stationary points of a univariate function, we equate its gradient to 0. Similarly, for functions of two variables $f(x, y)$, if we want to find stationary points, we look for points where its gradient, ∇f, is the zero vector, i.e.

$$\nabla f = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

In functions of two variables, the stationary points we often come across are maxima, minima and saddle points (so named because it looks like a horse saddle).

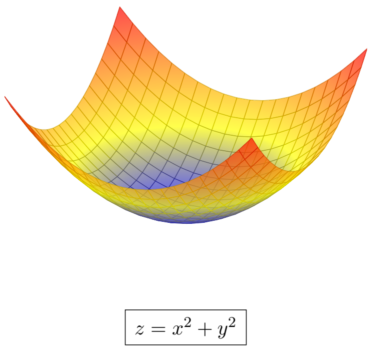


Figure 20.6: Minimum point at $(0, 0)$.

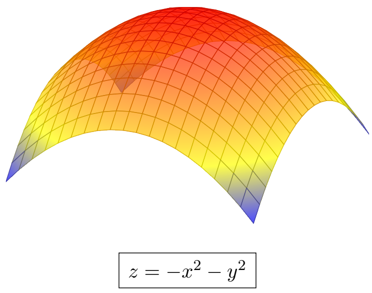


Figure 20.7: Minimum point at $(0, 0)$.

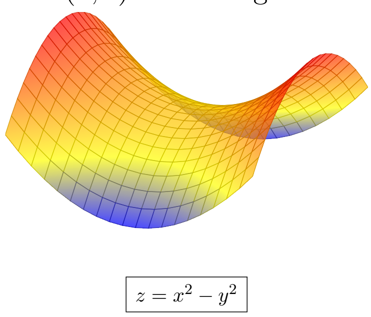


Figure 20.8: Saddle point at $(0, 0)$.

20.4.1 Global and Local Extrema

In optimization, we may distinguish between a **local extremum** (a collective term used to refer to the maximum and minimum) from a **global extremum**. Basically, a global maximum/minimum is the highest/lowest value which the function can achieve.

Local extrema are like the stationary points which we just discussed. For example, consider the following graph of $f(x, y) = x\mathrm{e}^{-x^2-y^2}$:

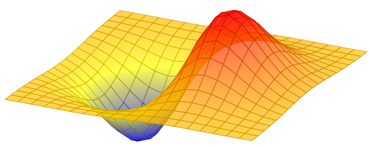


Figure 20.9

The intuitive idea behind local extrema is that when we move away from the maxima/minima in any direction, the value of the function will decrease/increase. However, this may not apply to global extrema. Consider the function $f(x, y) = x^2+y^2$ with domain $-2 \leq x \leq 2$, $-2 \leq y \leq 2$.

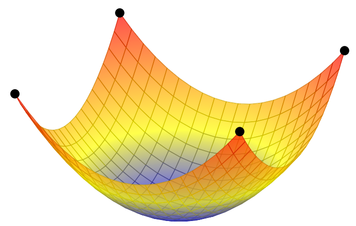


Figure 20.10

The global maxima occur at the corners of the domain. Note that these global maxima are also not stationary points.

Recipe 20.4.1 (Finding Global Extrema). To find the global extrema of a function, we must

- check all local extrema (set $\nabla f = \mathbf{0}$), and
- check for extrema along the boundary of the function's domain.

20.4.2 Second Partial Derivative Test

We can determine the nature of the stationary points by the second partial derivative test:

Proposition 20.4.2 (Second Partial Derivative Test). Let (a, b) be a stationary point of $f(x, y)$. Let

$$D = f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a, b)]^2 .$$

- If $D > 0$, and
  - $f_{xx}(a, b) > 0$ (or $f_{yy}(a, b) > 0$), then (a, b) is a minimum point.
  - $f_{xx}(a, b) < 0$ (or $f_{yy}(a, b) < 0$), then (a, b) is a maximum point.
- If $D < 0$, then (a, b) is a saddle point.
- If $D = 0$, the test is inconclusive.

The proof is similar to the proof of the second derivative test for univariate functions (see Proposition 16.3.5).

Proof. Consider the quadratic approximation $Q(x, y)$ of $f(x, y)$ at a stationary point (a, b). We have $f_x(a, b) = f_y(a, b) = 0$, hence

$$Q(x, y) = f(a, b) + \frac{1}{2}\left[f_{xx}(a, b)(x-a)^2 + 2f_{xy}(a, b)(x-a)(y-b) + f_{yy}(a, b)(y-b)^2\right].$$

Let

$$P(x, y) = f_{xx}(a, b)(x-a)^2 + 2f_{xy}(a, b)(x-a)(y-b) + f_{yy}(a, b)(y-b)^2.$$

We can view $P(x, y)$ as a quadratic in $(x-a)^2$. Consider the discriminant Δ of $P(x, y)$:

$$\begin{aligned}\Delta &= [2f_{xy}(a, b)(y-b)]^2 - 4f_{xx}(a, b)f_{yy}(a, b)(y-b)^2 \\ &= -4(y-b)^2\left(f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a, b)]^2\right).\end{aligned}$$

Let $D = f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a, b)]^2$. We make the following observations:

- If $D > 0$, then $\Delta < 0$.
  - If $f_{xx}(a,b) > 0$, then $P(x,y) > 0$ (since $f_{xx}(a,b)$ is the leading coefficient of $P(x,y)$). Thus, $Q(x,y) \geq f(a,b)$, whence (a,b) is a minimum point.
  - If $f_{xx}(a,b) < 0$, then $P(x,y) < 0$. Thus, $Q(x,y) \leq f(a,b)$, whence (a,b) is a maximum point.
- If $D < 0$, then $\Delta > 0$. This means that $P(x,y)$ has zeroes elsewhere other than (a,b), and it is sometimes positive and negative. Hence, (a,b) is a saddle point.
- If $D = 0$, then $\Delta = 0$. Hence, $P(x,y)$ has zeroes elsewhere other than (a,b), and it is either always > 0 or < 0 outside the zeroes. Thus, the stationary point could be a maximum, a minimum or even a saddle point; the test is inconclusive.

$\square$

21 Differential Equations

21.1 Definitions

Definition 21.1.1. A **differential equation** (DE) is an equation which involves one or more derivatives of a function y with respect to a variable x (i.e. y', y'', etc.). The **order** of a DE is determined by the highest derivative in the equation. The **degree** of a DE is the power of the highest derivative in the equation.

Example 21.1.2. The differential equation

$$x\left(\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right)^3 + x^2\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right) + y = 0$$

has order 2 and degree 3.

Observe that the equations $y = x^2 - 2$, $y = x^2$ and $y = x^2 + 10$ all satisfy the property $y' = 2x$ and are hence solutions of that DE. There are obviously many other possible solutions are we see that any equations of the form $y = x^2 + C$, where C is an arbitrary constant, will be a solution to the DE $y' = 2x$.

Definition 21.1.3. A **general solution** to a DE contains arbitrary constants, while a **particular solution** does not.

Hence, $y = x^2 + C$ is the general solution to the DE $y' = 2x$, while $y = x^2 - 2$, $y = x^2$ and $y = x^2 + 10$ are the particular solutions.

In general, the general solution of an nth order DE has n arbitrary constants.

21.2 Solving Differential Equations

In this section, we introduce methods to solve three special types of differential equations, namely

- separable DE,
- first-order linear DE, and
- second-order linear DE with constant coefficients.

We also demonstrate how to solve DEs using a given substitution, which is useful if the DE to be solved is not in one of the above three forms.

21.2.1 Separable Differential Equation

Definition 21.2.1. A **separable differential equation** is a DE that can be written in the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y).$$

Recipe 21.2.2 (Solving via Separation of Variables).

1. Separate the variables.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y) \implies \frac{1}{g(y)}\frac{\mathrm{d}y}{\mathrm{d}x} = f(x).$$

2. Integrate both sides with respect to x.

$$\int \frac{1}{g(y)}\frac{\mathrm{d}y}{\mathrm{d}x}\,dx = \int f(x)\,\mathrm{d}x \implies \int \frac{1}{g(y)}\,\mathrm{d}y = \int f(x)\,\mathrm{d}x.$$

Example 21.2.3 (Solving via Separation of Variables). Consider the separable DE

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

Separating variables,

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

Integrating both sides with respect to x, we get

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

Using the chain rule, we can rewrite the LHS as

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

Thus,

$$2\arctan y = \ln|x| + C.$$

This is the general solution to the given DE.

21.2.2 First-Order Linear Differential Equation

Definition 21.2.4. A **first-order linear differential equation** is a DE that can be written in the form

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

To solve a linear first-order DE, we first observe that the LHS looks like the product rule has been applied. This motivates us to multiply through by a new function $f(x)$ such that the LHS can be written as the derivative of a product:

$$f(x)\frac{\mathrm{d}y}{\mathrm{d}x} + f(x)p(x)y = f(x)q(x). \tag{1}$$

Recall that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)y\right] = f(x)\frac{\mathrm{d}y}{\mathrm{d}x} + f'(x)y.$$

Comparing this with (1), we want $f(x)$ to satisfy

$$f(x)p(x) = f'(x) \implies \frac{f'(x)}{f(x)} = p(x).$$

Observe that the LHS is simply the derivative of $\ln f(x)$. Integrating both sides, we get

$$\ln f(x) = \int p(x)\,\mathrm{d}x \implies f(x) = \exp\int p(x)\,\mathrm{d}x.$$

Going back to (1), we get

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[y\mathrm{e}^{\int p(x)\,\mathrm{d}x}\right] = q(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x}.$$

Once again, we get a separable DE, which we can solve easily:

$$y\mathrm{e}^{\int p(x)\,\mathrm{d}x} = \int q(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x}\,\mathrm{d}x.$$

This is the general solution to the DE.

Definition 21.2.5. The function $f(x) = \mathrm{e}^{\int p(x)\,\mathrm{d}x}$ is called the **integrating factor**, sometimes denoted I. F..

Note that we do not need to derive the integrating factor like above every time we solve a linear first-order DE. We can simply quote the result $\text{I. F.} = \mathrm{e}^{\int p(x)\,\mathrm{d}x}$. The following list is a summary of the steps we need to solve a linear first-order DE.

Recipe 21.2.6 (Solving via Integrating Factor)**.**

1. Multiply the DE through by the $\text{I. F.} = \mathrm{e}^{\int p(x)\,\mathrm{d}x}$.

$$\mathrm{e}^{\int p(x)\,\mathrm{d}x}\frac{\mathrm{d}y}{\mathrm{d}x} + \mathrm{e}^{\int p(x)\,\mathrm{d}x}p(x)y = \mathrm{e}^{\int p(x)\,\mathrm{d}x}q(x).$$

2. Express the LHS as the derivative of a product.

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[y\mathrm{e}^{\int p(x)\,\mathrm{d}x}\right] = \mathrm{e}^{\int p(x)\,\mathrm{d}x}q(x).$$

3. Integrating both sides with respect to x.

$$y\mathrm{e}^{\int p(x)\,\mathrm{d}x} = \int \mathrm{e}^{\int p(x)\,\mathrm{d}x}q(x)\,\mathrm{d}x.$$

Note that when finding the integrating factor, there is no need to include the arbitrary constant or consider $|x|$ when integrating $1/x$ with respect to x, as it does not contribute to the solution process in any way; the constants will cancel each other out.

Example 21.2.7 (Solving via Integrating Factor)**.** Consider the DE equation

$$x\frac{\mathrm{d}y}{\mathrm{d}x} + 3y = 5x^2.$$

Writing this in standard form,

$$\frac{\mathrm{d}y}{\mathrm{d}x} + \left(\frac{3}{x}\right)y = 5x.$$

The integrating factor is hence

$$\text{I. F.} = \mathrm{e}^{\int 3/x\,\mathrm{d}x} = \mathrm{e}^{3\ln x} = x^3.$$

Multiplying the integrating factor through the DE,

$$x^3\frac{\mathrm{d}y}{\mathrm{d}x} + 3x^2y = \frac{\mathrm{d}}{\mathrm{d}x}\left(x^3y\right) = 5x^4.$$

Integrating both sides with respect to x, we get the general solution

$$x^3 y = \int 5x^4 \, \mathrm{d}x = x^5 + C.$$

21.2.3 Second-Order Linear Differential Equations with Constant Coefficients

In this section, we look at second-order linear differential equations and constant coefficients, which has the general form

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

If $f(x) \equiv 0$, we call the DE **homogeneous**. Else, it is **non-homogeneous**. In general, a second-order DE will have two solutions.

Before looking at the methods to solve second-order DEs, we introduce two important concepts, namely the superposition principle and linear independence.

Theorem 21.2.8 (Superposition Principle). Let y_1 and y_2 be solutions to a linear, homogeneous differential equation. Then $Ay_1 + By_2$ is also a solution to the DE.

Proof. We consider the case where the DE has order 2, though the proof easily generalizes to higher orders.

Suppose y_1 and y_2 are solutions to

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

Substituting $y = Ay_1 + By_2$ into the DE, we get

$$\begin{aligned} &a\left(Ay_1'' + By_2''\right) + b\left(Ay_1' + By_2'\right) + c\left(Ay_1 + By_2\right) \\ &\quad = A\left(ay_1'' + by_1' + cy_1\right) + B\left(ay_2'' + by_2' + cy_2\right) \\ &\quad = 0. \end{aligned}$$

Hence, $Ay_1 + By_2$ satisfies the DE and is hence a solution. □

Definition 21.2.9. Two functions y_1 and y_2 are **linearly independent** if the only solution to

$$Ay_1 + By_2 = 0$$

is the trivial solution $A = B = 0$. If there exists non-zero solutions to A and B, then the two functions are **linearly dependent**.

We are now ready to solve second-order DEs.

Homogeneous Second-Order Linear Differential Equations with Constant Coefficients

Consider a homogeneous first-order linear differential equation with constant coefficients which has the form

$$a\frac{\mathrm{d}y}{\mathrm{d}x} + by = 0.$$

Using the method of integrating factor, we can show that the general solution is of the form

$$y = C\mathrm{e}^{-\frac{b}{a}x}.$$

We can extend this to the second-order case, i.e.

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

by looking for solutions of the form $y = \mathrm{e}^{mx}$, where m is a constant to be determined. Substituting $y = \mathrm{e}^{mx}$ into the differential equation, we get

$$am^2\mathrm{e}^{mx} + bm\mathrm{e}^{mx} + c\mathrm{e}^{mx} = 0.$$

Dividing by e^{mx}, we get the quadratic

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

This is known as the **characteristic equation** of the DE.

If we can solve for m in the characteristic equation, we can find the solution $y = \mathrm{e}^{mx}$. Since the characteristic equation is quadratic, it has, in general, two roots, say m_1 and m_2. We thus have the following three scenarios to consider:

- The roots are real and distinct.
- The roots are real and equal.
- The roots are complex conjugates.

Real and Distinct Roots If m_1 and m_2 are real and distinct, $y_1 = \mathrm{e}^{m_1 x}$ and $y_2 = \mathrm{e}^{m_2 x}$ will both be solutions to the DE. Hence, by the superposition principle, the general solution is

$$y = A\mathrm{e}^{m_1 x} + B\mathrm{e}^{m_2 x},$$

where A and B are constants.

Real and Equal Roots If the two roots are equal, i.e. $m_1 = m_2 = m$, then $y_1 = \mathrm{e}^{m_1 x}$ and $y_2 = \mathrm{e}^{m_2 x}$ are no longer linearly independent. Hence, we effectively only get one solution $y_1 = \mathrm{e}^{mx}$. To obtain the general solution, we have to find another solution that is not a constant multiple of e^{mx}. By intelligently guessing a solution, we see that $y_2 = x\mathrm{e}^{mx}$ satisfies the DE. Hence, by the superposition principle, the general solution is

$$y = A\mathrm{e}^{mx} + Bx\mathrm{e}^{mx} = (A + Bx)\mathrm{e}^{mx}.$$

Complex Roots If the two roots are complex, then they are conjugates, and we can write them as

$$m_1 = p + \mathrm{i}q, \quad m_2 = p - \mathrm{i}q.$$

Hence,

$$y_1 = \mathrm{e}^{(p+\mathrm{i}q)x} = \mathrm{e}^{px}\left(\cos qx + \mathrm{i}\sin qx\right)$$

and

$$y_2 = \mathrm{e}^{(p-\mathrm{i}q)x} = \mathrm{e}^{px}\left(\cos qx - \mathrm{i}\sin qx\right).$$

By the superposition principle, we get the general solution

$$\begin{aligned} y &= C\mathrm{e}^{px}\left(\cos qx + \mathrm{i}\sin qx\right) + D\mathrm{e}^{px}\left(\cos qx - \mathrm{i}\sin qx\right) \\ &= \mathrm{e}^{px}\left(A\cos qx + B\sin qx\right), \end{aligned}$$

where $A = C + D$ and $B = \mathrm{i}(C - D)$ are arbitrary constants.

In summary,

Recipe 21.2.10 (Homogeneous Second-Order Linear DE with Constant Coefficients)**.** To solve the second-order DE

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

1. Form the characteristic equation $am^2+bm+c=0$.
2. Find the roots m_1 and m_2 of this characteristic equation.
3. - If m_1 and m_2 are real and distinct, then
   $$y=A\mathrm{e}^{m_1x}+B\mathrm{e}^{m_2x}.$$
   - If m_1 and m_2 are real and equal, i.e. $m_1=m_2=m$, then
   $$y=(A+Bx)\mathrm{e}^{mx}.$$
   - If m_1 and m_2 are complex, i.e. $m_1=p+\mathrm{i}q$ and $m_2=p-\mathrm{i}q$, then
   $$y=\mathrm{e}^{px}\left(A\cos qx+B\sin qx\right).$$

Non-Homogeneous Second-Order Linear Differential Equations with Constant Coefficients

We now consider the non-homogeneous second-order linear DE with constant coefficients, which takes the form

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

In order to solve this DE, we apply the following result:

Theorem 21.2.11. If y_c is the general solution of

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

and y_p is a particular solution of

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

then

$$y=y_c+y_p$$

is the general solution to

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

Proof. We want to solve

$$ay''+by'+cy=f(x). \tag{1}$$

Let y_c be the solution to $ay''+by'+cy=0$. Then

$$ay_c''+by_c'+cy_c=0.$$

Let y_p be a particular solution to (1). Then

$$ay_p''+by_p'+cy_p=f(x).$$

Substituting $y = y_c + y_p$ into (1), we get

$$\begin{aligned} &a\left(y_c'' + y_p''\right) + b\left(y_c' + y_p'\right) + c\left(y_c + y_p\right) \\ &\quad = \left(ay_c'' + by_c' + cy_c\right) + \left(ay_p'' + by_p' + cy_p\right) \\ &\quad = 0 + f(x) = f(x). \end{aligned}$$

□

Note that y_c is called the **complementary function** while y_p is called the **particular integral** or **particular solution**.

We know how to solve the homogeneous DE, so getting y_c is easy. The hard part is getting a particular solution y_p. However, if we make some intelligent guesses, we can determine the general form of y_p. This is called the **method of undetermined coefficients**. We demonstrate this method with the following example:

Example 21.2.12 (Method of Undetermined Coefficients). Consider the differential equation

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 3\frac{\mathrm{d}y}{\mathrm{d}x} - 4y = 3 + 8x^2.$$

y_c can easily be obtained:

$$y_c = A\mathrm{e}^x + B\mathrm{e}^{-4x}.$$

Now, observe that $f(x) = 3 + 8x^2$ is a polynomial of degree 2. Thus, we guess that y_p is also a polynomial of degree 2, i.e. $y_p = Cx^2 + Dx + E$, where C, D and E are coefficients to be determined (hence the name "method of undetermined coefficients"). Substituting this into the DE yields

$$(2C) + 3(2Cx + D) - 4\left(Cx^2 + Dx + E\right) = 3 + 8x^2.$$

Comparing coefficients, we get the system

$$\begin{cases} -4C & = 8 \\ \quad 6C - 4D & = 0 \\ \quad 2C + 3D - 4E & = 3 \end{cases},$$

whence $C = -2$, $D = -3$ and $E = -4$. Thus, the particular solution is

$$y_p = -2x^2 - 3x - 4$$

and the general solution is

$$y = y_c + y_p = A\mathrm{e}^x + B\mathrm{e}^{-4x} - 2x^2 - 3x - 4.$$

In our syllabus, we are only required to solve non-homogeneous DEs where $f(x)$ is a polynomial of degree n (as above), of the form $p\mathrm{e}^{kx}$, or of the form $p\cos kx + q\sin kx$. The "guess" for y_p in each of the three cases is tabulated below:

$f(x)$	**"Guess" for y_p**
Polynomial of degree n	Polynomial of degree n
$p\mathrm{e}^{kx}$	$C\mathrm{e}^{kx}$
$p\cos kx + q\sin kx$	$C\cos kx + D\sin kx$

In the event where our "guess" for y_p appears in the complementary function y_c, we need to make some adjustments to our "guess" (similar to the case where $m_1 = m_2$ when solving a homogeneous DE). Typically, we multiply the guess by powers x until the guess no longer appears in the complementary function.

Example 21.2.13 (Adjusting y_p).

- If $ay'' + by' + cy = \mathrm{e}^{2x}$ has complementary function $y_c = A\mathrm{e}^{-5x} + B\mathrm{e}^{2x}$, we try $y_p = Cx\mathrm{e}^{2x}$.
- If $ay'' + by' + cy = \mathrm{e}^{2x}$ has complementary function $y_c = (A + Bx)\mathrm{e}^{2x}$, we try $y_p = Cx^2\mathrm{e}^{2x}$.

21.2.4 Solving via Substitution

Sometimes, we are given a DE that is not of the forms described in this section. We must then use the given substitution function to simplify the original DE into one of the standard forms. Similar to integration by substitution, all instances of the dependent variable (including its derivatives) must be substituted.

Recipe 21.2.14 (Solving via Substitution).

1. Differentiate the given substitution function.
2. Substitute into the original DE and simplify to obtain another DE that we know how to solve.
3. Obtain the general solution of the new DE with new dependent variables.
4. Express the solution in terms of the original variables.

Sample Problem 21.2.15. By using the substitution $y = ux^2$, find the general solution of the differential equation

$$x^2 \frac{\mathrm{d}y}{\mathrm{d}x} - 2xy = y^2, \quad x > 0.$$

Solution. From $y = ux^2$, we see that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2ux + x^2 \frac{\mathrm{d}u}{\mathrm{d}x}.$$

Substituting this into the original DE,

$$x^2 \left(2ux + x^2 \frac{\mathrm{d}u}{\mathrm{d}x}\right) - 2x\left(ux^2\right) = \left(ux^2\right)^2.$$

Simplifying, we get the separable DE

$$\frac{\mathrm{d}u}{\mathrm{d}x} = u^2,$$

which we can easily solve:

$$\int \frac{1}{u^2} \,\mathrm{d}u = \int 1 \,\mathrm{d}x \implies -\frac{1}{u} = x + C.$$

Re-substituting y back in, we have the general solution

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

□

21.3 Family of Solution Curves

Graphically, the general solution of a differential equation is represented by a family of solution curves which contains infinitely many curves as the arbitrary constant c can take any real number.

A particular solution of the differential equation is represented graphically by one member of that family of solution curves (i.e. one value of the arbitrary constant).

When sketching a family of curves, we choose values of the arbitrary constant that will result in qualitatively different curves. We also need to sketch sufficient members (usually at least 3) of the family to show all the general features of the family.

Example 21.3.1. The following diagram shows three members of the family of solution curves for the general solution $y = Ae^{x^2}$.

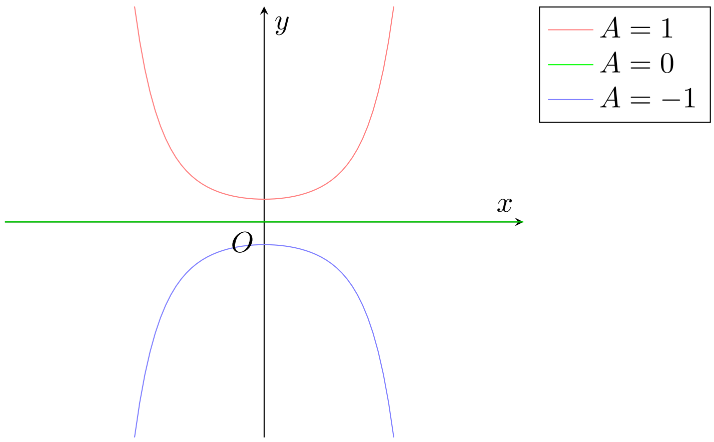


Figure 21.1

21.4 Approximating Solutions

Most of the time, a first-order differential equation of the general form $\mathrm{d}y/\mathrm{d}x = f(x, y)$ cannot be solved exactly and explicitly by analytical methods like those discussed in the earlier sections. In such cases, we can use numerical methods to approximate solutions to differential equations.

Different methods can be used to approximate solutions to a differential equation. A sequence of values $y_1, y_2, \dots$ is generated to approximate the exact solutions at the points $x_1, x_2, \dots$. It must be emphasized that the numerical methods do not generate a formula for the solution to the differential equation. Rather, they generate a sequence of approximations to the actual solution at the specified points.

In this section, we look at Euler's Method, as well as the improved Euler's Method.

21.4.1 Euler's Method

The key principle in Euler's method is the use of a linear approximation for the tangent line to the actual solution curve $y(t)$ to approximate a solution.

Derivation

Given an initial value problem

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(t, y), \quad y(t_0) = y_0,$$

we start at (t_0, y_0) on the solution curve as shown in the figure below. By the point-slope formula, the equation of the tangent line through (t_0, y_0) is given as

$$y - y_0 = \left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_{t=t_0}(t - t_0) = f(t_0, y_0)(t - t_0). \tag{1}$$

If we choose a step size of Δt on the t-axis, then $t_1 = t_0 + \Delta t$. Using (1) at $t = t_1$, we can obtain an approximate value y_1 from

$$y_1 = y_0 + (t_1 - t_0)f(t_0, y_0). \tag{2}$$

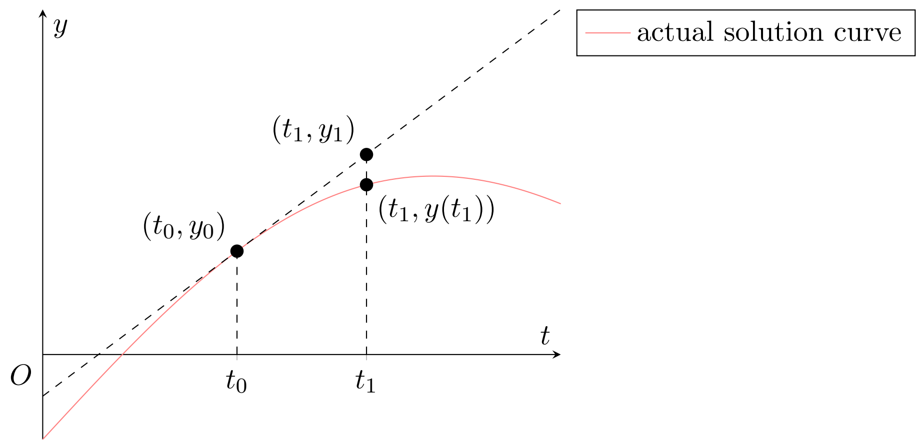


Figure 21.2

The point (t_1, y_1) on the tangent line is an approximation to the point $(t_1, y(t_1))$ on the actual solution curve. That is, $y_1 \approx y(t_1)$. From the above figure, it is observed that the accuracy of the approximation depends heavily on the size of Δt. Hence, we must choose an increment Δt which is “reasonably small”.

We can extend (2) further. In general, at $t = t_{n+1}$, it follows that

$$y_{n+1} = y_n + (t_{n+1} - t_n)f(t_n, y_n).$$

Recipe 21.4.1 (Euler's Method)**.** Euler's method, with step size Δt, gives the approximation

$$y(t_n) \approx y_{n+1} = y_n + (t_{n+1} - t_n)f(t_n, y_n).$$

Example 21.4.2 (Euler's Method)**.** Consider the initial value problem

$$\frac{\mathrm{d}y}{\mathrm{d}t} = 2y - 1, \quad y(0) = \frac{3}{2},$$

which can be verified to have solution $y = e^{2t} + 1/2$. Suppose we wish to approximate the value of $y(0.3)$ (which we know to be $\mathrm{e}^{2(0.3)} + 1/2 = 2.322$). Using Euler's method with step size $\Delta t = 0.1$, we get

$$\begin{aligned}
y_1 &= y_0 + \Delta t\,(2y_0 - 1) = 1.5 + 0.1\,[2(1.5) - 1] = 1.7\\
y_2 &= y_1 + \Delta t\,(2y_1 - 1) = 1.7 + 0.1\,[2(1.7) - 1] = 1.94\\
y_3 &= y_2 + \Delta t\,(2y_2 - 1) = 1.94 + 0.1\,[2(1.94) - 1] = 2.228
\end{aligned}$$

Hence, $y(0.3) \approx y_3 = 2.228$, which is a decent approximation (4.04% error).

Error in Approximations

Similar to the trapezium rule, the nature of the estimates given by Euler's method depends on the concavity of the actual solution curve.

- If the actual solution curve is concave upwards (i.e. lies above its tangents), the approximations are under-estimates.
- If the actual solution curve is concave downwards (i.e. lies below its tangents), the approximations are over-estimates.

Also note that the smaller the step size Δt, the better the approximations. However, in doing so, more calculations must be made. This is a situation that is typically of numerical methods: there is a trade-off between accuracy and speed.

21.4.2 Improved Euler's Method

In the previous section, we saw how Euler's method over- or under-estimates the actual solution curve due to the curve's concavity. The improved Euler's method address this.

Derivation

Suppose the actual solution curve is concave upward. Let T_0 and T_1 be the tangent lines at $t = t_0$ and $t = t_1$ respectively. Let the gradients of T_0 and T_1 be m_0 and m_1 respectively. We wish to find the optimal gradient m such that the line with gradient m passing through $(t_0, y(t_0))$ also passes through $(t_1, y(t_1))$.

Since the actual solution curve is concave upward, both T_0 and T_1 lie below the actual solution curve for all $t \in [t_0, t_1]$. This is depicted in the diagram below.

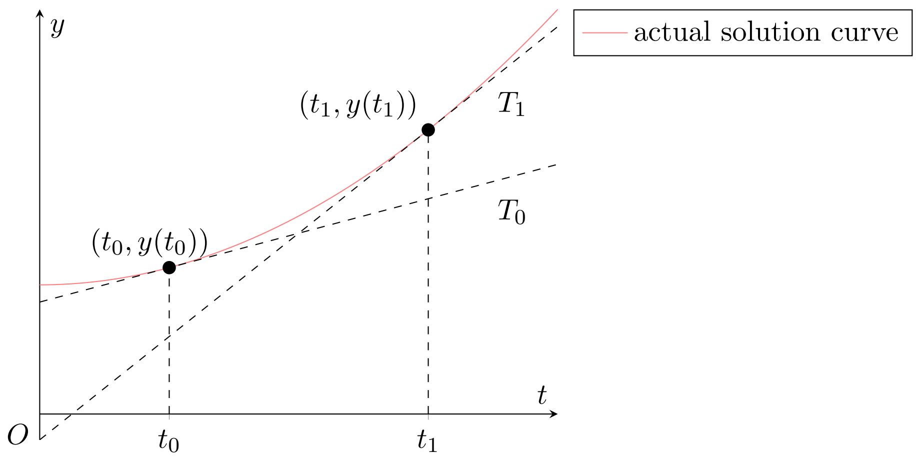


Figure 21.3

Now, observe what happens when we translate T_1 such that it passes through $(t_0, y(t_0))$:

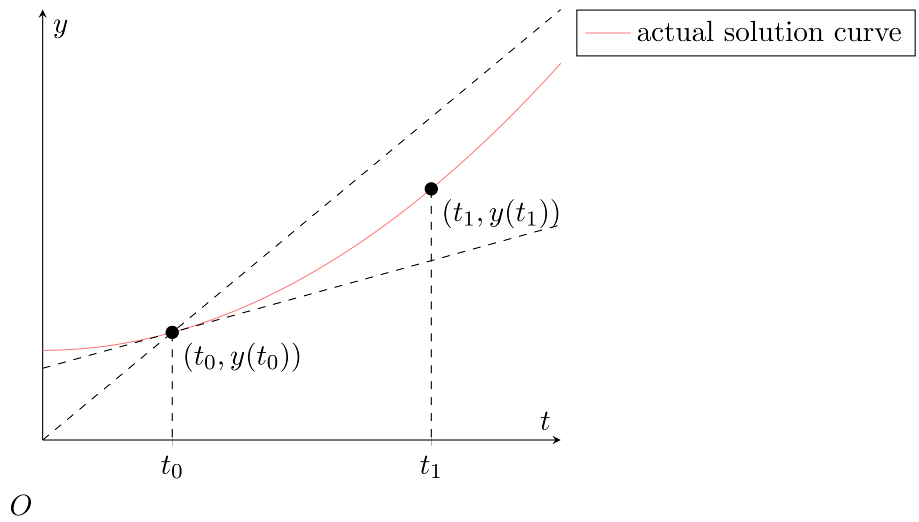


Figure 21.4

The translated T_1 is now overestimating the actual solution curve at $t = t_1$! Hence, the optimal gradient m is somewhere between m_0 and m_1. This motivates us to approximate m by taking the average of m_0 and m_1:

$$m \approx \frac{m_0 + m_1}{2}.$$

We now find m_0 and m_1. Note that

$$m_0 = f(t_0, y(t_0)) \quad \text{and} \quad m_1 = f(t_1, y(t_1)).$$

This poses a problem, as the value of $y(t_1)$ is not known to us. However, we can estimate it using the Euler method:

$$y(t_1) \approx \widetilde{y}_1 = y_0 + \Delta t f(t_0, y_0).$$

Note that we denote this approximation as $\widetilde{y}_1$. We thus have

$$m \approx \frac{m_0 + m_1}{2} = \frac{f(t_0, y_0) + f(t_1, \widetilde{y}_1)}{2}.$$

We are now ready to approximate $y(t_1)$. By the point-slope formula, the line with gradient m passing through (t_0, y_0) has equation

$$y - y_0 = m(t - t_0) \approx \frac{f(t_0, y_0) + f(t_1, \widetilde{y}_1)}{2}(t - t_0).$$

When $t = t_1$, we get

$$y(t_1) \approx y_1 = y_0 + \Delta t \left[\frac{f(t_0, y_0) + f(t_1, \widetilde{y}_1)}{2}\right]. \tag{1}$$

A similar derivation can be obtained when the actual solution curve is concave downwards.

Extending (1), we get the usual statement of the improved Euler's method:

Recipe 21.4.3 (Improved Euler's Method). The improved Euler's method, with step size Δt, gives the approximation

$$y_{n+1} = y_n + \Delta t \left[\frac{f(t_n, y_n) + f(t_{n+1}, \widetilde{y}_{n+1})}{2}\right],$$

where

$$\widetilde{y}_{n+1} = y_n + \Delta t f(t_n, y_n).$$

Definition 21.4.4. $\widetilde{y}_{n+1}$ is called the **predictor**, while y_{n+1} is called the **corrector**.

Example 21.4.5 (Improved Euler's Method). Consider the initial value problem

$$\frac{\mathrm{d}y}{\mathrm{d}t} = 2y - 1, \quad y(0) = \frac{3}{2},$$

which we previously saw in Example 21.4.2. Suppose we wish to approximate the value of $y(0.3)$. Using the improved Euler's method with step size $\Delta t = 0.1$,

$$\begin{aligned}
\widetilde{y}_1 &= y_0 + \Delta t f(t_0, y_0) = 1.7 \\
y_1 &= y_0 + \Delta t \left[\frac{f(t_0, y_0) + f(t_1, \widetilde{y}_1)}{2}\right] = 1.72 \\
\\
\widetilde{y}_2 &= y_1 + \Delta t f(t_1, y_1) = 1.964 \\
y_2 &= y_1 + \Delta t \left[\frac{f(t_1, y_1) + f(t_2, \widetilde{y}_2)}{2}\right] = 1.9884 \\
\\
\widetilde{y}_3 &= y_2 + \Delta t f(t_2, y_2) = 2.28608 \\
y_3 &= y_2 + \Delta t \left[\frac{f(t_2, y_2) + f(t_3, \widetilde{y}_3)}{2}\right] = 2.35848
\end{aligned}$$

Hence, $y(0.3) \approx y_3 = 2.35848$, which gives an error of 0.270%, much better than the 4.04% achieved by Euler's method.

21.4.3 Relationship with Approximations to Definite Integrals

Recall that solving differential equations analytically required us to integrate. It is thus no surprise that approximating solutions to differential equations is related to approximating the values of definite integrals. As we will see, the Euler method is akin to approximating definite integrals using a Riemann sum, while the improved Euler method is akin to using the trapezium rule.

Consider the differential equation $\frac{\mathrm{d}y}{\mathrm{d}t} = f(t, y)$. By the fundamental theorem of calculus, the area under the graph of $f(t, y)$ from $t = t_0$ to $t = t_1$ is given by

$$\int_{t_0}^{t_1} f(t, y) \, \mathrm{d}t = \int_{t_0}^{t_1} \frac{\mathrm{d}y}{\mathrm{d}t} \, \mathrm{d}t = y(t_1) - y(t_0). \tag{1}$$

Note that we know $y(t_0)$. Hence, the better the approximation of the integral, the better the approximation of $y(t_1)$, which is what we want.

We can approximate this integral using a Riemann sum with one rectangle. Note that this rectangle has width Δt and height $f(t_0, y_0)$. Hence,

$$\int_{t_0}^{t_1} f(t, y) \, \mathrm{d}t = y(t_1) - y(t_0) \approx \Delta t f(t_0, y_0).$$

Rewriting, we get the statement of the Euler method:

$$y(t_1) \approx y(t_0) + \Delta t f(t_0, y_0).$$

We now approximate the integral in (1) using the trapezium rule with 2 ordinates. Note that the area of this trapezium is given by $\frac{1}{2}\Delta t \left[f(t_0, y_0) + f(t_1, y_1)\right]$. Hence,

$$\int_{t_0}^{t_1} f(t, y) \, \mathrm{d}t = y(t_1) - y(t_0) \approx \Delta t \left[\frac{f(t_0, y_0) + f(t_1, y_1)}{2}\right].$$

Rewriting, we (almost) get the statement of the improved Euler method:

$$y(t_1) \approx y(t_0) + \Delta t \left[\frac{f(t_0, y_0) + f(t_1, y_1)}{2} \right].$$

Recall that generally, the trapezium rule is a much better approximation than a Riemann sum. Correspondingly, it follows that the improved Euler method is a much better approximation than the Euler method.

21.5 Modelling Populations with First-Order Differential Equations

Populations, however defined, generally change their magnitude as a function of time. The main goal here is to provide some mathematical models as to how these populations change, construct the corresponding solutions, analyse the properties of these solutions, and indicate some applications.

For the case of living biological populations, we assume that all environment and/or cultural factors operate on a timescale which is much longer than the intrinsic timescale of the population of interest. If this holds, then the mathematical model takes the following form of a simple population:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = f(P), \quad P(0) = p_0 \geq 0,$$

where $P(t)$ is the value of the population P at time t. The function $f(P)$ is what distinguishes one model from another.

We would expect the model to have the same structure

$$\frac{\mathrm{d}P}{\mathrm{d}t} = g(P) - d(P),$$

where $g(P)$ and $d(P)$ are the growth and decline factors respectively. Also, we assume $g(0) = d(0) = 0$, whence $f(0) = 0$. This is related to the **axiom of parenthood**, which states the "every organism must have parents; there is no spontaneous generation of organisms".

In this section, we will look at two common population growth models, namely the exponential growth model and the logistic growth model.

21.5.1 Exponential Growth Model

A biological population with plenty of food, space to grow, and no threat from predators, tend to grow at a rate that is proportional to the population. That is, in each unit of time, a certain percentage of the individuals produce new individuals (similar for death too). If reproduction (and death) takes place more or less continuously, then the growth rate is represented by

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP,$$

where k is the **proportionality constant**.

We know that all solutions of this differential equation have the form

$$P(t) = p_0 \mathrm{e}^{kt}.$$

As such, this model is known as the **exponential growth model**. Depending on the value of k, the model results in either an exponential growth, decay, or constant value function as seen in the diagram below.

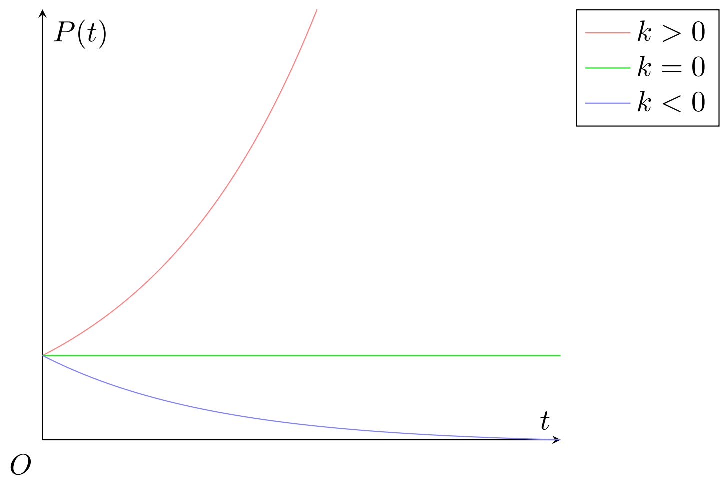


Figure 21.5

While the cases where $k \leq 0$ are possible to happen in real life, the case where $k > 0$ is not realistically possible as most populations are constrained by limitations of resources.

21.5.2 Logistic Growth Model

The following figure shows two possible courses for growth of a population. The red curve follows the exponential model, while the blue curve is constrained so that the population is always less than some number N. When the population is small relative to N, the two curves are identical. However, for the blue curve, when P gets closer to N, the growth rate drops to 0.

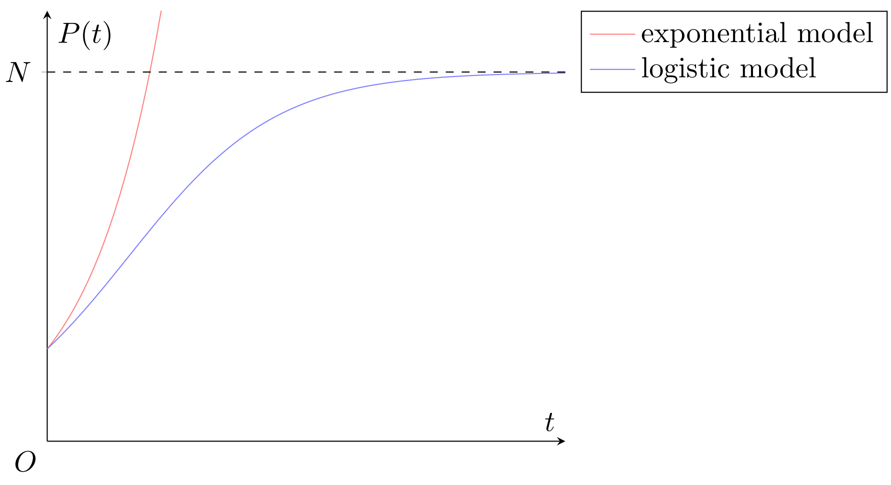


Figure 21.6

We may account for the growth rate declining to 0 by including in the model a factor $1 - P/N$, which is close to 1 (i.e. no effect) when P is much smaller than N, and close to 0 when P is close to N. The resulting model

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP\left(1 - \frac{P}{N}\right),$$

is called the **logistic growth model**. k is called the **intrinsic growth rate**, while N is called the **carrying capacity**.

Given the initial condition $P(0) = p_0$, the solution of the logistic equation is

$$P(t) = \frac{p_0 N}{[} p_0 + (N - p_0)\mathrm{e}^{-kt}].$$

Long-Term Behaviour

We now analyse the long-term behaviour of the model, which is determined by the value of P_0.

Notice that the derivative of the logistic growth model, $\mathrm{d}P/\mathrm{d}t = kP(1 - P/N)$, is 0 at $P = 0$ and $P = N$. Also notice that these are also solutions to the differential equation. These two values are the **equilibrium points** since they are constant solutions to the differential equation.

Consider the case where $k > 0$.

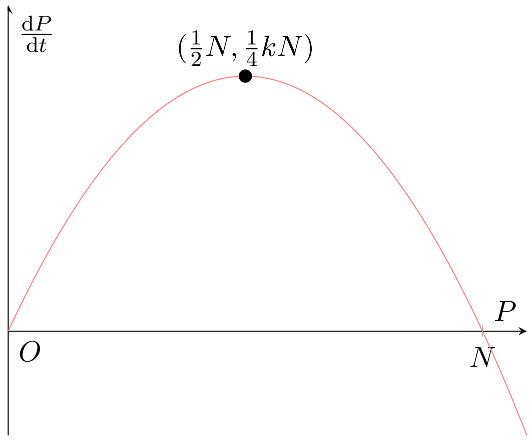


Figure 21.7

From the above diagram, we observe that

- if $0 < P_0 < N$, then P will increase towards N since $\mathrm{d}P/\mathrm{d}t > 0$.
- if $P_0 > N$, then P will decrease towards N since $\mathrm{d}P/\mathrm{d}t < 0$.

Since any population value in the neighbourhood of 0 will move away from 0, the equilibrium point at $P = 0$ is known as an **unstable equilibrium point**. On the contrary, since any population value in the neighbourhood of N will move towards N, the equilibrium point at $P = N$ is known as a **stable equilibrium point**.

Now consider the case where $k < 0$.

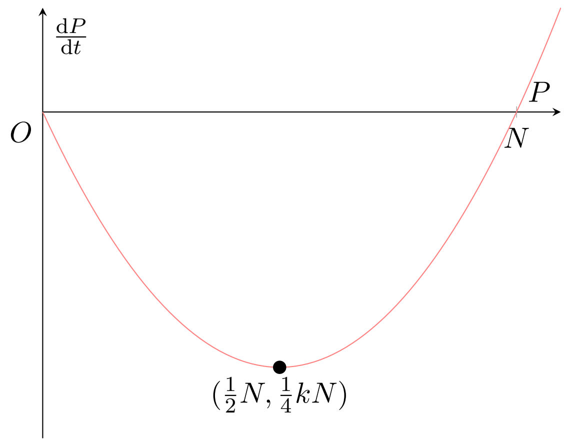


Figure 21.8

From the above diagram, we observe that

- if $0 < p_0 < N$, then P will decrease towards N since $\mathrm{d}P/\mathrm{d}t < 0$.
- if $p_0 > N$, then P will increase indefinitely since $\mathrm{d}P/\mathrm{d}t > 0$.

In this case, the equilibrium point at $P = 0$ is stable, while the equilibrium point at $P = N$ is unstable.

Thus, we see that what happens to the population in the long-run depends very much on the value of the initial population, P_0.

21.5.3 Harvesting

There are many single population systems for which harvesting takes place. **Harvesting** is a removal of a certain number of the population during each time period that the harvesting takes place. Below are some variants of the basic logistic model.

Constant Harvesting

The most direct way of harvesting is to use a strategy where a constant number, $H \geq 0$, of individuals are removed during each time period. For this situation, the logistic equation gets modified to the form

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP\left(1 - \frac{P}{N}\right) - H,$$

where H is known as the **harvesting rate**.

Observe that the equilibrium solutions to this modified logistic equation are:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP\left(1 - \frac{P}{N}\right) - H = 0 \implies P = \frac{N}{2} \pm \sqrt{\frac{N^2}{4} - \frac{NH}{k}}.$$

With the equilibrium solutions, we can do the same analysis above to determine the long-term behaviour of the model.

Variable Harvesting

The model

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP\left(1 - \frac{P}{N}\right) - HP$$

results by harvesting at a non-constant rate proportional to the present population P. The effect is to decrease the natural growth rate k by a constant amount H in the standard logistic model.

Restocking

The equation

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP\left(1 - \frac{P}{N}\right) - H\sin(\omega t)$$

models a logistic equation that is periodically harvested and restocked with maximal rate H. For sufficiently large p_0, the equation models a stable population that oscillates about the carrying capacity N with period $T = 2\pi/\omega$.

Part VI

Combinatorics

22 Permutations and Combinations

22.1 Counting Principles

Fact 22.1.1 (The Addition Principle)**.** Let E_1 and E_2 be two mutually exclusive events. If E_1 and E_2 can occur in n_1 and n_2 different ways respectively, then E_1 or E_2 can occur in $(n_1 + n_2)$ ways.

Fact 22.1.2 (The Multiplication Principle)**.** Consider a task S that can be broken down into two independent ordered stages S_1 and S_2. If S_1 and S_2 can occur in n_1 and n_2 ways respectively, then S_1 and S_2 can occur in succession in $n_1 n_2$ ways

Note that both the Addition and Multiplication Principles can be extended to any finite number of events.

22.2 Permutations

Definition 22.2.1. A **permutation** is an arrangement of a number of objects in which the **order is important**.

Example 22.2.2. ABC, BAC and CBA are possible permutations of the letters 'A', 'B' and 'C'.

Definition 22.2.3 (Factorial)**.** The **factorial** of a non-negative integer n is given by the recurrence relation

$$n! = n(n-1)!, \quad 0! = 1.$$

Equivalently,

$$n! = n(n-1)(n-2)\dots(3)(2)(1), \quad 0! = 1.$$

Proposition 22.2.4 (Permutations of Objects Taken from Sets of Distinct Objects)**.** The number of permutations of n distinct objects, taken r at a time without replacement, is given by

$$^{n}P_r = \underbrace{n(n-1)(n-2)\dots(n-r+1)}_{r \text{ consecutive integers}} = \frac{n!}{(n-r)!},$$

where $0 \leq r \leq n$.

Proof. Suppose we have n distinct objects that we want to fill up r ordered slots with. This operation can be done in r stages

- **Stage 1.** The number of ways to fill in the first slot is n.
- **Stage 2.** After filling in the first slot, the number of ways to fill in the second slot is $n-1$.
- **Stage 3.** After filling in the first and second slots, the number of ways to fill in the third slot is $n-2$.

This continues until we reach the last stage:

- **Stage r.** After filling all previous $r-1$ slots, the number of ways to fill in the last slot is $n-(r-1)=n-r+1$.

Thus, by the Multiplication Principle, the number of ways to fill up the r slots are

$$n(n-1)(n-2)\dots(n-r+1)=\frac{n!}{(n-r)!}.$$

□

Corollary 22.2.5 (Permutations of Distinct Objects in a Row)**.** The number of ways to arrange n distinct objects in a row, taken all at a time without replacement, is given by $n!$.

Proof. Take $r=n$. □

Proposition 22.2.6 (Permutations of Non-Distinct Objects in a Row)**.** The number of permutations of n objects in a row, taken all at a time without replacement, of which n_1 are of the 1st type, n_2 are of the 2nd type, …, n_k are of the kth type, where $n=n_1+n_2+\dots+n_k$, is given by

$$\frac{n!}{n_1!\,n_2!\dots n_k!}.$$

Proof. Let A_i be the set of arrangements where objects in the first i groups are now distinguishable, while objects in the remaining groups remain indistinguishable. For instance, A_1 is the set of arrangements of n objects in a row, of which n_2 are of the 2nd type, n_3 are of the 3rd type, …, n_k are of the kth type, while the objects previously of the 1st type are now distinct. We prove the above result by expressing $|A_0|$ in terms of $|A_k|$.

Suppose we make objects of the 1st type distinct. For each arrangement in A_0, the n_1 objects of the 1st type can be permuted among themselves in $n_1!$ ways. Hence,

$$|A_1|=n_1!\,|A_0|\,.$$

Next, suppose we make objects of the 2nd type distinct. For each arrangement in A_1, the n_2 objects of the 2nd type can be permuted among themselves in $n_2!$ ways. Hence,

$$|A_2|=n_2!\,|A_1|\,.$$

Continuing on, we see that

$$|A_k|=n_k!\,|A_{k-1}|=n_k!\,n_{k-1}!\,|A_{k-2}|=\dots=n_k!\,n_{k-1}!\dots n_1!\;|A_0|\,.$$

However, by definition, A_k is the set of arrangements of n distinct objects, which we know to be $n!$. Thus,

$$|A_0|=\frac{|A_k|}{n_1!\,n_2!\dots n_k!}=\frac{n!}{n_1!\,n_2!\dots n_k!}.$$

□

Remark. $\frac{n!}{n_1!n_2!\dots n_k!}$ is known as a **multinomial coefficient**, which is a generalization of the binomial coefficient and is related to the expansion of $(x_1+x_2+\dots+x_k)^n$.

Sample Problem 22.2.7. Find the number of different permutations of the letters in the word “BEEN”.

Solution. Note that there is 1 ‘B’, 2‘E’s and 1 ‘N’ in “BEEN”. Using the above result, the number of different permutations is given by

$$\frac{4!}{1!2!1!} = 12.$$

□

Proposition 22.2.8 (Circular Permutations)**.** The number of permutations of n distinct objects in a circle is given by $(n-1)!$.

Proof. Fix one object as the reference point. The remaining $n-1$ objects have $(n-1)!$ possible ways to be arranged in the remaining $n-1$ positions around the circle. □

Proposition 22.2.9 (Permutations of Objects Taken from Sets of Distinct Objects with Replacement)**.** The number of permutations of n distinct objects, taken r at a time with replacement, is given by n^r, where $0 \leq r \leq n$.

22.3 Combinations

Definition 22.3.1. A **combination** is a selection of objects from a given set where the order of selection does not matter.

Proposition 22.3.2 (Combinations of Objects Taken from Sets of Distinct Objects)**.** The number of combinations of n distinct objects, taken r at a time without replacement, is given by

$$^nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!},$$

where $0 \leq r \leq n$.

Proof. Observe the number of ways to choose r objects from n distinct objects is equivalent to the number of permutations of n objects, where r objects are of the first type (chosen) while $n-r$ objects are of the second type (not chosen). Using the formula derived above, we have

$$^nC_r = \frac{n!}{r!(n-r)!}.$$

□

Corollary 22.3.3. For integers r and n, where $0 \leq r \leq n$,

$$^nP_r = {}^nC_r \cdot r!.$$

Proof. Rearrange the above result. □

Corollary 22.3.4. For integers r and n, where $0 \leq r \leq n$,

$$^nC_r = {}^nC_{n-r}.$$

Proof. Observe that

$$\frac{n!}{r!(n-r)!}$$

is invariant under $r \mapsto n-r$. □

22.4 Methods for Solving Combinatorics Problems

Some problems involving permutations and combinations may involve restrictions. When dealing with such problems, one should consider the restrictions first. There are four basic strategies that can be employed to tackle these restrictions.

Recipe 22.4.1 (Fixing Positions). When certain objects must be at certain positions, place those objects first.

Sample Problem 22.4.2. How many ways are there to arrange the letters of the word "SOCIETY" if the arrangements start and end with a vowel?

Solution. We first address the restriction by placing the vowels at the start and end of the arrangement. Since there are 3 vowels in "SOCIETY", there are $3 \cdot 2 = 6$ ways to do so. Next, observe there are $5!$ ways to arrange the remaining 5 letters. Thus, by the Multiplication Principle, there are

$$6 \cdot 5! = 720$$

arrangements that satisfy the given restriction. □

Recipe 22.4.3 (Grouping Method). When certain objects must be placed together, group them together as one unit.

Sample Problem 22.4.4. Find the number of ways the letters of the word "COMBINE" can be arranged if all the consonants are to be together.

Solution. Consider the consonants 'C', 'M', 'B' and 'N' as one unit:

$$\boxed{\text{C} \quad \text{M} \quad \text{B} \quad \text{N}} \quad \boxed{\text{O}} \quad \boxed{\text{I}} \quad \boxed{\text{E}}.$$

- **Stage 1.** There are $4!$ ways to arrange the 4 units.
- **Stage 2.** There are $4!$ ways to arrange 'C', 'M', 'B' and 'N' within the group.

Hence, by the Multiplication Principle, the total number of arrangements is

$$4! \cdot 4! = 576.$$

□

Recipe 22.4.5 (Slotting Method). When certain objects are to be separated, we first arrange the other objects to form barriers before slotting in those to be separated.

Sample Problem 22.4.6. Find the number of ways the letters of the word "COMBINE" can be arranged if all the consonants are to be separated.

Solution. We begin by arranging the vowels, of which there are $3!$ ways to do so.

$$\uparrow \quad \boxed{\text{O}} \quad \uparrow \quad \boxed{\text{I}} \quad \uparrow \quad \boxed{\text{E}} \quad \uparrow.$$

Next, we slot the 4 consonants into the 4 gaps in between the vowels (i.e. where the arrows are). There are $4!$ ways to do so. Thus, by the Multiplication Principle, the total number of arrangements is

$$3! \cdot 4! = 144.$$

□

Recipe 22.4.7 (Complementary Method). If the direct method is too tedious, it is more efficient to count by taking all possibilities minus the complementary sets. This method can also be used for "at least/at most" problems.

Sample Problem 22.4.8. Find the number of ways the letters of the word "COMBINE" can be arranged if all the consonants are to be separated.

Solution. Note that, without restrictions, there are a total of 7! ways to arrange the letters in "COMBINE". From the previous example, we saw that the number of arrangements where all consonants are together is 576. Thus, by the complementary method, the number of arrangement where all consonants are separated is

$$\text{total} - \text{complementary} = 7! - 576 = 144,$$

which matches the answer given in the above example. □

23 Distribution Problems

In the previous chapter, we learnt how to count the number of ways to distribute distinct objects into distinct boxes:

Proposition 23.0.1. The number of ways of distributing r distinct objects into n distinct boxes such that each box can hold

- at most one object (assuming $r \leq n$) is ${}^nP_r$;
- any number of objects is n^r.

In this chapter, we focus mainly on counting the number of ways to distribute identical objects into distinct boxes.

23.1 The Bijection Principle

Theorem 23.1.1 (Bijection Principle)**.** Let A and B be finite sets. If there exists a bijection $f : A \to B$, then

$$|A| = |B|.$$

The bijection principle is particularly useful when enumerating A is hard, but enumerating B is easy.

Sample Problem 23.1.2. Determine the number of positive divisors of 12600.

Solution. Observe that $12600 = 2^3 \times 3^2 \times 5^2 \times 7^1$. Let A be the set of divisors of 12600. Let B be the set

$$B = \left\{(p, q, r, s) \in \mathbb{Z}^4 : 0 \leq p \leq 3 \text{ and } 0 \leq q \leq 2 \text{ and } 0 \leq r \leq 2 \text{ and } 0 \leq s \leq 1\right\}.$$

Let $f : B \to A$ be such that

$$f(p, q, r, s) = 2^p \times 3^q \times 5^r \times 7^s.$$

It is clear that f is bijective: by the Fundamental Theorem of Algebra, every divisor $d \in A$ is uniquely expressible as a product of prime powers of 2, 3, 5 and 7. Hence, by the bijective principle, we have

$$|A| = |B| = (3 + 1)(2 + 1)(2 + 1)(1 + 1) = 72,$$

i.e. 12600 has 72 divisors. □

One can easily generalize the above result:

Proposition 23.1.3. Let

$$n = \prod_{i=1}^{k} p_i^{e_i}$$

where p_i are distinct primes and e_i are non-negative integers. Then n has

$$\prod_{i=1}^{k} (e_i + 1)$$

positive divisors.

23.2 Identical Objects into Distinct Boxes

We first prove a standard result:

Proposition 23.2.1 (Stars and Bars)**.** The number of non-negative integer solutions to the equation $x_1 + \cdots + x_n = r$ is

$$\binom{r+n-1}{n-1} = \binom{r+n-1}{r}.$$

Proof. Let

$$A = \{(x_1, \ldots, x_n) \in \mathbb{N}_0 : x_1 + \cdots + x_n = r\}$$

be the set of all non-negative integer solutions to the above equation. Consider a row of $r+n-1$ objects. Let B be the set of all possible ways to colour $n-1$ of these $r+n-1$ objects red, and the remaining r objects blue. It is easy to see that

$$|B| = \binom{r+n-1}{n-1} = \binom{r+n-1}{r}.$$

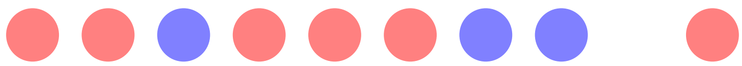


Figure 23.1: An example colouring, where $r = 2+3+1 = 6$ and $n = 4$.

We now establish a bijection between A and B. Consider the following procedure, starting with a solution $(x_1, \ldots, x_n) \in A$:

- Colour the first x_1 balls blue, and the next ball red.
- Colour the next x_2 balls blue, and the next ball red.

  $\vdots$

- Colour the next x_n balls blue.

It is easy to see that all $r+n-1$ balls will be coloured, and exactly $n-1$ balls will be red. Further, each solution $(x_1, \ldots, x_n) \in A$ uniquely determines a colouring in B and vice versa, i.e. the procedure is a bijection between A and B. By the bijection principle,

$$|A| = |B| = \binom{r+n-1}{n-1} = \binom{r+n-1}{r}.$$

□

The method of counting is commonly known as "stars and bars". We can think of the blue objects as "stars" (the objects we wish to distribute), and the red objects as "bars" (the dividers separating the objects).

Proposition 23.2.2 (Identical Objects into Distinct Boxes (Part I))**.** The number of ways of distributing r identical objects into n distinct boxes is given by

$$\binom{r+n-1}{n-1} = \binom{r+n-1}{r}.$$

Proof. Let x_i be the number of objects in the ith box. Since we have a total of r identical objects, we require

$$x_1 + x_2 + \cdots + x_n = r.$$

By stars and bars, we attain our desired result. □

Proposition 23.2.3 (Identical Objects into Distinct Boxes (Part II)). The number of ways of distributing r identical objects into n distinct boxes, such that each box has at least k objects, is given by

$$\binom{r-nk+n-1}{n-1}$$

Proof. Let $x_i + k$ be the number of objects in the ith box. Since each box has at least k objects, we have $x_i \leq 0$ for all $1 \leq i \leq n$. Since we have a total of r identical objects, we require

$$(x_1 + k) + (x_2 + k) + \cdots + (x_n + k) = r.$$

This equation simplifies to

$$x_1 + x_2 + \cdots + x_n = r - nk.$$

We hence seek the number of non-negative integer solutions to the above equation, which we know to be

$$\binom{r-nk+n-1}{n-1}$$

by stars and bars. □

Corollary 23.2.4. In the case where we require each box to be non-empty ($k = 1$), the number of distributions is given by

$$\binom{r-1}{n-1} = \binom{r-1}{r-n}.$$

23.3 Distinct Objects into Identical Boxes

Definition 23.3.1. A **Stirling number of the second kind** is defined to be the number of ways of distributing r distinct objects into n identical boxes such that no box is empty. It is denoted $S(r, n)$.

Proposition 23.3.2. For $0 < n < r$, we have the recurrence relation

$$S(r+1, n) = S(r, n-1) + nS(r, n),$$

with initial conditions $S(r, r) = 1$ for $r \geq 0$ and $S(r, 0) = S(0, r) = 0$ for $r > 0$.

Proof. Let A be an arbitrary object.

Case 1: *A is alone in a box.* There remains r distinct objects to be distributed into $n - 1$ identical boxes with no empty boxes. The number of ways to do so is $S(r, n-1)$.

Case 2: *A is not alone in a box.* We first distribute the other r distinct objects into n identical boxes such that no box is empty. This can be done in $S(r, n)$ ways. Then, we place A into one box. There are n boxes, thus by the multiplicative principle, the total number of ways in this case is $nS(r, n)$.

Altogether, the total number of ways to distribute $r+1$ distinct objects into n identical boxes such that no box is empty is given by

$$S(r+1, n) = S(r, n-1) + nS(r, n).$$

The initial conditions can easily be verified. □

Sample Problem 23.3.3. Find the number of ways to express 2730 as a product ab of two integers a and b, where $2 \geq a \geq b$.

Solution. Note that $2730 = 2 \times 3 \times 5 \times 7 \times 13$. The number of ways to express 2730 as a product ab is hence given by $S(5, 2) = 15$, as we have 5 distinct prime factors and 2 identical boxes (a and b). □

Proposition 23.3.4. The number of ways to distribute r distinct objects into n identical boxes with empty boxes allowed is given by

$$\sum_{k=1}^{n} S(r, k).$$

Proof. Suppose only k boxes are filled. There are $S(r, k)$ ways to distribute the objects into these k boxes. Enumerating over all possible cases, we see that the total possible distributions number

$$\sum_{k=1}^{n} S(r, k).$$

□

23.4 Identical Objects into Identical Boxes

Definition 23.4.1. The **partition** of a positive integer r into n parts is a set of n positive integers whose sum is r. We denote the number of different partitions of r into n parts with $P(r, n)$.

Proposition 23.4.2. We have the recurrence relation

$$P(r, n) = P(r - 1, n - 1) + P(r - n, n),$$

with conditions $P(r, 1) = 1$ for all $r \geq 1$, and $P(r, n) = 0$ if $n > r$.

Proof. Case 1: At least one box has exactly one object. We place one object in one box. We then distribute the remaining $r - 1$ objects into the remaining $n - 1$ boxes such that no boxes are empty. The number of ways this can be done is $P(r - 1, n - 1)$.

Case 2: All the boxes have more than one object. We place one object into each of the n boxes. We then distribute the remaining $r - n$ objects into the n boxes so that no boxes are empty. The number of ways this can be done is $P(r - n, n)$.

Altogether, we have

$$P(r, n) = P(r - 1, n - 1) + P(r - n, n)$$

as desired. □

24 Principle of Inclusion and Exclusion

Theorem 24.0.1 (Principle of Inclusion and Exclusion). Let A_1, A_2, …, A_n be finite sets. Then

$$\left|\bigcup_{k=1}^{n} A_k\right| = \sum_{\substack{I\subseteq[n]\\ I\neq\varnothing}} (-1)^{|I|+1} \left|\bigcap_{i\in I} A_i\right|.$$

Proof. Let $A = \bigcup_{k=1}^{n} A_k$ be the union of all n sets. Define the indicator function of a set A_i to be $\mathbf{1}_i : A \to \{0, 1\}$ such that

$$\mathbf{1}_i(x) = \begin{cases} 1, & x \in A_i, \\ 0, & x \notin A_i. \end{cases}$$

Consider now the function

$$F(x) = \prod_{i=1}^{n} \left[1 - \mathbf{1}_i(x)\right].$$

Observe that for all $x \in A$, we must have $x \in A_i$ for some $1 \le i \le n$, thus $F(x)$ is identically zero. We now expand $F(x)$:

$$F(x) = 1 + \sum_{\substack{I\subseteq[n]\\ I\neq\varnothing}} (-1)^{|I|} \prod_{i\in I} \mathbf{1}_i(x).$$

It is not too hard to see that $\prod_{i\in I} \mathbf{1}_i(x)$ is the indicator function of $\bigcap_{i\in I} A_i$. Summing over all $x \in A$, we hence obtain

$$\begin{aligned}
\sum_{x\in A} F(x) &= \sum_{x\in A} \left[1 + \sum_{\substack{I\subseteq[n]\\ I\neq\varnothing}} (-1)^{|I|} \prod_{i\in I} \mathbf{1}_i(x)\right] \\
&= |A| + \sum_{\substack{I\subseteq[n]\\ I\neq\varnothing}} (-1)^{|I|} \left(\sum_{x\in A} \prod_{i\in I} \mathbf{1}_i(x)\right) \\
&= \left|\bigcup_{k=1}^{n} A_k\right| + \sum_{I\subseteq[n]} (-1)^{|I|} \left|\bigcap_{i\in I} A_i\right|.
\end{aligned}$$

Since $F(x)$ is identically zero, we immediately obtain the desired result:

$$\left|\bigcup_{k=1}^{n} A_k\right| = \sum_{\substack{I\subseteq[n]\\ I\neq\varnothing}} (-1)^{|I|+1} \left|\bigcap_{i\in I} A_i\right|.$$

$\square$

A classic application of the Principle of Inclusion and Exclusion is counting the number of surjections between two finite sets.

Proposition 24.0.2. Let X and Y be finite sets with cardinality $|X| = m$ and $|Y| = n$, where $m \geq n$. Then the number of surjections from X to Y is given by

$$\sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n-k)^m.$$

Proof. For convenience, we number the elements of X and Y such that $X = [m]$ and $Y = [n]$. Let S be the set of mappings from X to Y, and A_i be the set of mappings from X to $Y \setminus \{i\}$, where $1 \leq i \leq n$. We see that for an arbitrary non-empty set of indices $I \subseteq [n]$ of size k,

$$\left| \bigcap_{i \in I} A_i \right| = \#\text{ (mappings from } m \text{ elements to } n-k \text{ elements)} = (n-k)^m.$$

Since there are $\binom{n}{k}$ possible sets of indices of size k, by the Principle of Inclusion and Exclusion,

$$\begin{aligned} \left| \bigcup_{k=1}^{n} A_k \right| &= \sum_{\substack{I \subseteq [n] \\ I \neq \varnothing}} (-1)^{|I|+1} \left| \bigcap_{i \in I} A_i \right|. \\ &= \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} (n-k)^m. \end{aligned}$$

This counts the number of mappings that are not surjective. For the number of mappings that are surjective, we simply take

$$\begin{aligned} |S| - \left| \bigcup_{k=1}^{n} A_k \right| &= n^m - \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} (n-k)^m \\ &= \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n-k)^m. \end{aligned}$$

□

Corollary 24.0.3. The Stirling numbers of the second kind are given by

$$S(m,n) = \frac{1}{n!} \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n-k)^m.$$

Proof. There are $S(m,n)$ ways to partition $[m]$ into n non-empty subsets. The number of ways to assign these n parts to a distinct value in $[n]$ is $n!$. Thus, the number of surjective functions from $[m]$ to $[n]$ is $n!S(m,n)$. Using the above result, we obtain

$$S(m,n) = \frac{1}{n!} \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n-k)^m.$$

□

Yet another famous application of the Principle of Inclusion and Exclusion is counting the number of derangements.

Definition 24.0.4. A **derangement** is a permutation $\pi : [n] \to [n]$ with no fixes point, i.e. for all $1 \leq i \leq n$, we have $\pi(i) \neq i$.

Proposition 24.0.5. The number of derangements $\pi : [n] \to [n]$ is given by

$$\sum_{k=0}^{n} (-1)^k \frac{n!}{k!}.$$

Proof. Let S be the set of all permutations of $[n]$, and let A_i be the set of all permutations that fix i. Note that $|S| = n!$, and for an arbitrary non-empty set of indices $I \subseteq [n]$ of size k,

$$\left| \bigcap_{i \in I} A_i \right| = \#(\text{permutations of } n-k \text{ elements}) = (n-k)!.$$

Since there are $\binom{n}{k}$ possible sets of indices of size k, by the Principle of Inclusion and Exclusion,

$$\begin{aligned} \left| \bigcup_{k=1}^{n} A_k \right| &= \sum_{\substack{I \subseteq [n] \\ I \neq \varnothing}} (-1)^{|I|+1} \left| \bigcap_{i \in I} A_i \right| \\ &= \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} (n-k)! \\ &= \sum_{k=1}^{n} (-1)^{k+1} \frac{n!}{k!}. \end{aligned}$$

This counts the number of permutations with fixed points. For the number of derangements, we simply take

$$\begin{aligned} |S| - \left| \bigcup_{k=1}^{n} A_k \right| &= n! - \sum_{k=1}^{n} (-1)^{k+1} \frac{n!}{k!} \\ &= \sum_{k=0}^{n} (-1)^k \frac{n!}{k!}. \end{aligned}$$

□

25 Probability

25.1 Basic Terminology

Definition 25.1.1. A statistical or random **experiment** (or trial) refers to a process that generates a set of observable outcomes, and can be repeated under the same set of conditions.

Definition 25.1.2. The **sample space** (or possibility space) S of an experiment is the set of all possible outcomes of the experiment.

Definition 25.1.3. An **event** E is a subset of S. The **complement** of E, denoted by E', is the event that E does not occur, i.e. $E' = S \setminus E$.

Definition 25.1.4. Given a subset $G \subseteq S$, the function $n(G)$ returns the **number of possible outcomes** in G.

25.2 Probability

Definition 25.2.1 (Classical Probability)**.** If the sample space S consists of a finite number of equally likely outcomes, then the probability of an event E occurring (a measure of the likelihood that E occurs) is denoted $\mathbb{P}[E]$ and is defined as

$$\mathbb{P}[E] = \frac{n(E)}{n(S)}.$$

Proposition 25.2.2 (Range of Probabilities)**.** For any event E,

$$\mathbb{P}[E] \in [0, 1].$$

Proof. Let the sample space be S. Since $E \subseteq S$, we have

$$0 \leq n(E) \leq n(S) \implies 0 \leq \frac{n(E)}{n(S)} \leq \frac{n(S)}{n(S)} \implies 0 \leq \mathbb{P}[E] \leq 1.$$

□

Corollary 25.2.3. Let A and B be any two events. If $A \subseteq B$, then $\mathbb{P}[A] \leq \mathbb{P}[B]$.

Proof. Identical as above. □

Definition 25.2.4. When $\mathbb{P}[E] = 0$, we say that E is an **impossible** event. When $\mathbb{P}[E] = 1$, we say that P is a **sure** event.

Proposition 25.2.5 (Probability of Complement)**.** For any event E,

$$\mathbb{P}[E] + \mathbb{P}\left[E'\right] = 1.$$

Proof. Let the sample space be S. By definition, $E' = S \setminus E$. Hence,

$$n(E') = n(S) - n(E) \implies \frac{n(E)}{n(S)} + \frac{n(E')}{n(S)} = \frac{n(S)}{n(S)} \implies \mathbb{P}[E] + \mathbb{P}\left[E'\right] = 1.$$

□

Definition 25.2.6. Let S be the sample space of a random experiment and A, B be any two events.

- The **intersection** of A and B, denoted by $A \cap B$, is the event that both A and B occur.
- The **union** of A and B, denoted by $A \cup B$, is the event that at least one occurs.

Proposition 25.2.7 (Inclusion-Exclusion Principle)**.** Let A and B be any two events in a sample space S. Then

$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B].$$

Proof. When we take the sum of the number of outcomes in events A and B, i.e. $n(A) + n(B)$, we will count the 'overlap', i.e. $n(A \cap B)$, twice. Hence,

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

Dividing throughout by $n(S)$ yields the desired result. □

Proposition 25.2.8 (Intersection of Complements)**.** Let A and B be any two events. Then

$$\mathbb{P}[A] = \mathbb{P}[A \cap B] + \mathbb{P}\left[A \cap B'\right].$$

Proof. By definition, $B' = S \setminus B$. Taking the intersection with A on both sides,

$$\mathbb{P}\left[A \cap B'\right] = \mathbb{P}[A \cap S] - \mathbb{P}[A \cap B] \implies \mathbb{P}[A \cap B] + \mathbb{P}\left[A \cap B'\right] = \mathbb{P}[A].$$

□

Proposition 25.2.9 ("Neither Nor")**.** Let A and B be any two events. Then

$$\mathbb{P}\left[A' \cap B'\right] = 1 - \mathbb{P}[A \cup B].$$

Proof. In layman terms, the above statement translates to

$$\mathbb{P}[\text{neither } A \text{ nor } B] = 1 - \mathbb{P}[A \text{ or } B],$$

which is clearly true. □

25.3 Mutually Exclusive Events

Definition 25.3.1. Two events A and B are said to be **mutually exclusive** if they cannot occur at the same time. Mathematically,

$$\mathbb{P}[A \cap B] = 0.$$

An equivalent criterion for mutual exclusivity is

$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]\,,$$

which can easily be derived from $\mathbb{P}[A \cap B] = 0$ via the inclusion-exclusion principle.

25.4 Conditional Probability and Independent Events

Proposition 25.4.1 (Conditional Probability)**.** The probability of an event A occurring, given that another event B has already occurred, is given by

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$

Proof. Since B has already occurred, the sample space is reduced to B. Hence,

$$\mathbb{P}[A \mid B] = \frac{n(A \cap B)}{n(B)}.$$

Dividing the numerator and denominator by $n(S)$ completes the proof. □

Corollary 25.4.2. The event (A, given B) is the complement of the event (not A, given B), i.e.

$$\mathbb{P}[A \mid B] + \mathbb{P}\big[A' \mid B\big] = 1.$$

Proof.

$$\mathbb{P}[A \mid B] + \mathbb{P}\big[A' \mid B\big] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} + \frac{\mathbb{P}[A' \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[B]}{\mathbb{P}[B]} = 1.$$

□

Definition 25.4.3 (Independent Events)**.** Let A and B be any two events. If either of the two occur without being affected by the other, then A and B are said to be **independent**. Mathematically,

$$\mathbb{P}[A \mid B] = \mathbb{P}[A]\,, \qquad \mathbb{P}[B \mid A] = \mathbb{P}[B]\,.$$

Proposition 25.4.4 (Multiplication Law)**.** A and B are independent events if and only if

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]\,.$$

Proof. Since $\mathbb{P}[A] = \mathbb{P}[A \cap B] \,/\, \mathbb{P}[B]$ and $\mathbb{P}[A \mid B] = \mathbb{P}[A]$,

$$\frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \mathbb{P}[A] \iff \mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]\,.$$

□

Proposition 25.4.5. If events A and B are independent, then so are the following pairs of events:

- A and B',
- A' and B,
- A' and B'.

Proof. We only prove that A' and B are independent. The proofs for the other pairs are almost identical.

Since A and B are independent events, we have $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]$. Now consider $\mathbb{P}[A' \cap B]$.

$$\mathbb{P}\left[A' \cap B\right] = \mathbb{P}[B] - \mathbb{P}[A \cap B] = \mathbb{P}[B] - \mathbb{P}[A]\,\mathbb{P}[B] = \mathbb{P}[B]\,[1 - \mathbb{P}[A]] = \mathbb{P}[B]\,\mathbb{P}\left[A'\right].$$

Hence, A' and B are independent. □

25.5 Common Heuristics used in Solving Probability Problems

Recipe 25.5.1 (Table of Outcomes)**.** Table of outcomes are useful as they serve as a systematic way of listing all the possible outcomes.

Sample Problem 25.5.2. Two fair dices are thrown. Find the probability that the sum of the two scores is odd and at least one of the two scores is greater than 4.

Solution. Consider the following table of outcomes.

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

From the table of outcomes, the required probability is clearly $\frac{10}{36}$. □

Recipe 25.5.3 (Venn Diagrams)**.** Venn diagrams are useful when we need to visualize how the events are interacting with each other.

Sample Problem 25.5.4. Let A and B be independent events. If $\mathbb{P}[A' \cap B'] = 0.4$, find the range of $\mathbb{P}[A \cap B]$.

Solution. Consider the following Venn diagram.

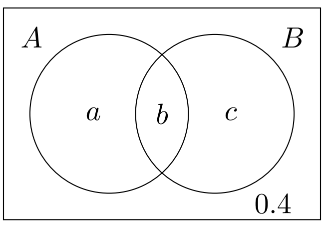


Figure 25.1

We see that

$$a + b + c = 0.6. \tag{$*$}$$

Further, since A and B are independent, we know

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B] \implies b = (a+b)(c+b) = (a+b)(0.6-a).$$

Expanding, we get a quadratic in a:

$$a^2 + (b-0.6)a + 0.4b = 0.$$

Since we want a to be real, the discriminant Δ is non-negative. Hence,

$$(b-0.6)^2 - 4(1)(0.4b) \geq 0 \implies b \leq 0.135 \quad \text{or} \quad b \geq 2.66.$$

Since $0 \leq b \leq 1$, we reject the latter. Thus, the range of $\mathbb{P}[A \cap B] = b$ is $[0, 0.135]$. □

Recipe 25.5.5 (Probability Trees). A probability tree is a useful tool for sequential events, or events that appear in stages. The number indicated on each branch represents the conditional probability of the event at the end node given that all the events at the previous nodes have occurred.

Sample Problem 25.5.6. Peter has a bag containing 6 black marbles and 3 white marbles. He takes out two marbles at random from the bag. Find the probability that he has taken out a black marble and a white marble.

Solution. Consider the following probability tree.

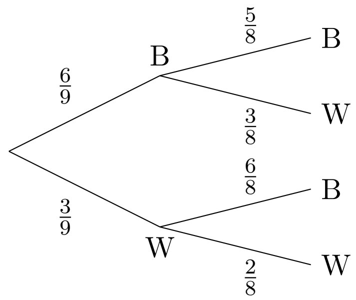


Figure 25.2

The required probability is thus

$$\left(\frac{6}{9}\right)\left(\frac{3}{8}\right) + \left(\frac{3}{9}\right)\left(\frac{6}{8}\right) = \frac{1}{2}.$$

□

Recipe 25.5.7 (Permutations and Combinations). Using combinatorial methods is useful when the most direct way to calculate $\mathbb{P}[E]$ is to find $n(E)$ and $n(S)$.

Sample Problem 25.5.8. A choir has 7 sopranos, 6 altos, 3 tenors and 4 basses. At a particular rehearsal, three members of the choir are chosen at random. Find the probability that exactly one bass is chosen.

Solution. Note that there are a total of 20 people in the choir. Hence, the number of ways to choose three members of the choir, without restriction, is given by ${}^{20}C_3$. Meanwhile, the number of ways to choose exactly one bass is given by ${}^{4}C_1 \cdot {}^{16}C_2$: first choose one bass out of the four, then choose 2 members out of the remaining 16. Thus, the required probability is

$$\frac{{}^{4}C_1 \cdot {}^{16}C_2}{{}^{20}C_3} = \frac{8}{19}.$$

□

Part VII

Statistics

26 Introduction to Statistics

Statistics is the art of learning from data. It is concerned with the collection of data, its subsequent description, and its analysis, which often leads to the drawing of conclusions.

Unlike other real-life problems that can be modelled with maths, the "answers" provided by statistics are never exact; there is always error. However, statistics allows us to *control* this error. Indeed, it is this precise control of statistical error that is at the heart of every statistical technique.

26.1 Samples and Populations

Definition 26.1.1. A **population** (or universe) is all possible subjects that meet certain criteria. It is the entire group of subjects that we are interested in studying.

We want to know something about a population, but there is a good chance that we can never get a very accurate picture of the population simply because it is constantly changing. Not only are populations often in a constant state of flux, practically speaking, we cannot always have access to an entire population for study. Time and cost often get in the way. As a result, we turn to a sample as a substitute of the entire population.

Definition 26.1.2. A **sample** is a subset of the population. A **random sample** is a sample that is representative of the population.

Example 26.1.3. If we were interested in the weight of all 12-year-old kids on Earth, then all the kids who meet the criteria (i.e. 12-year-old kids on Earth) would constitute the population.

However, realistically speaking, there is no way we can accurately weigh all 12-year-old kids on Earth. Instead, we could weigh a sample of 500 12-year-old kids from all around the globe, which would be representative of the population.

26.2 Two Categories of Statistics

Broadly speaking, the usage of statistics can be split into two categories: descriptive and inferential.

26.2.1 Descriptive Statistics

Descriptive statistics are used to summarize or describe data from samples and populations.

Suppose we are interested in the test results of a class of students. We could create a data distribution by listing the test scores of all students in the class and looking at it with the idea of getting some intuitive picture of how they are doing. Alternatively, we could simply calculate the mean of the students' test scores. The calculation of the mean represents the use of descriptive statistics, allowing us to summarize or describe our data.

26.2.2 Inferential Statistics

Using descriptive statistics, we can calculate the characteristics of a data set, e.g. mean, mode, etc. If this data set was collected from the entire population, we call such a characteristic a **parameter** of the population. This could be "mean test score of a cohort of students". However, if the data set was collected from a sample (i.e. not the entire population), we call the characteristic a **statistic**. This could be "mean test score of a class".

Because we are often not directly able to obtain a population parameter, we have to rely on sample data to make inferences about the population. This branch of statistics is known as inferential statistics – using sample statistics to make inferences about population parameters.

26.3 Measures of Central Tendency

A **central tendency** can be thought of as the "typical" value of a data set. There are three main measures of central tendency, namely the mean, median and mode.

26.3.1 Mean

Definition 26.3.1. The **mean** is the sum of all observations, divided by the total number of observations.

Mathematically, given n observations $x_1, x_2, x_3, \ldots, x_n$,

$$\text{Mean} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Here, a lower-case 'n' represents the sample size. We use the uppercase 'N' to represent the population size. It is essential to make it clear when we are referring to the mean of a sample or when we are referring to the mean of a population. To do so, statisticians use different symbols ($\bar{x}$ and μ):

$$\text{Sample mean} = \bar{x} = \frac{1}{n}\sum x, \quad \text{Population mean} = \mu = \frac{1}{N}\sum x.$$

We can also calculate the mean of a data set from its frequency table:

$$\text{Mean} = \frac{\sum xf}{\sum f},$$

where f represents the frequency of a value x.

Example 26.3.2. Suppose the test scores of students in a particular class has the following frequency table:

Test score, x	Frequency, f
12	2
13	3
15	6
16	5
17	4

Then, the mean test score can be calculated as

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{(12)(2) + (13)(3) + (15)(6) + (16)(5) + (17)(4)}{2 + 3 + 6 + 5 + 4} = 15.05.$$

Since the mean takes into account the entire sample data, it is very sensitive to outliers. Hence, the mean may be insufficient for data sets with outliers.

Example 26.3.3. Suppose now that another student in the class obtained a '1' on the test. The new mean can be calculated as

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{(1)(1) + (12)(2) + (13)(3) + (15)(6) + (16)(5) + (17)(4)}{1 + 2 + 3 + 6 + 5 + 4} = 14.14,$$

which is much less than the previous mean of 15.05.

26.3.2 Median

Definition 26.3.4. The **median** is the point in a distribution that divides the distribution into halves, i.e. the midpoint of a distribution.

Generally, for n values $x_1, x_2, \ldots, x_n$ arranged in ascending order,

$$\text{Median} = \begin{cases} x_{(n+1)/2}, & n \text{ odd,} \\ \frac{1}{2}\left(x_{n/2} + x_{n/2+1}\right), & n \text{ even} \end{cases}$$

Example 26.3.5. For the original data of 20 students, the set of data in ascending order is

$$12, 12, 13, 13, 13, 15, 15, 15, 15, 15, 15, 16, 16, 16, 16, 16, 17, 17, 17, 17.$$

The median is hence the average of the two middle values, i.e. $\frac{1}{2}(15 + 15) = 15$.

Unlike the mean, the median is not sensitive to outliers.

Example 26.3.6. For the data of 21 students (original 20 + one outlier), the set of data in ascending order is

$$1, 12, 12, 13, 13, 13, 15, 15, 15, 15, 15, 15, 16, 16, 16, 16, 16, 17, 17, 17, 17.$$

The median is hence the 11th value, 15.

26.3.3 Mode

Definition 26.3.7. The **mode** is the value that occurs the most frequently in a distribution.

In the previous examples, the mode for the original sample of 20 and the new sample of 21 are both 15.

A distribution containing the values 2, 3, 6, 1, 3, 7 and 7 would be referred to as a **bimodal distribution** because it has two modes – 3 and 7. A distribution with a single mode is called **unimodal**. If each value appears the same number of times, the distribution has no mode.

The mode, unlike the mean, is not affected by outliers. It is easy to state as it does not require any calculation. However, it is a crude measure of central tendency as it ignores a substantial part of the data and is thus usually not very representative and useful.

26.3.4 Bonus: Relationship with L^p-norms

So far, we have motivated the introduction and use of the mean, median and mode to counter the shortcomings of the other measures. While this is sufficient for understanding why (and when) we should care about certain measures of central tendency, there is a more fundamental property that these three measures have in common.

Recall that we introduced a *central tendency* as the "typical" value of a data set. Intuitively, a measure of central tendency minimizes the total "distance" between any data point and itself. One method to measure this "distance" is the L^p-norm.

Definition 26.3.8. Let $p \geq 1$. The **L^p-norm** of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, denoted $\|\mathbf{x}\|_p$, is defined as

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

Example 26.3.9. When $p = 2$, we recover the Euclidean norm:

$$\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

In our case, we can take x_i to be the values of our data set. Now, consider an n-dimensional vector $\mathbf{c} = (c, c, \ldots, c)$. Then $\|\mathbf{x} - \mathbf{c}\|_p$ measures the total "distance" between c and any data point. Thus, the value of c that minimizes $\|\mathbf{x} - \mathbf{c}\|_p$ will be a measure of central tendency.

We now show that the mean, median and mode correspond to the cases where $p = 2$, 1 and 0 respectively.

Proposition 26.3.10. The mean minimizes $\|\mathbf{x} - \mathbf{c}\|_2$.

Proof. By definition,

$$\|\mathbf{x} - \mathbf{c}\|_2 = \left(\sum_{i=1}^{n} (x_i - c)^2 \right)^{1/2}.$$

Differentiating this with respect to c,

$$\frac{\mathrm{d}}{\mathrm{d}c} \|\mathbf{x} - \mathbf{c}\|_2 = -\left(\sum_{i=1}^{n} (x_i - c)^2 \right)^{-1/2} \sum_{i=1}^{n} (x_i - c).$$

For stationary points, we want $\frac{\mathrm{d}}{\mathrm{d}c} \|\mathbf{x} - \mathbf{c}\|_2 = 0$. Hence,

$$\sum_{i=1}^{n} (x_i - c) = 0 \implies \sum_{i=1}^{n} x_i - cn = 0 \implies c = \frac{1}{n} \sum_{i=1}^{n} x_i,$$

which is exactly the definition of the mean. It is an exercise for the reader to show that this stationary point is a minimum. □

Proposition 26.3.11. The median minimizes $\|\mathbf{x} - \mathbf{c}\|_1$.

Proof. By definition,

$$\|\mathbf{x} - \mathbf{c}\|_1 = \sum_{i=1}^{n} |x_i - c|.$$

Without loss of generality, suppose $x_1 \leq x_2 \leq \cdots \leq x_n$. For $\|\mathbf{x}-\mathbf{c}\|_1$ to be minimized, there must exist a $k \geq 1$ such that $x_k \leq c$ for all $i \leq k$ and $x_k \geq c$ for all $i > k$. Then

$$\|\mathbf{x}-\mathbf{c}\|_1 = \sum_{i=1}^{k}(c - x_i) + \sum_{i=k+1}^{n}(x_i - c).$$

Differentiating this with respect to c,

$$\frac{\mathrm{d}}{\mathrm{d}c}\|\mathbf{x}-\mathbf{c}\|_1 = 2k - n.$$

Setting this equal to 0 yields $k = n/2$. That is, half of the data values are less than c, while the other half are greater than c. Thus, c is the median. □

Proposition 26.3.12. The mode minimizes $\|\mathbf{x}-\mathbf{c}\|_0$.

Proof. While the L^p norm is not defined for $p = 0$, we can take the appropriate limit to get

$$\|\mathbf{x}-\mathbf{c}\|_0 = \lim_{p\to 0}\left(\sum_{i=1}^{n}|x_i - c|^p\right)^{1/p} = \sum_{i=1}^{n}|x_i - c|^0,$$

where we take $0^0 = 0$. Clearly, to minimize $\|\mathbf{x}-\mathbf{c}\|_0$, we must have $c = x_i$ for as many i possible. It follows that c must be the mode. □

26.4 Measures of Spread

Suppose that the original 20 test scores come from students from a particular class, and that there is another class of 20 whose test score has the following frequency distribution table:

Test score, x	Frequency, f
9	2
10	2
13	4
15	2
16	2
17	3
18	2
20	1
21	2

The mean test score of both classes are the same (15.05). However, the second class clearly has a wider spread of test scores.

Measures of central tendencies do not give any indication of these differences in spread, so it is necessary to devise some other measures to summarize the spread of data.

26.4.1 Range and Interquartile Range

Definition 26.4.1. The **range** is the difference between the maximum and minimum values in the set of data.

Example 26.4.2. The first class has a range of $17 - 12 = 5$, while the second class has a range of $21 - 9 = 12$. Hence, the test scores for the second class are more diverse as compared to that for the first class.

Note, however, that the range is usually not a good measure of dispersion as it only considers the extreme values which may be atypical of the rest of the distribution and does not give any information about the distribution of the values in between. For instance, if we include the outlier in the first class, the range becomes $17 - 1 = 16$.

For this reason, we typically consider the interquartile range instead.

Definition 26.4.3. The **interquartile range** is the difference between the first and third quartiles, i.e. $Q_3 - Q_1$.

Recall that the nth percentile of a distribution is the value such that $n\%$ of the data is less than or equal to that number. The first and third quartiles are hence the 25th and 75th percentile respectively. Note that the second quartile (50th percentile) is simply the median.

Example 26.4.4. The first class has interquartile range $16 - 14 = 2$, while the second class has interquartile range $17.5 - 13 = 4.5$.

If we include the outlier in the first class, then the interquartile range becomes $16 - 13 = 3$, which is a much smaller change compared to that of the range.

Again, the interquartile range may not be a good measure of dispersion as it only takes into account the two specific percentiles.

26.4.2 Variance and Standard Deviation

One of the main reasons for using the interquartile range in preference to the range as a measure of spread is that it takes some account of how the interior values are spread rather than concentrating on the spread of the extreme values. The interquartile range, however, does not take into account of the spread of all the data values and so, in some sense, it is still an inadequate measure. An alternative measure of spread, which takes into account of all the values, can be devised by finding how far each data value is from the mean.

This can be represented mathematically with the formula

$$\text{Mean distance} = \frac{1}{n} \sum |x - \bar{x}| .$$

Unfortunately, a formula involving the modulus sign is awkward to handle algebraically. This can be avoided by squaring each of the quantities $x - \bar{x}$, leading to the expression

$$\frac{1}{n} \sum (x - \bar{x})^2$$

as a measure of spread. We call this quantity the **variance** of the distribution.

If the data values $x_1, \ldots, x_n$ have units associated with them, then the variance will be measured in units^2. This can be avoided by taking the positive square root of the variance. The positive square root of the variance is known as the **standard deviation**, and it always has the same units as the original data values, i.e.

$$\text{Standard deviation} = \sqrt{\frac{1}{n} \sum (x - \bar{x})^2}.$$

When referring to the standard deviation of the population, we use the symbol σ. Hence, the population variance is denoted by σ^2.

In its given form, the variance of a data set is tedious to calculate. Fortunately, an alternative formula is easier to use is available:

Proposition 26.4.5.

$$\text{Variance} = \frac{1}{n}\sum x^2 - \bar{x}^2.$$

Proof. We have

$$\text{Variance} = \frac{1}{n}\sum (x-\bar{x})^2 = \frac{1}{n}\sum \left(x^2 - 2x\bar{x} + \bar{x}^2\right) = \frac{1}{n}\sum x^2 - \frac{2\bar{x}\sum x}{n} + \frac{\bar{x}^2\sum 1}{n}.$$

Observe that $\frac{1}{n}\sum x = \bar{x}$ and $\sum 1 = n$. Thus,

$$\text{Variance} = \frac{1}{n}\sum x^2 - 2\bar{x}^2 + \bar{x}^2 = \frac{1}{n}\sum x^2 - \bar{x}^2.$$

$\square$

27 Discrete Random Variables

27.1 Random Variables

Definition 27.1.1. A **random variable** is a variable whose possible values are numerical outcomes of a random experiment.

Random variables are typically denoted by capital letters such as X or Y.
There are two types of random variables: discrete and continuous.

Definition 27.1.2. A **discrete random variable** is a random variable that assumes countable values $x_1, x_2, \ldots, x_n$ (can be infinite).

Examples of discrete random variables include the number that shows on the toss of a fair die ($X = 1, 2, \ldots, 6$), and the number of times a fair die is thrown until a '6' is obtained ($Y = 1, 2, \ldots,$ to infinity).

In this chapter, we will only discuss discrete random variables. We will deal more with continuous random variables in §28.

27.2 Properties

27.2.1 Probability Distribution

Since the values of a random variable are determined by chance, there is a distribution associated with them. We call this a probability distribution.

Definition 27.2.1. A **probability distribution** describes all possible values of the random variable and their corresponding probabilities. It assigns a probability value to each possible outcome in the sample space.

A probability distribution of a discrete random variable can be given in the form of a table, a graph or a mathematical formula.

Note that the particular values of a random variable are denoted by lower-case letters. For instance, the particular values of a random variable X are denoted by x.

Example 27.2.2. A single fair 6-sided die is thrown. Let X be the random variable representing the number of dots showing on the die. Note that the possible values of X are $x = 1, 2, 3, 4, 5, 6$.

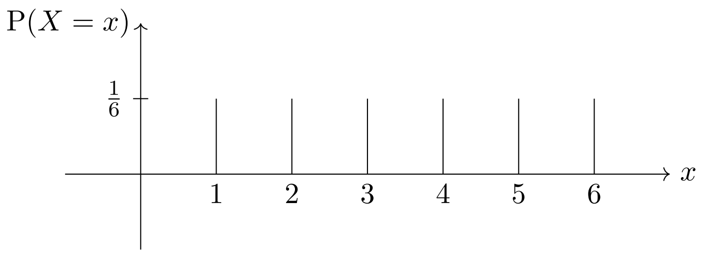
The probability distribution associated with X can be given in table form:

x	1	2	3	4	5	6
$\mathbb{P}[X = x]$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

or expressed as a formula:

$$\mathbb{P}[X = x] = \frac{1}{6}, \quad x \in \{1, 2, 3, 4, 5, 6\},$$

or expressed as a graph:



From the above example, the discrete random variable X takes on only countable values, and that if we sum all probabilities, we get a total of 1. In fact, these are conditions that all discrete random variables must satisfy.

Condition 27.2.3 (Discrete Random Variable). For X to be a discrete random variable,

- X can take only countable values (finite or infinitely many), and
- X has a probability distribution such that $0 \leq \mathbb{P}[X = x] \leq 1$ for all x and

$$\sum_x \mathbb{P}[X = x] = 1.$$

27.2.2 Expectation

Recall that in descriptive statistics, the mean of a sample can be calculated as

$$\text{Mean} = \frac{\sum xf}{n},$$

where x is a data value and f is its frequency. In the case of a discrete random variable X, we can think of x as a particular value of X, and f/n as the probability that x occurs (i.e. how "frequently" x occurs). Thus,

$$\text{Mean} = \sum_x x\,\mathbb{P}[X = x]\,.$$

We call this "mean" the expectation of X.

Definition 27.2.4. The **expectation**, or **expected value**, of X, denoted as $\mathbb{E}[X]$ or μ, is given by

$$\mathbb{E}[X] = \sum_x x\,\mathbb{P}[X = x]\,.$$

Example 27.2.5. A single fair 6-sided die is thrown. Let X be the random variable representing the number of dots showing on the die. Note that the possible values of X are $x = 1, 2, 3, 4, 5, 6$. Since $\mathbb{P}[X = x] = \frac{1}{6}$ for all possible values of x, the expectation of X is given by

$$\mathbb{E}[X] = \sum_{x=1}^{6} x\,\mathbb{P}[X = x] = \frac{1}{6}\sum_{x=1}^{6} x = 3.5.$$

Note that the phrase "expected value of X" refers to the long-term weighted average value of a random variable X and is not a typical value that X can take. In fact, a random variable might never be equal to its "expected value". For instance, in the above example, a 6-sided dice will clearly never roll a value of 3.5.

We can generalize the notion of expectation to other functions involving X.

Definition 27.2.6. Let $f(X)$ be any function of the discrete random variable X. Then

$$\mathbb{E}[f(X)] = \sum_x f(x)\,\mathbb{P}[X = x]\,.$$

For instance, $\mathbb{E}[10X] = \sum 10x\,\mathbb{P}[X = x]$, and $\mathbb{E}\big[X^2 - 4\big] = \sum(x^2 - 4)\,\mathbb{P}[X = x]$. From the definition of $\mathbb{E}[f(X)]$, one can easily prove the following results:

Proposition 27.2.7 (Properties of Expectation). For a real constant a,

- $\mathbb{E}[a] = a$,
- $\mathbb{E}[aX] = a\,\mathbb{E}[x]$,
- $\mathbb{E}[f_1(X) + f_2(X)] = \mathbb{E}[f_1(X)] + \mathbb{E}[f_2(X)]$, where f_1 and f_2 are functions of X.

In fact, the last property is a direct consequence of the linearity of the expectation with respect to multiple random variables:

Proposition 27.2.8 (Linearity of Expectation). Let X and Y be random variables (dependent or independent), and let a and b be real constants. Then

$$\mathbb{E}[aX \pm bY] = a\,\mathbb{E}[X] \pm b\,\mathbb{E}[Y]\,.$$

27.2.3 Variance

Recall that in descriptive statistics, the variance of a sample can be calculated as

$$\text{Variance} = \frac{1}{n}\sum f\,(x - \bar{x})^2\,,$$

where f is the frequency of a data value x and $\bar{x}$ is the mean of the sample. In the context of discrete random variables, $\mathbb{P}[X = x]$ corresponds to f/n, while μ corresponds to $\bar{x}$. Thus,

$$\text{Variance} = \sum (x - \mu)^2\,\mathbb{P}[X = x] = \mathbb{E}\big[(x - \mu)^2\big]\,.$$

Definition 27.2.9. The **variance** of a random variable X, denoted by $\mathrm{Var}[X]$ or σ^2, is defined as the expectation of the squared deviation of X from the mean μ. Mathematically,

$$\mathrm{Var}[X] = \mathbb{E}\Big[(X - \mu)^2\Big]\,.$$

As motivated above, we can rewrite $\mathrm{Var}[X]$ solely in terms of expectations:

Proposition 27.2.10.

$$\mathrm{Var}[X] = \mathbb{E}\big[X^2\big] - \mathbb{E}[X]^2\,.$$

Proof.

$$\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}\Big[(X - \mu)^2\Big] \\
&= \mathbb{E}\big[X^2 - 2\mu X + \mu^2\big] \\
&= \mathbb{E}\big[X^2\big] - 2\mu\,\mathbb{E}[X] + \mu^2 \\
&= \mathbb{E}\big[X^2\big] - 2\,\mathbb{E}[X]^2 + \mathbb{E}[X]^2 \\
&= \mathbb{E}\big[X^2\big] - \mathbb{E}[X]^2\,.
\end{aligned}$$

□

Compare this with the alternative expression for the variance used in descriptive statistics:

$$\text{Variance} = \frac{1}{n}\sum fx^2 - \left(\frac{1}{n}\sum fx\right)^2.$$

A small value for the variance indicates that most of the values that X can take are clustered about the mean. Conversely, a higher value for the variance indicates that the values that X can take are spread over a larger range about the mean.

Correspondingly, the **standard deviation**, which is the positive square root of the variance, is denoted by σ, i.e.

$$\sigma = \sqrt{\operatorname{Var}[X]}.$$

From the definition of variance, one can easily prove the following properties:

Proposition 27.2.11 (Properties of Variance)**.** Given that a and b are real constants,

- $\operatorname{Var}[a] = 0$,
- $\operatorname{Var}[aX] = a^2 \operatorname{Var}[X]$,
- $\operatorname{Var}[aX + b] = a^2 \operatorname{Var}[X]$.

Proof. It suffices to prove the last statement. Applying the formula $\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$, we have

$$\begin{aligned}
\operatorname{Var}[aX + b] &= \mathbb{E}\left[(aX + b)^2\right] - \mathbb{E}[aX + b]^2 \\
&= \mathbb{E}\left[a^2X^2 + 2abX + b^2\right] - (a\,\mathbb{E}[X] + b)^2 \\
&= a^2\,\mathbb{E}\left[X^2\right] + 2ab\,\mathbb{E}[X] + b^2 - a^2\,\mathbb{E}[X]^2 - 2ab\,\mathbb{E}[X] - b^2 \\
&= a^2\left[\mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2\right] \\
&= a^2 \operatorname{Var}[X].
\end{aligned}$$

□

Another important property is the variance of more than one random variable. In fact, the property $\operatorname{Var}[aX + b] = a^2 \operatorname{Var}[X]$ is a direct consequence of the statement below:

Proposition 27.2.12 (Variance of More Than One Random Variable)**.** If X and Y are two *independent* variables, then

$$\operatorname{Var}[aX \pm bY] = a^2 \operatorname{Var}[X] + b^2 \operatorname{Var}[Y].$$

Notice that the sign on the RHS is always a '+' regardless of the sign on the LHS. Intuitively, we expect deviations to increase when combining more observations together, not reduce it.

27.3 Binomial Distribution

Consider an experiment which has two possible outcomes, one we term "success" and the other "failure". A binomial situation arises when n independent trials of such experiments are performed.

Examples of such experiments are:

- Tossing a fair coin 6 times (consider obtaining a head on a single toss as "success" and obtaining a tail as "failure").
- Shooting a target 5 times (consider hitting the bull's eye in each shot as "success" and not hitting the bull's eye as "failure").

Condition 27.3.1 (Binomial Model)**.** The conditions for a binomial model are:

- a finite number, n, trials are carried out,
- the trials are independent,
- the outcome of each trial is either a "success" or a "failure", and
- the probability of success, p, is the same for each trial.

Definition 27.3.2. Let the random variable X be the number of trials, out of n trials, that are successful. If the above conditions are met, then X is said to follow a **binomial distribution** with n number of trials and probability of success p, written as

$$X \sim \mathrm{B}(n, p).$$

Example 27.3.3. Recall the example of tossing a fair coin 6 times. This experiment clearly fits a binomial model:

- There are 6 tosses – i.e. a finite number of trials.
- Given that the tosses likely take place one after another, the outcome of one toss will not affect the outcome of another toss – i.e. the trials are independent.
- Each toss only results in a head or tail – i.e. only two possible outcomes, a "success" or "failure".
- The probability of obtaining heads remains the same at 0.5 for each toss – i.e. the probability of success remains unchanged.

27.3.1 Probability Distribution

Proposition 27.3.4 (Probability Distribution of Binomial Distribution)**.** Let the random variable $X \sim \mathrm{B}(n, p)$. Then

$$\mathbb{P}[X = x] = \binom{n}{x} p^x (1-p)^{n-x}.$$

Proof. The event $X = x$ represents obtaining x successes (and $n - x$ failures) out of n total trials. The probability of x successes is simply p^x, while the probability of $n - x$ failures is $(1-p)^{n-x}$. Since there are ${}^nC_x$ ways to choose the x successes from n total trials, the probability of having exactly x successes, i.e. $\mathbb{P}[X = x]$, is

$$\mathbb{P}[X = x] = \binom{n}{x} p^x (1-p)^{n-x}.$$

□

27.3.2 Expectation and Variance

Proposition 27.3.5 (Expectation of Binomial Distribution)**.** For $X \sim \mathrm{B}(n, p)$,

$$\mathbb{E}[X] = np.$$

Proof. Since probabilities sum to 1, we have

$$\sum_{r=0}^{n} \mathbb{P}[X = r] = \sum_{r=0}^{n} \binom{n}{r} p^r (1-p)^{n-r} = 1.$$

Differentiating this with respect to p, we have

$$\sum_{r=0}^{n} \binom{n}{r} \left[rp^{r-1}(1-p)^{n-r} - (n-r)p^r(1-p)^{n-r-1} \right] = 0.$$

We can expand the LHS as

$$\frac{1}{p}\sum_{r=0}^{n} r\binom{n}{r}p^r(1-p)^r - \frac{n}{1-p}\sum_{r=0}^{n}\binom{n}{r}p^r(1-p)^{n-r} + \frac{1}{1-p}\sum_{r=0}^{n} r\binom{n}{r}p^r(1-p)^{n-r} = 0.$$

Rewriting this in terms of $\mathbb{P}[X=r]$ yields

$$\frac{1}{p}\underbrace{\sum_{r=0}^{n} r\,\mathbb{P}[X=r]}_{\mathbb{E}[X]} - \frac{n}{1-p}\underbrace{\sum_{r=0}^{n}\mathbb{P}[X=r]}_{1} + \frac{1}{1-p}\underbrace{\sum_{r=0}^{n} r\,\mathbb{P}[X=r]}_{\mathbb{E}[X]} = 0.$$

Thus,

$$\frac{1}{p}\,\mathbb{E}[X] - \frac{n}{1-p} + \frac{1}{1-p}\,\mathbb{E}[X] = 0 \implies \mathbb{E}[X] = np.$$

□

Proposition 27.3.6 (Variance of Binomial Distribution)**.** For $X \sim \mathrm{B}(n,p)$,

$$\mathrm{Var}[X] = np(1-p).$$

Proof. One can use a similar trick (differentiating $\mathbb{E}[X] = np$) to obtain

$$\mathbb{E}\left[X^2\right] = np(1-p+np).$$

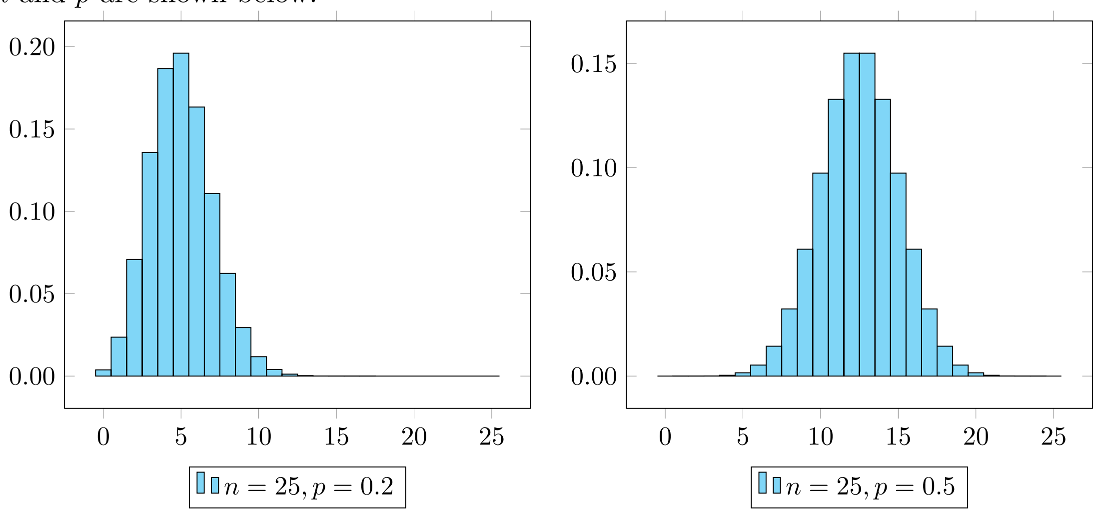
Thus,

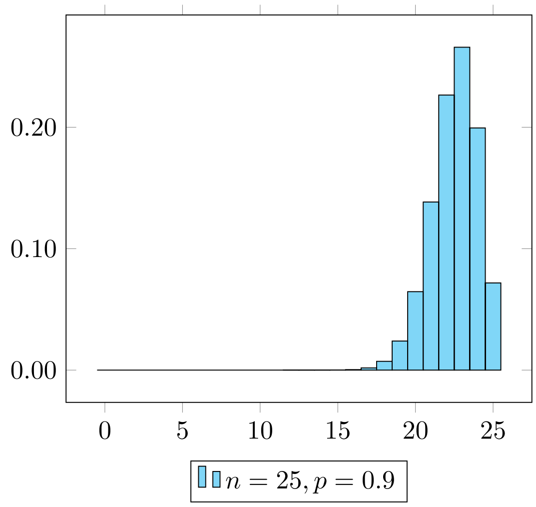
$$\mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = np(1-p+np) - (np)^2 = np(1-p).$$

□

27.3.3 Graphs of Probability Distribution

Given that $X \sim \mathrm{B}(n,p)$, the graphs of the probability distribution of X for various values of n and p are shown below.





Notice that

- when p is low, the graph is skewed to the left, i.e. probabilities are larger for lower values of X,
- when p is high, the graph is skewed to the right, i.e. probabilities are larger for higher values of X, and
- when $p = 0.5$, we get a symmetrical distribution.

Also note that a binomial distribution can only have 1 or 2 modes. In addition, if there are 2 modes, they must be adjacent to each other, i.e. they differ by 1.

27.4 Poisson Distribution

Definition 27.4.1. Let X be the number of occurrences of a particular event over an interval of time (or space) t. Let λ be the mean rate of occurrence per unit time. Then X is said to follow a **Poisson distribution** with parameter λt, written as

$$X \sim \text{Po}(\lambda t).$$

Remark. Typically, we assume t to be the unit time interval, in which case we simply write $X \sim \text{Po}(\lambda)$.

For X to follow the Poisson distribution, the following conditions must also be fulfilled:

Condition 27.4.2 (Poisson Model)**.**

- Events must be independent.
- Events occur singly (i.e. the chances of 2 or more occurrences at precisely the same point in time (or space) is negligible) and randomly.
- Events occur at a constant average rate, i.e. for a given interval of time (or space), the mean number of occurrences is proportional to the length of the interval.

Such a model is also called a Poisson process.
Situations where a Poisson model could be used include:

- the number of car accidents on a stretch of road on a random day, and
- the number of raisins per 10 cm^3 of a chocolate bar.

27.4.1 Probability Distribution

Proposition 27.4.3 (Probability Distribution of Poisson Distribution)**.** Let $X \sim \text{Po}(\lambda t)$. Then

$$\mathbb{P}[X = x] = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^x}{x!}, \quad x \in \mathbb{N}_0.$$

We will present two proofs/derivations for the probability distribution of the Poisson distribution. The first proof (adapted from a note by Cowan) involves infinitesimals and differential equations, while the second proof (adapted from a blog post) uses a measure-theoretic argument.

Proof 1 (Differential Equations). Suppose X is the number of occurrences of an event over some time interval t. We can divide this interval into infinitely short subintervals Δt. For convenience, let $\mathbb{P}[x; t]$ be the probability that exactly x events happen in the time interval t.

Since λ is the mean rate of occurrence, we have

$$\mathbb{P}[1; \Delta t] = \lambda \Delta t.$$

Additionally, since Δt is infinitely short, we can assume that either one event occurs, or no event occurs, i.e.

$$\mathbb{P}[0; \Delta t] = 1 - \mathbb{P}[1; \Delta t] = 1 - \lambda \Delta t.$$

We now wish to find an expression for $\mathbb{P}[x; t]$. To do so, we first consider $\mathbb{P}[0; t]$. Suppose we extend the time interval t by Δt. Since events occur independently and randomly, we must have

$$\mathbb{P}[0; t + \Delta t] = \mathbb{P}[0; t] \, \mathbb{P}[0; \Delta t] = \mathbb{P}[0; t] \left(1 - \lambda \Delta t\right).$$

We can rearrange this to get

$$-\lambda \, \mathbb{P}[0; t] = \frac{\mathbb{P}[0; t + \Delta t] - \mathbb{P}[0; t]}{\Delta t} = \frac{\mathrm{d}}{\mathrm{d}t} \mathbb{P}[0; t] \, .$$

$\mathbb{P}[0; t]$ thus satisfies the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{P}[0; t] = -\lambda \, \mathbb{P}[0; t] \, ,$$

which has solution

$$\mathbb{P}[0; t] = C \mathrm{e}^{-\lambda t}.$$

Since no event can happen in a time interval of 0 seconds, we have

$$\mathbb{P}[0; 0] = 1 \implies C = 1.$$

Thus,

$$\mathbb{P}[0; t] = \mathrm{e}^{-\lambda t}. \tag{1}$$

We now consider $\mathbb{P}[x; t + \Delta t]$, where $x \neq 0$. If x events have occurred in a time interval of $t + \Delta t$, one of two things must have occurred:

- There were x events in the first t seconds, but none in the last Δt.
- There were $x - 1$ events in the first t seconds, and one in the last Δt.

We hence have

$$\begin{aligned}\mathbb{P}[x; t + \Delta t] &= \mathbb{P}[x; t]\, \mathbb{P}[0; \Delta t] + \mathbb{P}[x-1; t]\, \mathbb{P}[1; \Delta t] \\ &= \mathbb{P}[x; t]\,(1 - \lambda \Delta t) + \mathbb{P}[x-1; t]\, \lambda \Delta t.\end{aligned}$$

Rearranging, we get a differential equation involving $\mathbb{P}[x; t]$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\, \mathbb{P}[x; t] + \lambda\, \mathbb{P}[x; t] = \lambda\, \mathbb{P}[x-1; t]\,.$$

Multiplying through by the integrating factor $\mathrm{e}^{\lambda t}$, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{\lambda t}\, \mathbb{P}[x; t]\right] = \lambda \mathrm{e}^{\lambda t}\, \mathbb{P}[x-1; t]\,. \tag{2}$$

We now induct on (2) to get an expression for $\mathbb{P}[x; t]$. We claim that

$$\mathbb{P}[x; t] = \frac{(\lambda t)^x}{x!} \mathrm{e}^{-\lambda t}.$$

We have already shown that this holds for the $x = 0$ case. Now, substituting $x + 1$ into (2), we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{\lambda t}\, \mathbb{P}[x+1; t]\right] = \lambda \mathrm{e}^{\lambda t}\, \mathbb{P}[x; t] = \lambda \mathrm{e}^{\lambda t}\left[\frac{(\lambda t)^x}{x!} \mathrm{e}^{-\lambda t}\right] = \frac{\lambda^{x+1} t^x}{x!}.$$

Integrating and simplifying, we get

$$\mathbb{P}[x+1; t] = \mathrm{e}^{-\lambda t} \int \frac{\lambda^{x+1} t^x}{x!}\, \mathrm{d}t = \frac{(\lambda t)^{x+1}}{(x+1)!} \mathrm{e}^{-\lambda t} + C \mathrm{e}^{-\lambda t}.$$

Since $\mathbb{P}[x+1; 0] = 0$, we have $C = 0$, whence

$$\mathbb{P}[x+1; t] = \frac{(\lambda t)^{x+1}}{(x+1)!} \mathrm{e}^{-\lambda t}.$$

This closes the induction, and we conclude that

$$\mathbb{P}[X = x] = \mathbb{P}[x; t] = \frac{(\lambda t)^x}{x!} \mathrm{e}^{-\lambda t}.$$

□

Proof 2 (Measure Theory). Suppose x events occur in the time interval $[0, t)$, and let their times be given by the unordered x-tuple $(t_1, t_2, \ldots, t_x)$. Without loss of generality, we take $0 \leq t_1 < t_2 < \cdots < t_x < t$. Let S_x be the set of all such x-tuples. Since λ is the mean rate of events per unit time, we define the measure μ such that $\mu([0, 1)) = \lambda$.

Consider the set $T = [0, t)^x$ of all ordered x-tuples. Its measure is given by

$$\mu(T) = \mu([0, t)^x) = (t\mu([0, 1))^x = (\lambda t)^x.$$

Define the equivalence relation $\sim$ on T such that any two x-tuples $u = (u_1, u_2, \ldots, u_x)$ and $v = (v_1, v_2, \ldots, v_x)$ in T,

$$u \sim v \iff \{u_1, u_2, \ldots, u_x\} = \{v_1, v_2, \ldots, v_x\}\,.$$

Then the quotient set $T/\sim$ is exactly S_x. Furthermore, since $\sim$ partitions T into equivalence classes of size $x!$, it follows that

$$\mu(S_x) = \frac{\mu(T)}{x!} = \frac{(\lambda t)^x}{x!}.$$

Now consider the sample space S, which is given by

$$S = \bigcup_{x=0}^{\infty} S_x.$$

Since all S_x are disjoint, the measure of S is simply

$$\mu(S) = \sum_{x=0}^{\infty} \mu(S_x) = \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} = \mathrm{e}^{\lambda t}.$$

Thus, the probability that exactly x events occur in time t is given by the ratio

$$\frac{\mu(S_x)}{\mu(S)} = \frac{(\lambda t)^x/x!}{\mathrm{e}^{\lambda t}} = \mathrm{e}^{-\lambda t}\frac{(\lambda t)^x}{x!}.$$

□

Let X and Y measure the number of events E and F over some time interval. Then $X+Y$ counts the event $G = X+Y$ over the same time interval. Intuitively, $X+Y$ should follow a Poisson distribution since it satisfies the three conditions (27.4.2):

- G is independent: Since X and Y both follow a Poisson distribution, E and F must both occur independently. Since X and Y are independent of each other, E and F are also independent of each other. Thus, G occurs independently.
- G occurs singly and randomly.
- G occurs at a constant average rate: Since E occurs with constant random rate λ_1, and F occurs with constant random rate λ_2, we expect G to also occur with constant random rate $\lambda_1 + \lambda_2$.

We can prove this statement more rigorously using the probability distribution of a Poisson random variable:

Proposition 27.4.4 (Sum of Independent Poisson Random Variables is a Poisson Random Variable)**.** Let $X \sim \mathrm{Po}(\lambda_1)$, $Y \sim \mathrm{Po}(\lambda_2)$ be independent random variables. Then $X+Y \sim \mathrm{Po}(\lambda_1 + \lambda_2)$

Proof. Consider the event $X+Y = n$. This can only happen if $X = m$ and $Y = n-m$. Thus,

$$\mathbb{P}[X+Y=n] = \sum_{m=0}^{n} \mathbb{P}[X=m \text{ and } Y=n-m].$$

Since X and Y are independent, we can split the summands into products:

$$\mathbb{P}[X+Y=n] = \sum_{m=0}^{n} \mathbb{P}[X=m]\,\mathbb{P}[Y=n-m].$$

Using the probability distribution we derived earlier,

$$\mathbb{P}[X+Y=n] = \sum_{m=0}^{n}\left[\mathrm{e}^{-\lambda_1}\frac{\lambda_1^m}{m!}\right]\left[\mathrm{e}^{-\lambda_2}\frac{\lambda_2^{n-m}}{(n-m)!}\right] = \frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{n!}\sum_{m=0}^{n}\frac{n!}{m!(n-m)!}\lambda_1^m\lambda_2^{n-m}.$$

Observe that the sum is simply the binomial expansion of $(\lambda_1+\lambda_2)^n$. Thus,

$$\mathbb{P}[X+Y=n] = \mathrm{e}^{-(\lambda_1+\lambda_2)} = \frac{(\lambda_1+\lambda_2)^n}{n!},$$

which is exactly the probability distribution of a Poisson random variable with parameter $\lambda_1 + \lambda_2$. □

27.4.2 Expectation and Variance

Proposition 27.4.5 (Expectation of Poisson Distribution)**.** Let $X \sim \mathrm{Po}(\lambda t)$. Then $\mathbb{E}[X] = \lambda t$.

Recall that we defined λ as the mean rate of occurrence per unit time. Since we measure X over a time interval of length t, the mean number of events, $\mathbb{E}[X]$, is simply λt. We can verify this with the following calculation:

Proof.

$$\begin{aligned}\mathbb{E}[X] &= \sum_{x=0}^{\infty} x\,\mathbb{P}[X = x] = \sum_{x=1}^{\infty} x\,\mathbb{P}[X = x] = \sum_{x=1}^{\infty} x\mathrm{e}^{-\lambda t}\frac{(\lambda t)^x}{x!} \\ &= \lambda t\mathrm{e}^{-\lambda t}\sum_{x=1}^{\infty}\frac{(\lambda t)^{x-1}}{(x-1)!} = \lambda t\mathrm{e}^{-\lambda t}\sum_{x=0}^{\infty}\frac{(\lambda t)^x}{x!} = \lambda t\mathrm{e}^{-\lambda t}\mathrm{e}^{\lambda t} = \lambda t.\end{aligned}$$

□

Proposition 27.4.6 (Variance of Poisson Distribution)**.** Let $X \sim \mathrm{Po}(\lambda t)$. Then $\mathrm{Var}[X] = \lambda t$.

Proof 1. Consider $\mathbb{E}\left[X^2 - X\right] = \mathbb{E}[X(X-1)]$.

$$\begin{aligned}\mathbb{E}[X(X-1)] &= \sum_{x=0}^{\infty} x(x-1)\,\mathbb{P}[X = x] = \sum_{x=2}^{\infty} x(x-1)\,\mathbb{P}[X = x] \\ &= (\lambda t)^2\mathrm{e}^{-\lambda t}\sum_{x=2}^{\infty}\frac{(\lambda t)^{x-2}}{(x-2)!} = (\lambda t)^2\mathrm{e}^{-\lambda t}\mathrm{e}^{\lambda t} = (\lambda t)^2.\end{aligned}$$

Thus, $\mathbb{E}\left[X^2\right] = \mathbb{E}\left[X^2 - X\right] + \mathbb{E}[X] = (\lambda t)^2 + \lambda t$, from which it follows

$$\mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \lambda t.$$

□

Proof 2. Partition the time interval on which we measure X into n equal subdivisions. Let Y_i measure the number of events that occur in the ith subdivision. As $n \to \infty$, each Y_i approaches a point, in which case Y_i follows a Bernoulli distribution with probability of success $p = \mathbb{E}[Y_i] = \lambda t/n$. Thus,

$$\mathrm{Var}[Y_i] = p(1-p) = \frac{\lambda t}{n}\left(1 - \frac{\lambda t}{n}\right).$$

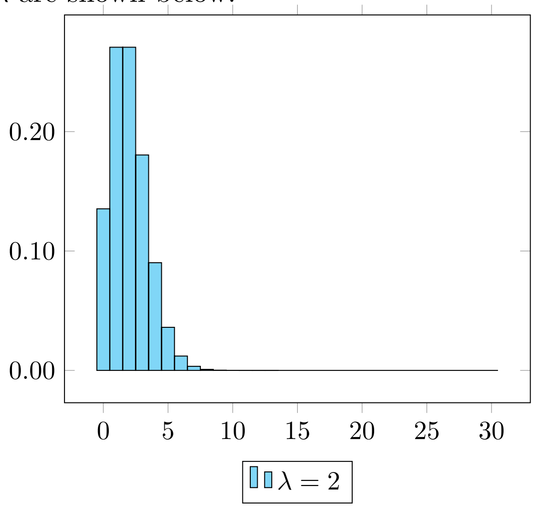
Since the events occur independently, the variance of X is simply the sum of the variances of Y_i. We thus obtain

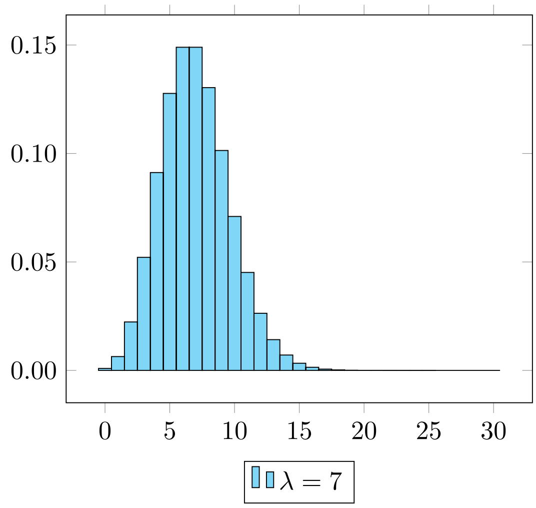
$$\mathrm{Var}[X] = \lim_{n\to\infty}\sum_{i=1}^{n}\mathrm{Var}[Y_i] = \lim_{n\to\infty}\sum_{i=0}^{n}\frac{\lambda t}{n}\left(1 - \frac{\lambda t}{n}\right) = \lim_{n\to\infty} n\left(\frac{\lambda t}{n}\right)\left(1 - \frac{\lambda t}{n}\right) = \lambda t.$$

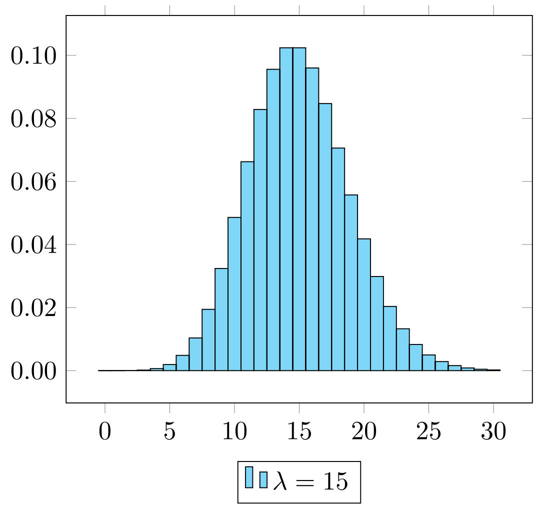
□

27.4.3 Graphs of Probability Distributions

Given that $X \sim \mathrm{Po}(\lambda)$, the graphs of the probability distribution of X for various values of λ are shown below:







27.4.4 Poisson Distribution as an Approximation to the Binomial Distribution

Proposition 27.4.7. If $X \sim \mathrm{B}(n, p)$ and n is large ($n > 50$) and p is small ($p < 0.1$), then X can be approximated by $\mathrm{Po}(\lambda)$, where $\lambda = np$.

Proof. We know that

$$\mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k}. \tag{1}$$

Since n is large relative to k, we have

$$\binom{n}{k} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} \approx \frac{n^k}{k!}. \tag{2}$$

Note also that

$$(1-p)^{n-k} = \mathrm{e}^{(n-k)\ln(1-p)}.$$

Since p is small, we have $\ln(1-p) \approx -p$. Since n is large relative to k, we have $n - k \approx n$. Thus,

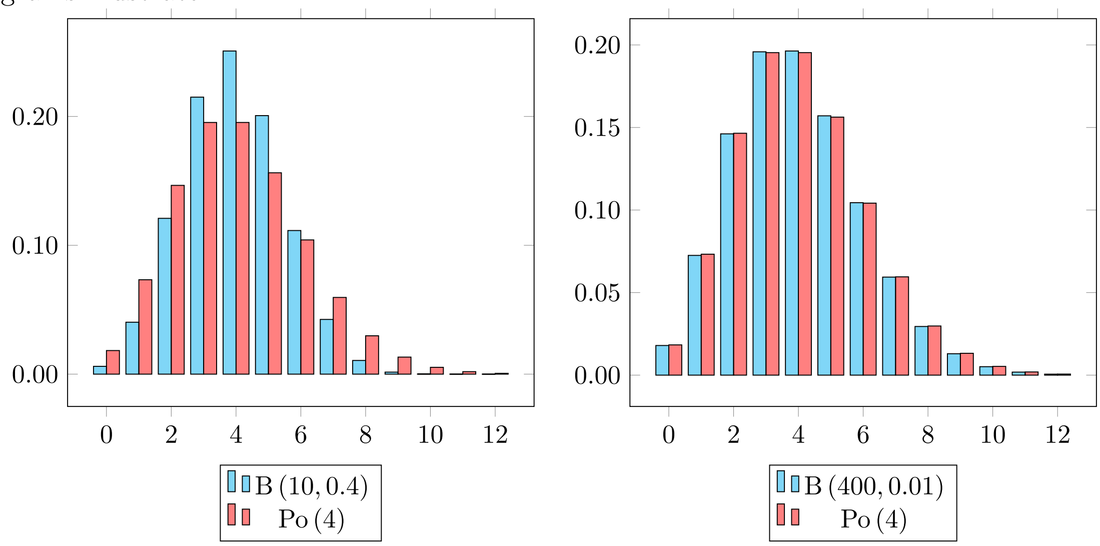
$$(1-p)^{n-k} \approx \mathrm{e}^{-pn}. \tag{3}$$

Substituting (2), (3) and $\lambda = pn$ into (1), we get the approximation

$$\mathbb{P}[X = k] \approx \frac{n^k}{k!} p^k \mathrm{e}^{-pn} = \mathrm{e}^{-\lambda} \frac{n^k}{k!} \left(\frac{\lambda}{n}\right)^k = \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!}.$$

Thus, X is approximately a Poisson distribution where $X \sim \mathrm{Po}(\lambda)$, where $\lambda = np$. $\square$

The approximation gets better as n gets larger and p gets smaller, as the following diagrams illustrate.



This relationship between the binomial and Poisson distributions is particularly useful when we wish to find the sum of two binomial distributions. Consider two random variables $X_1 \sim \mathrm{B}(n_1, p_1)$ and $X_2 \sim \mathrm{B}(n_2, p_2)$, and let $Y = X_1 + X_2$. If we stick with binomial distributions, finding $\mathbb{P}[Y = k]$ would be a nightmare, as we would have to enumerate through all possible cases and calculate many terms:

$$\mathbb{P}[Y = k] = \sum_{i=0}^{k} \mathbb{P}[X_1 = i]\, \mathbb{P}[X_2 = k - i].$$

However, if we use approximate X_1 and X_2 using the Poisson distribution, i.e. $X_1 \sim \mathrm{Po}(\lambda_1)$ and $X_2 \sim \mathrm{Po}(\lambda_2)$, we immediately have $Y \sim \mathrm{Po}(\lambda_1 + \lambda_2)$, and we can easily approximate $\mathbb{P}[Y = k]$:

$$\mathbb{P}[Y = k] \approx \mathrm{e}^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^k}{k!}.$$

27.5 Geometric Distribution

Definition 27.5.1. Let X be the number of trials up to and including the first success. Then X follows a **geometric distribution** with probability of success p, denoted $X \sim \mathrm{Geo}(p)$.

Condition 27.5.2 (Conditions for Geometric Distribution)**.** The conditions for a geometric model are:

- The trials are independent.
- There are only two possible outcomes to each trial, which we will call "success" and "failure".
- The probability of "success", p, is the same for each trial.

Note that the geometric model requires the same conditions as the binomial model, with the exception that the number of trials need not be finite. Intuitively, one could be extremely unlucky and keep on failing.

Situations where the geometric model could be applied to include:

- The number of cards drawn from a pack (with replacement) before an ace is drawn.
- The number of times a fisherman casts a line into a river before he catches a fish.

27.5.1 Probability Distribution

Proposition 27.5.3 (Probability Distribution of Geometric Distribution)**.** Let $X \sim \text{Geo}(p)$. Then

$$\mathbb{P}[X = x] = (1-p)^{x-1} p, \quad x \in \mathbb{Z}^+.$$

Proof. By definition, the event $X = x$ can only occur if the previous $x - 1$ trials are failures (which occur with probability $1 - p$), and the xth trial is a success (which occur with probability p). Thus,

$$\mathbb{P}[X = x] = (1-p)^{x-1} p.$$

□

The geometric distribution has the following useful property:

Proposition 27.5.4. Let $X \sim \text{Geo}(p)$. Then

$$\mathbb{P}[X > x] = (1-p)^x.$$

Proof 1. The event $X > x$ is equivalent to the event that the first x trials were all failures. Thus, $\mathbb{P}[X = x] = (1-p)^x$. □

Proof 2 (Probability Distribution). We have

$$\mathbb{P}[X > x] = \sum_{k=x+1}^{\infty} \mathbb{P}[X = k] = \sum_{k=x+1}^{\infty} (1-p)^{k-1} p.$$

This is simply an infinite geometric series with common ratio $1-p$ and first term $(1-p)^x p$. Thus,

$$\mathbb{P}[X > x] = \frac{(1-p)^x p}{1-(1-p)} = (1-p)^x.$$

□

This actually implies a much stronger property about the geometric distribution:

Definition 27.5.5. A random variable X is said to be **memoryless** if

$$\mathbb{P}[X > s + t \mid X > t] = \mathbb{P}[X > s]$$

for all non-negative s, t.

Proposition 27.5.6 (Geometric Distribution is Memoryless)**.** Let $X \sim \text{Geo}(p)$. Then X is memoryless.

Proof.

$$\begin{aligned}\mathbb{P}[X > s+t \mid X > t] &= \frac{\mathbb{P}[X > s+t \text{ and } X > t]}{\mathbb{P}[X > t]} = \frac{\mathbb{P}[X > s+t]}{\mathbb{P}[X > t]} \\ &= \frac{(1-p)^{s+t}}{(1-p)^t} = (1-p)^s = \mathbb{P}[X > s]\,.\end{aligned}$$

□

Intuitively, this means that having s more observations before a success does not depend on there already being t observations of failure. In other words, the "waiting time" for a success does not depend on how much "time" has already passed.

27.5.2 Expectation and Variance

Proposition 27.5.7 (Expectation of Geometric Distribution)**.** Let $X \sim \mathrm{Geo}(p)$. Then

$$\mathbb{E}[X] = \frac{1}{p}.$$

Proof 1. Intuitively, since each trial has probability of success p, we expect p successes for every 1 trial. This is equivalent to 1 success every $1/p$ trials. Hence, $\mathbb{E}[X] = 1/p$. □

Of course, we can prove this fact more rigorously:

Proof 2 (Probability Distribution).

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k\,\mathbb{P}[X = k] = p\sum_{k=1}^{\infty} k(1-p)^{k-1}.$$

Recall that the Maclaurin series of $(1-x)^{-2}$ is

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}.$$

Substituting $1-p$ for x, we get

$$\mathbb{E}[X] = \frac{p}{p^2} = \frac{1}{p}.$$

□

Proof 3 (Memoryless Property). The first trial can result in one of two outcomes:

- The first trial is a success (occurs with probability p). If this happens, the process stops, and $X = 1$.
- The first trial is a failure (occurs with probability $1-p$). If this happens, the process effectively "restarts" (memoryless property). The expected number of trials in this case becomes $\mathbb{E}[1+X] = 1 + \mathbb{E}[X]$.

The expectation of X can thus be calculated as:

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{P}[\text{success}]\,(\#\text{ trials if success}) + \mathbb{P}[\text{failure}]\,(\#\text{ trials if failure}) \\ &= (p)(1) + (1-p)\,\mathbb{E}[1+X]\end{aligned}$$

Simplifying, we have $\mathbb{E}[X] = 1/p$. □

Proposition 27.5.8 (Variance of Geometric Distribution)**.** Let $X \sim \mathrm{Geo}(p)$. Then

$$\mathrm{Var}[X] = \frac{1-p}{p^2}.$$

Proof 1 (Probability Distribution). Recall that

$$\sum_{k=1}^{\infty} x^k = \frac{1}{1-x}.$$

Differentiating this twice with respect to x, we get

$$\sum_{k=1}^{\infty} k(k-1)x^{k-2} = \frac{2}{(1-x)^3} \implies \sum_{k=1}^{\infty} \left(k^2 - k\right) x^{k-1} = \frac{2x}{(1-x)^3}. \tag{1}$$

Now consider $\mathbb{E}\left[X^2\right] - \mathbb{E}[X]$:

$$\mathbb{E}\left[X^2\right] - \mathbb{E}[X] = \sum_{k=1}^{\infty} \left(k^2 - k\right) \mathbb{P}[X = k] = p \sum_{k=1}^{\infty} \left(k^2 - k\right) (1-p)^{k-1}.$$

Using (1) with $x = 1 - p$,

$$\mathbb{E}\left[X^2\right] - \mathbb{E}[X] = p\left[\frac{2(1-p)}{p^3}\right] = \frac{2-2p}{p^2}.$$

Thus,

$$\mathbb{E}\left[X^2\right] = \frac{2-2p}{p^2} + \frac{1}{p} = \frac{2-p}{p^2} \implies \mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{1-p}{p^2}.$$

□

Proof 2 (Memoryless Property). Following the memoryless property proof above, we have

$$\begin{aligned}
\mathbb{E}\left[X^2\right] &= \mathbb{P}[\text{success}]\,(\#\text{ trials if success})^2 + \mathbb{P}[\text{failure}]\,(\#\text{ trials if failure})^2 \\
&= (p)(1)^2 + (1-p)\,\mathbb{E}\left[(1+X)^2\right] \\
&= p + (1-p)\left[1 + \frac{2}{p} + \mathbb{E}\left[X^2\right]\right]
\end{aligned}$$

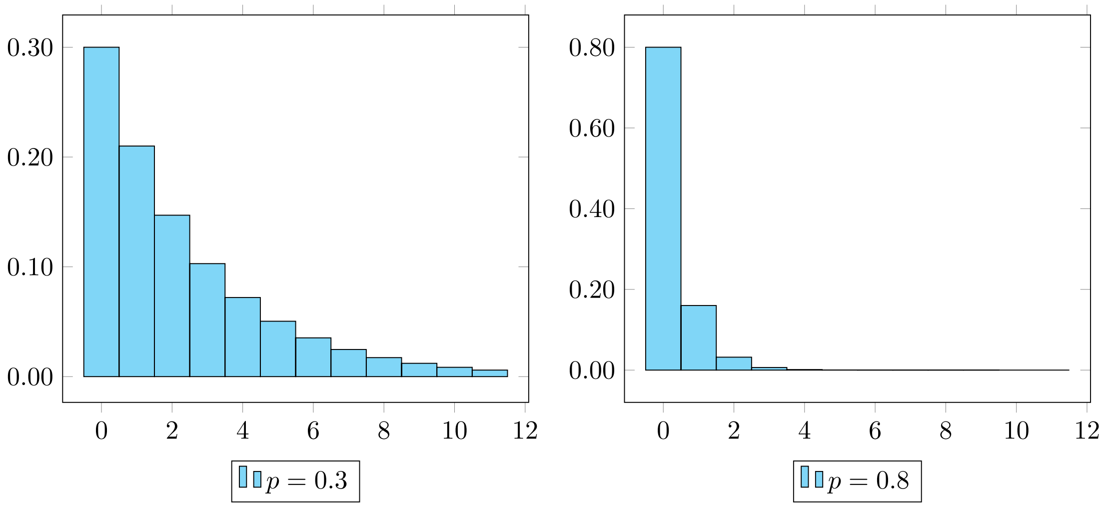
Simplifying, we have

$$\mathbb{E}\left[X^2\right] = \frac{2-p}{p^2} \implies \mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{1-p}{p^2}.$$

□

27.5.3 Graphs of Probability Distribution

Given that $X \sim \mathrm{Geo}(p)$, the graphs of the probability distribution of X for various values of p are shown below:



All geometric distributions show this type of skewness (extreme positive skewness).

28 Continuous Random Variables

In the previous chapter, we saw how a discrete random variable assumes countable values. If we want a random variable to take on uncountably many values, then we must turn to continuous random variables instead.

Definition 28.0.1. A **continuous random variable** is a random variable that can take on any value in a given interval.

Since the value of a continuous random variable is uncountable, it can only take on an interval of values, not a specific value.

An example of continuous random variables is the volume of beverage (in ml) in a 500 ml bottle ($100 \le X \le 200$, $200 \le X \le 300$, etc.)

28.1 Discrete to Continuous

In the previous chapter, we saw how we could represent the probability distribution of a discrete random variable using a table. For instance, the probability distribution of the outcome of a single throw of a 6-sided dice is given by the following table:

x	1	2	3	4	5	6
$\mathbb{P}[X = x]$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

We can try to specify the distribution of a continuous random variable in the same way. Consider the lengths, in millimetres, of 50 leaves that have fallen from a particular tree. We can illustrate the distribution of the lengths using a histogram:

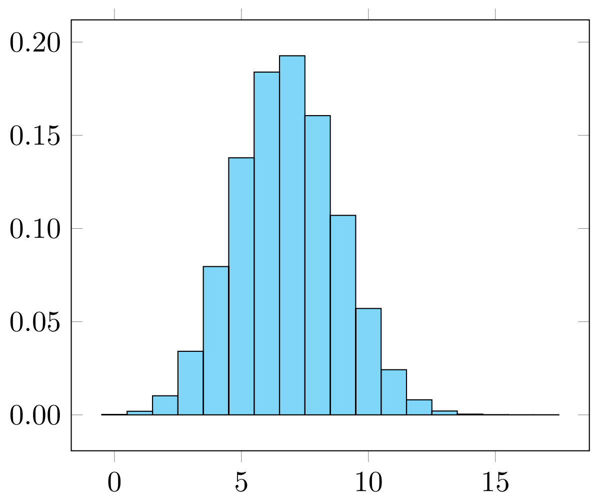


Figure 28.1: A histogram of the lengths of leaves.

Here, the vertical axis represents the frequency density of lengths in a particular interval, hence the total area of the histogram is 1. This property also allows us to find the probability that a length is in a given interval: simply sum up the area of the rectangles in the given interval.

Notice that if we want the probability of a certain length, e.g. $L = 6.3$ cm, the answer would be zero. Though it is theoretically possible for L to be 6.3 cm exactly (i.e. $6.30000\ldots$), the probability is actually zero. This means that

$$\mathbb{P}[6 < L < 7] = \mathbb{P}[6 \le L < 7] = \mathbb{P}[6 < L \le 7] = \mathbb{P}[6 \le L \le 7]\,.$$

That is, whether we include the bounds of the interval does not affect the probability that L falls within the interval.

The probabilities calculated from the histogram could be used to model the length of a tree leaf. However, the model is crude, because of the limited amount of data, and the small number of classes in which the leaves are grouped into, resulting in the "steps" in the histogram.

The model could be further refined by repeating the process of collecting more data and reducing the class width. If this process were to be continued indefinitely, then the outline of the histogram would become a smooth curve:

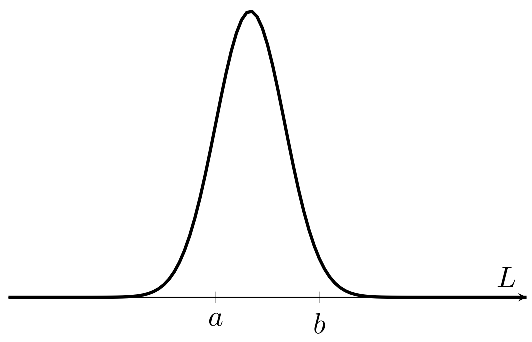


Figure 28.2: Smooth curve after repeating process infinitely.

The probability of the length of a leaf lying between a and b is given by the area under the curve between a and b.

28.2 Properties

28.2.1 Probability Density Function

We have seen how the outline of a histogram may approach a smooth curve when we allow the sample size to increase with correspondingly narrower class widths.

Definition 28.2.1. The curve is the graph of the **probability density function** (pdf in short), and the function is usually denoted by the small letter f. It describes mathematically how the unit of probability is distributed over the range of x-values.

Note that $f(x)$ *does not* represent the probability. It is the area under $f(x)$ that represents probability.

The probability density function $f(x)$ of a continuous random X has the following properties:

Fact 28.2.2 (Properties of pdf)**.**

- $f(x)$ is non-negative (since we cannot have negative probabilities):
$$\forall x: \quad f(x) \geq 0.$$
- The total area under the graph is 1 (since the probability must sum to 1):
$$\int_{-\infty}^{\infty} f(x) = 1.$$
- Probability is given by the area under $f(x)$:
$$\mathbb{P}[a < X < b] = \int_a^b f(x)\,\mathrm{d}x.$$
- The boundary of an interval does not affect probability:
$$\mathbb{P}[a < X < b] = \mathbb{P}[a \leq X < b] = \mathbb{P}[a < X \leq b] = \mathbb{P}[a \leq X \leq b]\,.$$
- If f has a maximum when $x = M$, then M is the mode.
- If $\mathbb{P}[X \leq m] = \int_{-\infty}^{m} f(x)\,\mathrm{d}x = 1/2$, then m is the median. If f is symmetric about the line $x = x_0$, then m is simply x_0.

Note that $f(x)$ need not be continuous; it only needs to be non-negative and have a total area of 1. For instance, the piecewise function

$$f(x) = \begin{cases} x, & 0 \leq x \leq 1, \\ 2 - x, & 1 < x \leq 2, \\ 0, & \text{otherwise} \end{cases}$$

is a valid probability density function.

28.2.2 Cumulative Distribution Function

Definition 28.2.3. The **cumulative distribution function** $F(x)$ is often referred to as the distribution function, or as the cdf. The function is defined by

$$F(x) = \mathbb{P}[X \leq x] = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$

Example 28.2.4. Let the continuous random variable X have pdf $f(x)$ given by

$$f(x) = \begin{cases} \mathrm{e}^{-x}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Let the cdf of X be $F(x)$. For $x \leq 0$, we clearly have $F(x) = 0$. For $x > 0$, we have

$$F(x) = F(0) + \int_0^x f(t)\,\mathrm{d}t = 0 + \int_0^x \mathrm{e}^{-t}\,\mathrm{d}t = \left[-\mathrm{e}^{-t}\right]_0^x = 1 - \mathrm{e}^{-x}.$$

Thus,

$$F(x) = \begin{cases} 0, & x \leq 0, \\ 1 - \mathrm{e}^{-x}, & x > 0. \end{cases}$$

The cdf of a continuous random variable X has the following properties:

Fact 28.2.5 (Properties of cdf)**.**

- By the fundamental theorem of calculus, we have

$$\frac{\mathrm{d}}{\mathrm{d}x} F(x) = f(x).$$

- The lower and upper limits of $F(x)$ are 0 and 1 respectively:

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} F(x) = 1.$$

- F is a non-decreasing function, i.e. $a \leq b$ implies $F(a) \leq F(b)$.
- F is a continuous function, even if f is discontinuous.
- $\mathbb{P}[a < X < b] = F(b) - F(a)$.
- The median m satisfies $F(m) = 1/2$.

28.2.3 Expectation and Variance

Definition 28.2.6. For a continuous random variable X with pdf f, the **expectation** of X is given by

$$\mu = \mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x.$$

For a general function g, we calculate $\mathbb{E}[g(X)]$ as

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\,\mathrm{d}x.$$

Note that if f is symmetric about the line $x = c$, then $\mathbb{E}[X] = c$.
Using the above definitions, we can easily calculate the variance of X:

Definition 28.2.7. The **variance** of X, denoted $\mathrm{Var}[X]$, is given by

$$\mathrm{Var}[X] = \mathbb{E}\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,\mathrm{d}x.$$

However, it is usually easier to calculate $\mathrm{Var}[X]$ using

$$\mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2.$$

Note that all results of expectation and variance algebra (see §27.2.2 and §27.2.3) continue to hold:

Fact 28.2.8 (Properties of Expectation and Variance)**.** For a continuous random variable X and constants a and b,

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]\,,$$

where Y is any continuous random variable. If Y is also independent with X, then

$$\operatorname{Var}[aX+bY] = a^2 \operatorname{Var}[X] + b^2 \operatorname{Var}[Y]\,.$$

The proofs of the two facts are similar to the discrete case.

28.2.4 Distribution of a Function of a Random Variable

Suppose we have a continuous random variable Y that is given as a function of another continuous random variable X, i.e. $Y = g(X)$. If we know that cdf of X, we can easily find the pdf and cdf of Y using the following method:

Recipe 28.2.9 (Finding pdf and cdf of Y)**.** Let X be a continuous random variable with pdf f_X. If $Y = g(X)$ (i.e. Y depends on X), then

$$F_Y(y) = \mathbb{P}[Y \le y] = \mathbb{P}[g(X) \le y]\,.$$

Then, to obtain the pdf of Y, we differentiate $F_Y(y)$ with respect to y.

Sample Problem 28.2.10. Let X have pdf

$$f_X(x) = \begin{cases} \frac{2}{\pi}, & 0 \le x \le \frac{\pi}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

Find the pdf of Y, where $Y = \sin X$.

Solution. Integrating f_X, we obtain the cdf of X:

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \frac{2}{\pi}x, & 0 \le x \le \frac{\pi}{2}, \\ 1, & x > \frac{\pi}{2}. \end{cases}$$

Now consider $F_Y(y)$:

$$\begin{aligned} F_Y(y) &= \mathbb{P}[Y \le y] = \mathbb{P}[\sin X \le y] = \mathbb{P}[X \le \arcsin y] \\ &= \begin{cases} 0, & \arcsin y < 0, \\ \frac{2}{\pi}\arcsin y, & 0 \le \arcsin y \le \frac{\pi}{2}, \\ 1, & \arcsin y > \frac{\pi}{2} \end{cases} = \begin{cases} 0, & y < 0, \\ \frac{2}{\pi}\arcsin y, & 0 \le y \le 1, \\ 1, & y > 1. \end{cases} \end{aligned}$$

Differentiating, we obtain the pdf of Y:

$$f_Y(y) = \begin{cases} \frac{2}{\pi\sqrt{1-y^2}}, & 0 \le y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

□

28.3 Uniform Distribution

Definition 28.3.1. If the continuous random variable X is equally likely to lie anywhere in the interval $[a, b]$, where a and b are constants, then X follows a **uniform distribution**, denoted $X \sim \mathrm{U}(a, b)$.

28.3.1 Density and Distribution Functions

Proposition 28.3.2. The probability density function of $X \sim \mathrm{U}(a, b)$ is

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{otherwise.} \end{cases}$$

Proof. Since X is equally likely to lie anywhere in the interval $[a, b]$, we know its pdf has the form

$$f(x) = \begin{cases} c, & a \leq x \leq b, \\ 0, & \text{otherwise,} \end{cases}$$

where c is a constant. Since the sum of probabilities is 1,

$$1 = \int_{-\infty}^{\infty} f(x) \, \mathrm{d}x = \int_a^b c \, \mathrm{d}x = c(b - a).$$

Thus, $c = 1/(b - a)$, as desired. □

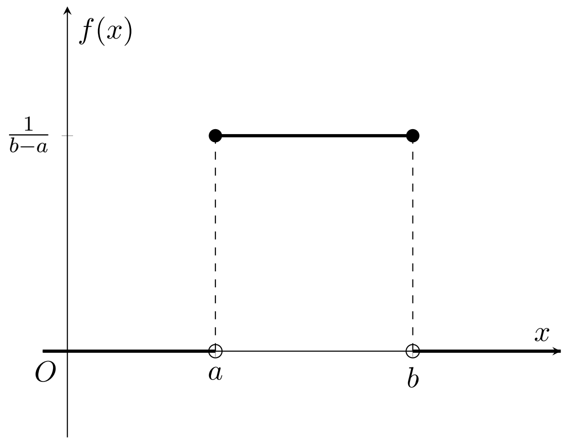


Figure 28.3: The probability density function $f(x)$.

Proposition 28.3.3. The cumulative density function of $X \sim \mathrm{U}(a, b)$ is

$$F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases}$$

Proof. Clearly, $F(x) = 0$ for all $x < a$. For $a \leq x \leq b$, we have

$$F(x) = F(0) + \int_a^x f(t) \, \mathrm{d}t = 0 + \int_a^x \frac{1}{b-a} \, \mathrm{d}t = \frac{x-a}{b-a}.$$

For $x > b$, we clearly have $F(x) = 1$. Thus,

$$F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases}$$

$\square$

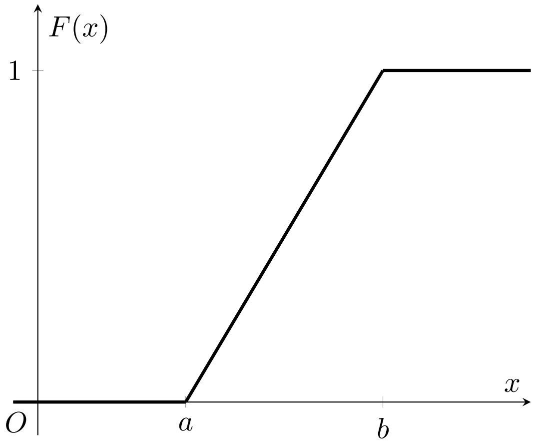


Figure 28.4: The cumulative distribution function $F(x)$.

28.3.2 Expectation and Variance

Proposition 28.3.4. If $X \sim \mathrm{U}(a, b)$, then $\mathbb{E}[X] = (a+b)/2$.

Proof. The pdf of X is symmetric about $x = (a+b)/2$. Thus, $(a+b)/2$ is the mean. $\square$

Proposition 28.3.5. If $X \sim \mathrm{U}(a, b)$, then $\mathrm{Var}[X] = (b-a)^2/12$.

Proof. Consider $\mathbb{E}\left[X^2\right]$:

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} \frac{x^2}{b-a}\,\mathrm{d}x = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b = \frac{(b-a)^2}{3}.$$

Thus,

$$\mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{(b-a)^2}{3} - \left(\frac{b-a}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

$\square$

28.4 Exponential Distribution

Definition 28.4.1. Let the continuous random variable X be the "waiting times" between successive events in a Poisson process with mean rate λ. Then X follows an **exponential distribution** with parameter λ, written $X \sim \mathrm{Exp}(\lambda)$.

As its definition suggests, the exponential distribution is often used to model waiting times. Some situations where the exponential model is applicable include:

- time between telephone calls or accidents,
- the length of time until an electronic device fails,
- the time required to wait for the first emission of a particle from a radioactive source.

28.4.1 Density and Distribution Functions

Proposition 28.4.2. The probability density function of $X \sim \mathrm{Exp}(\lambda)$ is given by

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

and the cumulative distribution function of X is given by

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - \mathrm{e}^{-\lambda x}, & x \geq 0. \end{cases}$$

Proof. Consider a Poisson process with mean rate λ. Let Y be the number of events occurring in a time interval of length x, i.e. $Y \sim \mathrm{Po}(\lambda x)$. Let X be the random variable denoting the "waiting time" between successive such random events.

Since X is the amount of time until the next event occurs, the event $X > x$ is equivalent to no events happening in a time interval of x. In other words, $X > x$ is equivalent to $Y = 0$. Hence,

$$\mathbb{P}[X > x] = \mathbb{P}[Y = 0] = \frac{(\lambda x)^0}{0!}\mathrm{e}^{-\lambda x} = \mathrm{e}^{-\lambda x}$$

Hence, for $x \geq 0$, the cdf of X is given by

$$F(x) = \mathbb{P}[X \leq x] = 1 - \mathbb{P}[X > x] = 1 - \mathrm{e}^{-\lambda x}.$$

Also, since the "waiting time" cannot be negative, we have

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - \mathrm{e}^{-\lambda x}, & x \geq 0. \end{cases}$$

Differentiating, we obtain the pdf of X:

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

□

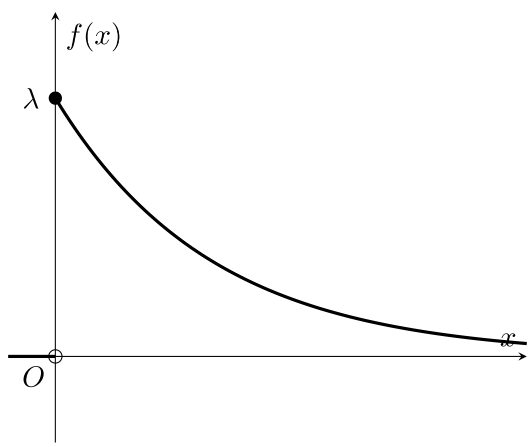


Figure 28.5: The probability density function $f(x)$.

Proposition 28.4.3. The exponential distribution is memoryless.

Proof. Let $X \sim \operatorname{Exp}(\lambda)$. We have

$$\begin{aligned}\mathbb{P}[X > a+b \mid X > a] &= \frac{\mathbb{P}[X > a+b \text{ and } X > a]}{\mathbb{P}[X > a]} = \frac{\mathbb{P}[X > a+b]}{\mathbb{P}[X > a]} \\ &= \frac{\mathrm{e}^{\lambda(a+b)}}{\mathrm{e}^{\lambda a}} = \mathrm{e}^{-\lambda b} = \mathbb{P}[X > b]\,.\end{aligned}$$

Thus, the probability that one has to "wait" another b units of time does not depend on the time already spent "waiting", i.e. X is memoryless. □

28.4.2 Expectation, Variance and Median

Proposition 28.4.4. If $X \sim \operatorname{Exp}(\lambda)$, then $\mathbb{E}[X] = 1/\lambda$.

Proof. We have

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_0^{\infty} \lambda x \mathrm{e}^{-\lambda x}\,\mathrm{d}x.$$

Integrating by parts, we get

$$\mathbb{E}[X] = \left[-x\mathrm{e}^{-\lambda x} - \frac{\mathrm{e}^{-\lambda x}}{\lambda}\right]_0^{\infty} = \frac{1}{\lambda}.$$

□

Proposition 28.4.5. If $X \sim \operatorname{Exp}(\lambda)$, then $\operatorname{Var}[X] = 1/\lambda^2$.

Proof. We have

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \int_0^{\infty} \lambda x^2 \mathrm{e}^{-\lambda x}\,\mathrm{d}x.$$

Integrating by parts, we get

$$\mathbb{E}\left[X^2\right] = \left[-x^2\mathrm{e}^{-\lambda x}\right]_0^{\infty} + 2\int_0^{\infty} x\mathrm{e}^{-\lambda x}\,\mathrm{d}x = 0 + \frac{2}{\lambda}\,\mathbb{E}[X] = \frac{2}{\lambda^2}.$$

Thus,

$$\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right) = \frac{1}{\lambda^2}.$$

□

Proposition 28.4.6. The median of $X \sim \operatorname{Exp}(\lambda)$ is $\ln 2/\lambda$.

Proof. Let m be the median. Then $F(m) = 1/2$. Hence,

$$\frac{1}{2} = F(m) = 1 - \mathrm{e}^{-\lambda m} \implies \mathrm{e}^{\lambda m} = 2 \implies m = \frac{\ln 2}{\lambda}.$$

□

28.5 Normal Distribution

Definition 28.5.1. The probability density function of a continuous random variable X that follows a **normal distribution** with mean μ and standard deviation σ, written $X \sim \mathrm{N}\big(\mu, \sigma^2\big)$, is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

The normal distribution arises in many different situations. For instance, the normal distribution can be used to model various characteristics of a model, e.g. heights, weights, and even test scores. The reason why the normal distribution is such a good fit for modelling population-sized data sets is due to a very important theorem called the **Central Limit Theorem**, which we will learn in a later chapter.

28.5.1 Properties

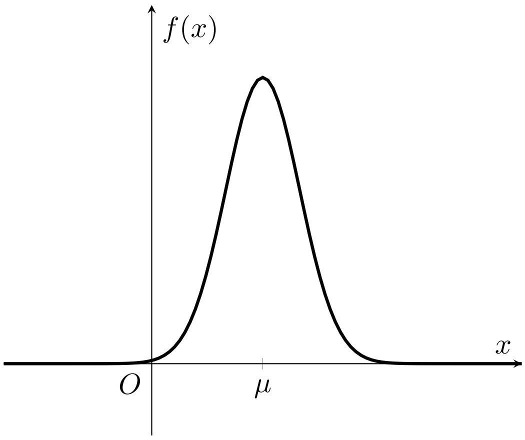


Figure 28.6: The pdf of a normal distribution.

As exemplified by the figure above, a normal curve has the following properties:

- It is bell-shaped.
- The mean, median and mode are all equal (symmetric about $x = \mu$, maximum at $x = \mu$).
- It approaches the x-axis as $x \to \pm\infty$.

Note also that the shape of the normal curve is completely determined by two parameters, namely the mean μ and the standard deviation σ. The following figures show how the mean and the standard deviation affect the shape of the normal curve:

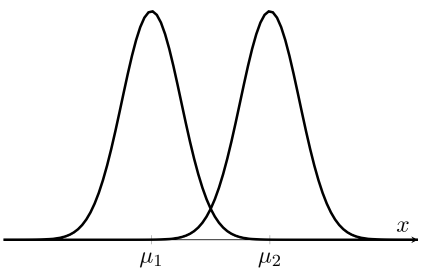


Figure 28.7: Varying μ.

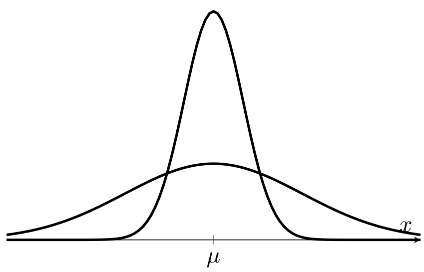


Figure 28.8: Varying σ.

Increasing μ has the same effect as translating the normal distribution curve in the positive x-direction. Meanwhile, increasing σ has the effect of flattening the normal distribution curve, i.e. the area under the curve about μ becomes less concentrated, or more dispersed.

In a normal distribution, about 68.3%, 95.4% and 99.7% of the values of x are expected to lie within ± 1, ± 2 and ± 3 standard deviations from the mean of X respectively.

Perhaps the most important property of the normal distribution is that the sum or difference of normal distributions is also a normal distribution.

Proposition 28.5.2. If X and Y are two *independent* random variables such that $X \sim \mathrm{N}\left(\mu_1, \sigma_1^2\right)$ and $Y \sim \mathrm{N}\left(\mu_2, \sigma_2^2\right)$, then their sum and differences also follow a normal distribution:

$$aX + bY \sim \mathrm{N}\left(a\mu_1 \pm b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2\right).$$

28.5.2 Standard Normal Distribution

Definition 28.5.3. A random variable Z is said to follow a standard normal distribution if $Z \sim \mathrm{N}(0, 1)$, i.e. Z has mean 0 and variance 1.

Suppose $X \sim \mathrm{N}\left(\mu, \sigma^2\right)$. Then the random variable defined by $Z = (X - \mu)/\sigma$ follows a standard normal distribution. The process of converting $X \sim \mathrm{N}\left(\mu, \sigma^2\right)$ into $Z \sim \mathrm{N}(0, 1)$ is known as **standardization** and can be viewed as a transformation on the normal curve of X.

Standardization is typically used to compare different random variables that follow normal distributions, such as test scores for different subjects.

Definition 28.5.4. Let $X \sim \mathrm{N}\left(\mu, \sigma^2\right)$, and let x be an observation of X. Then the normalized score of x, called a **z-score**, measures the position of a score from the mean where its distance from the mean is measured in standard deviations. Mathematically,

$$z = \frac{x - \mu}{\sigma}.$$

As the definition suggests, the higher the z-score, the better x is relative to its distribution. For instance, if $z = 1$, then x is 1 standard deviation above the mean, while if $z = -2$, then x is 2 standard deviations below the mean.

Sample Problem 28.5.5. In the final year examination, a student obtains a score of 70 for Chemistry and 65 for Mathematics. If the cohort's scores for Chemistry and Mathematics follows $\mathrm{N}\left(60, 10^2\right)$ and $\mathrm{N}\left(57, 4^2\right)$ respectively, which subject did the student do better in?

Solution. Normalizing the student's Chemistry score, we get a z-score of

$$z_1 = \frac{X - \mu}{\sigma} = \frac{70 - 60}{10} = 1.$$

Normalizing the student's Mathematics score, we get a z-score of

$$z_2 = \frac{X - \mu}{\sigma} = \frac{65 - 57}{4} = 2.$$

We see that the student has a higher z-score for Mathematics than for Chemistry. Thus, even though the student obtained a higher score for Chemistry, he did better in Mathematics when compared against his peers. □

The standard normal distribution is also used for various scoring systems, such as PSLE T-scores, IQ scores and SAT scores.

28.5.3 Normal Distribution as an Approximation

Previously, we saw how the binomial distribution, under certain conditions, could be approximated to the Poisson distribution. Similarly, the normal distribution can be used to approximate both the binomial and Poisson distributions when certain conditions are satisfied.[1]

However, unlike the case of binomial to Poisson, which is a discrete-to-discrete approximation, approximately either the binomial or Poisson distribution to the normal distribution is a discrete-to-continuous change. We hence introduce the idea of a "continuity correction". Intuitively, what this means is that $\mathbb{P}[X = k]$ (in the discrete case) is taken to be $\mathbb{P}[k - 0.5 < X < k + 0.5]$ (in the continuous case). For instance, $\mathbb{P}[X = 16] = \mathbb{P}[15.5 < X < 16.5]$, and $\mathbb{P}[2 < X \leq 20] = \mathbb{P}[2.5 < X < 20.5]$.

Approximating the Binomial Distribution

Proposition 28.5.6. If $X \sim \mathrm{B}(n, p)$ and n is sufficiently large such that $\mu = np > 5$ and $n(1 - p) > 5$, then X can be approximated by $\mathrm{N}(np, np(1 - p))$, taking into account the continuity correction.

If p is close to 0.5, the binomial distribution is almost symmetrical. Thus, the approximation by a normal distribution (which is symmetrical) gets better as p gets closer to 0.5.

Consider the following figure, where $X \sim \mathrm{B}(15, 0.5)$. We can approximate the distribution X with a normal distribution with mean $np = 7.5$ and variance $np(1 - p) = 3.75$.

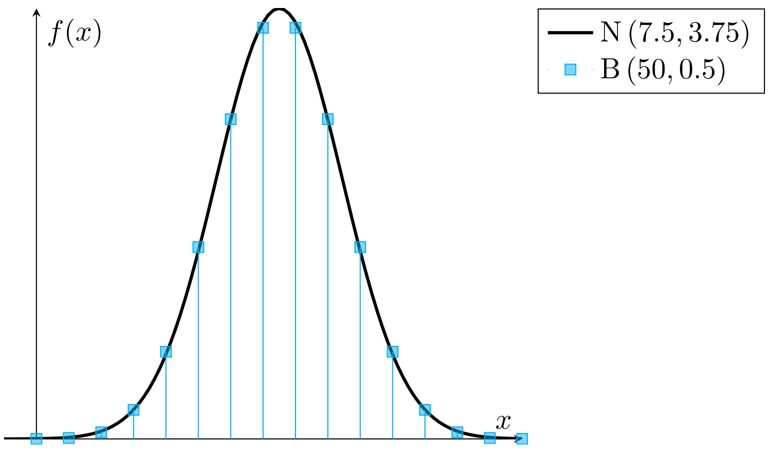


Figure 28.9: Approximating the binomial distribution.

Approximating the Poisson Distribution

Proposition 28.5.7. If $X \sim \mathrm{Po}(\lambda)$ such that $\lambda > 10$, then X can be approximated by $\mathrm{N}(\lambda, \lambda)$, taking into account the continuity correction.

[1] This is a consequence of the Central Limit Theorem, which we introduced earlier in the section.

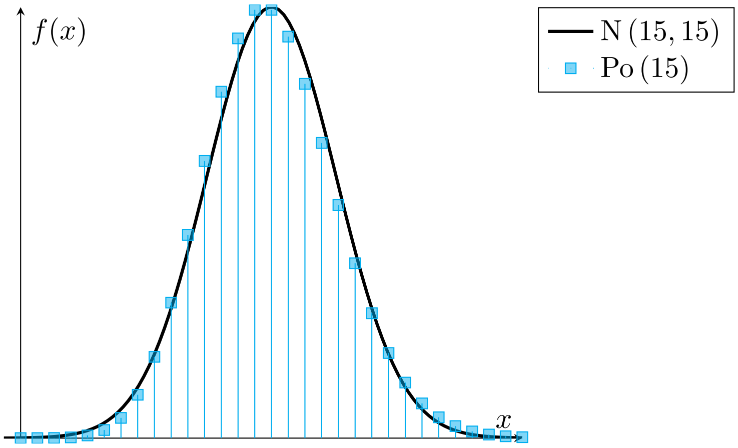


Figure 28.10: Approximating the Poisson distribution.

29 Sampling

29.1 Random Sampling

In §26, we saw how we cannot always have access to an entire population for study. Hence, we often turn to a sample to make inferences about the characteristics of the population.

A central notion about samples is the idea of them being representative of the population. We use the phrase **random sample** to denote such samples. We can think of random samples as a "fair" or "unbiased" sample; every member of the population has an equal, non-zero probabilities of getting sampled. On the other hand, a **non-random sample** is biased and are not representative of the sample; every member of the population does not have an equal chance of getting sampled.

29.1.1 Simple Random Sampling

Simple random sampling is a method of selecting n members from a population of size N such that each possible sample of that size has the same chance of being chosen.

One procedure for obtaining a simple random sample is the following:

Recipe 29.1.1 (Simple Random Sampling)**.**

1. Make a list of all N members of the population. This is called the **sampling frame**.
2. Assign each member of the population a different number.
3. For each member of the population, place a corresponding numbered ball in a bag.
4. Draw n balls from the bag, without replacement. The balls should be chosen at random.
5. The numbers on the ball identify the chosen members of the population.

29.2 Sample Mean

We now look at the first objective of obtaining a random sample: calculating probabilities relating to the sample mean.

Definition 29.2.1. If $X_1, X_2, \ldots, X_n$ is a random sample of n independent observations from a population, then the sample mean $\overline{X}$ is defined as

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

Note that the sample mean $\overline{X}$ is also a random variable since it varies depending on the samples taken.

Proposition 29.2.2. Let the population mean be μ and the population variance be σ^2. Then the sample mean $\overline{X}$ has expectation μ and variance σ^2/n.

Proof. We have

$$\begin{aligned}
\mathbb{E}\left[\overline{X}\right] &= \mathbb{E}\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] = \frac{\mathbb{E}[X_1 + X_2 + \cdots + X_n]}{n} \\
&= \frac{\mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n]}{n} = \frac{n\,\mathbb{E}[X]}{n} = \mathbb{E}[X] = \mu
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{Var}\left[\overline{X}\right] &= \operatorname{Var}\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] = \frac{1}{n^2}\operatorname{Var}[X_1 + X_2 + \cdots + X_n] \\
&= \frac{\operatorname{Var}[X_1] + \operatorname{Var}[X_2] + \cdots + \operatorname{Var}[X_n]}{n^2} = \frac{n\operatorname{Var}[X]}{n^2} = \frac{\sigma^2}{n}.
\end{aligned}$$

□

Definition 29.2.3. The standard deviation of $\overline{X}$, $\sigma/\sqrt{n}$, is known as the **standard error** of the mean.

Observe that as n increases, the standard error of the sample mean decreases. This aligns with our intuition: as n increases, we are effectively sampling a larger proportion of the population, so our statistic (the sample mean) should tend towards the parameter (the population mean).

29.2.1 The Central Limit Theorem

If sampling is done from a normal population, then the sample mean will also follow a normal distribution.

Proposition 29.2.4. If $X \sim \mathrm{N}(\mu, \sigma^2)$, then

$$\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \text{ exactly.}$$

However, if the population does not follow a normal distribution, then the sample mean also does not follow a normal distribution. However, if the sample size is large, then the distribution of the sample mean will be approximately normal. This result is known as the Central Limit Theorem.

Theorem 29.2.5 (Central Limit Theorem)**.** If X does not follow a normal distribution, with $\mathbb{E}[X] = \mu$ and $\operatorname{Var}[X] = \sigma^2$, and n is large (typically $n \geq 30$), then

$$\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \text{ approximately.}$$

Here, we are assuming that the samples $X_1, X_2, \ldots, X_n$ are independent and identically distributed. Further, the variance σ^2 must be finite.

Note that the condition $n \geq 30$ is only a guideline. Depending on the context, the distribution of the sample mean can still be approximated using a normal distribution with a smaller sample size.

29.3 Estimation

In many cases, we are concerned with two population parameters, namely, the population mean (μ) and population variance (σ^2). So far, we have studied the distribution of the sample mean assuming complete knowledge of these parameters. In most situations, however, it is difficult to compute these parameters. Hence, we will often need to use sample statistics to help us estimate the population parameters.

29.3.1 Estimators and Estimates

Definition 29.3.1. An **estimator** is a method for estimating the quantity of interest. An **estimate** is a numerical estimate of the quantity of interest that results from the use of a particular estimator.

Example 29.3.2. Suppose our quantity of interest is the mean height μ of all male adults in Singapore. Suppose we take a random sample of 100 adult mean in Singapore and measure their heights.

Using this data, we can compute the sample average, $\overline{x}$ of the heights. That is, the sample mean random variable, $\overline{X} = \frac{1}{100}(X_1 + \cdots + X_{100})$, is an estimator that provides an estimate of our quantity of interest. For instance, if $\overline{x} = 170$ cm, then 170 cm is the estimate of μ provided by the "sample average" estimator.

Another strategy could be to use the "sample median" of the heights as an estimator. Suppose the sample median is 169 cm. Then 169 cm is the estimate of μ provided by the "sample median" estimator.

29.3.2 Unbiased Estimators

As illustrated by the above example, there are many estimators we can use to estimate μ. However, we would want to choose the estimator that performs the best. Logically, a good estimator should be *unbiased*. That is, the expected value of the estimator should be equal to the true value of the quantity it estimates.

Definition 29.3.3. If a population has an unknown parameter θ and T is a statistic derived from a random sample taken from the population, then T is an **unbiased estimator** for θ if and only if $\mathbb{E}[T] = \theta$.

Population Mean

Proposition 29.3.4. The sample mean $\overline{X} = \frac{1}{n}\sum x$ is an unbiased estimator for the population mean μ.

Proof. Previously, we showed that $\mathbb{E}\left[\overline{X}\right] = \mu$. Hence, by definition, $\overline{X}$ is an unbiased estimator for μ. □

Population Variance

Proposition 29.3.5. Let $\overline{x}$ be the sample mean. Then

$$s^2 = \frac{1}{n-1}\sum (x-\overline{x})^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{1}{n}\left(\sum x\right)^2\right]$$

is an unbiased estimator for the population variance σ^2.

Proof. We first show that the two forms of s^2 are equivalent:

$$\sum (x-\overline{x})^2 = \sum \left(x^2 - 2x\overline{x} + \overline{x}^2\right) = \sum x^2 - 2\overline{x}\sum x + n\overline{x}^2$$
$$= \sum x^2 - 2\left(\frac{1}{n}\sum x\right)\left(\sum x\right) + n\left(\frac{1}{n}\sum x\right)^2 = \sum x^2 - \frac{1}{n}\left(\sum x\right)^2.$$

Dividing throughout by $n-1$ gives us the desired equality. In fact, we can go one step further and write s^2 as

$$s^2 = \frac{1}{n-1}\left(\sum x^2 - n\overline{x}^2\right).$$

This is the form of σ^2 we will work with.

Before we process, we note that

$$\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 \implies \mathbb{E}\left[X^2\right] = \mu^2 + \sigma^2.$$

Similarly,

$$\operatorname{Var}\left[\overline{X}\right] = \mathbb{E}\left[\overline{X}^2\right] - \mathbb{E}\left[\overline{X}\right]^2 \implies \mathbb{E}\left[\overline{X}^2\right] = \mu^2 + \frac{\sigma^2}{n}.$$

Now consider $\mathbb{E}\left[S^2\right]$:

$$\mathbb{E}\left[S^2\right] = \mathbb{E}\left[\frac{1}{n-1}\left(\sum X^2 - n\overline{X}^2\right)\right] = \frac{1}{n-1}\left(\sum \mathbb{E}\left[X^2\right] - n\,\mathbb{E}\left[\overline{X}^2\right]\right)$$
$$= \frac{1}{n-1}\left[n\left(\mu^2+\sigma^2\right) - n\left(\mu^2 + \frac{\sigma^2}{n}\right)\right] = \sigma^2.$$

Hence, s^2 is an unbiased estimator for the population variance σ^2. $\square$

Note that the presence of $n-1$ in the denominator reflects the *degrees of freedom* we have when calculating s^2. We will elaborate more on this in the next chapter.

Corollary 29.3.6. If c is a constant, then

$$s^2 = \frac{1}{n-1}\left[\sum (x-c)^2 - \frac{1}{n}\left(\sum (x-c)\right)^2\right].$$

This is particularly useful when the sample data is given in summarized form.

Population Proportion

Definition 29.3.7. A **population proportion** p is a parameter that describes the percentage of individuals in a population that exhibit a certain property that we wish to investigate. Mathematically,

$$p = \frac{X}{N},$$

where X is the number of "successes" in the population (individuals who exhibit the property), and N is the population size. The sample proportion P_S is defined similarly:

$$P_S = \frac{X_S}{n},$$

where X_S is the number of "successes" in the sample.

Example 29.3.8. Suppose we wish to investigate the number of Singaporean citizens aged 35 years or older. The associated population parameter P is then calculated as

$$P = \frac{\text{number of Singaporean citizens aged 35 years or older}}{\text{total number of Singaporean citizens}}.$$

If we obtain a sample of 1000 Singapore citizens, of whom 750 are aged 35 years or older, then the observed sample proportion, which we denote $\hat{p}$, is simply $\hat{p} = 750/1000$.

Proposition 29.3.9. The sample proportion P_S is an unbiased estimator for the population proportion p.

Proof. Consider a population in which the proportion of "success" is p. If a random variable of size n is taken from this population, and X_S is the random variable denoting the number of "successes" in this sample, then

$$X_S \sim \mathrm{B}(n, p).$$

The expected value of P_S is thus

$$\mathbb{E}[P_S] = \mathbb{E}\left[\frac{X_S}{n}\right] = \frac{\mathbb{E}[X_S]}{n} = \frac{np}{n} = p.$$

Thus, P_S is an unbiased estimator for p. □

We can use the same idea to calculate $\mathrm{Var}[P_S]$:

$$\mathrm{Var}[P_S] = \mathrm{Var}\left[\frac{X_S}{n}\right] = \frac{\mathrm{Var}[X_S]}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

Hence, for large n, by the Central Limit Theorem, we have the following approximation:

$$P_S \sim \mathrm{N}\left(p, \frac{p(1-p)}{n}\right) \text{ approximately.}$$

The distribution of P_S is known as the **sampling distribution of the sample proportion** and its standard deviation, $\sqrt{p(1-p)/n}$, is known as the **standard error of proportion**.

30 Confidence Intervals

30.1 Definition

So far, we have seen how we can estimate an unknown population parameter from a random sample. For instance, if the parameter we seek to estimate is the mean μ, we can employ an unbiased estimator, i.e. the sample mean $\overline{x}$, to get a rough value for μ. This is what we call a **point estimate**. However, a point estimate does not provide any information about the uncertainty present. To this end, it is more desirable to obtain an interval estimate.

Definition 30.1.1. An **interval estimate** of an unknown population parameter is a random interval constructed so that it has a given probability of including the parameter.

This leads us to the notion of a confidence interval.

Definition 30.1.2. Given a fixed value $\alpha \in [0, 1]$ (known as the **level of significance**), a $100(1-\alpha)$**% confidence interval** for an unknown population parameter θ is any interval (a, b) such that

$$\mathbb{P}[a < \theta < b] = 1 - \alpha.$$

As an example, let us take $\alpha = 0.05$. If we can find a method of calculating the limits a and b, this means that in the long run, if we repeatedly take samples, then the calculated interval (a, b) will contain the population parameter θ for 95% of the samples taken. Equivalently, the probability of obtaining a random sample for which the corresponding interval contains θ is 0.95.

Note however, that for a particular sample, we do not know whether this is one of the samples for which θ is in the sample. Our "confidence" in the interval comes from the fact that we are using a formula which gives a correct result *most of the time*.

We can express the above notions diagrammatically:

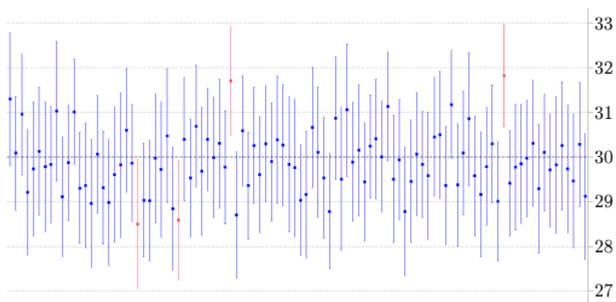


Figure 30.1: One hundred 95% confidence intervals for μ ($= 30$) computed from 100 different samples. Confidence intervals coloured red do not contain μ.[1]

[1]Source: https://amsi.org.au/ESA_Senior_Years/SeniorTopic4/4h/4h_2content_10.html

30.2 Population Mean

In this section, we explore interval estimates for the population mean μ.

Recall that for a significance level of α, we wish to find an interval (a, b) such that

$$\mathbb{P}[a < \mu < b] = 1 - \alpha.$$

To make our lives easier, we impose the restriction that the confidence interval be symmetric about μ, that is, the interval should be of the form $(\mu - E, \mu + E)$, where E is the **margin of error**. However, we obviously do not know μ, so we make use of the next best thing available: $\overline{x}$, to get something of the form

$$(\overline{x} - E, \overline{x} + E).$$

We thus wish to find the value of E such that

$$\mathbb{P}[\overline{x} - E < \mu < \overline{x} + E] = 1 - \alpha. \tag{30.1}$$

Depending on the situation, μ will be distributed differently, so E will differ accordingly.

There are four cases we will consider, with their respectively subsection numbers labelled in the table below:

<table>
<tr><td rowspan="2">σ^2</td><td rowspan="2">n</td><td colspan="2">Population Distribution</td></tr>
<tr><td>Normal</td><td>Unknown</td></tr>
<tr><td rowspan="2">Known</td><td>Large</td><td rowspan="2">§30.2.1</td><td>§30.2.2</td></tr>
<tr><td>Small</td><td></td></tr>
<tr><td rowspan="2">Unknown</td><td>Large</td><td colspan="2">§30.2.3</td></tr>
<tr><td>Small</td><td>§30.2.4</td><td></td></tr>
</table>

30.2.1 Normally Distributed Population with Known Variance

Suppose our population is normally distributed with unknown mean μ and known variance σ^2, so $X \sim \mathrm{N}(\mu, \sigma^2)$. In the previous chapter, we learnt that

$$\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right),$$

where n is the sample size. If we standardize this, we get

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}},$$

where Z is the standard normal distribution $\mathrm{N}(0, 1)$. Manipulating (30.1), we get

$$\mathbb{P}[\overline{x} - E < \mu < \overline{x} + E] = \mathbb{P}\left[-\frac{E}{\sigma/\sqrt{n}} < \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} < \frac{E}{\sigma/\sqrt{n}}\right] = 1 - \alpha.$$

But we recognize the middle expression as Z, so we really have

$$\mathbb{P}\left[-\frac{E}{\sigma/\sqrt{n}} < Z < \frac{E}{\sigma/\sqrt{n}}\right] = 1 - \alpha.$$

Because Z is symmetric about 0, we can finally isolate E:

$$\mathbb{P}\left[0 < Z < \frac{E}{\sigma/\sqrt{n}}\right] = \frac{1 - \alpha}{2} \implies \mathbb{P}\left[Z < \frac{E}{\sigma/\sqrt{n}}\right] = 1 - \frac{\alpha}{2}.$$

We now introduce some notation regarding z-values.

Definition 30.2.1. Given a probability $c \in [0,1]$, the **critical value** z_c is defined as

$$\mathbb{P}[Z < z_c] = c,$$

i.e. it acts as an "inverse" to the standard normal distribution.

With this notation, we can isolate our margin of error E:

$$\frac{E}{\sigma/\sqrt{n}} = z_{1-\frac{\alpha}{2}} \implies E = z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}.$$

We thus obtain the following result:

Proposition 30.2.2. If X is normally distributed and has known variance σ^2, then the symmetric $100(1-\alpha)\%$ confidence interval for μ is given by

$$\left(\overline{x} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \overline{x} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right).$$

The two limiting values that define the interval are known as the **$100(1-\alpha)\%$ lower and upper confidence limits**, sometimes writen as

$$\overline{x} \pm z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}.$$

Graphically, the area under $\mathrm{N}(\overline{x}, \sigma^2)$ over the confidence interval is $1-\alpha$:

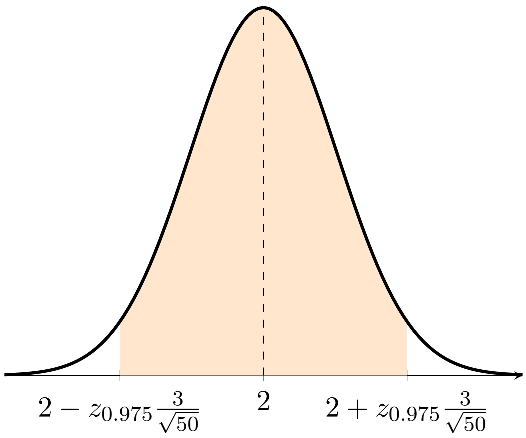


Figure 30.2: An illustration of a 95% confidence interval for $\overline{x} = 2$, $\sigma = 3$ and $n = 50$.

Sample Problem 30.2.3. After a rainy night, 12 worms surfaced on the lawn. Their lengths, measured in cm, were:

$$9.5, 9.5, 11.2, 10.6, 9.9, 11.1, 10.9, 9.8, 10.1, 10.2, 10.9, 11.0.$$

Assuming that this sample came from a normal population with variance 4, calculate a 99% confidence interval for the mean length of all worms in the garden.

Solution. Let X cm be the length of a worm. We have $\sigma = 2$ and $n = 12$. From the sample, we calculate $\overline{x} = 10.392$. Feeding this into the above expression, we see that a 99% confidence interval for the mean length of all worms in the garden is

$$\left(10.392 - z_{0.995} \frac{2}{\sqrt{12}}, 10.392 + z_{0.995} \frac{2}{\sqrt{12}}\right) = (8.90, 11.9).$$

□

30.2.2 Large Sample Size from Any Population with Known Variance

In the case where the sample size is large ($n \geq 30$), we can invoke the Central Limit Theorem, regardless of the distribution of the population. If X has variance σ^2, then we know from the previous chapter that

$$\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \text{ approximately.}$$

By a similar argument as in §30.2.1, we obtain the following (more general) result:

Proposition 30.2.4. If X has known variance σ^2 and the sample size is large ($n \geq 30$), then the symmetric $100(1-\alpha)\%$ confidence interval for μ is given by

$$\left(\overline{x} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{x} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right).$$

30.2.3 Large Sample Size from Any Population with Unknown Variance

In most practical situations, it is likely that both the mean and variance are unknown. Provided that the sample size is large ($n \geq 30$), by the Central Limit Theorem, we can say that the distribution of $\overline{X}$ is approximately normal. In place of the unknown population variance σ^2, we use s^2, the unbiased estimate of the population variance as an approximation. Hence,

$$\overline{X} \sim \mathrm{N}\left(\mu, \frac{s^2}{n}\right) \text{ approximately.}$$

Just like before, we get the following result:

Proposition 30.2.5. If X has unknown variance but the sample size is large ($n \geq 30$), then the symmetric $100(1-\alpha)\%$ confidence interval for μ is given by

$$\left(\overline{x} - z_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \overline{x} + z_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right).$$

30.2.4 Normally Distributed Population with Unknown Variance and Small Sample Size

Before looking at confidence intervals of μ when the sample size is small, we first need to consider the Student's t-distribution.

The t-distribution

The crucial statistic in the construction of a confidence interval for the mean of a normal distribution is Z, given by

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}.$$

In §30.2.3, when σ was unknown, we were able to σ by s by virtue of the large sample size, which allowed us to approximate $\overline{X}$ with a normal distribution.

In the present case, however, we do not have such a luxury. Now, when σ is replaced by S, the random variable

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

can no longer be apporixmated by a normal distribution. Here, T depends on two random variables: namely $\overline{X}$ and S, the random variable corresponding to s. Note that the value of T varies from sample to sample not only because of the variation in $\overline{X}$ as in the case of Z, but also because of the variation in S.

For samples of size n, it can be shown that

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim \mathrm{t}(n-1).$$

Note that this requires $X_1, \ldots, X_n$ to have independent and identical normal distributions.

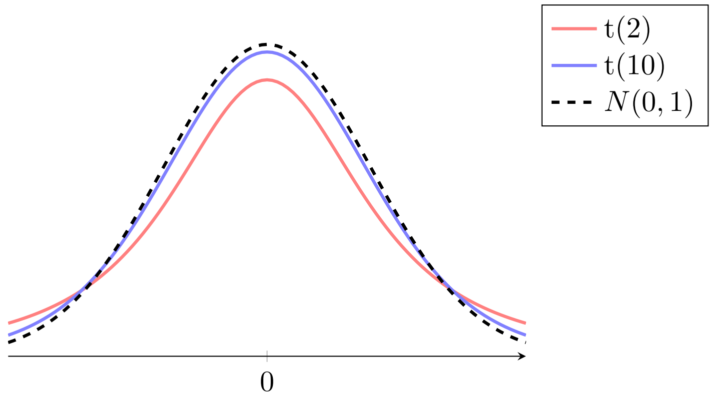


Figure 30.3: The t-distribution when $\nu = 2$ and $\nu = 10$. Observe that as ν increases, $\mathrm{t}(\nu)$ approaches $\mathrm{N}(0,1)$ in distribution.

The distribution of T is a member of a family of distributions known as t-distributions. All t-distributions are symmetric about 0 and have a single parameter, ν, which is a positive integer known as the **degrees of freedom** of the distribution. We notate this as $\mathrm{t}(\nu)$. As $\nu \to \infty$, the corresponding $\mathrm{t}(\nu)$ distribution approaches the standard normal distribution Z. In fact, when $\nu \geq 30$, the difference between the two is negligible, which explains why the normal distribution could continue to be used for cases where n was large in §30.2.3.

Why does T have $n-1$ degrees of freedom? Let us begin by introducing an informal definition of a degree of freedom.

Definition 30.2.6 (Informal)**.** The **degrees of freedom** of a statistic is the number of independent bits of information that are used in estimating the statistic.

In the present case, we initially have a total of n bits of information, namely our n observations $(X_1, \ldots, X_n)$. In order to estimate the value of our T statistic, we must first determine the value of the sample mean $\overline{X}$ and variance S. In an ideal world, both $\overline{X}$ and S would be allowed to vary independently. Unfortunately, S depends on the observed value of $\overline{X}$:[2]

$$s^2 = \frac{1}{n-1} \sum (x - \overline{x})^2 .$$

That is to say, we must estimate $\overline{X}$ in order to estimate S. We hence treat $\overline{x}$ as a constant, which we calculate as

$$\overline{x} = \frac{x_1 + \cdots + x_n}{n}.$$

But this effectively imposes a constraint on $x_1, \ldots, x_n$; if we somehow forgot our initial n observations after calculating $\overline{x}$, we would only need to remember $n-1$ observations. We thus have $n-1$ independent bits of information, so our degrees of freedom is $n-1$.

[2]Of course, we could have used the calculated value of s^2 to estimate $\overline{x}$. After working through the algebra, one will find that we still end up with $n-1$ degrees of freedom.

Confidence Interval using t-distribution

Suppose X is normally distributed with mean μ and unknown variance. Then

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim \mathrm{t}(\nu - 1),$$

where S is estimated by s. Once again, employing a similar argument as in §30.2.1, we obtain the following result:

Proposition 30.2.7. If X is normally distributed with unknown variance, then the symmetric $100(1-\alpha)\%$ confidence interval for μ is given by

$$\left(\overline{x} - t_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}, \overline{x} + t_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right).$$

Here t_c is the critical value for the t-distribution and is given by $PT < t_c = c$.

30.2.5 Summary

The following table shows the appropriate margin of error to be used in different scenarios when finding confidence intervals for the population mean. For conciseness, we use $c = 1 - \frac{\alpha}{2}$. Cells with gray backgrounds indicate an approximation.

<table>
<tr><th rowspan="2">σ^2</th><th rowspan="2">n</th><th colspan="2">Population Distribution</th></tr>
<tr><th>Normal</th><th>Unknown</th></tr>
<tr><td rowspan="2">Known</td><td>Large</td><td rowspan="2">$z_c \frac{\sigma}{\sqrt{n}}$</td><td>$z_c \frac{\sigma}{\sqrt{n}}$</td></tr>
<tr><td>Small</td><td></td></tr>
<tr><td rowspan="2">Unknown</td><td>Large</td><td colspan="2">$z_c \frac{s}{\sqrt{n}}$</td></tr>
<tr><td>Small</td><td>$t_c \frac{s}{\sqrt{n}}$</td><td></td></tr>
</table>

30.3 Population Parameter

Suppose we wish to find p, the proportion of "successes" in a population. For a large sample size n,

$$P_S \sim \mathrm{N}\left(p, \frac{p(1-p)}{n}\right) \text{ approximately,}$$

where P_S is the sample proportion. Standardizing, we see that

$$Z = \frac{P_S - p}{\sqrt{p(1-p)/n}}.$$

Notice the parallels with what we obtained in §30.2.1! Indeed, we can once again repeat our argument to obtain the following result:

Proposition 30.3.1. Given a sample proportion $\hat{s}$, the symmetric $100(1-\alpha)\%$ confidence interval for p is given by

$$\left(\hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}, \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\right).$$

Sample Problem 30.3.2. In a random sample of 400 carpet shops, it was discovered that 136 of them sold carpets at below the list prices recommended by the manufacturer. Calculate a 90% confidence interest for the proportion of shops that sell below list price.

Solution. Let p be the population proportion, and let the sample proportion be $P_S \sim \mathrm{N}(p, p(1-p)/n)$. We have $\hat{p} = 136/400$, so a 90% confidence interval for p is

$$\left(\frac{136}{400} - z_{0.95}\sqrt{\frac{\frac{136}{400}\left(1-\frac{136}{400}\right)}{400}}, \frac{136}{400} + z_{0.95}\sqrt{\frac{\frac{136}{400}\left(1-\frac{136}{400}\right)}{400}}\right) = (0.30104, 0.37896)\,.$$

□

31 Hypothesis Testing (Parametric)

Hypothesis testing is a statistical procedure used to determine if the data supports a particular assumption (hypothesis) about the population. In this chapter, we will examine various statistical tests employed in *parametric* hypothesis testing. Here, "parametric" means that we are given (or assuming) that the observed data have well-known distributions, such as the normal distribution. If we cannot make such assumptions, we will use a *non-parametric* test, which is covered in the next chapter.

31.1 An Introductory Example

Let us look at a simple example. The manufacturer of a beverage claims that each bottle they produce contains 500 ml of beverage on average. However, a consumer believes that the mean volume is actually smaller than claimed. To investigate this, the consumer takes a random sample of 30 bottles and finds that the mean volume of beverage in these 30 bottles is 498 ml.

The sample mean is certainly lower than the manufacturer's claim, but how low is too low? To answer this, we perform a hypothesis test.

Let X ml the volume of beverage in each bottle, and let the mean of X be μ, where μ is unknown. Assume that the standard deviation $\sigma = 5$, so that $X \sim \mathrm{N}(\mu, 25)$.

First, a hypothesis is made that $\mu = 500$ ml. This is known as the **null hypothesis**, H_0, and is written

$$\mathrm{H}_0: \quad \mu = 500.$$

Since it is suspected that the mean volume is *lower than* the claimed 500 ml, we establish the **alternative hypothesis**, H_1, which is that the mean is *lesser than* 500 ml. This is written

$$\mathrm{H}_1: \quad \mu < 500.$$

To carry out the test, the focus moves from X, the volume of liquid in each can, to the distribution of $\overline{X}$, the *mean* volume of a sample of 30 cans. In this test, $\overline{X}$ is known as the **test statistic** and its distribution is needed. Luckily for us, because we assumed that $X \sim \mathrm{N}(\mu, 25)$, we know from previous chapters that $\overline{X} \sim \mathrm{N}(\mu, 25/30)$.

The hypothesis test starts by assuming the null hypothesis is true, so $\mu = 500$. Under H_0,

$$\overline{X} \sim \mathrm{N}\left(500, \frac{25}{30}\right).$$

The result of the test depends on the whereabouts in the sampling distribution of the observed sample mean of $\overline{x} = 498$. We need to find out whether $\overline{x}$ is close to 500 or far away from 500. If $\overline{x}$ is close to 500, then it is likely that $\overline{x}$ comes from a distribution with mean 500, so there would not be enough evidence to say that the mean volume has decreased. On the other hand, if the $\overline{x}$ is far away from 500, than it is unlikely that $\overline{x}$ comes from a distribution with mean 500, so the mean μ is then likely to be lower than 500.

To quantify this "closeness", we can look at the **probability value** (also called **p-value**) associated with the test statistic $\overline{X}$. In our case, the p-value is $\mathbb{P}\big[\overline{X} \leq 498\big]$. A large p-value will indicate that if H_0: $\mu = 500$ is true, then obtaining a value of $\overline{x} = 498$ is likely and hence a reasonable variation we should allow. However, a small p-value will indicate that obtaining a value of $\overline{x} = 498$ is a rare event if H_0 is true, and hence, perhaps μ isn't 500, but something else (in this case, less than 500).

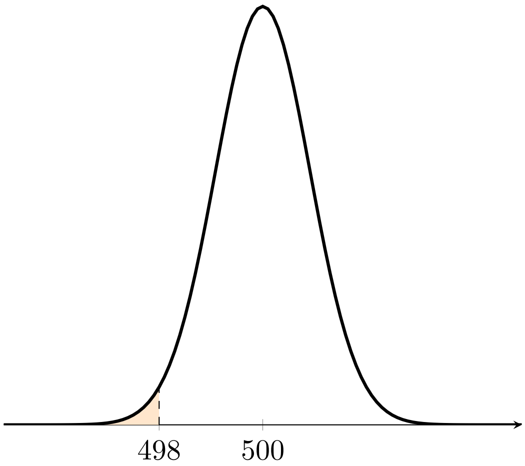


Figure 31.1: The p-value $\mathbb{P}\left[\overline{X} \leq 489\right]$ is given by the shaded area.

Note that whenever we use the test-statistic or p-value in this example, both are associated with the left tail of the distribution. This is because we began with the suspicion that μ was *lower* than claimed. This type of test is called a 1-tail (left tail) test.

To determine if the p-value is small enough, we introduce a cut-off point, c, known as the **critical value**, which indicates the boundary of the region where values of $\overline{x}$ would be considered *too far away* from 500 ml and therefore would be unlikely to occur. This region is known as the **critical/rejection region**. The probability corresponding to this critical region will then become the upper probability limit of what we will consider to imply that an unlikely or rare event has occurred. This probability, α, is called the **significance level** of the test. In general for a left tail test at the α level, the critical value c is fixed so that $\mathbb{P}\left[\overline{X} \leq c\right] = \alpha$ and the critical region is $\overline{x} \leq c$. In practice, to avoid being influenced by sample readings, it is important that α is decided before any samples values are taken.

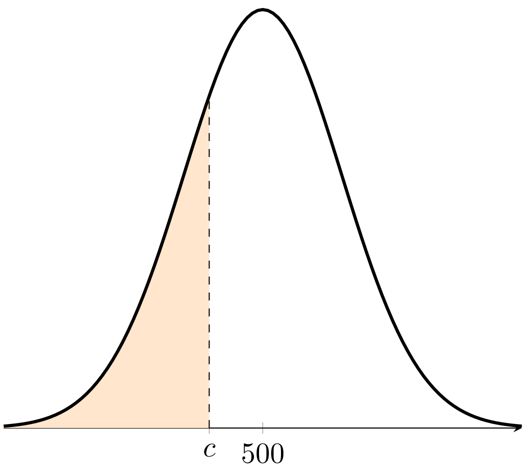


Figure 31.2: The critical region for $\alpha = 0.25$.

The hypothesis test then involves finding whether the sample value $\overline{x}$ lies in the critical region, or whether the p-value is smaller than the significance level α. If $\overline{x}$ lies in the critical region or if the p-value $\leq \alpha$, then a decision is taken that $\overline{x}$ is too far away from the mean associated with H_0 to have come from a distribution with this mean, hence we reject H_0 in favour of H_1. Else, if $\overline{x}$ lies outside the critical region or if the p-value $> \alpha$, we do not reject H_0. For a significance level of α, if the null hypotheses $H_0$is rejected, then the result is said to be **significant at the α level**.

To complete our example, suppose that a significance level of 1% is chosen. Since $\overline{X} \sim \mathrm{N}(500, 25/30)$, we can work out the critical value or the p-value.

Critical Value Approach Using G.C.,

$$\mathbb{P}\left[\overline{X} \leq c\right] = 0.01 \implies c = 497.88$$

Since $\overline{x} = 498$ lies outside the critical region ($\overline{x} = 498 > 497.88 = c$), we do not reject H_0 and conclude there is insufficient evidence at the 1% significance level than the mean volume of beverage in each bottle is lesser than 500 ml.

p-Value Approach The p-value of our sample is

$$\mathbb{P}\left[\overline{X} \leq 498\right] = 0.14230.$$

Since the p-value is greater than our significance level ($0.14230 > 0.01 = \alpha$), we do not reject H_0 and conclude there is insufficient evidence at the 1% significance level than the mean volume of beverage in each bottle is lesser than 500 ml.

31.2 Terminology

31.2.1 Formal Definitions of Statistical Terms

Definition 31.2.1. The **level of significance** of a hypothesis test, denoted by α, is defined as the probability of rejecting H_0 when H_0 is true.

Definition 31.2.2. The **p-value** is the probability of getting a test statistic as extreme or more extreme than the observed value. Equivalently, it is the lowest significance level at which H_0 is rejected.

31.2.2 Types of Tests

Suppose that the null hypothesis is H_0: $\mu = \mu_0$.[1]

There are three types of tests we can use, depending on what our alternative hypothesis looking for:

- If H_1 is looking for an increase in μ, we employ a 1-tail (right tail) test.
- If H_1 is looking for a decrease in μ, we employ a 1-tail (left tail) test.
- If H_1 is looking for a change (either increase or decrease) in μ, we employ a 2-tail test.

[1] In the introductory example, we saw how H_0 was defined to be the "status quo". However, this is not always the case. Given two hypotheses P and $\neg P$, the null hypothesis is the one that contains the *equality case*. For instance, if $P : \mu > 500$, then we take $\neg P : \mu \leq 500$ to be our null hypothesis, in which case we write H_0: $\mu = 500$ and H_1: $\mu > 500$.

1-Tail (Right Tail) Test

In a 1-tail (right tail) test, H_1: $\mu > \mu_0$. Both the critical region and p-value are in the right tail, with $\alpha = \mathbb{P}\big[\overline{X} \geq c\big]$ and the p-value $= \mathbb{P}\big[\overline{X} \geq \overline{x}\big]$.

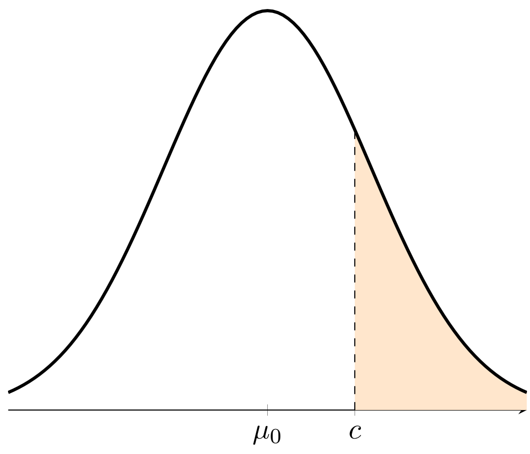


Figure 31.3: The critical region for a right tail test.

1-Tail (Left Tail) Test

In a 1-tail (left tail) test, H_1: $\mu < \mu_0$. Both the critical region and p-value are in the left tail, with $\alpha = \mathbb{P}\big[\overline{X} \leq c\big]$ and the p-value $= \mathbb{P}\big[\overline{X} \leq \overline{x}\big]$.

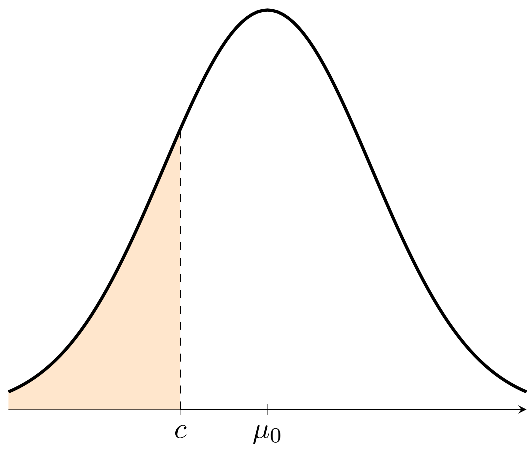


Figure 31.4: The critical region for a left tail test.

2-Tail Test

In a 2-tail test, H_1: $\mu \neq \mu_0$. The critical region and the p-value are in two parts. The critical value is given by any one of the following expressions

$$\alpha = \mathbb{P}\big[\overline{X} \leq c_1\big] + \mathbb{P}[\overline{x} \geq c_2] = 2\,\mathbb{P}\big[\overline{X} \leq c_1\big] = 2\,\mathbb{P}\big[\overline{X} \geq c_2\big]\,,$$

while the p-value is given by

$$p\text{-value} = \begin{cases} 2\,\mathbb{P}\big[\overline{X} \leq \overline{x}\big]\,, & \text{if } \overline{x} < \mu_0, \\ 2\,\mathbb{P}\big[\overline{X} \geq \overline{x}\big]\,, & \text{if } \overline{x} > \mu_0. \end{cases}.$$

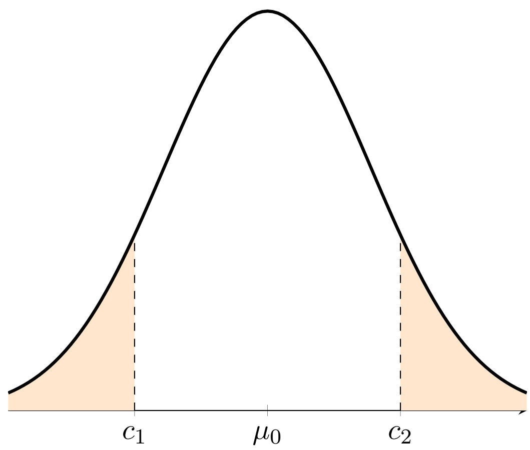


Figure 31.5: The critical region for a two-tail test.

31.2.3 Procedure

Below is a general framework for performing a hypothesis test.

Recipe 31.2.3 (Hypothesis Testing)**.**

(a) State the null hypothesis, H_0, and the alternative hypothesis, H_1.

(b) State the level of significance, α.

(c) Consider the distribution of the test statistic, assuming that H_0 is true.

(d) **Critical Value Approach.** Calculate the critical value based on α, and the test statistic value based on the sample data. Reject H_0 if the value of the test statistic falls in the critical region. Otherwise, do not reject H_0.

p-**Value Approach.** Calculate the p-value based on the sample data. Reject H_0 if the p-value $\leq \alpha$. Otherwise, do not reject H_0.

(e) Write down the conclusion in the context of the question.

Apart from step 3, the other steps are purely procedural. Hence, the most crucial step is to decide the test statistic. This is what we will focus on in the next few sections.

31.3 Population Mean

For hypothesis tests on the population mean, the test statistic is the sample mean $\overline{X}$. Similar to what we saw in §30.2, the following table shows the appropriate distribution to consider for different scenarios. Cells with gray backgrounds indicate an approximation.

σ^2	n	Population Distribution	
		Normal	Unknown
Known	Large	$\overline{X} \sim \mathrm{N}\left(\mu_0, \dfrac{\sigma^2}{n}\right)$	$\overline{X} \sim \mathrm{N}\left(\mu_0, \dfrac{\sigma^2}{n}\right)$
	Small		
Unknown	Large	$\overline{X} \sim \mathrm{N}\left(\mu_0, \dfrac{s^2}{n}\right)$	
	Small	$\dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim \mathrm{t}(n-1)$	

When our test statistic follows a normal distribution, we say that we perform a **z-test**. If instead our test statistic follows a t-distribution, we say that we perform a **t-test**.

Sample Problem 31.3.1. The lengths of metal bars produced by a particular machine are normally distributed with mean 420 cm and standard deviation 15 cm. After changing the machine specifications, a sample of 20 metal bars is taken and the length of each bar is measured. The result shows that the sample mean is 413 cm. Is there evidence, at the 5% significance level, that there is a change in the mean length of the metal bars?

Solution. Let X cm be the length of a metal bar after the machine specifications were changed. Our hypotheses are H_0: $\mu = 420$ and H_1: $\mu \neq 420$. We perform a 2-tail z-test at the 5% significance level. Under H_0, our test statistic is $\overline{X} \sim \mathrm{N}\left(420, 15^2/20\right)$. From the sample, $\overline{x} = 413$. Using G.C., the p-value is 0.0309, which is less than our significance level of 5%. Thus, we reject H_0 and conclude there is sufficient evidence at the 5% significance level that there is a change in the mean length of the metal bars. $\square$

31.3.1 Connection With Confidence Intervals

The testing of H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$ at a significance level $100\alpha\%$ is equivalent to computing a symmetric $100(1-\alpha)\%$ confidence interval for μ. If μ_0 is outside the confidence interval, H_0 is rejected. If μ_0 is within the confidence interval, H_0 is not rejected.

Sample Problem 31.3.2. In a study on the mathematical competencies of 15-year-old Singaporean students, the following PISA test results for a sample of 17 students is such that its sample mean is 565 with a sample standard deviation of 50. Find a 95% confidence interval for the population mean of the results of students for the PISA test. Hence, state the conclusion of a hypothesis test, at the 5% significance level, that tests if the mean of the test results for the Singaporean students differs from 600.

Solution. Let X be the random variable denoting the PISA test results of a 15-year-old Singaporean student. Our test statistic is

$$\frac{\overline{X} - 565}{S/\sqrt{17}} \sim \mathrm{t}(16).$$

From the sample, $s = 50$, so a symmetric 95% confidence interval for μ is $(539.29, 590.71)$. Since 600 is outside the confidence interval, we reject the null hypothesis that $\mu = 600$ and conclude there is sufficient evidence at a 5% significance level that the mean of the test results differ from 600. □

31.4 Difference of Population Means

In this section, we explore the distributions of the differences of population means. This is typically used when we are interested in comparing the population means from two populations. There is a major distinction we must make when we encounter such bivariate data:

Definition 31.4.1. If the data occurs in pairs, we say they are **paired**. Else, we say they are **unpaired**.

Example 31.4.2. Suppose we measure the blood pressure of a number of hospital patients before and after some treatment aimed at reducing blood pressure. Two values will be recorded from each patient, hence the data is paired.

However, if we measure the blood pressure of two groups of patients, one receiving treatment in Hospital A and the other in Hospital B, the data is unpaired.

There are some guidelines we can use to distinguish between paired and unpaired data:

- If the two samples are of unequal size, then they are unpaired.
- For data to be paired, there must be a reason to associate a particular measurement in one sample with a measurement in the other sample. If there is no reason to pair measurements in this way, the data is treated as unpaired.

31.4.1 Unpaired Samples

Let X_1 and X_2 be two random variables with random sample sizes n_1 and n_2, mean μ_1 and μ_2. In comparing the two populations, we typically set up our null hypothesis as H_0: $\mu_1 - \mu_2 = \mu_0$ with a one- or two-sided alternative hypothesis, similar to the single-value case discussed in the previous section.

When comparing unpaired data, one key assumption we typically make is that X_1 and X_2 are *independent*, as this allows us to formulate our test statistics nicely.

Known Population Variance

Suppose X_1 and X_2 have known variances σ_1^2 and σ_2^2 respectively. If X_1 and X_2 are normally distributed, then

$$\overline{X_1} \sim \mathrm{N}\left(\mu_1, \frac{\sigma_1^2}{n_1}\right) \quad \text{and} \quad \overline{X_2} \sim \mathrm{N}\left(\mu_2, \frac{\sigma_2^2}{n_2}\right).$$

If X_1 and X_2 are not normally distributed, then for large samples ($n_1, n_2 \geq 30$), by the Central Limit Theorem, we can approximate $\overline{X_1}$ and $\overline{X_2}$ using a normal distribution:

$$\overline{X_1} \sim \mathrm{N}\left(\mu_1, \frac{\sigma_1^2}{n_1}\right) \quad \text{and} \quad \overline{X_2} \sim \mathrm{N}\left(\mu_2, \frac{\sigma_2^2}{n_2}\right) \text{ approximately.}$$

Our test statistic is thus

$$\overline{X_1} - \overline{X_2} \sim \mathrm{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right),$$

and we proceed with the two-sample z-test.

Sample Problem 31.4.3. A random sample of size 100 is taken from a population with variance $\sigma_1^2 = 40$. Its sample mean $\overline{x_1}$ is 38.3. Another random sample of size 80 is taken from a population with variance $\sigma_2^2 = 30$. Its sample mean $\overline{x_2}$ is 40.1. Assuming that the two populations are independent, test, at the 5% level, whether there is a difference in the population means μ_1 and μ_2.

Solution. Our hypotheses are H_0: $\mu_1 - \mu_2 = 0$ and H_1: $\mu_1 - \mu_2 \neq 0$. Under H_0, our test statistic is

$$\overline{X_1} - \overline{X_2} \sim \mathrm{N}\left(0, \frac{40}{100} + \frac{30}{80}\right).$$

From the sample, $\overline{x_1} = 38.3$ and $\overline{x_2} = 40.1$. Using G.C., the p-value is 0.040888, which is less than our significance level of 5%. Thus, we reject H_0 and conclude there is sufficient evidence at the 5% level that there is a difference in the two population means. □

Unknown Population Variance with Large Sample Size

If we do not know the population variances of X_1 and X_2, we instead use the unbiased estimates s_1^2 and s_2^2. For large samples ($n_1, n_2 \geq 30$), we have, by the Central Limit Theorem, the following test-statistic:

$$\overline{X_1} - \overline{X_2} \sim \mathrm{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) \text{ approximately.}$$

If we know further that the two populations have common variance[2], i.e. $\sigma_1^2 = \sigma_2^2$, the **pooled variance**

$$s_p^2 = \frac{(n-1)\,s_1^2 + (n_2 - 1)\,s_2^2}{(n_1 - 1) + (n_2 - 2)} = \frac{\sum (x_1 - \overline{x_1})^2 + \sum (x_2 - \overline{x_2})^2}{(n_1 - 1) + (n_2 - 1)}$$

would provide a more precise estimate of the population variance. Our test statistic is hence

$$\overline{X_1} - \overline{X_2} \sim \mathrm{N}\left(\mu_1 - \mu_2, s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right) \text{ approximately.}$$

Either way, we proceed with the two-sample z-test.

Sample Problem 31.4.4. Two statistics teachers, Mr Tan and Mr Wee, argue about their abilities at golf. Mr Tan claims that with a number 7 ion he can hit the ball, on average, at least 10 m further than Mr Wee. Denoting the distance Mr Tan hits the ball by $(100 + c)$ m, the following results were obtained:

$$n_1 = 40, \quad \sum c = 80, \quad \sum (c - \overline{c})^2 = 1132.$$

Denoting the distance Mr Wee hits the ball by $(100 + t)$ m, the following results were obtained:

$$n_2 = 35, \quad \sum t = -175, \quad \sum \left(t - \overline{t}\right)^2 = 1197.$$

If the distances for both teachers have a common variance, test whether there is any evidence at the 1% level, to support Mr Tan's claim.

Solution. Let X_1 and X_2 be the random variable denoting the distance, in m, for Mr Tan and Mr Wee, with population mean μ_1 and μ_2 respectively. From the data, we have

$$\overline{x_1} = 100 + \frac{80}{40} = 102 \quad \text{and} \quad \overline{x_2} = 100 + \frac{-175}{35} = 95,$$

[2]As a rule of thumb, the assumption $\sigma_1 = \sigma_2$ is considered reasonable if $1/2 \leq s_1/s_2 \leq 2$.

so the pooled variance is

$$s_p^2 = \frac{\sum (x_1 - \overline{x_1})^2 + \sum (x_2 - \overline{x_2})^2}{(n_1 - 1) + (n_2 - 1)} = \frac{1132 + 1197}{(30 - 1) + (35 - 1)} = 31.90.$$

We now perform a two-sample z-test at the 1% level. Our hypotheses are H_0: $\mu_1 - \mu_2 = 10$ and $\mu_1 - \mu_2 < 10$. Under H_0, our test statistic is

$$\overline{X_1} - \overline{X_2} \sim \mathrm{N}\left(\mu_1 - \mu_2, s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right) = \mathrm{N}(10, 1.70915).$$

Using G.C., the p-value is 0.0109, which is greater than our significance level of 1%. Thus, we do not reject H_0 and conclude there is insufficient evidence to suppose Mr Tan's claim. □

Unknown Population Variance with Small Sample Size

If the random sample sizes are not large, then the normal distribution is no longer a reasonable approximation to the distribution of the test statistic. In order to progress, we must have the following assumptions:

- X_1 and X_2 have independent, normal distributions.
- X_1 and X_2 have a common variance.

With these assumptions, it can be shown that the test statistic T given by

$$T = \frac{\left(\overline{X_1} - \overline{X_2}\right) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_1}}} \sim \mathrm{t}((n_1 - 1) + (n_2 - 1)),$$

where

$$S_p^2 = \frac{(n - 1)\, S_1^2 + (n_2 - 1)\, S_2^2}{(n_1 - 1) + (n_2 - 2)}$$

is the pooled variance (unbiased estimate of the common variance). Note that there we lose 2 degrees of freedom since we use both $\overline{x_1}$ and $\overline{x_2}$ to estimate s_1^2 and s_2^2.

Sample Problem 31.4.5. The heights (measured to the nearest cm) of a random sample of six policemen from country A were found to be

176, 180, 179, 181, 183, 179.

The heights (measured to the nearest cm) of a random sample of eleven policemen from country B have the following data:

$$\sum y = 1991, \quad \sum (y - \overline{y})^2 = 54.$$

Test, at the 5% level, the hypothesis that policemen from country A are shorter than policemen from country B. State any assumptions that are needed for this test.

Solution. Let X_A and X_B be the height in cm of a policeman from country A and B, with population mean μ_A and μ_B respectively. We assume that X_A and X_B have independent, normal distributions, and they share a common variance. Our hypotheses are H_0: $\mu_A - \mu_B = 0$ and H_1: $\mu_A - \mu_B < 0$. Under H_0, our test statistic is

$$T = \frac{\overline{X_A} - \overline{X_B}}{S_p \sqrt{\frac{1}{6} + \frac{1}{11}}} \sim \mathrm{t}(15).$$

From the sample,

$$\overline{x_A} = 179.67 \quad \text{and} \quad \overline{x_B} = \frac{1991}{11} = 81.$$

The unbiased estimates of each sample variance is

$$s_A^2 = 5.4667 \quad \text{and} \quad s_B^2 = \frac{1}{10}\sum(y-\overline{y})^2 = 5.4.$$

Thus, the pooled variance is

$$s_p^2 = \frac{(6-1)(5.4667)+(11-1)(5.4)}{(6-1)+(11-1)} = 5.4222.$$

Using G.C., the p-value is 0.139, which is greater than our significance level of 5%. Thus, we do not reject H_0 and conclude there is insufficient evidence to claim that policemen from country A are shorter than policemen from country B. □

31.4.2 Paired Samples

If the given data is paired, then the two populations are no longer independent, hence we cannot use any of the tests previously discussed. Instead, we will now consider the difference $D = X_1 - X_2$, which is calculated for each matched pair. Writing μ_D for the mean of the distribution of differences between the paired values, our null hypothesis is H_0: $\mu_D = \mu_0$ with a one-sided or two-sided H_1 as appropriate.

Notice that by working with the differences, we have effectively reduced our problem into a single sample situation, so the usual hypothesis test considerations for a single sample mean applies. For instance, if D can be presumed to be normally distributed, or if n is sufficiently large that the Central Limit Theorem can be applied to approximate D to have a normal distribution, then

$$\overline{D} \sim \mathrm{N}\left(\mu_D, \frac{s_D^2}{n}\right),$$

and we proceed with a paired-sample z-test. Alternatively, if D can be presumed to have a normal distribution, but n is small, then the test statistic

$$T = \frac{\overline{D}-\mu_D}{S_D/\sqrt{n}} \sim \mathrm{t}(n-1)$$

can be used. In this case, we proceed with a paired-sample t-test.

31.5 χ^2 Tests

31.5.1 The χ^2 Distribution

The χ^2 distribution is a continuous distribution with a positive integer parameter ν.

Definition 31.5.1. The sum of the squares of ν independent standard normal random variables $Z_1, \ldots, Z_\nu$ is distributed according to a **χ^2 distribution ν degrees of freedom**, denoted χ^2_ν.

$$Z_1^2 + \cdots + Z_\nu^2 \sim \chi^2_\nu.$$

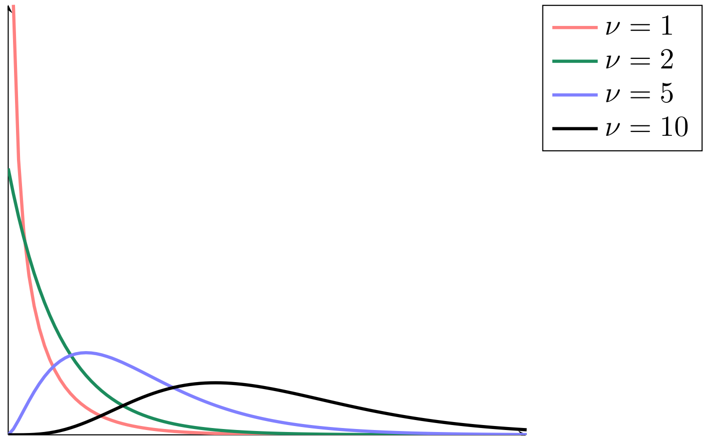


Figure 31.6: The χ^2_ν distribution for varying values of ν.

The χ^2 distribution has a reverse "J"-shape for $\nu = 1, 2$, and is positively skewed for $\nu > 2$. As ν increase, the distribution becomes more symmetric. For large ν, the distribution is approximately normal.

Fact 31.5.2 (Properties of the χ^2 Distribution)**.**

- A χ^2_ν distribution has mean ν and variance 2ν.
- A χ^2_ν distribution has mode $\nu - 2$ for $\nu \geq 2$.
- If U and V are independent random variables such that $U \sim \chi^2_u$ and $V \sim \chi^2_v$, then $U + V \sim \chi^2_{u+v}$.

31.5.2 χ^2 Goodness-of-Fit Test

Previously, we have always assumed that a particular type of distribution is appropriate for the data given and have focused on estimating and testing hypotheses about the parameter of the distribution. In this section, the focus changes to the distribution itself, and we ask "Does the data support the assumption that a particular type of distribution is appropriate?"

As a motivating example, suppose we roll a six-sided die 60 times and obtain the following observed frequencies:

Outcome	1	2	3	4	5	6
Observed frequency, O	4	7	16	8	8	17

In this sample, there seems to be a rather large number of 3's and 6's. Is this die fair, or is it biased? With a fair die, the expected frequencies would each be $60/6 = 10$.

Outcome	1	2	3	4	5	6
Expected frequency, E	10	10	10	10	10	10

The question is thus whether the observed frequencies O and the expected frequencies E are reasonably close or unreasonably different. An obvious comparison would be the differences $(O - E)$:

Outcome	1	2	3	4	5	6
Observed frequency, O	4	7	16	8	8	17
Expected frequency, E	10	10	10	10	10	10
Difference, $O - E$	-6	-3	6	-2	-2	7

The larger the magnitude of the differences, the more the observed data differs from the model that the die was fair.

Suppose we now roll a second die 660 times and obtain the following results:

Outcome	1	2	3	4	5	6
Observed frequency, O	104	107	116	108	108	117
Expected frequency, E	110	110	110	110	110	110
Difference, $O - E$	-6	-3	6	-2	-2	7

This time, the observed and expected frequencies seem close, yet the differences $O - E$ are the same as before. We see that it is not just the size of $O - E$ that matters, but also its relative size to the expected frequency $(O - E)/E$.

Combining the ideas, the goodness-of-fit for an outcome i is measured using

$$(O_i - E_i) \cdot \frac{O_i - E_i}{E_i} = \frac{(O_i - E_i)^2}{E_i}.$$

The smaller this quantity is, the better the fit. An aggregate measure of goodness-of-fit of the model is thus given by the χ^2 statistic:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}.$$

As the name suggests, this test statistic follows a χ^2 distribution.

Observe that if $\chi^2 = 0$, there is exact agreement between O_i and E_i, so the model is a perfect fit. If $\chi^2 > 0$, then O_i and E_i do not agree exactly. The larger the value of χ^2, the greater the discrepancy.

For the test, we define H_0 as our sample having the expected probabilities of the various categories. The alternative hypothesis H_1 will be that H_0 is incorrect, i.e. the sample does not have the expected probabilities of the various categories. We use the χ^2 test statistic, which generally follows a χ^2_{m-1-k} distribution, where m is the number of categories being compared, and k is the number of parameters estimated from the data.

Example 31.5.3. Suppose we wish to test if a given set of data fits a Poisson model. If we are not given the mean rate λ, we can estimate it using $\overline{x} \approx \lambda$. In doing so, we lose one degree of freedom, so the resulting χ^2 test statistic will follow a χ^2_{m-2} distribution.

Example 31.5.4. To formalize our motivating example, we define H_0: the die is fair, and H_1: the die is not fair. We take a 2.5% level of significance. Our test statistic is

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{6-1} = \chi^2_5.$$

From the sample, the individual contributions are given by

Outcome	1	2	3	4	5	6
O_i	4	7	16	8	8	17
E_i	10	10	10	10	10	10
$(O_i - E_i)^2/E_i$	3.6	0.9	3.6	0.4	0.4	4.9

The test statistic value is thus

$$3.6 + 0.9 + 3.6 + 0.4 + 0.4 + 0.9 = 13.8.$$

Using G.C., the p-value is

$$\mathbb{P}\left[\chi^2 \geq 13.8\right] = 0.016931.$$

Since the p-value is less than our significance level of 2.5%, we reject H_0 and conclude there is sufficient evidence at the 2.5% significance level that the die is not fair.

Small Expected Frequencies

The distribution of $\sum(O_i - E_i)^2/E_i$ is discrete. The continuous χ^2 distribution is simply a convenient approximation which becomes less accurate as the expected frequencies become smaller. Generally, the approximation may be used only when *all expected frequencies are less than 5.* If a category has an expected frequency less than 5, we must combine it with other categories. This combination may be done in any sensible grounds, but should be done without reference to the observed frequencies to avoid bias.

Sample Problem 31.5.5. A random sample of 40 observations on the discrete random variable X is summarized below:

x	0	1	2	3	4	≥ 5
Frequency	4	14	9	7	6	0

Test, at the 5% significance level, whether X has a Poisson distribution with mean equal to 2.

Solution. Our hypotheses are H_0: the data is consistent with a Po(2) model, and H_1: the data is inconsistent with a Po(2) model. From the given data, the observed and expected frequencies are

x	0	1	2	3	4	≥ 5
O_i	4	14	9	7	6	0
E_i	5.4143	10.821	10.827	7.2179	3.6089	2.1061

The last two categories have expected frequencies less than 5, so we combine them into a single category:

x	0	1	2	3	≥ 4
O_i	4	14	9	7	6
E_i	5.4143	10.821	10.827	7.2179	5.7151

Our test statistic is

$$\sum \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{5-1}.$$

Using G.C., the p-value is 0.80373, which is larger than our 5% significance level, thus we do not reject H_0 and conclude there is insufficient evidence that the data is inconsistent with a Po(2) model. □

In general, we have the following procedure:

Recipe 31.5.6 (χ^2 Goodness-of-Fit Test).

1. State hypotheses and significance level.
2. Compute expected frequencies under H_0.
3. Combine any categories if there are expected frequencies under 5.
4. Determine the degrees of freedom and state the test statistic.
5. Calculate the p-value.
6. State the conclusion of the test in context.

31.5.3 χ^2 Test for Independence

Suppose we record data concerning two categorical variables for a sample of individuals chosen randomly from a population. It is convenient to display the data in the form of a **contingency table**. Here is an example which shows information on voting:

	Party A	Party B	Party C	Total
Male	313	124	391	828
Female	344	158	388	890
Total	657	282	779	1718

Sample data of this type are collected in order to answer interesting questions about the behaviour of the population, such as "Are there differences in the way males and females vote?" If there are differences, then the variables "vote" and "gender" are said to be **associated**, else they are **independent**.

To test for independence between variables, we employ a χ^2 test for independence. Our null hypothesis is that the variables are independent, while our alternative hypothesis is that the variables are associated.

Under the null hypothesis, the best estimate of the population proportion voting for Party A is 657/1718. The expected number of males voting for Party A would thus be $828 \times 657/1718 = 316.64$, and the number of females would be $890 \times 657/1718 = 340.36$. These expected frequencies, E_i, are calculated using the formula

$$E = \frac{\text{row total} \times \text{column total}}{\text{grand total}}.$$

Doing this for all combination of party and gender, we get the following table of expected frequencies:

	Expected Frequencies		
	Party A	Party B	Party C
Male	316.64	135.91	375.44
Female	340.36	146.09	403.56

The test statistic $\sum(O_i - E_i)^2/E_i$ is computed and compared with the relevant χ^2 distribution. For a contingency table with r rows and c columns, the degrees of freedom

ν is given by

$$\nu = (r-1)(c-1),$$

since we only need $(r-1)(c-1)$ values to completely determine the entire table (try it!). In our case, $\nu = (2-1)(3-1) = 2$.

In general, we have the following procedure:

Recipe 31.5.7 (χ^2 Test for Independence)**.**

1. State hypotheses and significance level.
2. Compute expected frequencies under H_0 and tabulate them.
3. Combine any rows/columns if there are expected frequencies under 5.
4. Determine the degrees of freedom and state the test statistic.
5. Calculate the p-value.
6. State the conclusion of the test in context.

32 Hypothesis Testing (Non-Parametric)

Previously, we examined tests that require certain assumptions about the underlying distribution from which the data arises. Tests which do not require such assumptions are called *non-parametric*. Note that non-parametric tests are generally less powerful than the equivalent parametric tests, especially if the assumptions required by the parametric tests can be justified.

32.1 Sign Test

32.1.1 Single Sample

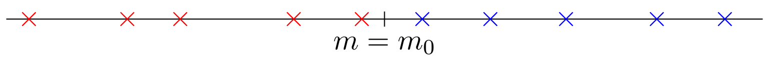
Consider a random sample of size n from a population which has a continuous distribution with median m. We are interested in whether the median m takes on a particular value m_0. That is, we are interested in testing the null hypothesis

$$\mathrm{H}_0 : m = m_0$$

against any of the possible alternative hypotheses:

$$\mathrm{H}_1 : m > m_0 \quad \mathrm{H}_1 : m < m_0 \quad \mathrm{H}_1 : m \neq m_0.$$

Define K_+ to be the number of data values greater than m_0, and K_- to be the number of data values smaller than m_0. Under H_0, we expect about the same number of data values that are greater than m_0 and less than m_0.

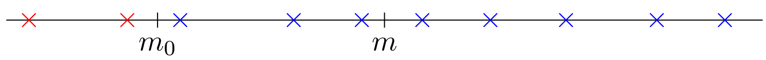


Hence, our test statistic is either

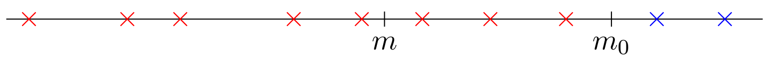
$$K_+ \sim \mathrm{B}\left(n, \frac{1}{2}\right) \quad \text{or} \quad K_- \sim \mathrm{B}\left(n, \frac{1}{2}\right),$$

depending on which is more convenient. For now, we take K_+ to be our test statistic.

If we test H_0 against H_1: $m > m_0$, then we reject H_0 if the observed number of data values greater than m_0 is too large, i.e. $k_+ \geq c_+$ for some critical value c_+. Alternatively, we can consider the p-value, which is given by $\mathbb{P}[K_+ \geq k_+]$. If this p-value is smaller than our significance level α, we reject H_0.



If we test H_0 against H_1: $m < m_0$, then we reject H_0 if the observed k_+ is too small. Alternatively, if the p-value $\mathbb{P}[K_+ \leq k_+]$ is smaller than our significance level α, we reject H_0.



Lastly, if we test H_0 against H_1: $m \neq m_0$, then we reject H_0 if the observed k_+ is too small or too large. In this case, the p-value is given by

$$2\min\{\mathbb{P}[K_+ \geq k_+]\,, \mathbb{P}[K_+ \leq k_+]\}\,.$$

Note that we choose the shorter tail since we want the more "extreme" end.

To summarize,

H_0	$m = m_0$		
H_1	$m > m_0$	$m < m_0$	$m \neq m_0$
p**-value** (K_+)	$\mathbb{P}[K_+ \geq k_+]$	$\mathbb{P}[K_+ \leq k_+]$	$2\min\{\mathbb{P}[K_+ \geq k_+]\,, \mathbb{P}[K_+ \leq k_+]\}$
p**-value** (K_-)	$\mathbb{P}[K_- \leq k_-]$	$\mathbb{P}[K_- \geq k_-]$	$2\min\{\mathbb{P}[K_- \geq k_-]\,, \mathbb{P}[K_- \leq k_-]\}$

In the case where there are zeroes, we discard them and reduce the sample size accordingly.

Sample Problem 32.1.1. The lifetimes of a random sample of candles, measured in minutes are

$$354, \quad 358, \quad 348, \quad 342, \quad 352, \quad 335, \quad 364, \quad 345, \quad 360, \quad 341.$$

The manufacturer claims that the median lifetime is at least 360 minutes. Use a sign test, at the 5% significance level, to test whether the manufacturer's claim is justified.

Solution. Let m be the population median. Our hypotheses are H_0: $m = 360$ and H_1: $m < 360$. We take a 5% level of significance. Subtracting the observed data values by the postulated median $m = 360$ and writing down the signs, we obtain

$$-, \quad -, \quad -, \quad -, \quad -, \quad -, \quad +, \quad -, \quad 0, \quad -.$$

Let K_+ be the number of data values greater than 360. Discarding the zero, we have, under H_0, $K_+ \sim \mathrm{B}(9, 1/2)$. From the sample, $k_+ = 1$. The p-value is hence $\mathbb{P}[K_+ \leq 1] = 0.0195$. Since the p-value is smaller than our 5% significance level, we reject H_0 and conclude there is sufficient evidence at the 5% level that the manufacturer's claim is not justified. □

32.1.2 Paired Sample

By considering the difference in population medians, the sign test can be used for paired samples, as demonstrated in the example below.

Sample Problem 32.1.2. Students in a school take a mock examination before taking the actual A-level examination. The marks for a particular subject, in both the mock and actual examinations, by a random sample of 13 students are shown below.

Candidate Number	1	2	3	4	5	6	7	8	9	10	11	12	13
Mock Exam Mark	40	65	53	79	87	42	80	63	51	82	27	71	29
Actual Exam Mark	45	68	47	75	88	60	77	69	60	88	30	73	35

Test, at the 5% level, whether the candidates did better in the actual A-level than in the mock examination for this subject.

Solution. Let m be the population median mark difference of (Actual $-$ Mock). Our hypotheses are H_0: $m = 0$ and H_1: $m > 0$. We take a 5% level of significance. Subtracting matched pairs of (Actual $-$ Mock) and writing down the signs, we obtain

$$+, \quad +, \quad -, \quad -, \quad +, \quad +, \quad -, \quad +, \quad +, \quad +, \quad +, \quad +, \quad +.$$

Let K_+ be the number of data values greater than 0. Under H_0, $K_+ \sim \mathrm{B}(13, 1/2)$. From the sample, $k_+ = 10$. The p-value is hence $\mathbb{P}[K_+ \geq 10] = 0.0461$, which is greater than our 5% significance level. Hence, we reject H_0 and conclude there is sufficient evidence at the 5% level that the students did better in the actual A-level examination. □

32.1.3 Large Sample

Let $X \sim \mathrm{B}(n, 1/2)$. For large n ($n \geq 30$), we can approximate X with a normal distribution via the Central Limit Theorem:

$$X \sim \mathrm{N}\left(\frac{n}{2}, \frac{n}{4}\right) \text{ approximately.}$$

This is useful when conducting a sign test with a large sample.

32.2 Wilcoxon Matched-Pair Signed Rank Test

When testing paired samples, one drawback of using the sign test is that it only takes into account the sign of the differences between paired values. To see how this might be problematic, consider the following set of differences:

Magnitude of Difference	7	2	6	4	22	15	5	1	12	16
Sign of Difference	+	−	+	+	+	+	+	−	+	+

We see that negative differences are very small (e.g. -1, -2) as compared to some of the positive differences (e.g. 22, 16).

The Wilcoxon matched-pair signed rank test improves on the sign test by considering the magnitude of the differences. This is done by ranking the magnitudes of the differences in ascending order, starting with rank 1. For instance, the ranks for the above example are given by

Magnitude of Difference	7	2	6	4	22	15	5	1	12	16
Sign of Difference	+	−	+	+	+	+	+	−	+	+
Rank	6	2	5	3	10	8	4	1	7	9

Let P be the sum of the ranks corresponding to the positive differences and let Q be the sum of the ranks corresponding to the negative differences. Let m be the population median. Our null hypothesis is H_0: $m = 0$. From here, the main idea is

- If we test H_1: $m > 0$, we reject H_0 if Q is too small, i.e. $q \leq c_-$ for some critical value c_-.
- If we test H_1: $m < 0$, we reject H_0 if P is too small, i.e. $p \leq c_+$ for some critical value c_+.
- If we test H_1: $m \neq 0$, we reject H_0 when either P or Q is too small.

In all cases above, we can either choose our test statistic T to be either P or Q. Typically, we take T to be the smaller of two, as demonstrated above.

For small n, the critical value can be found in the provided formula list.

Sample Problem 32.2.1. Eight strands of wires were tested for their breaking points and then were retested after they were rusted. The breaking points were recorded as follows:

Non-Rusted	9.4	8.1	6.6	9.9	8.7	8.3	7.0	7.5
Rusted	7.2	5.4	7.1	8.1	7.0	7.9	8.5	6.2

Carry out a Wilcoxon matched-pair signed rank test at the 5% level of significance to determine whether, on average, the rusted wires have lower breaking points.

Solution. Let m be the population median difference of (Non-Rusted − Rusted). Our hypotheses are H_0: $m = 0$ and H_1: $m > 0$. We take a 5% significance level.

Non-Rusted	9.4	8.1	6.6	9.9	8.7	8.3	7.0	7.5
Rusted	7.2	5.4	7.1	8.1	7.0	7.9	8.5	6.2
NR − R	2.2	2.7	−0.5	1.8	1.7	0.4	−1.5	1.3
Rank	8	7	2	6	5	1	4	3

Let P be the sum of ranks corresponding to positive differences, and let Q be the sum of ranks corresponding to negative differences. Let T be the smaller of the two. From the above table, we see that $p = 6$ and $q = 30$, so $t = 6$. From the formula list, we reject H_0 if $t \leq 5$. Since $t = 6 > 5$, we do not reject H_0 and conclude there is insufficient evidence at the 5% level that the rusted wires have lower breaking points. □

32.2.1 Large Sample

For large n ($n > 20$), the test statistic T can be approximated with a normal distribution via the Central Limit Theorem:

$$T \sim \mathrm{N}\left(\frac{n(n+1)}{4}, \frac{n(n+1)(2n+1)}{24}\right) \text{ approximately.}$$

With this approximation, we can calculate the appropriate p-value. Note that T can either be P or Q.

32.3 Comparison of the Tests

The sign test and the Wilcoxon matched-pair signed rank test do not always produce the same results.

The advantage of the Wilcoxon matched-pair signed rank test compared to the sign test is that it takes into account the magnitude of the differences of the matched observations as well as the signs of the difference. Thus, it is a more powerful test than the sign test.

However, one disadvantage of the Wilcoxon matched-pair signed rank test compared to the sign test is that it requires an additional assumption that the distribution of the differences must be symmetric about the median zero.

33 Correlation and Regression

Correlation and regression are statistical methods that examine the relationship between two quantitative variables.

Correlation is concerned with quantifying the (linear) *relationship* between two variables. Informally, it allows us to tell how strongly two variables move with each other. For instance, suppose we measure the heights and weights of a group of people. Intuitively, we would expect taller people to be heavier, hence there is a positive correlation between height and weight.

Regression, on the other hand, is concerned with quantifying how a change in one variable will affect the other variable. That is, regression predicts the value of a variable based on the value of the other variable. Reusing our previous example, regression allows us to predict the height of a person that weighs 70 kg.

33.1 Independent and Dependent Variables

When performing correlation and regression analysis, we need two sets of data, one for each variable. The resulting data is called bivariate data. A set of n bivariate data can be expressed using ordered pairs (x_i, y_i), where x and y are the two variables.

Definition 33.1.1. In a bivariate relationship, the **independent variable** is the one that does not rely on changes in another variable, while the **dependent variable** is the one that depends on or changes in response to the independent variable.

Informally, the independent variable is the variable we can "control" in an experiment, allowing us to vary its value to observe the resulting change in the value of the dependent variable.

Recipe 33.1.2. To determine if there exists an independent/dependent relationship between two variables x and y, we look at

- The context of the question – Does one variable depend on the other?
- Key phrases in the question, e.g. "investigate how A depends on B" means that B is likely the independent variable and A the dependent variable.
- Fixed or controlled variable in an experiment – If a variable is manipulated in fixed increments, it is likely to be independent variable.

Note however, that not all bivariate relationships have an independent and dependent variable. For instance, consider the following example:

Example 33.1.3. Six newly-born babies were randomly selected. Their head circumference x cm, and body length, y cm were measured by the paediatrician and tabulated.

x	31	32	33.5	34	35.5	36
y	45	49	47	50	53	51

All three heuristics for determining the independent/dependent relationship between x and y are not applicable. Hence, we say there is no clear independent and dependent variables, and we assume that no such relationship exists between the two variables.

33.2 Scatter Diagram

A scatter diagram is obtained when each pair of data value (x_i, y_i) from a set of bivariate sample $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ is plotted as a point on an x-y graph.

Recipe 33.2.1 (Drawing a Scatter Diagram). When drawing a scatter diagram, note that

- data points should be marked with a cross ($\times$);
- axes need not start from 0;
- axes need to be labelled according to context;
- the range of data values and the relative scale of the axes need to be indicated;
- the relative position of the points should be accurate.

Example 33.2.2. The number of employees, y, who stay back and continue in the office t minutes after 5 pm on a particular day in a company is recorded. The results are shown in the table.

t	15	30	45	60	75	90	105
y	30	19	15	13	12	11	10

Plotting the above points, we get our scatter diagram:

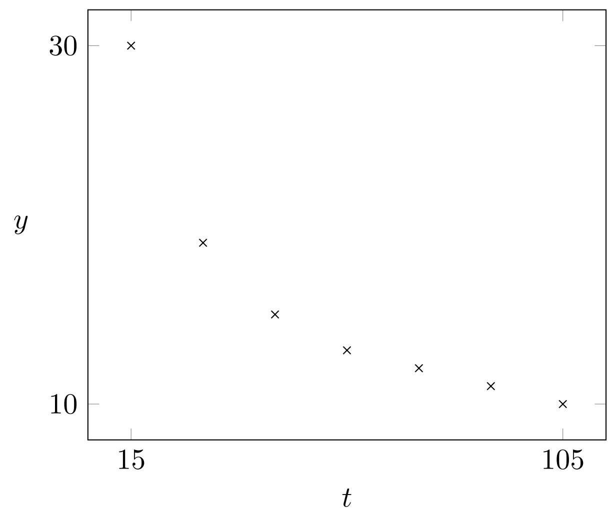


Figure 33.1

33.2.1 Interpreting Scatter Diagrams

There are four main relationships we can observe on a scatter diagram:

- Positive linear relationship – As x increases, y increases.
- Negative linear relationship – As x increases, y decreases.
- Curvilinear relationship – The points seem to lie on a curve.
- No clear relationship – The points seem to be randomly scattered.

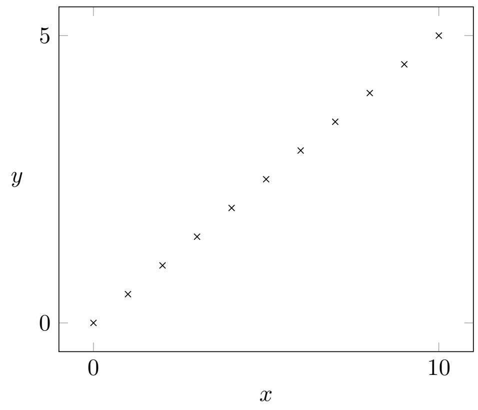


Figure 33.2: Positive linear relationship.

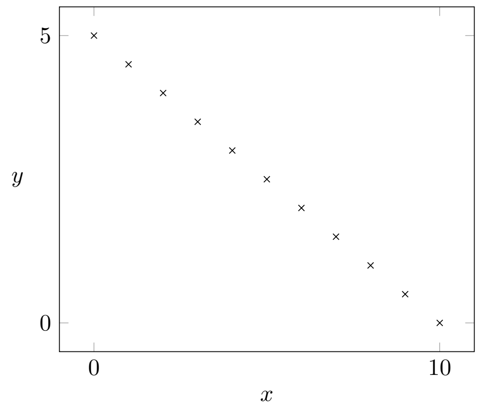


Figure 33.3: Negative linear relationship.

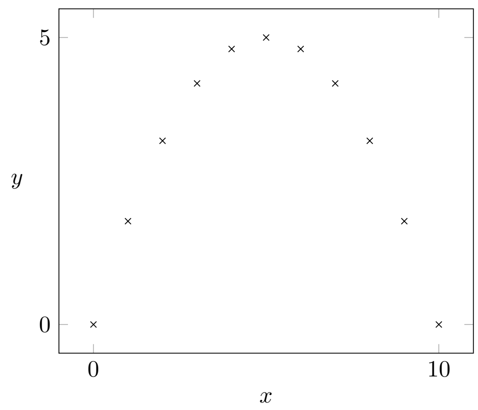


Figure 33.4: Curvilinear linear relationship.

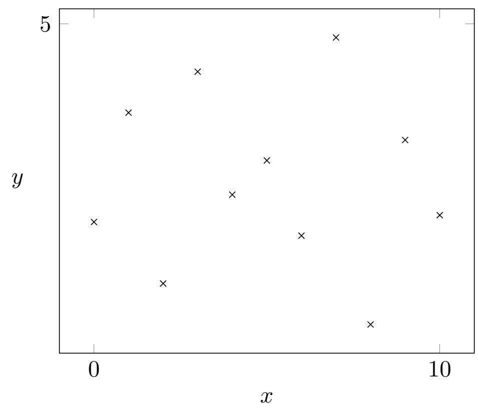


Figure 33.5: No clear relationship.

33.3 Product Moment Correlation Coefficient

As mentioned in the introduction, correlation refers to the relationship between two variables. We can quantify this relationship by the product moment correlation coefficient.

Definition 33.3.1. The **product moment correlation coefficient**, denoted r, for a sample of bivariate data, is given by

$$r = \frac{\sum (x - \overline{x})(y - \overline{y})}{\sqrt{\sum (x - \overline{x})^2}\sqrt{\sum (y - \overline{y})^2}}.$$

We can manipulate r to get rid of $\overline{x}$ and $\overline{y}$:

$$r = \frac{\sum xy - \frac{1}{n}\sum x \sum y}{\sqrt{\sum x^2 - \frac{1}{n}\left(\sum x\right)^2}\sqrt{\sum y^2 - \frac{1}{n}\left(\sum y\right)^2}},$$

where n is the number of ordered pairs in the sample.

33.3.1 Characteristic of r

r can only take on values between -1 and 1. A summary of the value(s) of r and the associated linear correlation is given below.

Value of r	**Linear Correlation**	**Observation on Scatter Diagram**
$r = 1$	Perfect positive linear correlation	The points all lie on a straight line with positive gradient
$r \approx 1$	Strong positive linear correlation	The points lie close to a straight line with positive gradient
$0 < r < 1$	Positive linear correlation	Most points lie in a band with positive gradient
$r = 0$	No linear correlation	No pattern or non-linear pattern
$-1 < r < 0$	Negative linear correlation	Most points lie in a band with negative gradient
$r \approx -1$	Strong negative linear correlation	The points lie close to a straight line with negative gradient
$r = -1$	Perfect negative linear correlation	The points all lie on a straight line with negative gradient

To understand why this is the case, consider the sign of r. Looking at the definition of r, it is clear that

$$r > 0 \iff \sum (x - \overline{x})(y - \overline{y}) > 0.$$

Likewise,

$$r < 0 \iff \sum (x - \overline{x})(y - \overline{y}) < 0.$$

Consider now the following figure:

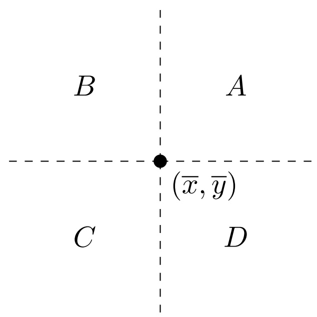


Figure 33.6

Consider quadrant A. Any data point (x, y) within this quadrant will satisfy $x > \overline{x}$ and $y > \overline{y}$, so $(x - \overline{x})(y - \overline{y}) > 0$. Similar analysis reveals that

$$(x - \overline{x})(y - \overline{y}) = \begin{cases} > 0 & \text{for quadrants } A \text{ and } C, \\ < 0 & \text{for quadrants } B \text{ and } D. \end{cases}$$

Thus, if the overall sum is positive, the points must have been largely scattered within quadrants A and C, which we visually interpret as a "positive gradient". Likewise, if the overall sum is negative, the points must have been largely scattered within quadrants B and D, which we interpret as a "negative gradient". Lastly, if the overall sum is near 0, the points must have been scattered randomly throughout all four quadrants, so there is no linear relationship between the variables.

33.3.2 Importance of Scatter Diagram

The value of r should always be interpreted together with a scatter diagram where possible. The value of r can be affected by outliers and can give a misleading conclusion on the linear correlation of two variables. For instance, the following two sets of bivariate data differ only by one data point, yet they have drastically different product moment correlation coefficients:

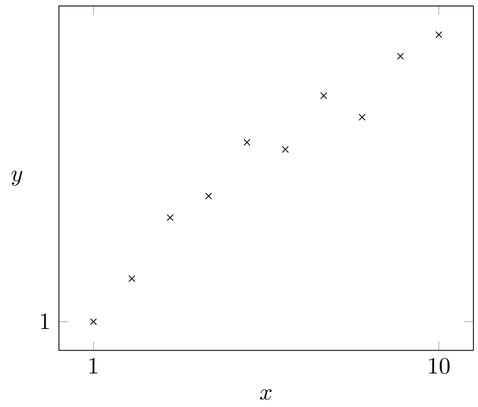


Figure 33.7: $r = 0.975$.

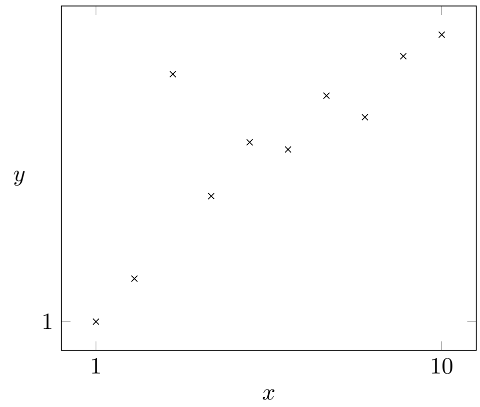


Figure 33.8: $r = 0.821$.

Thus, the scatter diagram should always be used in the interpretation of correlation, as it not only shows the pattern trend between the variables, but it also reveals the existence of any outliers which may have affected the value of r.

33.3.3 Correlation and Causation

A strong or perfect linear correlation between two variables does not necessarily imply one directly causes the other; correlation does not imply causation.

33.4 Predicting or Estimating Using Regression Line

In statistical studies, when it is observed that a significant linear correlation exists between two variables of study, best-fit lines or regression lines are often obtained in order to make predictions or estimations relating to x and/or y. For bivariate data, there are two possible regression lines that we can draw:

- regression line of y on x, or
- regression line of x on y.

33.4.1 Regression Line of y on x

Let (x_i, y_i) for $i = 1, \ldots, n$ be a set of n observed data points.

Definition 33.4.1. The **vertical residual**, denoted v_i, is the deviation between the actual and predicted y-values.

$$v_i = y_i - (a + bx_i)$$

for some constants a and b.

We can think of a vertical residual as the (signed) *vertical* distance between an observed data point (x_i, y_i) and the line $y = a + bx$.

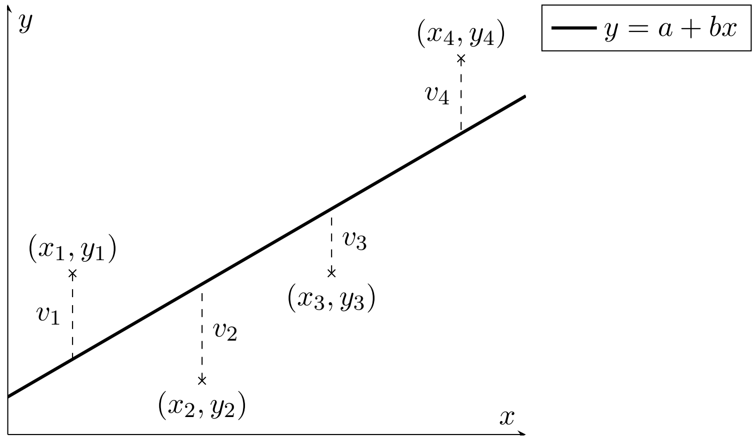


Figure 33.9: The vertical residuals as vertical distances between actual and observed values.

Definition 33.4.2. The **least-squares regression line of y on x** is obtained by finding the values of a and b in $y = a + bx$ that minimizes the sum of the squares of the vertical residuals, S:

$$S = \sum_{i=1}^{n} v_i^2 = \sum_{i=1}^{n} [y_i - (a + bx_i)]^2 .$$

The values of a and b that minimize S is called the **least-squares estimates** of a and b. b is also sometimes called the **regression coefficient**.

The following result can be shown using functions of two variables (see Assignment B11 Problem 3):

Proposition 33.4.3. The regression line of y on x is given by $y - \overline{y} = b(x - \overline{x})$ where

$$b = \frac{\sum (x - \overline{x})(y - \overline{y})}{\sum (x - \overline{x})^2} = \frac{\sum xy - n(\overline{x})(\overline{y})}{\sum x^2 - n(\overline{x})^2}.$$

Observe that the regression line of y on x passes through the **mean point** $(\overline{x}, \overline{y})$.

33.4.2 Regression Line of x on y

The regression line of x on y is similar. In this case, however, we are concerned with *horizontal* deviations instead.

Definition 33.4.4. The **horizontal residual**, denoted h_i, is the deviation between the actual and predicted x-values.

$$h_i = y_i - (c + dx_i)$$

for some constants c and d.

Analogous to v_i, we can think of a horizontal residual as the (signed) *horizontal* distance between an observed data point (x_i, y_i) and the line $x = c + dy$.

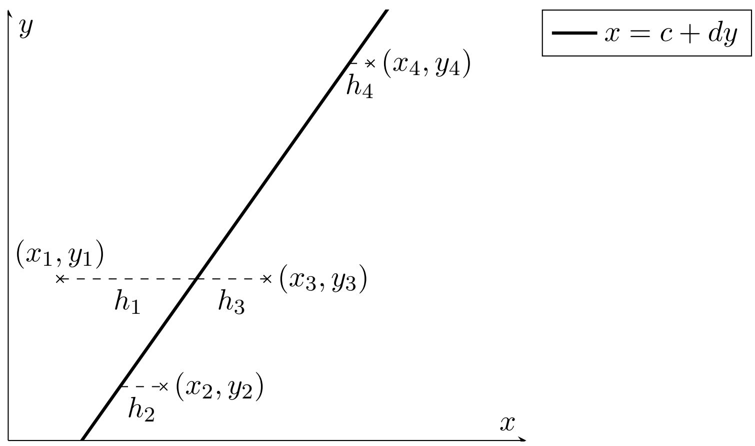


Figure 33.10: The horizontal residuals as horizontal distances between actual and observed values.

Definition 33.4.5. The **least-squares regression line of x on y** is obtained by finding the values of c and d in $x = c + dy$ that minimizes the sum of the squares of the horizontal residuals, S:

$$S = \sum_{i=1}^{n} h_i^2 = \sum_{i=1}^{n} [x_i - (c + dy_i)]^2 .$$

Problem 1. The regression line of x on y is given by $x - \overline{x} = d(y - \overline{y})$, where

$$d = \frac{\sum (x - \overline{x})(y - \overline{y})}{\sum (y - \overline{y})^2} = \frac{\sum xy - n(\overline{x})(\overline{y})}{\sum y^2 - n(\overline{y})^2}.$$

As in the y on x case, we call d the regression coefficient. Note that $1/d$, and not d, is the gradient of the regression line. Observe that the regression line of x on y also passes through the mean point $(\overline{x}, \overline{y})$.

33.4.3 Determining Which Regression to Use

If there is an independent variable x, we use the regression line y on x regardless of whether we are predicting or estimating y or x, and vice versa when y is the independent variable.

However, if there is no clear dependent-independent relationship, we determine the independent variable based on the given value. For example, if we are given the value of x, we use the regression line y on x.

33.4.4 Interpolation and Extrapolation

Definition 33.4.6. An estimate is said to be an **interpolation** if it is within the given range of values of data. Else, it is an **extrapolation**.

Extrapolation of the sample should be used with caution as the relationship between x and y may not be linear beyond a certain point.

33.4.5 Reliability of an Estimate

There are three criteria we typically use when commenting on the reliability of an estimate:

- Appropriateness of the regression line used – The correct regression line should be used for the estimate to be reliable.

- Strength of linear correlation – $|r|$ should be close to 1 for the estimate to be reliable.
- Interpolation or extrapolation – Interpolation is likely to give a more reliable estimate than extrapolation.

For an estimate to be reliable, all three criteria should be satisfied. If at least one of the criteria is not satisfied, we deem the estimate to be unreliable.

33.5 Transformations to Linearize Bivariate Data

The relationship between two variables involved, x and y, may not always be linear. Thus, it would be inappropriate to use the regression lines relating to x and y to make estimations. However, non-linear relationships can be transformed into a linear form by a process usually called **transformation to linearity**. The table below shows some examples:

Original Equations	Transformed Equations	Linearly-related Expressions
$y = a + bx^2$	-	y vs x^2
$y = ab^x$	$\ln y = \ln a + x \ln b$	$\ln y$ vs x
$y = ax^b$	$\ln y = \ln a + b \ln x$	$\ln y$ vs $\ln x$

Sometimes, we are given a scatter diagram and are tasked with comparing two or more proposed models and determine which model is a better fit. In such a scenario, we simply state which equation fits the shape of the scatter plot better. If there is more than one possibility, we can compute the product moment correlation coefficient for each model and "break the tie" by choosing the model with $|r|$ closest to 1.

33.6 Bonus: A Probabilistic Approach to Linear Regression

In an ideal world, our variables will be exactly related by the model $y = a + bx$. However, in the real world, whenever we observe a data point, our readings will contain some error ϵ, so our observations are actually modelled by $y = a + bx + \epsilon$. In real life, these errors are caused by thousands of different factors. We can hence think of ϵ as the sum of many independent random variables. But by the Central Limit Theorem, it follows that ϵ is distributed normally, so

$$\epsilon \sim \mathrm{N}\left(0, \sigma^2\right).$$

Suppose now that we obtain an observation, (x_i, y_i). Since $\epsilon_i = y_i - (a + bx_i)$, the probability of observing this data point is given by

$$\mathbb{P}[(x_i, y_i)] = \mathbb{P}[\epsilon = \epsilon_i] = \mathbb{P}[\epsilon = y_i - (a + bx_i)] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - (a + bx_i))^2}{2\sigma^2}\right).$$

If we make n *independent* observations, then the overall probability of observing all n data points is simply the product of each individual probability:

$$\mathbb{P}[\text{data}] = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - (a + bx_i))^2}{2\sigma^2}\right).$$

It is now natural to define the "best" model ($y = a + bx$) as the one that maximizes the probability of observing our data. That is, we wish to find a and b that maximizes

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - (a + bx_i))^2}{2\sigma^2}\right).$$

Since the logarithm is monotonic, we can convert our objective function from a product into a sum:

$$\underset{a,b}{\operatorname{argmax}} \ln \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i-(a+bx_i))^2}{2\sigma^2}\right) = \underset{a,b}{\operatorname{argmax}} \sum_{i=1}^{n} \left(-\frac{(y_i-(a+bx_i))^2}{2\sigma^2}\right),$$

where we ignored the constant terms contributed by $1/\sqrt{2\pi}\sigma$ since they do not affect the location of the maxima. We can further ignore the $1/\sigma^2$ term since it is a constant factor. Lastly, flipping the sign changes our objective into a minimization problem, so we get

$$\underset{a,b}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i-(a+bx_i))^2 .$$

But this is exactly the objective of the least-squares regression line of x on y we introduced earlier!

33.7 Bonus: r and Vectors

Suppose we have two sets of data, say $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$. Let $\overline{x}$ and $\overline{y}$ denote their respective means. Recall that the product moment correlation coefficient r between these two samples is given by

$$r = \frac{\sum (x-\overline{x})(y-\overline{y})}{\sqrt{\sum (x-\overline{x})^2 \sum (y-\overline{y})^2}}.$$

Observe that the definition of r resembles the definition of the cosine of an angle between two vectors! Indeed, if we define

$$\mathbf{x} = \begin{pmatrix} x_1-\overline{x} \\ \vdots \\ x_n-\overline{x} \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} y_1-\overline{y} \\ \vdots \\ y_n-\overline{y} \end{pmatrix},$$

then we can simply express r as

$$r = \frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}|^2 |\mathbf{y}|^2} = \cos\theta,$$

where θ is the angle between the two vectors $\mathbf{x}$ and $\mathbf{y}$.[1] Similarly, we can rewrite the regression coefficients b and d vectorially:

$$b = \frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}|^2} \quad \text{and} \quad d = \frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{y}|^2}.$$

If we manipulate the above two expressions, we see that

$$b = \frac{\hat{\mathbf{x}}\cdot\mathbf{y}}{|\mathbf{x}|} \quad \text{and} \quad d = \frac{\mathbf{x}\cdot\hat{\mathbf{y}}}{|\mathbf{y}|}.$$

[1] We can think of these two vectors as the "deviation" between the sample data and their respectively means. Indeed, it is not too hard to see that the sample variances are given by $s_X^2 = \frac{1}{n-1}|\mathbf{x}|^2$ and $s_Y^2 = \frac{1}{n-1}|\mathbf{y}|^2$. The scaled dot product $\frac{1}{n-1}(\mathbf{x}\cdot\mathbf{y})$ also has a special name, called the "sample covariance", typically denoted s_{XY}^2, so the product moment correlation coefficient can be expressed more succinctly as

$$r = \frac{s_{XY}^2}{s_X s_Y}.$$

Now observe that the numerator of b is exactly the length of projection of $\mathbf{y}$ onto $\mathbf{x}$. Similarly, the numerator of d is exactly the length of projection of $\mathbf{x}$ on $\mathbf{y}$.

That is to say, b measures the ratio between the vector projection of $\mathbf{y}$ onto $\mathbf{x}$, and similarly for d:

$$b = \frac{\text{length of projection of } \mathbf{y} \text{ onto } \mathbf{x}}{\text{length of } \mathbf{x}} \quad \text{and} \quad d = \frac{\text{length of projection of } \mathbf{x} \text{ onto } \mathbf{y}}{\text{length of } \mathbf{y}}.$$

This aligns with our intuition of b and d: If the two samples share a strong linear correlation, we would expect the regression lines of y on x and x on y to be roughly the same. Indeed, $\mathbf{x}$ and $\mathbf{y}$ are roughly multiples of each other, say $\mathbf{x} \approx \lambda \mathbf{y}$ for some λ, so

$$b \approx \frac{|\lambda \mathbf{y}|}{|\mathbf{y}|} = |\lambda| \quad \text{and} \quad d \approx \frac{|\mathbf{x}|}{|\lambda \mathbf{x}|} = \frac{1}{|\lambda|} \implies b \approx \frac{1}{d}.$$

But b and $1/d$ represent the gradients of the regression lines of y on x and of x on y respectively, so the two lines have roughly equivalent gradients, i.e. the two lines are roughly the same.

Part VIII

Mathematical Proofs and Reasoning

34 Propositional Logic

Mathematics is a deductive science, where from a set of basic axioms, we prove more complex results. To do so, we often restate a sentence into **statements**, which are mathematical expressions. One important axiom that all statements obey is the law of the excluded middle.

Axiom 34.0.1 (Law of the Excluded Middle)**.** The **law of the excluded middle** states that either a statement or its negation is true. Equivalently, a statement cannot be both true and false, nor can it be neither true nor false.

34.1 Statements

34.1.1 Forming Statements

We call a sentence such as "x is even" that depends on the value of x a "statement about x". We can denote this statement more compactly as $P(x)$. For instance, $P(5)$ is the statement "5 is even", while $P(72)$ is the statement "72 is even", and so forth. We can also write $P(x)$ more compactly as P.

We now introduce some operations of statements, namely the negation, conjunction and disjunction operations.

Definition 34.1.1. The **negation** of a statement P, denoted $\neg P$, is false when P is true, and true when P is false. In a truth table,

P	$\neg P$
T	F
F	T

Example 34.1.2. If $P(x)$ is the statement "x is even", then $\neg P(x)$ is the statement "x is odd".

Definition 34.1.3. The **conjunction** of two statements P and Q, denoted $P \wedge Q$, has truth table

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

Example 34.1.4. If P is the statement "I like cats", and Q is the statement "I like dogs", then $P \wedge Q$ is the statement "I like cats and dogs".

Definition 34.1.5. The **disjunction** of two statements P and Q, denoted $P \vee Q$, has truth table

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

Example 34.1.6. If P is the statement "I like cats", and Q is the statement "I like dogs", then $P \vee Q$ is the statement "I like cats or dogs or both".

Proposition 34.1.7 (De Morgan's Law)**.** Let P and Q be statements. Then

$$\neg (P \wedge Q) \iff (\neg P) \vee (\neg Q)$$

and

$$\neg (P \vee Q) \iff (\neg P) \wedge (\neg Q).$$

Proof. Consider the following truth tables:

P	Q	$P \wedge Q$	$P \vee Q$	$\neg(P \wedge Q)$	$\neg(P \vee Q)$	$\neg P$	$\neg Q$	$(\neg P) \wedge (\neg Q)$	$(\neg P) \vee (\neg Q)$
T	T	T	T	F	F	F	F	F	F
T	F	F	T	T	F	F	T	F	T
F	T	F	T	T	F	T	F	F	T
F	F	F	F	T	T	T	T	T	T

We see that the truth table of $\neg (P \wedge Q)$ is equivalent to that of $(\neg P) \vee (\neg Q)$, thus the statements are equivalent.

Similarly, the truth table of $\neg (P \vee Q)$ is equivalent to that of $(\neg P) \wedge (\neg Q)$, thus the statements are equivalent. □

Example 34.1.8. Let P be the statement "I like cats", and Q be the statement "I like dogs". Then $\neg(P \wedge Q)$ is "It is not the case that I like both cats and dogs", while $(\neg P) \vee (\neg Q)$ is "I do not like cats, or I do not like dogs, or I do not like both". Clearly, the two statements are equivalent.

34.1.2 Conditional and Biconditional Statements

In this section, we examine how statements are linked together to form more complicated statements. The first type of statement we will examine is the conditional statement.

Definition 34.1.9. A **conditional statement** has the form "if P then Q". Here, P is the **hypothesis** and Q is the **conclusion**, denoted by $P \implies Q$. This statement is defined to have the truth table

P	Q	$P \implies Q$
T	T	T
T	F	F
F	T	T
F	F	T

In words, the statement $P \implies Q$ also reads:

- P **implies** Q.
- P is a **sufficient condition** for Q.
- Q is a **necessary condition** for P.
- P **only if** Q.

To justify the truth table of $P \implies Q$, consider the following example:

Example 34.1.10 (Conditional Statement). Suppose I say

"If it is raining, then the floor is wet."

We can write this as $P \implies Q$, where P is the statement "it is raining" and Q is the statement "the floor is wet".

- Suppose both P and Q are true, i.e. it is raining, and the floor is wet. It is reasonable to say that I am telling the truth, whence $P \implies Q$ is true.
- Suppose P is true but Q is false, i.e. it is raining, and the floor is not wet. Clearly, I am not telling the truth; the floor would be wet if I was. Hence, $P \implies Q$ is false.
- Suppose P is false, i.e. it is not raining. Notice that the hypothesis of my claim is not fulfilled; I did not say anything about the floor when it is not raining. Hence, I am not lying, so $P \implies Q$ is true whenever P is false.

Examples of conditional statements in mathematics include

- If $|x - 1| < 4$, then $-3 < x < 5$.
- If a function f is differentiable, then f is continuous.

We now look at biconditional statements. As the name suggests, a biconditional statement comprises two conditional statements: $P \implies Q$ and $Q \implies P$. The conditional statement is much stronger than the conditional statement.

Definition 34.1.11. A **biconditional statement** has the form "P if and only if", denoted $P \iff Q$. This statement is defined to have the truth table

P	Q	$P \iff Q$
T	T	T
T	F	F
F	T	F
F	F	T

When $P \iff Q$ is true, we say that P and Q are **equivalent**, i.e. $P \equiv Q$.

An equivalent definition of $P \iff Q$ is the statement

$$(P \implies Q) \quad \text{and} \quad (Q \implies P).$$

This allows us to easily justify the truth table of $P \iff Q$:

P	Q	$P \implies Q$	$Q \implies P$	$P \iff Q$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

Examples of conditional statements in mathematics include

- A triangle ABC is equilateral if and only if its three angles are congruent.
- a is a rational number if and only if $2a + 4$ is rational.

34.1.3 Quantifiers

We now introduce two important symbols, namely the universal quantifier ($\forall$) and the existential quantifier ($\exists$)

Definition 34.1.12. Let $P(x)$ be a statement about x, where x is a member of some set S (i.e. S is the **domain** of x). Then the notation

$$\forall x \in S,\ P(x)$$

means that $P(x)$ is true for every x in the set S. The notation

$$\exists x \in S,\ P(x)$$

means that there exists at least one element of x of S for which $P(x)$ is true.

Example 34.1.13. Let $P(x)$ be the statement "x is even". Clearly, the statement

$$\forall x \in \mathbb{Z},\ P(x)$$

is not true; not all integers are even. However, the statement

$$\exists x \in \mathbb{Z},\ P(x)$$

is true, because we can find an integer that is even (e.g. $x = 8$).

Note that a statement $P(x)$ does not necessarily have to mention x. For instance, we could define $P(x)$ as the statement "5 is even". Compare this with how a function $f(x)$ does not necessarily have to "mention" x, e.g. we could have $f(x) = 5$.

Proposition 34.1.14. The negation of a universal statement is an existential statement, and vice versa.

$$\neg\left(\forall x \in D,\, P(x)\right) \quad \iff \quad \exists x \in D,\, \neg P(x).$$

Proof. We prove that the negation a universal statement is an existential statement. Observe that a universal statement is equivalent to a conjunction of many statements:

$$\forall x \in D,\, P(x) \quad \iff \quad P(x_1) \wedge P(x_2) \wedge \ldots,$$

where $D = \{x_1, x_2, \ldots\}$. Using De Morgan's laws, we can easily negate the above statements:

$$\neg\left(\forall x \in D,\, P(x)\right) \quad \iff \neg P(x_1) \vee \neg P(x_2) \vee \ldots.$$

However, the last statement is equivalent to the existential statement

$$\exists x \in D,\, \neg P(x).$$

Thus,

$$\neg\left(\forall x \in D,\, P(x)\right) \quad \iff \quad \exists x \in D,\, \neg P(x).$$

Using a similar argument, one can prove that the negation of an existential statement is a universal statement, i.e.

$$\neg\left(\exists x \in D,\, P(x)\right) \quad \iff \quad \forall x \in D,\, \neg P(x).$$

□

Example 34.1.15. Let D be the set of all students in a class, and let $P(x)$ be "x likes durian". Then the statement $\forall x \in D,\, P(x)$ reads as "everyone in the class likes durian". Intuitively, its negation would be "someone in the class does not like durian", which we can write as $\exists x \in D,\, \neg P(x)$.

34.1.4 Types of Statements

Most of the statements we will encounter can be grouped into three classes, namely axioms, definitions and theorems.

Definition 34.1.16.

- An **axiom** is a mathematical statement that does not require proof.
- A **definition** is a true mathematical statement that gives the precise meaning of a word or phrase that represents some object, property or other concepts.
- A **theorem** is a true mathematical statement that can be proven mathematically.

34.2 Proofs

Mathematical proofs are convincing arguments expressed in mathematical language, i.e. a sequence of statements leading logically to the conclusion, where each statement is either an accepted truth, or an assumption, or a statement derived from previous statements. Occasionally there will be the clarifying remark, but this is just for the reader and has no logical bearing on the structure of the proof.

Definition 34.2.1. A **proof** is a deductive argument for a mathematical statement, showing that the stated assumptions logically guarantee the conclusion.

There are three main types of proofs: direct proof, proof by contrapositive and proof by contradiction.

34.2.1 Direct Proof

A **direct proof** is an approach to prove a conditional statement $P \implies Q$. It is a series of valid arguments that starts with the hypothesis P, and ends with the conclusion Q.

As an example, we will prove the following statement:

Statement 34.2.2. For all $n \in \mathbb{Z}^+$, both n and n^2 have the same parity.

Proof. Since n can only be either odd or even, we just need to consider the following cases:

Case 1. Suppose n is even. By definition, there exists some $k \in \mathbb{Z}$ such that $n = 2k$. Then

$$n^2 = (2k)^2 = 4k^2 = 2\left(2k^2\right) = 2a,$$

where $a = 2k^2$. Since a is an integer, it follows from our definition that n^2 is even. Hence, n and n^2 have the same parity.

Case 2. Suppose n is odd. By definition, there exists some $h \in \mathbb{Z}$ such that $n = 2h+1$. Then

$$n^2 = (2h+1)^2 = 4h^2 + 4h + 1 = 2\left(2h^2 + 2h\right) + 1 = 2b + 1,$$

where $b = 2h^2 + 2h$. Since b is an integer, it follows from our definition that n^2 is odd. Hence, n and n^2 have the same parity. $\square$

34.2.2 Proof by Contrapositive

Suppose we wish to prove $P \implies Q$. Occasionally, the hypothesis P is more complicated than the conclusion Q, which is not desirable. In such a scenario, we can choose to prove the statement via the **contrapositive**, i.e. prove that $\neg Q \implies \neg P$. This typically simplifies the proof, since our hypothesis $\neg Q$ is now simpler.

We now show the equivalence between $P \implies Q$ and $\neg Q \implies \neg P$.

Proposition 34.2.3. Let P and Q be statements. Then

$$P \implies Q \iff \neg Q \implies \neg P.$$

Proof. Consider the following truth table:

P	Q	$P \implies Q$	$\neg Q$	$\neg P$	$\neg Q \implies \neg P$
T	T	T	F	F	T
T	F	F	T	F	F
F	T	T	F	T	T
F	F	T	T	T	T

Since $P \implies Q$ and $\neg Q \implies \neg P$ have the same truth table, they are equivalent. $\square$

As an example, we will prove the following statement using the contrapositive.

Statement 34.2.4. For any real numbers x and y, if $x^2y + xy^2 < 30$, then $x < 2$ or $y < 3$.

Proof. Since the hypothesis is much more complicated than the conclusion, we are motivated to use the contrapositive.

Suppose $x > 2$ and $y > 3$ (this is the negation of $x < 2$ or $y < 3$). Then $x^2y > (2)^2(3) = 12$ and $xy^2 > (2)(3)^2 = 18$. Thus, $x^2y + xy^2 > 12 + 18 = 30$. (this is the negation of $x^2y + xy^2 < 30$). Thus, by the contrapositive, the statement is true. $\square$

34.2.3 Proof by Contradiction

A **proof by contradiction** is a proving technique where we want to prove that a statement is true by assuming that it is false, and arrive at a contradiction. That is, to prove a statement P, we can

1. Assume $\neg P$.
2. Derive a contradiction, or absurdity.
3. Conclude that $\neg P$ is false, which implies P is true.

A classic example of a proof by contradiction is the irrationality of $\sqrt{2}$.

Statement 34.2.5. $\sqrt{2}$ is irrational.

Proof. Seeking a contradiction, suppose $\sqrt{2}$ is rational. Write $\sqrt{2} = a/b$, where a and b are coprime integers with $b \neq 0$. Squaring, we get

$$2 = \frac{a^2}{b^2} \implies a^2 = 2b^2. \tag{1}$$

Thus, a^2 is even, which implies a is even. Hence, $a = 2k$ for some integer k. Substituting this back into (1), we get

$$(2k)^2 = 2b^2 \implies b^2 = 2k^2,$$

whence b^2 is even, which implies b is also even. Thus, both a and b have a factor of 2, contradicting our assumption that a and b are coprime. Thus, our assumption that $\sqrt{2}$ is rational is false, whence $\sqrt{2}$ is irrational. $\square$

34.2.4 Induction

Induction is typically used to prove statements of the form "$P(n)$ is true for all $n \in \mathbb{Z}^+$". There are several variants of induction.

Principle of Mathematical Induction

The basic form of mathematical induction requires two steps:

- Showing that $P(0)$ is true, and
- Proving that $P(k) \implies P(k+1)$ for some $k \in \mathbb{Z}^+$.

With these two statements, we see that

$$P(0) \implies P(1) \implies P(2) \implies P(3) \implies \ldots,$$

i.e. $P(n)$ is true for all $n \in \mathbb{Z}^+$.

Of course, the base case need not always be $n = 0$. If we wish to prove that $P(n)$ holds for $n = m, m+1, m+2, \ldots$ for some integer m, our base case becomes $n = m$, so we have to verify that $P(m)$ holds.

Intuitively, we can think of induction as a ladder. The base case acts as the first rung, while the statement $P(k) \implies P(k+1)$ enables us to climb the ladder rung by rung.

A classic example of an inductive proof is to verify that the first n natural numbers sum to $n(n+1)/2$.

Statement 34.2.6. For n a natural number, $1+2+\cdots+n = n(n+1)/2$.

Proof. Let $P(n)$ be the statement $1+2+\cdots+n = n(n+1)/2$. We induct on n.

The base case $P(1)$ is trivial, since $1 = (1)(2)/2$. Suppose that $P(k)$ holds for some natural number k. Consider the sum of the first $k+1$ natural numbers. By our **induction hypothesis**, we see that

$$1+2+\cdots+k+(k+1) = \frac{k(k+1)}{2} + (k+1) = \frac{(k+1)((k+1)+1)}{2},$$

so $P(k+1)$ also holds. By the principle of mathematical induction, it follows that $P(n)$ holds for all natural numbers n. □

Principle of Strong Induction

Another common variant of induction is *strong* induction. Like before, it involves showing two steps:

- Showing that $P(0)$ is true, and
- If $P(0)$, $P(1)$, ..., $P(k)$ are true, then so is $P(k+1)$.

Here, the inductive step is replaced with a *stronger* hypothesis that requires all the terms before $P(k+1)$ to be true, as demonstrated in the following example:

Statement 34.2.7. All integers greater than 1 are either a prime or a product of primes.

Proof. Let $P(n)$ be the statement "n is either a prime or a product of primes". We induct on n. The base case $n = 2$ is trivial (2 itself is a prime). Now suppose $P(2)$ to $P(k)$ are true for some integer $k \geq 2$. If $k+1$ is prime, then $P(k+1)$ is trivially true. Else, $k+1$ must be composite, so we can write $k+1 = ab$, for some $2 \leq a, b \leq k$. But by our induction hypothesis, both a and b are either primes or a product of primes, hence ab itself is a product of primes, so $P(k+1)$ is true. This closes the induction. □

We can also use multiple base cases for strong induction:

- Showing that the base cases $P(0)$, $P(1)$, ..., $P(m)$ are true, and
- Proving that if $P(k)$, $P(k+1)$, ..., $P(k+m)$ are true, then $P(k+m+1)$ is true.

All Horses are the Same Colour

Caution must be exercised when proving a statement inductively. Consider now the following "proof" that purports to show that all horses share the same colour.

Statement 34.2.8. All horses are the same colour.

Proof. Let $P(n)$ be the statement "A group of n horses have the same colour". We induct on n. $P(1)$ is trivial. Suppose that $P(k)$ is true for some integer $k \geq 1$. Consider now a group of $k+1$ horses.

- First, exclude horse $k+1$. Horses 1 to k are a group of k horses, so by our induction hypothesis, they must all be of the same colour.
- Next, exclude horse 1. Horses 2 to $k+1$ form another group of k horses, so they must also all be of the same colour.

Hence, horse $k+1$ must have been the same colour as the non-excluded horses, i.e. all $k+1$ horses share the same colour, so $P(k+1)$ holds. Thus, by the principle of mathematical induction, $P(n)$ is true for all integers $n \geq 1$, so all horses are the same colour. $\square$

Of course, we know that the claim is wrong, so we must have made an error somewhere in the proof. As an exercise, find the flaw in the proof. (Hint: consider the inductive step $P(1) \implies P(2)$.)

34.2.5 Counter-Example

In the case where we wish to prove a statement false, we can find a counter-example. In providing a counter-example, it must fulfil the hypothesis, but not the conclusion. That is, to show that $P \implies Q$ is false, we must show that P is true but Q is false.

Example 34.2.9 (Counter-Example)**.** Consider the statement $c \mid ab$, then $c \mid a$ or $c \mid b$, where $a, b, c \in \mathbb{Z}^+$. We can easily find a counter-example to this statement, e.g. $a = 3 \times 37$, $b = 7 \times 37$, $c = 3 \times 7$.

35 Number Theory

35.1 Congruence

Definition 35.1.1. Let two integers a and b (with $b \neq 0$). If there exists some integer n such that $a = bn$, we say

- b divides a, and
- a is divisible by b.

We write this as $b \mid a$.

Proposition 35.1.2. For $a, b, c \in \mathbb{Z}$, if $a \mid b$ and $a \mid c$, then $a \mid (b \pm c)$.

Proof. From our definition, we there exists integers x and y such that $b = ax$ and $c = ay$. Hence,

$$b \pm c = ax \pm ay = a\left(x \pm y\right).$$

Since $x \pm y$ is an integer, $a \mid (b \pm c)$. □

Definition 35.1.3 (Congruence Modulo). Let $a, b, n \in \mathbb{Z}$ with $n > 0$. We say that a is **congruent** to b **modulo** n, denoted as

$$a \equiv b \pmod{n},$$

iff n divides $a - b$. Equivalently, $a = b + nk$ for some $k \in ZZ$.

Example 35.1.4. $25 \equiv 7$ modulo 3 since $25 - 7 = 18$ is a multiple of 3.

Proposition 35.1.5 (Congruence is an Equivalence Relation). Let $a, b, n \in \mathbb{Z}$.

- Congruence is reflexive, i.e. $a \equiv a$ modulo n.
- Congruence is symmetric, i.e. if $a \equiv b$ then $b \equiv a$ (modulo n).
- Congruence is transitive, i.e. if $a \equiv b$ and $b \equiv c$, then $a \equiv c$ (all modulo n).

Proof. Trivial. □

Proposition 35.1.6. For all integers a, b, c, d, k, n, with $n > 1$, suppose $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$. Then

- $a \pm c \equiv b \pm d \pmod{n}$.
- $a \cdot c \equiv b \cdot d \pmod{n}$.
- $a + k \equiv b + k \pmod{n}$.
- $ka \equiv kb \pmod{n}$.
- $a^m \equiv b^m \pmod{n}$ for all $m \in \mathbb{Z}^+$.

In other words, congruence modulo preserves addition, subtraction, multiplication, and exponentiation. Take not that congruence modulo does NOT always preserve division. That is, if $c \mid a$ and $d \mid b$, it is not always true that

$$\frac{a}{c} \equiv \frac{b}{d} \pmod{n}.$$

We now state an important result that formalizes our notion of remainders when dividing integers.

Lemma 35.1.7 (Euclid's Division Lemma)**.** Let $n \in \mathbb{Z}^+$. Then for any $m \in \mathbb{Z}$, there exists a unique integer r with $0 \leq r < n$ such that

$$m \equiv r \pmod{n}.$$

Equivalently, there exists an integer q such that

$$m = qn + r.$$

We will prove this statement for $m, n > 0$. We can take $m > n$ since if $0 < m < n$, we can simply take $q = 0$ and $r = m$.

Proof. We prove that such an r exists, and show that it must be unique.

Existence. Let q be the largest number such that $m \geq nq$ and let $r = m - nq \geq 0$. Seeking a contradiction, suppose $r \geq n$, i.e. $r = n + d$ for $d \geq 0$. Then

$$m = nq + r = nq + (n + d) = n(q + 1) + d \geq n(q + 1),$$

contradicting the maximality of q. Hence, $0 \leq r < n$, i.e. r exists.

Uniqueness. Suppose there exist r_1, r_2, with $0 \leq r_1, r_2 < n$ such that

$$m = q_1 n + r_1 = q_2 n + r_2.$$

Then $r_1 = (q_2 - q_1)n + r_2$. Since $0 \leq r_1, r_2 < n$, we must have $r_1 = r_2$. Hence, r must be unique. This concludes the proof. □

Lemma 35.1.8 (Euclid's Lemma)**.** Let p be prime. If p divides ab, then p divides a or p divides b.

Proof. Let

$$a = \prod_{i=1}^{k} p_i^{n_i}, \quad b = \prod_{j=1}^{l} q_j^{m_j},$$

where p_i and q_j are primes, while n_i and m_j are positive integers. Then

$$p \mid ab = \prod_{i=1}^{k} p_i^{n_i} \prod_{j=1}^{l} q_j^{m_j}.$$

By the uniqueness of prime decomposition, either $p = p_i$ for some $i = 1, \ldots, k$ (in which case $p \mid a$), or $p = q_j$ for some $j = 1, \ldots, l$ (in which case $p \mid b$). Hence, either $p \mid a$ or $p \mid b$. □

Theorem 35.1.9. There are infinitely many primes.

Proof. Seeking a contradiction, suppose there are finitely many primes $p_1, p_2, \ldots, p_n$. Consider

$$a = p_1 p_2 \ldots p_n + 1.$$

Since $a > p_1, p_2, \ldots p_n$, by our hypothesis, a cannot be a prime, i.e. a is composite. Hence, it must have a prime factorization. Without loss of generality, suppose p_1 be a prime factor of a. Then $p_1 \mid a$. However,

$$p_1 \mid a - 1 = p_1 p_2 \ldots p_n$$

too. Hence, by divisibility rules, p_1 must divide the difference between a and $a - 1$, i.e.

$$p_1 \mid [a - (a - 1)] = 1,$$

which implies that $p_1 = 1$. This is a contradiction, since 1 is not a prime. Thus, there must be infinitely many primes. $\square$

EXERCISES

Part IX

Group A

A1 Equations and Inequalities

Tutorial A1

Problem 1. Determine whether each of the following systems of equations has a unique solution, infinitely many solutions, or no solutions. Find the solutions, where appropriate.

(a) $\begin{cases} a + 2b - 3c = -5 \\ -2a - 4b - 6c = 10 \\ 3a + 7b - 2c = -13 \end{cases}$

(b) $\begin{cases} x - y + 3z = 3 \\ 4x - 8y + 32z = 24 \\ 2x - 3y + 11z = 4 \end{cases}$

(c) $\begin{cases} x_1 + x_2 = 5 \\ 2x_1 + x_2 + x_3 = 13 \\ 4x_1 + 3x_2 + x_3 = 23 \end{cases}$

(d) $\begin{cases} 1/p + 1/q + 1/r = 5 \\ 2/p - 3/q - 4/r = -11 \\ 3/p + 2/q - 1/r = -6 \end{cases}$

(e) $\begin{cases} 2\sin\alpha - \cos\beta + 3\tan\gamma = 3 \\ 4\sin\alpha + 2\cos\beta - 2\tan\gamma = 2 \\ 6\sin\alpha - 3\cos\beta + \tan\gamma = 9 \end{cases}$, where $0 \leq \alpha \leq 2\pi$, $0 \leq \beta \leq 2\pi$, and $0 \leq \gamma < \pi$.

Solution.
Part (a). Unique solution: $a = -9$, $b = 2$, $c = 0$.
Part (b). No solution.
Part (c). Infinitely many solutions: $x_1 = 8 - t$, $x_2 = t - 3$, $x_3 = t$.
Part (d). Solving, we obtain

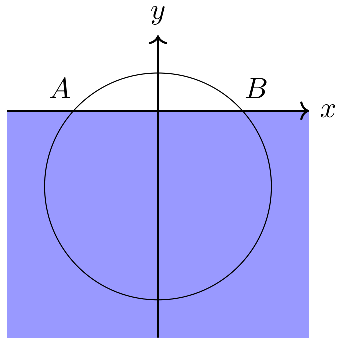
$$\frac{1}{p} = 2, \quad \frac{1}{q} = -3, \quad \frac{1}{r} = 6.$$

There is hence a unique solution: $p = 1/2$, $q = -1/3$, $r = 1/6$.
Part (e). Solving, we obtain

$$\sin\alpha = 1, \quad \cos\beta = -1, \quad \tan\gamma = 0.$$

There is hence a unique solution: $\alpha = \pi/2$, $\beta = \pi$, $\gamma = 0$.

Problem 2. The following figure shows the circular cross-section of a uniform log floating in a canal.



With respect to the axes shown, the circular outline of the log can be modelled by the equation

$$x^2 + y^2 + ax + by + c = 0.$$

A and B are points on the outline that lie on the water surface. Given that the highest point of the log is 10 cm above the water surface when AB is 40 cm apart horizontally, determine the values of a, b and c by forming a system of linear equations.

Solution. Since $AB = 40$, we have $A(-20, 0)$ and $B(20, 0)$. We also know $(0, 10)$ lies on the circle. Substituting these points into the given equation, we have the following system of equations:

$$\begin{cases} -20a \qquad + c = -400 \\ \;\; 20a \qquad + c = -400 \\ \qquad\quad 10b + c = -100 \end{cases}$$

Solving, we obtain $a = 0$, $b = 30$, $c = -400$.

* * * * *

Problem 3. Find the exact solution set of the following inequalities.

(a) $x^2 - 2 \geq 0$

(b) $4x^2 - 12x + 10 > 0$

(c) $x^2 + 4x + 13 < 0$

(d) $x^3 < 6x - x^2$

(e) $x^2(x-1)(x+3) \geq 0$

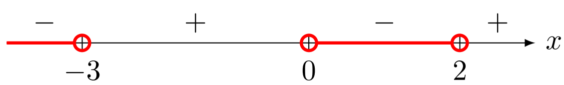
Solution.

Part (a). Note that $x^2 - 2 \geq 0 \implies x \leq -\sqrt{2}$ or $x \geq \sqrt{2}$. The solution set is thus $\{x \in \mathbb{R} : x \leq -\sqrt{2} \text{ or } x \geq \sqrt{2}\}$.

Part (b). Completing the square, we see that $4x^2 - 12x + 10 > 0 \implies (x - \frac{3}{2})^2 + \frac{1}{4} > 0$. Since $(x - \frac{3}{2})^2 \geq 0$, all $x \in \mathbb{R}$ satisfy the inequality, whence the solution set is $\mathbb{R}$.

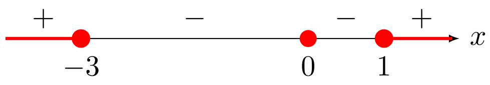
Part (c). Completing the square, we have $x^2 + 4x + 13 < 0 \implies (x+2)^2 + 9 < 0$. Since $(x+2)^2 \geq 0$, there is no solution to the inequality, whence the solution set is $\varnothing$.

Part (d). Note that $x^3 < 6x - x^2 \implies x(x+3)(x-2) < 0$.



The solution set is thus $\{x \in \mathbb{R} \colon x < -3 \text{ or } 0 < x < 2\}$.

Part (e).



The solution set is thus $\{x \in \mathbb{R} \colon x \leq -3 \text{ or } x = 0 \text{ or } x \geq 1\}$.

* * * * *

Problem 4. Find the exact solution set of the following inequalities.

(a) $|3x + 5| < 4$

(b) $|x - 2| < 2x$

Solution.

Part (a). If $3x + 5 < 4$, then $x < -\frac{1}{3}$. If $-(3x + 5) < 4$, then $x > -3$. Combining both inequalities, we have $-3 < x < -\frac{1}{3}$. Thus, the solution set is $\left\{x \in \mathbb{R} \colon -3 < x < -\frac{1}{3}\right\}$.

Part (b). If $x - 2 < 2x$, then $x > -2$. If $-(x - 2) < 2x$, then $x > \frac{2}{3}$. Combining both inequalities, we have $x > \frac{2}{3}$. Thus, the solution set is $\left\{x \in \mathbb{R} \colon x > \frac{2}{3}\right\}$.

* * * * *

Problem 5. It is given that $p(x) = x^4 + ax^3 + bx^2 + cx + d$, where a, b, c and d are constants. Given that the curve with equation $y = p(x)$ is symmetrical about the y-axis, and that it passes through the points with coordinates $(1, 2)$ and $(2, 11)$, find the values of a, b, c and d.

Solution. We know that $(1, 2)$ and $(2, 11)$ lie on the curve. Since $y = p(x)$ is symmetrical about the y-axis, we have that $(-1, 2)$ and $(-2, 11)$ also lie on the curve. Substituting these points into $y = p(x)$, we obtain the following system of equations:

$$\begin{cases} a + b + c + d = 1 \\ a - b + c - d = -1 \\ 8a + 4b + 2c + d = -5 \\ 8a - 4b + 2c - d = 5 \end{cases}$$

Solving, we obtain $a = 0$, $b = -2$, $c = 0$, $d = 3$.

* * * * *

Problem 6. Mr Mok invested \$50,000 in three funds A, B and C. Each fund has a different risk level and offers a different rate of return.

In 2016, the rates of return for funds A, B and C were 6%, 8%, and 10% respectively and Mr Mok attained a total return of \$3,700. He invested twice as much money in Fund A as in Fund C. How much did he invest in each of the funds in 2016?

Solution. Let a, b and c be the amount of money Mr Mok invested in Funds A, B and C respectively, in dollars. We thus have the following system of equations.

$$\begin{cases} a + b + c = 50000 \\ \frac{6}{100}a + \frac{8}{100}b + \frac{10}{100}c = 3700 \\ a = 2c \end{cases}$$

Solving, we have $a = 30000$, $b = 5000$ and $c = 15000$. Thus, Mr Mok invested \$30,000, \$5,000 and \$15,000 in Funds A, B and C respectively.

* * * * *

Problem 7. Solve the following inequalities with exact answers.

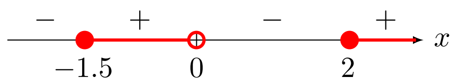
(a) $2x - 1 \geq \frac{6}{x}$

(b) $x - \frac{1}{x} < 1$

(c) $-1 < \frac{2x+3}{x-1} < 1$

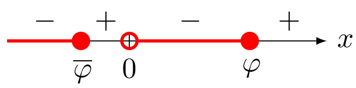
Solution.
Part (a). Note that $x \neq 0$.

$$2x - 1 \geq \frac{6}{x} \implies x^2(2x-1) \geq 6x \implies x\left(2x^2 - x - 6\right) \geq 0 \implies x(2x+3)(x-2) \geq 0.$$



Thus, $-\frac{3}{2} \leq x < 0$ or $x \geq 2$.
Part (b). Note that $x \neq 0$.

$$x - \frac{1}{x} < 1 \implies x^3 - x < x^2 \implies x\left(x^2 - x - 1\right) < 0 \implies x(x-\varphi)(x-\overline{\varphi}) < 0.$$



Thus, $x \leq \bar{\varphi}$ or $0 < x \leq \varphi$.
Part (c).

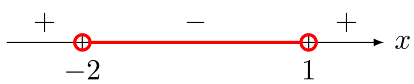
$$-1 < \frac{2x+3}{x-1} < 1 \implies -3 < \frac{5}{x-1} < -1 \implies -\frac{3}{5} < \frac{1}{x-1} < -\frac{1}{5} \implies -4 < x < -\frac{2}{3}.$$

* * * * *

Problem 8. Without using a calculator, solve the inequality $\frac{x^2+x+1}{x^2+x-2} < 0$.

Solution. Observe that $x^2 + x + 1 = \left(x + \frac{1}{2}\right)^2 + \frac{3}{4} > 0$. The inequality thus reduces to $\frac{1}{x^2+x-2} < 0$.

$$\frac{1}{x^2 + x - 2} < 0 \implies x^2 + x - 2 < 0 \implies (x-1)(x+2) < 0.$$



Hence, $-2 < x < 1$.

Problem 9. Solve the following inequalities using a graphical method.

(a) $|3x+1| < (4x+3)^2$

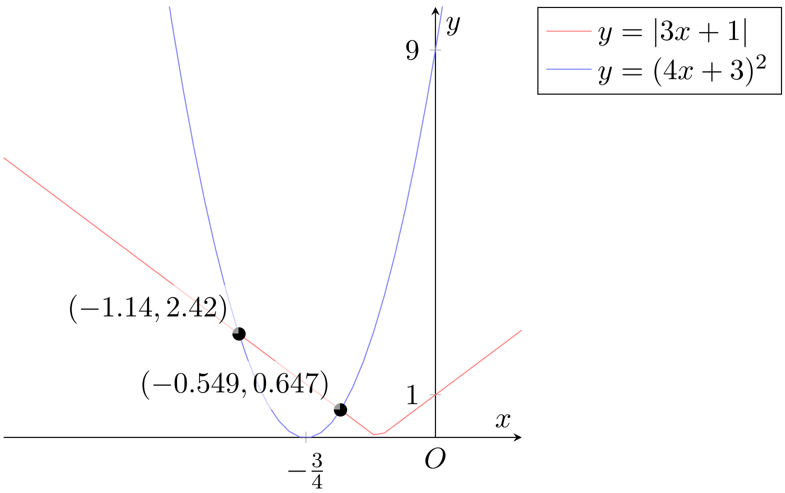
(b) $|3x+1| \geq |2x+7|$

(c) $|x-2| \geq x+|x|$

(d) $5x^2+4x-3 > \ln(x+1)$

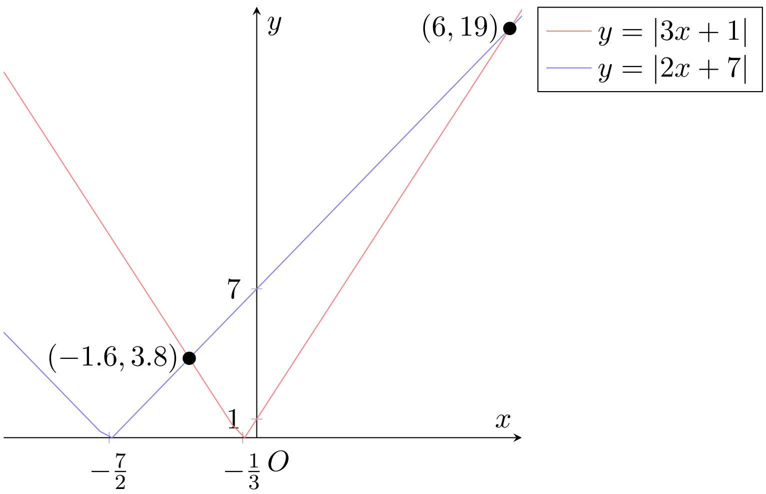
Solution.

Part (a).



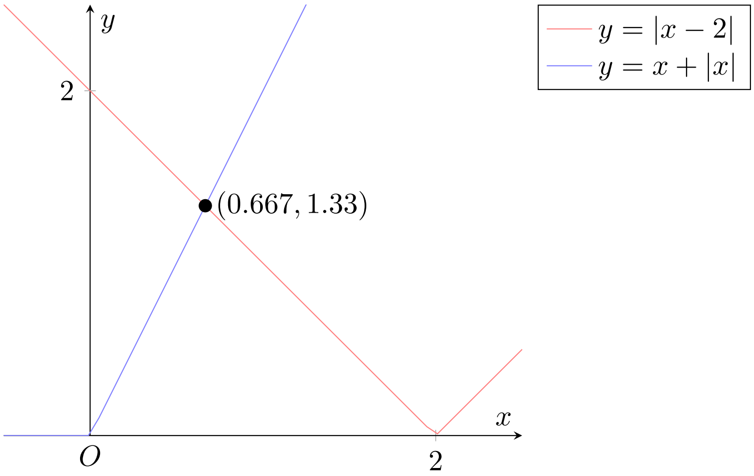
Thus, $x < -1.14$ or $x > -0.549$.

Part (b).



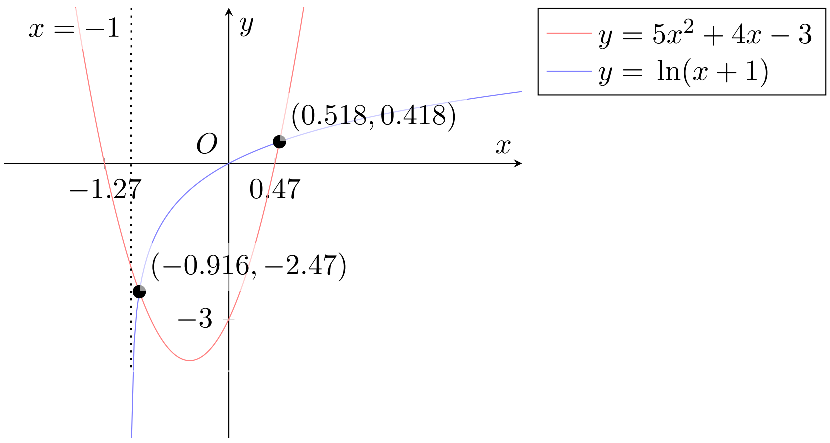
Thus, $x \leq -1.6$ or $x \geq 6$.

Part (c).



Thus, $x \leq 0.667$.

Part (d).

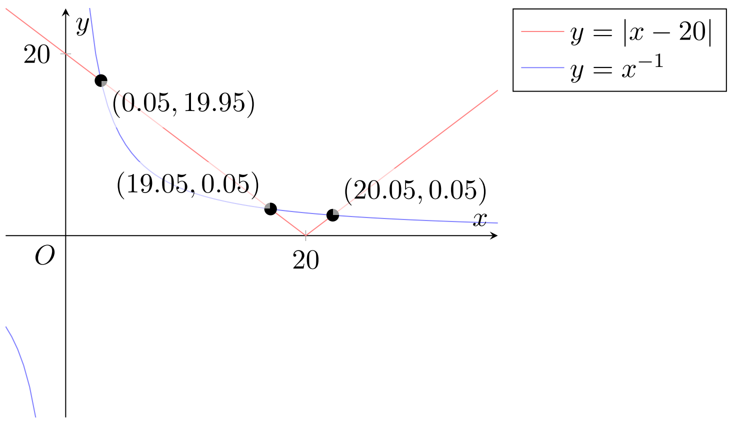


Thus, $-1 < x < -0.916$ or $x > 0.518$.

* * * * *

Problem 10. Sketch the graphs of $y = |x-20|$ and $y = \frac{1}{x}$ on the same diagram. Hence or otherwise, solve the inequality $|x-20| < \frac{1}{x}$, leaving your answers correct to 2 decimal places.

Solution.



Thus, $0 < x < 0.05$ or $19.95 < x < 20.05$.

* * * * *

Problem 11. Solve the inequality $\frac{x-9}{x^2-9} \leq 1$. Hence, solve the inequalities

(a) $\frac{|x|-9}{x^2-9} \leq 1$

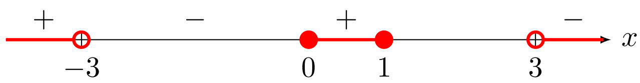
(b) $\frac{x+9}{x^2-9} \geq -1$

Solution. Note that $x^2 - 9 \neq 0 \implies x \neq \pm 3$.

$$\frac{x-9}{x^2-9} \leq 1 \implies (x-9)\left(x^2-9\right) \leq \left(x^2-9\right)^2.$$

Expanding and factoring, we get

$$x^4 - x^3 - 9x^2 + 9x = x(x+3)(x-1)(x-3) \geq 0.$$



Thus, $x < -3$ or $0 \leq x \leq 1$ or $x > 3$.

Part (a). Consider the substitution $x \mapsto |x|$. Then

$$|x| < -3 \text{ or } 0 \leq |x| \leq 1 \text{ or } |x| > 3.$$

This immediately gives us $x < -3$ or $-1 \leq x \leq 1$ or $x > 3$.

Part (b). Consider the substitution $x \mapsto -x$. Then

$$-x < -3 \text{ or } 0 \leq -x \leq 1 \text{ or } -x > 3.$$

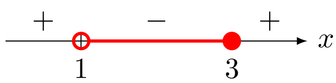
This immediately gives us $x < -3$ or $-1 \leq x \leq 0$ or $x > 3$.

* * * * *

Problem 12. Solve the inequality $\frac{x-5}{1-x} \geq 1$. Hence, solve $0 < \frac{1-\ln x}{\ln x - 5} \leq 1$.

Solution. Note that $x \neq 1$.

$$\frac{x-5}{1-x} \geq 1 \implies (x-5)(1-x) \geq (1-x)^2 \implies 2x^2 - 8x + 6 \leq 0 \implies 2(x-1)(x-3) \leq 0.$$



Thus, $1 < x \leq 3$.

Consider the substitution $x \mapsto \ln x$. Taking reciprocals, we have our desired inequality $0 < \frac{1-\ln x}{\ln x - 5} \leq 1$. Hence,

$$1 < \ln x \leq 3 \implies e < x \leq e^3.$$

* * * * *

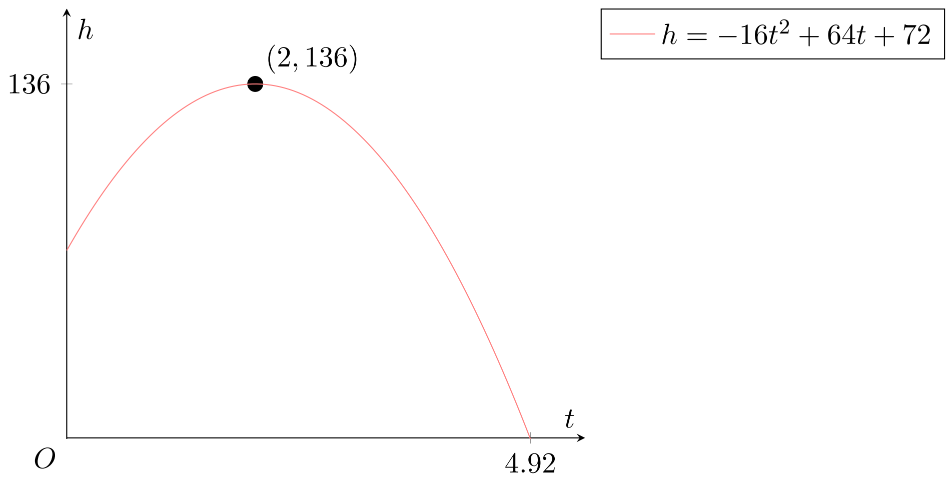
Problem 13. A small rocket is launched from a height of 72 m from the ground. The height of the rocket in metres, h, is represented by the equation $h = -16t^2 + 64t + 72$, where t is the time in seconds after the launch.

(a) Sketch the graph of h against t.

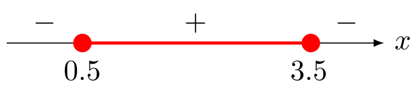
(b) Determine the number of seconds that the rocket will remain at or above 100 m from the ground.

Solution.

Part (a).



Part (b). Note that $-16t^2 + 64t + 72 \geq 100 \implies -4(2t-1)(2t-7) \geq 0$.

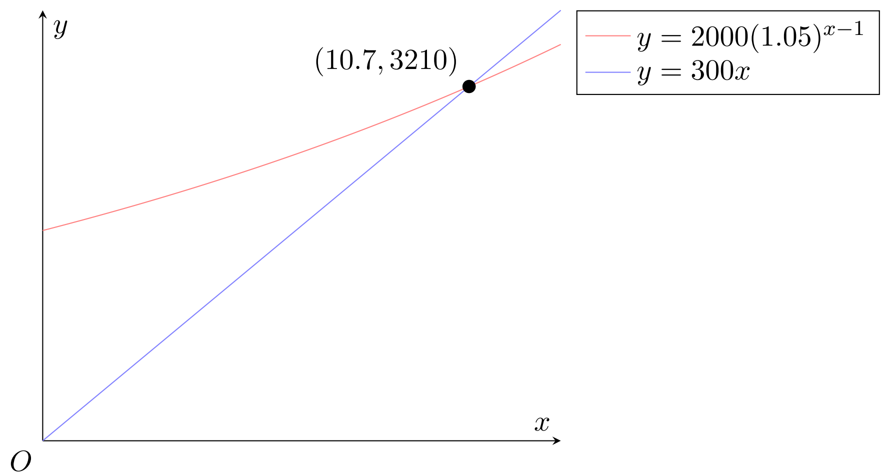


Thus, the rocket will remain at or above 100 m from the ground for 3 seconds.

* * * * *

Problem 14. Xinxin, a new graduate, starts work at a company with an initial monthly pay of \$2,000. For every subsequent quarter that she works, she will get a pay increase of 5%, leading to a new monthly pay of $2000(1.05)^{n-1}$ dollars in the nth quarter, where n is a positive integer. She also gives a regular donation of \$$300n$ in the nth quarter that she works. However, she will stop the donation when her monthly pay falls below the donation amount. At which quarter will this first happen?

Solution. Consider the curves $y = 2000(1.05)^{x-1}$ and $y = 300x$.



Hence, Xinxin will stop donating in the 11th quarter.

Self-Practice A1

Problem 1. On joining ABC International School, each of the 200 students is placed in exactly one of the four performing arts groups: Choir, Chinese Orchestra, Concert Band and Dance. The following table shows some information about each of the performing arts groups:

Performing Arts Group	**Choir**	**Chinese Orchestra**	**Concert Band**	**Dance**
Membership Fee (per student per month)	\$15	\$20	\$20	\$18
Instructor Fee (per student per month)	\$50	\$60	\$75	\$40
Costume Fee (one-time payment per student)	\$45	?	\$40	\$60
No. of Training Hours	5	6	8	7

In a typical month, the school collects a total of \$3,721 for membership fee from the students, and pays the instructors a total sum of \$11,830 (assuming that this sum of money is fully paid by the students). As for the training in a typical week, students from Chinese Orchestra and Concert Band spend in total 431 hours more than their peers in Choir and Dance. Find the enrolment in each of the performing arts groups.

Hence, find the costume fee paid by each student from Chinese Orchestra if a vendor charges a total of \$9,440 for all the costumes for the four performing arts groups.

Solution. Let a, b, c, d be the number of students in Choir, Chinese Orchestra, Concert Band and Dance respectively. From the given information, we have the following equations:

$$\begin{cases} a + b + c + d = 200 \\ 15a + 20b + 20c + 18d = 3721 \\ 50a + 60b + 75c + 40d = 11830 \\ -5a + 6b + 8c - 7d = 431 \end{cases}$$

Using G.C., we obtain the unique solution

$$a = 43, \quad b = 65, \quad c = 60, \quad d = 32.$$

Let the Chinese Orchestra's custom fee (per student) be x. From the given information, we have the following equation:

$$45a + xb + 40c + 60d = 9440.$$

Hence,

$$x = \frac{9440 - 45a - 40c - 60d}{b} = 49.$$

Thus, the costume fee paid by each student from Chinese Orchestra is \$49.

* * * * *

Problem 2. Solve the inequality $(x+2)^2\left(x^2+2x-8\right) \geq 0$.

Solution. Since $(x+2)^2 \geq 0$, we can remove it from the inequality, keeping in mind that $x = -2$ is a solution. We are hence left with

$$x^2 + 2x - 8 = (x+4)(x-2) \geq 0.$$

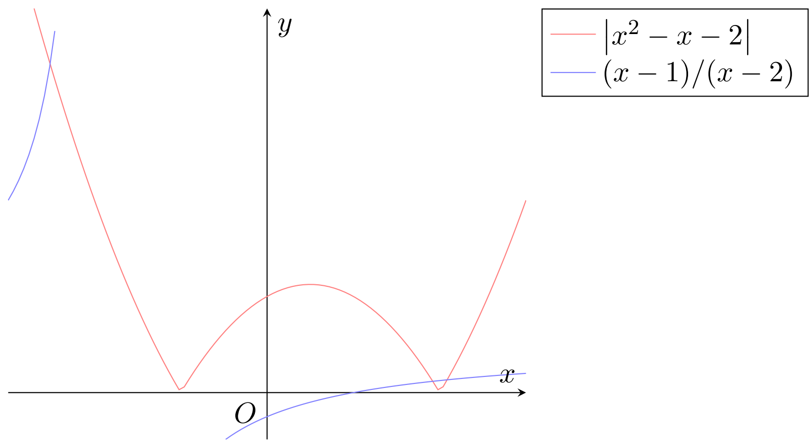
Since this quadratic is concave up, we clearly have $x \le -4$ or $x \ge 2$. Altogether, we have

$$x \le -4 \text{ or } x = -2 \text{ or } x \ge 2.$$

* * * * *

Problem 3. By using a graphical method, solve the inequality $\left|x^2 - x - 2\right| \ge \frac{x-1}{x+2}$.

Solution.



From the graph, the x-coordinates of the intersection points are -2.51, 1.92 and 2.09. Hence,

$$x \le -2.51 \quad \text{and} \quad -2 \le x \le 1.92 \quad \text{and} \quad x \ge 2.09.$$

* * * * *

Problem 4. Show that $x^2 + 2x + 3$ is always positive for all real values of x. Hence, solve the inequality $\frac{x^2+2x+3}{3+2x-x^2} \le 0$. Deduce the solution set of the inequality $\frac{x^2+2|x|+3}{3+2|x|-x^2} \le 0$.

Solution. Note that the discriminant of $x^2 + 2x + 3 = 0$ is $\Delta = 2^2 - 4(1)(3) = -8 < 0$. Since the y-intercept is positive $(3 > 0)$, it follows that $x^2 + 2x + 3$ is always positive for real x.

Consider the inequality $\frac{x^2+2x+3}{3+2x-x^2} \le 0$. Since $x^2 + 2x + 3$ is always positive, it suffices to solve $3 + 2x - x^2 \le 0$. Observe that the roots of $3 + 2x - x^2 = 0$ are $x = 3$ and $x = -1$. Since $3 + 2x - x^2$ is concave down, we have

$$x \le -1 \quad \text{or} \quad x \ge 3.$$

Replacing x with $|x|$, we get $|x| \le -1$ (no solutions) and $|x| \ge 3$, whence $x \le -3$ or $x \ge 3$. The solution set is thus

$$\{x \in \mathbb{R} : x \le -3 \text{ or } x \ge 3\}.$$

Problem 5. Without use of a graphing calculator, solve the inequality $\frac{2x^2-7x+6}{x^2-x-2} \geq 1$. Deduce the range of values of x such that

(a) $\frac{2(\ln x)^2-7\ln x+6}{(\ln x)^2-\ln x-2} > 1$

(b) $\frac{2-7x+6x^2}{1-x-2x^2} \geq 1$

Solution. Moving all terms to one side, we get

$$\frac{2x^2-7x+6}{x^2-x-2} \geq 1 \implies \frac{x^2-6x+8}{x^2-x-2} \geq 0.$$

Note that x^2-6x+8 factors as $(x-2)(x-4)$ while x^2-x-2 factors as $(x-2)(x+1)$. Hence,

$$\frac{x-4}{x+1} \geq 0 \implies (x-4)(x+1) \geq 0.$$

Thus, we clearly have

$$x < -1 \quad \text{or} \quad x \geq 4.$$

Note that $x \neq -1$ since $x^2-x-2 \neq 0$.

Part (a). Replacing x with $\ln x$, we get

$$\ln x < -1 \quad \text{or} \quad \ln x \geq 4,$$

whence

$$0 \leq x < \mathrm{e}^{-1} \quad \text{or} \quad x \geq \mathrm{e}^4.$$

Part (b). Replacing x with $1/x$, we get

$$\frac{1}{x} < -1 \quad \text{or} \quad \frac{1}{x} \geq 4.$$

Hence,

$$-1 < x < 0 \quad \text{or} \quad 0 < x \leq \frac{1}{4}.$$

Note that $x = 0$ also satisfies the inequality $(2 \geq 1)$. Hence,

$$-1 < x \leq \frac{1}{4}.$$

* * * * *

Problem 6. It is given that $y = \frac{x^2+x+1}{x-1}$, $x \in \mathbb{R}$, $x \neq 1$. Without using a calculator, find the set of values that y can take.

Solution. Clearing denominators, we have

$$y(x-1) = x^2+x+1 \implies x^2+(1-y)x+(1+y) = 0.$$

Since we are interested in the set of values that y can take, we want this quadratic to have roots. Hence, the discriminant Δ should be non-negative:

$$\Delta = (1-y)^2 - 4(1+y) = y^2 - 6y - 3 \geq 0.$$

Completing the square,

$$(y-3)^2 \geq 12 \implies |y-3| \geq \sqrt{12} = 2\sqrt{3}.$$

Hence,

$$y \leq 3 - 2\sqrt{3} \text{ or } y \geq 3 + 2\sqrt{3},$$

whence the solution set is

$$\left\{ y \in \mathbb{R} : y \leq 3 - 2\sqrt{3} \text{ or } y \geq 3 + 2\sqrt{3} \right\}.$$

* * * * *

Problem 7 (). Solve for x, in terms of a, the inequality

$$\left| x^2 - 3ax + 2a^2 \right| < \left| x^2 + 3ax - a^2 \right|,$$

where $x \in \mathbb{R}$, $a \neq 0$.

Solution. Squaring both sides, we get

$$\left(x^2 - 3ax + 2a^2 \right)^2 < \left(x^2 + 3ax - a^2 \right)^2.$$

Collecting terms on one side,

$$\left(x^2 + 3ax - a^2 \right)^2 - \left(x^2 - 3ax + 2a^2 \right)^2 = 3a\,(2x - a)\left(2x^2 + a^2 \right) > 0.$$

Clearly, $2x^2 + a^2 > 0$ for all x. We are hence left with $a(2x - a) > 0$.

Case 1. If $a > 0$, then $2x - a > 0$, whence $x > a/2$.

Case 2. If $a < 0$, then $2x - a < 0$, whence $x < a/2$.

* * * * *

Problem 8 (). Find constants a, b, c and d such that $1 + 2^3 + 3^3 + \cdots + n^3 = an^4 + bn^3 + cn^2 + dn$.

Solution 1. Substituting $n = 1, 2, 3, 4$ into the equation, we get the system

$$\begin{cases} a + b + c + d = 1 \\ 16a + 8b + 4c + 2d = 9 \\ 81a + 27b + 9c + 3d = 36 \\ 256a + 64b + 16c + 4d = 100 \end{cases}.$$

Solving, we have

$$a = \frac{1}{4}, \quad b = \frac{1}{2}, \quad c = \frac{1}{4}, \quad d = 0.$$

Solution 2. Recall that

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

Now observe that

$$(k+1)^3 - 1 = \sum_{k=1}^{n} \left[(k+1)^3 - k^3 \right] = \sum_{k=1}^{n} \left(3k^2 + 3k + 1 \right).$$

Rearranging, we obtain

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Similarly, we have

$$(k+1)^4 - 1 = \sum_{k=1}^{n} \left[(k+1)^4 - k^4\right] = \sum_{k=1}^{n} \left(4k^3 + 6k^2 + 4k + 1\right),$$

whence we obtain, upon rearranging,

$$\sum_{k=1}^{n} k^3 = \frac{n^4 + 2n^3 + n^2}{4}.$$

Comparing coefficients, we have

$$a = \frac{1}{4}, \quad b = \frac{1}{2}, \quad c = \frac{1}{4}, \quad d = 0.$$

* * * * *

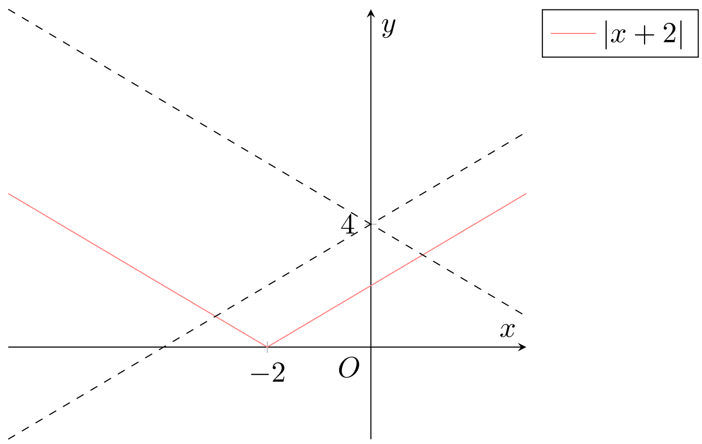
Problem 9 (🌶).

(a) By means of a sketch, or otherwise, state the range of values of a for which the equation $|x+2| = ax + 4$ has two distinct real roots.

(b) Solve the inequality $|x+2| < ax + 4$.

Solution.

Part (a).



Consider the figure above. Clearly, for 2 distinct roots (i.e. 2 distinct intersection points), we need $-1 < a < 1$.

Part (b). Note that the x-coordinate of the point of intersection between $y = ax + 4$ and $y = x + 2$ is:

$$x + 2 = ax + 4 \implies x = \frac{-2}{a-1}.$$

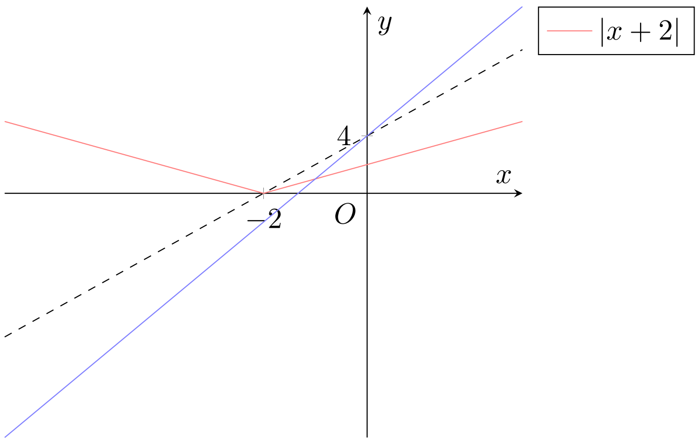
Similarly, the x-coordinate of the point of intersection between $y = ax+4$ and $y = -(x+2)$ is:

$$x + 2 = ax + 4 \implies x = \frac{-6}{a+1}.$$

Now consider the inequality $|x+2| < ax + 4$.

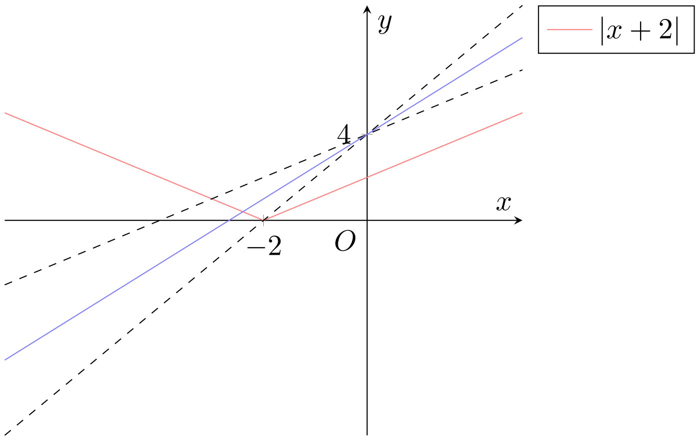
Case 1: $a > 2$. $y = ax + 4$ only intersects the line $y = x + 2$. Hence,

$$x > \frac{-2}{a-1}.$$



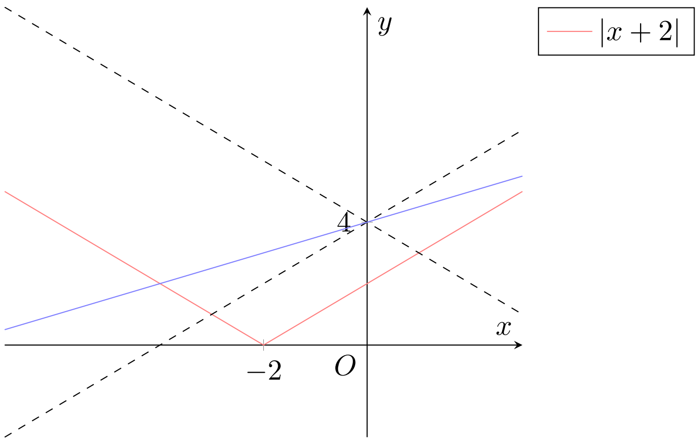
Case 2: $1 \leq a \leq 2$. $y = ax + 4$ only intersects the line $y = -(x + 2)$. Hence,

$$x \geq \frac{-6}{a-1}.$$



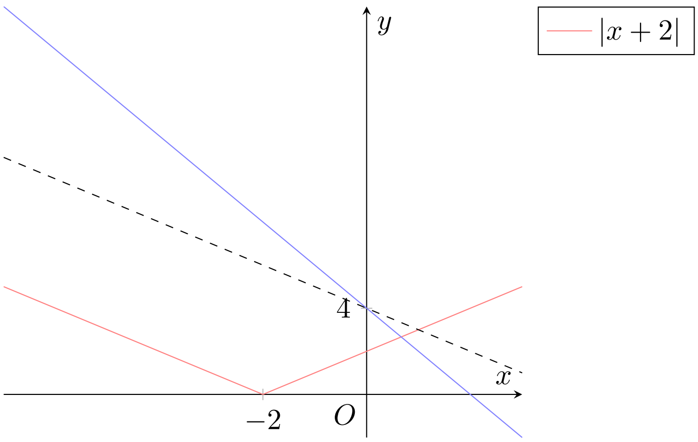
Case 3: $-1 < a < 1$. $y = ax + 4$ intersects both $y = x + 2$ and $y = -(x + 2)$. Hence,

$$\frac{-6}{a-1} < x < \frac{-2}{a-1}.$$



Case 4: $a \leq -1$. $y = ax + 4$ only intersects the line $y = x + 2$. Hence,

$$x \leq \frac{-2}{a-1}.$$



y
$|x+2|$
4
x
-2
O

Assignment A1

Problem 1. A traveller just returned from Germany, France and Spain. The amount (in dollars) that he spent each day on housing, food and incidental expenses in each country are shown in the table below.

Country	Housing	Food	Incidental Expenses
Germany	28	30	14
France	23	25	8
Spain	19	22	12

The traveller's records of the trip indicate a total of \$391 spent for housing, \$430 for food and \$180 for incidental expenses. Calculate the number of days the traveller spent in each country.

He did his account again and the amount spent on food is \$337. Is this record correct? Why?

Solution. Let g, f and s represent the number of days the traveller spent in Germany, France and Spain respectively. From the table, we obtain the following system of equations:

$$\begin{cases} 23f + 28g + 19s = 391 \\ 25f + 30g + 22s = 430 \\ 8f + 14g + 12s = 180 \end{cases}$$

This gives the unique solution $g = 4$, $f = 8$ and $s = 5$. The traveller thus spent 4 days in Germany, 8 days in France and 5 days in Spain.

Consider the scenario where the amount spent on food is \$337.

$$\begin{cases} 23f + 28g + 19s = 391 \\ 25f + 30g + 22s = 337 \\ 8f + 14g + 12s = 180 \end{cases}$$

This gives the unique solution $g = 66$, $f = -27$ and $s = -44$. The record is hence incorrect as f and s must be positive.

* * * * *

Problem 2.

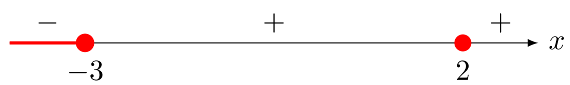
(a) Solve algebraically $x^2 - 9 \geq (x+3)\left(x^2 - 3x + 1\right)$.

(b) Solve algebraically $\frac{7-2x}{3-x^2} \leq 1$.

Solution.

Part (a).

$$\begin{aligned} & x^2 - 9 \geq (x+3)\left(x^2 - 3x + 1\right) \\ \Longrightarrow\ & (x+3)(x-3) \geq (x+3)\left(x^2 - 3x + 1\right) \\ \Longrightarrow\ & (x+3)\left(x^2 - 4x + 4\right) \leq 0 \\ \Longrightarrow\ & (x+3)(x-2)^2 \leq 0 \end{aligned}$$



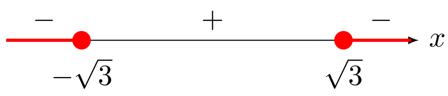
Thus, $x \leq -3$ or $x = 2$.

Part (b). Note that $3 - x^2 \neq 0 \implies x \neq \pm\sqrt{3}$.

$$\begin{aligned} \frac{7-2x}{3-x^2} &\leq 1 \\ \implies \frac{7-2x}{3-x^2} - \frac{3-x^2}{3-x^2} &\leq 0 \\ \implies \frac{x^2-2x+4}{3-x^2} &\leq 0 \end{aligned}$$

Observe that $x^2 - 2x + 4 = (x-1)^2 + 3 > 0$. Dividing through by $x^2 - 2x + 4$, we obtain

$$\begin{aligned} \frac{1}{3-x^2} &\leq 0 \\ \implies 3 - x^2 &\leq 0 \end{aligned}$$



Thus, $x < -\sqrt{3}$ or $x > \sqrt{3}$.

* * * * *

Problem 3.

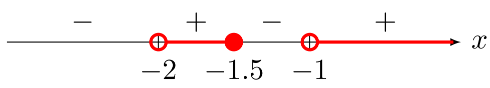
(a) Without using a calculator, solve the inequality $\frac{3x+4}{x^2+3x+2} \geq \frac{1}{x+2}$.

(b) Hence, deduce the set of values of x that satisfies $\frac{3|x|+4}{x^2+3|x|+2} \geq \frac{1}{|x|+2}$.

Solution.

Part (a). Note that $x^2 + 3x + 2 \neq 0$ and $x + 2 \neq 0$, whence $x \neq -1, -2$.

$$\begin{aligned} \frac{3x+4}{x^2+3x+2} &\geq \frac{1}{x+2} \\ \implies \frac{3x+4}{(x+2)(x+1)} &\geq \frac{1}{x+2} \\ \implies (3x+4)(x+2)(x+1) &\geq (x+2)(x+1)^2 \\ \implies (x+2)(x+1)(2x+3) &\geq 0 \end{aligned}$$



Thus, $-2 < x \leq -\frac{3}{2}$ or $x > -1$.

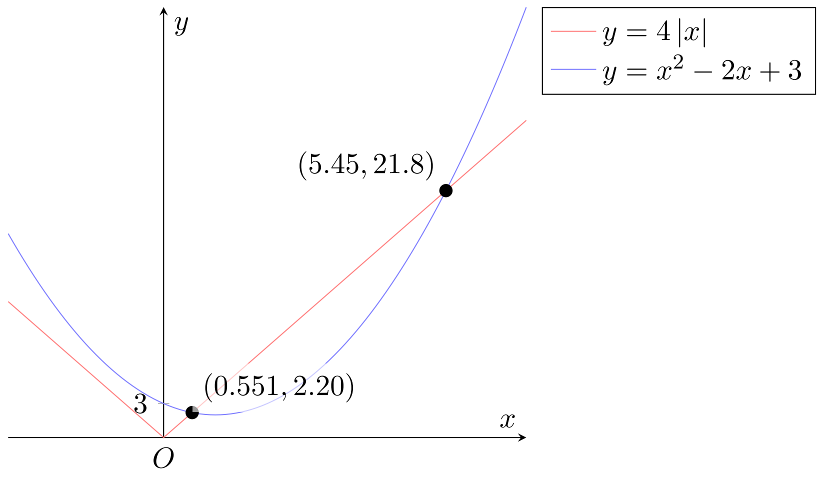
Part (b). Observe that $|x|^2 = x^2$. Hence, with the map $x \mapsto |x|$, we obtain

$$-2 < |x| \leq -\frac{3}{2} \text{ or } |x| > -1.$$

Since $|x| \geq 0$, we have that $|x| > -1$ is satisfied for all real x. Hence, the solution set is $\mathbb{R}$.

Problem 4. On the same diagram, sketch the graphs of $y = 4|x|$ and $y = x^2 - 2x + 3$. Hence or otherwise, solve the inequality $4|x| \geq x^2 - 2x + 3$.

Solution.



From the graph, we see that $0.551 \leq x \leq 5.45$.

A2 Numerical Methods of Finding Roots

Tutorial A2

Problem 1. Without using a graphing calculator, show that the equation $x^3+2x^2-2=0$ has exactly one positive root.

This root is denoted by α and is to be found using two different iterative methods, starting with the same initial approximation in each case.

(a) Show that α is a root of the equation $x = \sqrt{\frac{2}{x+2}}$, and use the iterative formula $x_{n+1} = \sqrt{\frac{2}{x_n+2}}$, with $x_1 = 1$, to find α correct to 2 significant figures.

(b) Use the Newton-Raphson method, with $x_1 = 1$, to find α correct to 3 significant figures.

Solution. Let $f(x) = x^3+2x^2-2$. Observe that for all $x > 0$, we have $f'(x) = 3x^2+4x > 0$. Hence, $f(x)$ is strictly increasing on $(0, \infty)$. Since $f(0)f(1) = (-2)(1) < 0$, it follows that $f(x)$ has exactly one positive root.

Part (a). We know $f(\alpha) = 0$. Hence,

$$\alpha^3 + 2\alpha^2 - 2 = 0 \implies \alpha^2(\alpha+2) = 2 \implies \alpha^2 = \frac{2}{\alpha+2} \implies \alpha = \sqrt{\frac{2}{\alpha+2}}.$$

Note that we reject the negative branch since $\alpha > 0$. We hence see that α is a root of the equation $x = \sqrt{\frac{2}{x+2}}$. Using the iterative formula $x_{n+1} = \sqrt{\frac{2}{x_n+2}}$ with $x_1 = 1$, we have

n	x_n
1	1
2	0.81650
3	0.84268
4	0.83879

Hence, $\alpha = 0.84$ (2 s.f.).

Part (b). Using the Newton-Raphson method ($x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$) with $x_1 = 1$, we have

n	x_n
1	1
2	0.857143
3	0.839545
4	0.839287
5	0.839287

Hence, $\alpha = 0.839$ (3 s.f.).

Problem 2.

(a) Show that the tangent at the point $(\mathrm{e}, 1)$ to the graph $y = \ln x$ passes through the origin, and deduce that the line $y = mx$ cuts the graph $y = \ln x$ in two points provided that $0 < m < 1/\mathrm{e}$.

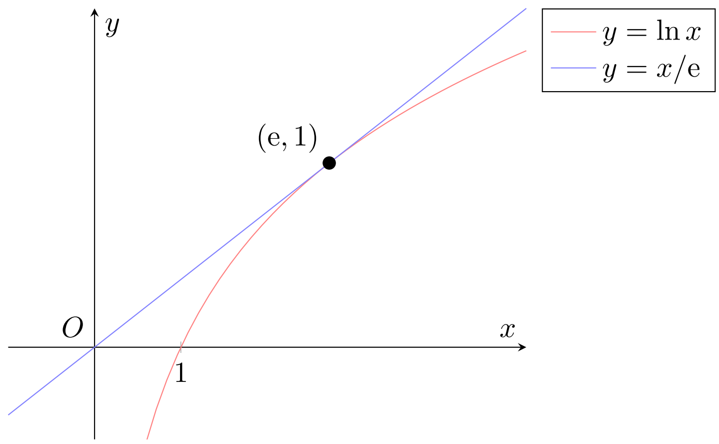
(b) For each root of the equation $\ln x = x/3$, find an integer n such that the interval $n < x < n + 1$ contains the root. Using linear interpolation, based on $x = n$ and $x = n + 1$, find a first approximation to the smaller root, giving your answer to 1 decimal place. Using your first approximation, obtain, by the Newton-Raphson method, a second approximation to the smaller root, giving your answer to 2 decimal places.

Solution.

Part (a). Note that the derivative of $y = \ln x$ at $x = \mathrm{e}$ is $1/\mathrm{e}$. Using the point slope formula, we see that the equation of the tangent at the point $(\mathrm{e}, 1)$ is given by

$$y - 1 = \frac{x - \mathrm{e}}{\mathrm{e}} \implies y = \frac{x}{\mathrm{e}}.$$

Since $x = 0$, $y = 0$ is clearly a solution, the tangent passes through the origin. From the graph below, it is clear that for $y = mx$ to intersect $y = \ln x$ twice, we must have $0 < m < 1/\mathrm{e}$.



Part (b). Consider $f(x) = x/3 - \ln x$. Let α and β be the smaller and larger root to $f(x) = 0$ respectively. Observe that $f(1)f(2) = (1)(-0.03) < 0$ and $f(4)f(5) = (-0.05)(0.06) < 0$. Thus, for the smaller root α, $n = 1$, while for the larger root β, $n = 4$.

Let x_1 be the first approximation to α. Using linear interpolation, we have

$$x_1 = \frac{f(2) - 2f(1)}{f(2) - f(1)} = 1.9 \text{ (1 d.p.)}$$

Note that $f'(x) = 1/3 - 1/x$. Using the Newton-Raphson method ($x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$), we have

n	x_n
1	1.9
2	1.85585
3	1.85718

Hence, $\alpha = 1.86$ (2 d.p.).

Problem 3. Find the exact coordinates of the turning points on the graph of $y = f(x)$ where $f(x) = x^3 - x^2 - x - 1$. Deduce that the equation $f(x) = 0$ has only one real root α, and prove that α lies between 1 and 2. Use the Newton-Raphson method applied to the equation $f(x) = 0$ to find a second approximation x_2 to α, taking x_1, the first approximation, to be 2. With reference to a graph of $y = f(x)$, explain why all further approximations to α by this process are always larger than α.

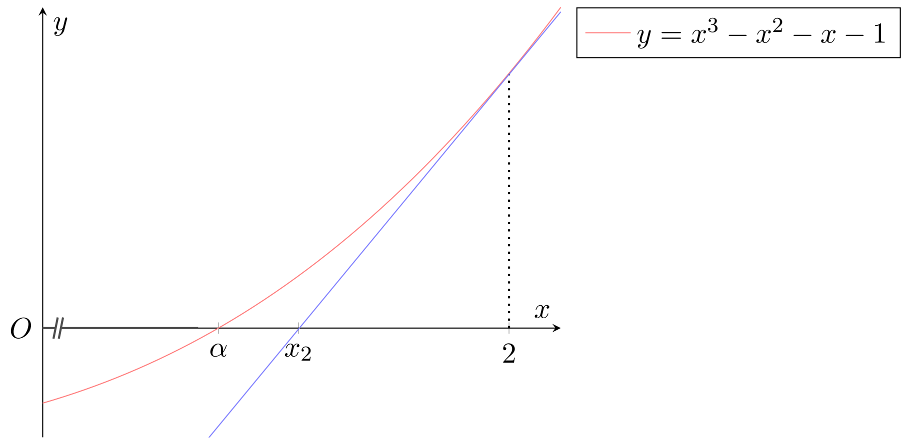
Solution. For turning points, $f'(x) = 0$.

$$f'(x) = 0 \implies 3x^2 - 2x - 1 = 0 \implies (3x+1)(x-1) = 0.$$

Hence, $x = -1/3$ or $x = 1$. When $x = -1/3$, we have $y = -0.815$, giving the coordinate $(-1/3, -0.815)$. When $x = 1$, we have $y = -2$, giving the coordinate $(1, -2)$.

Observe that $f(x)$ is strictly increasing for all $x > 1$. Further, since both turning points have a negative y-coordinate, it follows that $y < 0$ for all $x \leq 1$. Since $f(1)f(2) = (-2)(1) < 0$, the equation $f(x) = 0$ has only one real root.

Using the Newton-Raphson method with $x_1 = 2$, we have $x_2 = x_1 - f(x_1)/f'(x_1) = 13/7$.



Since x_2 lies on the right of α, the Newton-Raphson method gives an over-estimation given an initial approximation of 2. Thus, all further approximations to α will also be larger than α.

* * * * *

Problem 4. A curve C has equation $y = x^5 + 50x$. Find the least value of $\mathrm{d}y/\mathrm{d}x$ and hence give a reason why the equation $x^5 + 50x = 10^5$ has exactly one real root. Use the Newton-Raphson method, with a suitable first approximation, to find, correct to 4 decimal places, the root of the equation $x^5 + 50x = 10^5$. You should demonstrate that your answer has the required accuracy.

Solution. Since $y = x^5 + 50x$, we have $\mathrm{d}y/\mathrm{d}x = 5x^4 + 50$. Since $x^4 \geq 0$ for all real x, the minimum value of $\mathrm{d}y/\mathrm{d}x$ is 50.

Let $f(x) = x^5 + 50x$. Since $\min \mathrm{d}f/\mathrm{d}x = 50 > 0$, it follows that $f(x)$ is strictly increasing. Hence, $f(x)$ will intersect only once with the line $y = 10^5$, whence the equation $x^5 + 50x = 10^5$ has exactly one real root.

Observe that $f(9)f(10) = (-40901)(50) < 0$. Thus, there must be a root in the interval $(9, 10)$. We now use the Newton-Raphson method ($x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$) with $x_1 = 9$ as the first approximation.

n	x_n
1	9
2	10.2178921
3	10.0017491
4	9.9901221
5	9.9899912
6	9.9899900

Thus, the root is 9.9900 (4 d.p.).

Observe that $f(9.98995)f(9.99005) = (-2.00)(3.00) < 0$. Hence, the root lies in the interval $(9.98995, 9.99005)$ whence the calculated root has the required accuracy.

* * * * *

Problem 5.

(a) A function f is such that $f(4) = 1.158$ and $f(5) = -3.381$, correct to 3 decimal places in each case. Assuming that there is a value of x between 4 and 5 for which $f(x) = 0$, use linear interpolation to estimate this value.

For the case when $f(x) = \tan x$, and x is measured in radians, the value of $f(4)$ and $f(5)$ are as given above. Explain, with the aid of a sketch, why linear interpolation using these values does not give an approximation to a solution of the equation $\tan x = 0$.

(b) Show, by means of a graphical argument or otherwise, that the equation $\ln(x-1) = -2x$ has exactly one real root, and show that this root lies between 1 and 2.

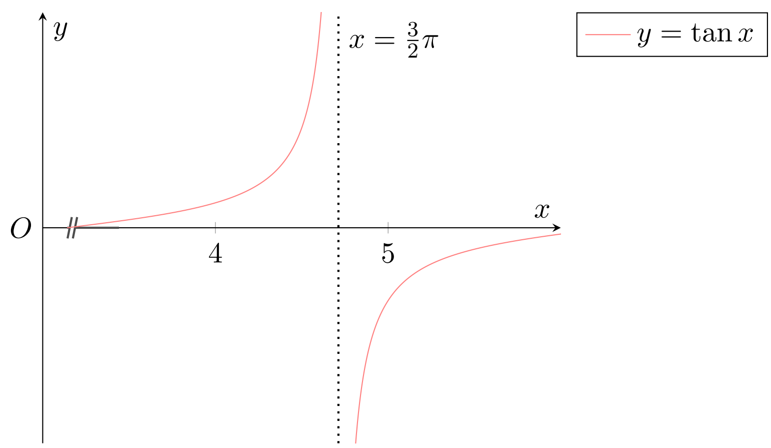
The equation may be written in the form $\ln(x-1) + 2x = 0$. Show that neither $x = 1$ nor $x = 2$ is a suitable initial value for the Newton-Raphson method in this case.

The equation may also be written in the form $x - 1 - \mathrm{e}^{-2x} = 0$. For this form, use two applications of the Newton-Raphson method, starting with $x = 1$, to obtain an approximation to the root, giving 3 decimal places in your answer.

Solution.

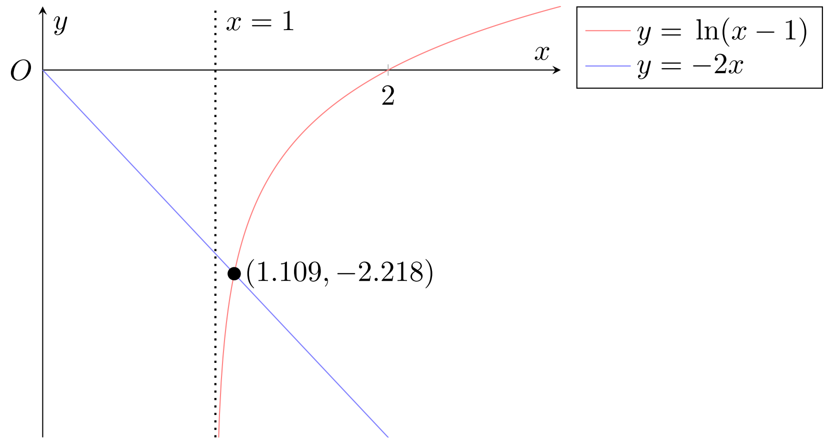
Part (a). Let the root of $f(x) = 0$ be α. Using linear interpolation on the interval $[4, 5]$, we have

$$\alpha = \frac{4f(5) - 5f(4)}{f(5) - f(4)} = 4.255 \text{ (3 d.p.)}.$$



Since $\tan x$ has a vertical asymptote at $x = 3\pi/2$, it is not continuous on $[4, 5]$. Thus, linear interpolation diverges when applied to the equation $\tan x = 0$.

Part (b).



Since there is only one intersection between the graphs $y = \ln(x-1)$ and $y = -2x$, there is only one real root to the equation $\ln(x-1) = -2x$. Furthermore, since $y = -2x$ is negative for all $x > 0$ and $y = \ln(x-1)$ is negative only when $1 < x < 2$, it follows that the root must lie between 1 and 2.

Let $f(x) = \ln(x-1) + 2x$. Then $f'(x) = \frac{1}{x-1} + 2$. Note that the Newton-Raphson method is given by $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$.

Since $f'(1)$ is undefined, an initial approximation of $x_1 = 1$ cannot be used for the Newton-Raphson method, which requires a division by $f'(1)$.

Using the Newton-Raphson method with the initial approximation $x_2 = 2$, we see that $x_2 = 1$. Once again, because $f'(1)$ is undefined, $x_1 = 2$ is also not a suitable initial value.

Let $g(x) = x - 1 - \mathrm{e}^{-2x}$. Then $g'(x) = 1 + 2\mathrm{e}^{-2x}$. Using the Newton-Raphson method with the initial approximation $x_1 = 1$, we have

n	x_n
1	1
2	1.106507
3	1.108857

Hence, $x = 1.109$ (3 d.p.).

* * * * *

Problem 6. The equation $x = 3\ln x$ has two roots α and β, where $1 < \alpha < 2$ and $4 < \beta < 5$. Using the iterative formula $x_{n+1} = F(x_n)$, where $F(x) = 3\ln x$, and starting with $x_0 = 4.5$, find the value of β correct to 3 significant figures. Find a suitable $F(x)$ for computing α.

Solution. Using the iterative formula $x_{n+1} = F(x_n)$, we have

n	x_n	n	x_n
0	4.5	5	4.53175
1	4.51223	6	4.53333
2	4.52038	7	4.53437
3	4.52579	8	4.53506
4	4.52937	9	4.53551

Hence, $\beta = 4.54$ (3 s.f.).

Note that $x = 3\ln x \implies x = \mathrm{e}^{x/3}$. Observe that $\mathrm{d}(\mathrm{e}^{x/3})/\mathrm{d}x = \frac{1}{3}\mathrm{e}^{x/3}$, which is between -1 and 1 for all $1 < x < 2$. Thus, the iterative formula $x_{n+1} = F(x_n)$ will converge, whence $F(x) = \mathrm{e}^{x/3}$ is suitable for computing α.

* * * * *

Problem 7. Show that the cubic equation $x^3 + 3x - 15 = 0$ has only one real root. This root is near $x = 2$. The cubic equation can be written in any one of the forms below:

(a) $x = \frac{1}{3}(15 - x^3)$

(b) $x = \frac{15}{x^2+3}$

(c) $x = (15 - 3x)^{1/3}$

Determine which of these forms would be suitable for the use of the iterative formula $x_{r+1} = F(x_r)$, where $r = 1, 2, 3, \ldots$.

Hence, find the root correct to 3 decimal places.

Solution. Let $f(x) = x^3 + 3x - 15$. Then $f'(x) = 3x^2 + 3 > 0$ for all real x. Hence, f is strictly increasing. Since f is continuous, $f(x) = 0$ has only one real root.

Part (a). Let $g_1(x) = \frac{1}{3}(15 - x^3)$. Then $g_1'(x) = -x^2$. For values of x near 2, $|g_1'(x)| > 1$. Hence, the iterative formula $x_{n+1} = g_1(x_n)$ will diverge and $g_1(x)$ is unsuitable.

Part (b). Let $g_2(x) = \frac{15}{x^2+3}$. Then $g_2'(x) = \frac{-30x}{(x^2+3)^2}$. For values of x near 2, $|g_2'(x)| > 1$. Hence, the iterative formula $x_{n+1} = g_2(x_n)$ will diverge and $g_2(x)$ is unsuitable.

Part (c). Let $g_3(x) = (15 - 3x)^{1/3}$. Then $g_3'(x) = -(15 - 3x)^{-2/3}$. For values of x near 2, $|g_3'(x)| < 1$. Hence, the iterative formula $x_{n+1} = g_3(x_n)$ will converge and $g_3(x)$ is suitable.

Using the iterative formula $x_{r+1} = g_3(x_r)$, we get

r	x_r
1	2
2	2.080084
3	2.061408
4	2.065793
5	2.064765

Hence, $x = 2.065$ (3 d.p.).

* * * * *

Problem 8. The equation of a curve is $y = f(x)$. The curve passes through the points $(a, f(a))$ and $(b, f(b))$, where $0 < a < b$, $f(a) > 0$ and $f(b) < 0$. The equation $f(x) = 0$ has precisely one root α such that $a < \alpha < b$. Derive an expression, in terms of a, b, $f(a)$ and $f(b)$, for the estimated value of α based on linear interpolation.

Let $f(x) = 3\mathrm{e}^{-x} - x$. Show that $f(x) = 0$ has a root α such that $1 < \alpha < 2$, and that for all x, $f'(x) < 0$ and $f''(x) > 0$. Obtain an estimate of α using linear interpolation to 2 decimal places, and explain by means of a sketch whether the value obtained is an over-estimate or an under-estimate.

Use one application of the Newton-Raphson method to obtain a better estimate of α, giving your answer to 2 decimal places.

Solution. Using the point-slope formula, the equation of the line that passes through both $(a, f(a))$ and $(b, f(b))$ is

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

Note that $(\alpha, 0)$ is approximately the solution to the above equation. Thus,

$$0 - f(a) \approx \frac{f(a) - f(b)}{a - b}(\alpha - a) \implies \alpha \approx \frac{bf(a) - af(b)}{f(a) - f(b)}.$$

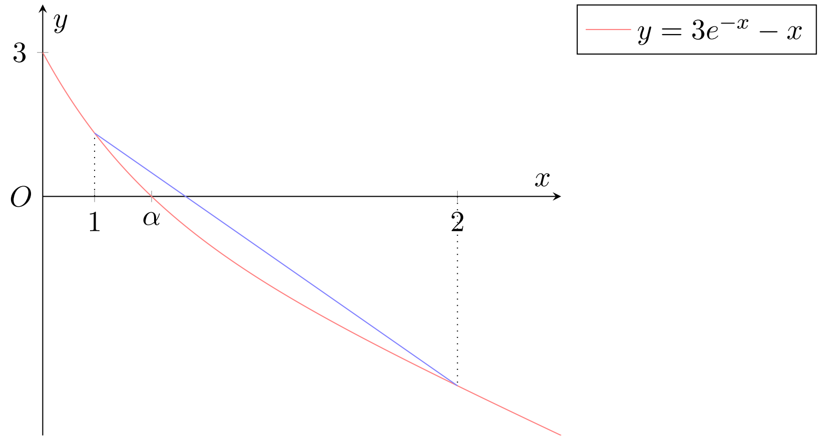
Since $f(x)$ is continuous, and $f(1)f(2) = (0.10)(-1.6) < 0$, there exists a root $\alpha \in (1, 2)$.

Note that $f'(x) = -3e^{-x} - 1$ and $f''(x) = 3e^{-x}$. Since $e^{-x} > 0$ for all x, we have that $f'(x) < 0$ and $f''(x) > 0$ for all x.

Using linear interpolation on the interval $(1, 2)$, we have

$$\alpha = \frac{2f(1) - f(2)}{f(1) - f(2)} = 1.06 \text{ (2 d.p.)}.$$

Since $f'(x) < 0$ and $f''(x) > 0$, we know that $f(x)$ is strictly decreasing and is concave upwards. $f(x)$ hence has the following shape:



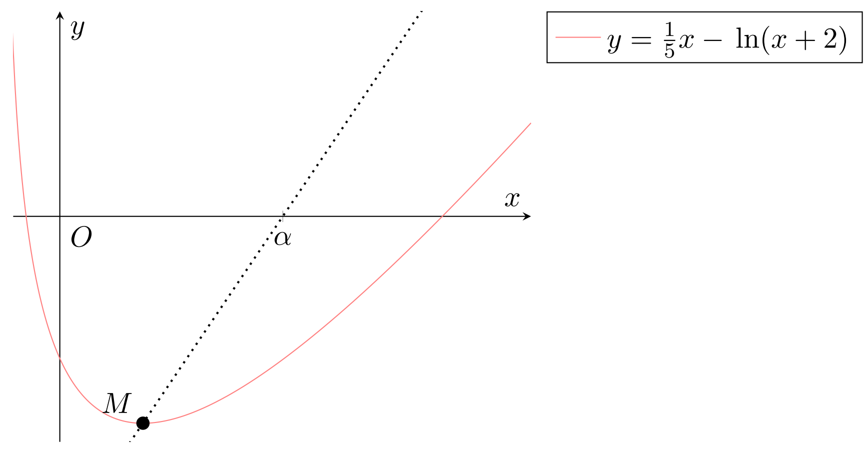
From the graph, we see that the value obtained is an over-estimate.

Using the Newton-Raphson method with the initial approximation $x_1 = 1.06$, we get

$$\alpha = x_1 - \frac{f(x_1)}{f'(x_1)} = 1.05 \text{ (2 d.p.)}.$$

* * * * *

Problem 9.



The diagram shows a sketch of the graph $y = x/3 - \ln(x + 2)$. Find the x-coordinate of the minimum point M on the graph, and verify that y is positive when $x = 20$.

Show that the gradient of the curve is always less than $1/5$. Hence, by considering the line through M having gradient $1/5$, show that the positive root of the equation $x/3 - \ln(x+2) = 0$ is greater than 8.

Use linear interpolation, once only, on the interval $[8, 20]$, to find an approximate value a for this positive root, giving your answer to 1 decimal place.

Using a as an initial value, carry out one application of the Newton-Raphson method to obtain another approximation to the positive root, giving your answer to 2 decimal places.

Solution. For stationary points, $y' = 0$.

$$y' = 0 \implies \frac{1}{5} - \frac{1}{x+2} \implies x = 3.$$

By the second derivative test, we see that $y''(x) = \frac{1}{(x+2)^2} > 0$. Hence, the x-coordinate of M is 3. Substituting $x = 20$ into the equation of the curve gives $y = 4 - \ln 22 = 0.909 > 0$.

We know that $y' = 1/5 - 1/(x+2)$, hence $y' < 1/5$ for all $x > -2$. Since the domain of the curve is $x > -2$, y' is always less than $1/5$.

Let $(\alpha, 0)$ be the coordinates of the root of the line through M having gradient $\frac{1}{5}$. We know that the coordinates of M are $(3, 3/5 - \ln 5)$. Taking the gradient of the line segment joining M and $(\alpha, 0)$, we get

$$\frac{(3/5 - \ln 5) - 0}{3 - \alpha} = \frac{1}{5} \implies \alpha = 5 \ln 5 = 8.05 > 8.$$

Since the gradient of the curve is always less than $1/5$, α represents the lowest possible value of the positive root of the curve. Hence, the positive root of the equation $x/5 - \ln(x+2) = 0$ is greater than 8.

Let $f(x) = x/5 - \ln(x+2)$. Using linear interpolation on the interval $[8, 20]$, we have

$$\alpha = \frac{8f(20) - 20f(8)}{f(20) - f(8)} = 13.2 \text{ (1 d.p.)}.$$

Using the Newton-Raphson method with the initial approximation $x_1 = 13.2$, we have

$$\alpha = x_1 - \frac{f(x_1)}{f'(x_1)} = 13.81 \text{ (2 d.p.)}.$$

* * * * *

Problem 10.

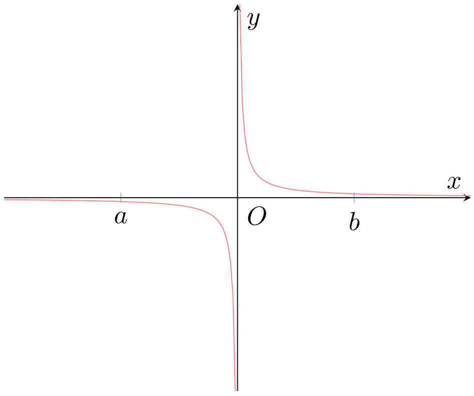
(a) The function f is such that $f(a)f(b) < 0$, where $a < b$. A student concludes that the equation $f(x) = 0$ has exactly one root in the interval (a, b). Draw sketches to illustrate two distinct ways in which the student could be wrong.

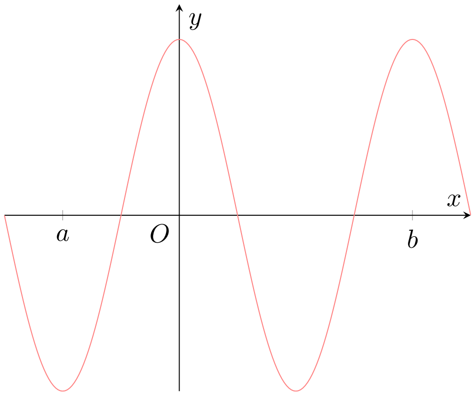
(b) The equation $\sec^2 x - \mathrm{e}^2 = 0$ has a root α in the interval $[1.5, 2.5]$. A student uses linear interpolation once on this interval to find an approximation to α. Find the approximation to α given by this method and comment on the suitability of the method in this case.

(c) The equation $\sec^2 x - \mathrm{e}^x = 0$ also has a root β in the interval $(0.1, 0.9)$. Use the Newton-Raphson method, with $f(x) = \sec^2 x - \mathrm{e}^x$ and initial approximation 0.5, to find a sequence of approximations $\{x_1, x_2, x_3, \ldots\}$ to β. Describe what is happening to x_n for large n, and use a graph of the function to explain why the sequence is not converging to β.

Solution.

Part (a).





Part (b). Let $f(x) = \sec^2 x - e^x$. Using linear interpolation on the interval $[1.5, 2.5]$,

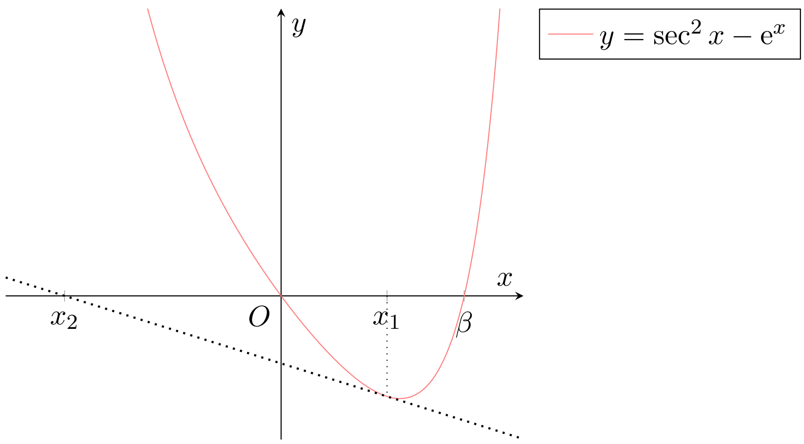
$$a = \frac{1.5f(2.5) - 2.5f(1.5)}{f(2.5) - f(1.5)} = 1.06 \text{ (2 d.p.)}.$$

$\sec^2 x$ is not continuous on the interval $[1.5, 2.5]$ due to the presence of an asymptote at $x = \pi/2$. Hence, linear interpolation is not suitable in this case.

Part (c). We know $f'(x) = 2\sec^2 x \tan x - e^x$. Using the Newton-Raphson method with the initial approximation $x_1 = 0.5$,

r	x_r
1	0.5
2	-1.02272
3	-0.75526
4	-0.40306
5	-0.09667
6	-0.00466
7	-0.00000

As $n \to \infty$, $x_n \to 0^-$.



From the above graph, we see that the initial approximation of $x_1 = 0.5$ is past the turning point. Hence, all subsequent approximations will converge to the root at 0 instead of the root at β. Thus, the sequence does not converge to β.

Problem 11. The function f is given by $f(x) = \sqrt{1-x^2} + \cos x - 1$ for $0 \le x \le 1$. It is known, from graphical work, that the equation $f(x) = 0$ has a single root $x = \alpha$.

(a) Express g(x) in terms of x, where $g(x) = x - \frac{f(x)}{f'(x)}$.

A student attempts to use the Newton-Raphson method, based on the form $x_{n+1} = g(x_n)$, to calculate the value of α correct to 3 decimal places.

(b) (i) The student first uses an initial approximation to α of $x_1 = 0$. Explain why this will be unsuccessful in finding a value for α.

(ii) The student next uses an initial approximation to α of $x_1 = 1$. Explain why this will also be unsuccessful in finding a value for α.

(iii) The student then uses an initial approximate to α of $x_1 = 0.5$. Investigate what happens in this case.

(iv) By choosing a suitable value for x_1, use the Newton-Raphson method, based on the form $x_{n+1} = g(x_n)$, to determine α correct to 3 decimal places.

Solution.

Part (a). We know $f'(x) = \frac{-x}{\sqrt{1-x^2}} - \sin x$. Hence,

$$g(x) = x - \frac{\sqrt{1-x^2} + \cos x - 1}{\frac{-x}{\sqrt{1-x^2}} - \sin x}.$$

Part (b).

Part (b)(i). Observe that $f'(0) = 0$. Hence, $g(0)$ is undefined. Thus, starting with an initial approximation of $x_1 = 0$ will be unsuccessful in finding a value for α.

Part (b)(ii). Observe that $\sqrt{1-x^2}$ is 0 when $x = 1$. Hence, $f'(0)$ is undefined. Thus, $g(0)$ is also undefined. Hence, starting with an initial approximation of $x_1 = 1$ will also be unsuccessful in finding a value for α.

Part (b)(iii). When $x_1 = 0.5$, we have $x_2 = g(x_1) = 1.20$. Since $g(x)$ is only defined for $0 \le x \le 1$, $x_3 = g(x_2)$ is undefined. Hence, an initial approximation of $x_1 = 0.5$ will also be unsuccessful in finding a value for α.

Part (b)(iv). Using the Newton-Raphson method with $x_1 = 0.9$, we have

r	x_r
1	0.9
2	0.92019
3	0.91928
4	0.91928

Thus, $\alpha = 0.919$ (3 d.p.).

Self-Practice A2

Problem 1.

(a) Sketch on the same diagram the graphs of $y = x - 1$ and $y = ke^{-3x}$, where $-1 < k < 0$. State the number of real roots of the equation $ke^{-3x} - (x - 1) = 0$.

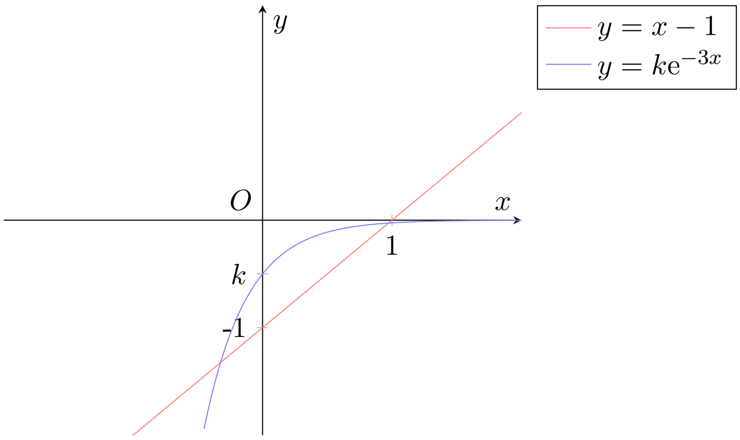
For the case $k = 1$, sketch appropriate graphs to show that the equation $e^{-3x} - (x - 1) = 0$ has exactly one real root. Denoting this real root by α, find the integer n such that the interval $[n - 1, n]$ contains α. Use linear interpolation, once only, on this interval to find an estimate for α, giving your answer correct to 2 decimal places.

(b) Let $f(x) = e^{-3x} - (x - 1)$. By considering the signs of $f'(x)$ and $f''(x)$ for all real values of x, explain with the aid of a simple diagram whether the value of α obtained in (a) is an over-estimate or an under-estimate.

(c) Taking the value of α obtained in (a) as the initial value, apply the Newton-Raphson method to find the value of α correct to 3 decimal places.

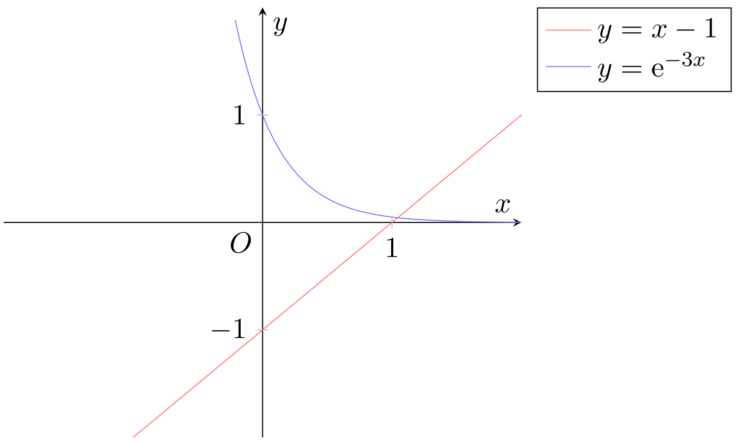
Solution.

Part (a).



There are 2 real roots to $ke^{-3x} - (x - 1) = 0$ when $-1 < k < 0$.

Note that $e^{-3x} - (x - 1) = 0$ is equivalent to $e^{-3x} = x - 1$. We hence plot $y = e^{-3x}$ and $y = x - 1$.



Since the two curves only intersect at one point, there is only one root to the equation $e^{-3x} - (x - 1) = 0$.

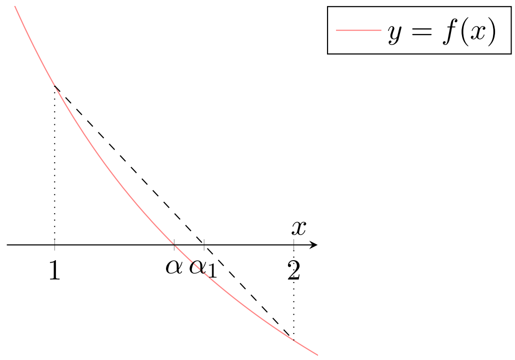
From the graph, $\alpha \in (1, 2)$, so $n = 2$. Let $f(x) = \mathrm{e}^{-3x} - (x - 1)$. Using linear interpolation on the interval $(1, 2)$, we obtain

$$\alpha_1 = \frac{2f(1) - 1f(2)}{f(1) - f(2)} = 1.05 \text{ (2 d.p.)}.$$

Part (b). For all $x \in \mathbb{R}$, we have

$$f'(x) = -3\mathrm{e}^{-3x} - 1 < 0 \quad \text{and} \quad f''(x) = 9\mathrm{e}^{-3x} > 0.$$

Thus, the graph of $y = f(x)$ is decreasing and convex.



From the above figure, we see that the estimate given by linear interpolation is an overestimate.

Part (c). The recursive formula given by the Newton-Raphson method is

$$\alpha_{n+1} = \alpha_n - \frac{f(\alpha_n)}{f'(\alpha_n)} = \alpha_n - \frac{\mathrm{e}^{-3\alpha_n} - (\alpha_n - 1)}{-3\mathrm{e}^{-3\alpha_n} - 1}.$$

Using the initial estimate $\alpha_1 = 1.05$, we have $\alpha_2 = 1.044$ (3 d.p.). Indeed, since f is continuous on $(1.0435, 1.0444)$, and

$$f(1.0435)f(1.0444) = -1.6 \times 10^{-7} < 0,$$

by the Intermediate Value Theorem, we conclude that $\alpha \in (1.0435, 1.0444)$, thus $\alpha = 1.044$ (3 d.p.).

* * * * *

Problem 2. The equation $f(x) = 0$ where $f(x) = \frac{1}{x} - 2 + \ln x$ has exactly two real roots α and β.

Verify that the larger root β lies between 6 and 7 and use one application of linear interpolation on the interval $[6, 7]$ to estimate this root, giving your answer correct to 2 decimal places.

Sketch the graph of $y = f(x)$, stating clearly the coordinates of the turning point. Using the graph of $y = f(x)$, deduce the integer N such that the interval $[N - 1, N]$ contains the smaller root α.

An attempt to calculate the smaller root α is made. Explain why neither $x = 0$ nor $x = 1$ is a suitable initial value for the Newton-Raphson method in this case.

Taking $x = 0.3$ as the initial value, use the Newton-Raphson method to find a second approximation to the root α, giving your answer correct to three decimal places.

Solution. Since f is continuous over $(6,7)$ and

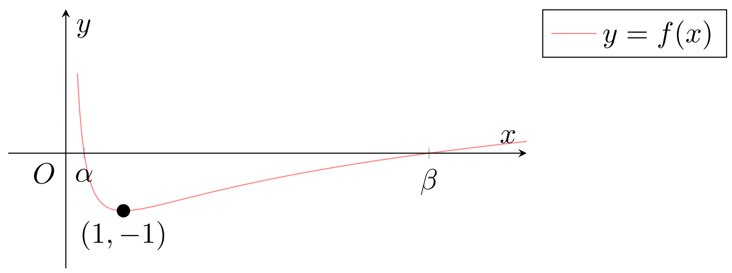
$$f(6)f(7) = -0.00369 < 0,$$

by the Intermediate Value Theorem, there exists a root β within $(6,7)$. Further,

$$f'(x) = -\frac{1}{x^2} + \frac{1}{x}$$

is positive for $x > 6$, so $f(x) > f(\beta) = 0$ for all $x > \beta$, whence β is the only root greater than 6, i.e. β is the largest root. Using linear interpolation on the interval $[6,7]$, we have the estimate

$$\beta_0 = \frac{6f(7) - 7f(6)}{f(7) - f(6)} = 6.32 \text{ (2 d.p.)}.$$



From the graph, $\alpha \in (0,1)$, hence $N = 1$.

The Newton-Raphson method gives the following recursion:

$$\alpha_{n+1} = \alpha_n - \frac{f(\alpha_n)}{f'(\alpha_n)} = \alpha_n - \frac{\frac{1}{\alpha_n} - 2 + \ln \alpha_n}{-\frac{1}{\alpha^2} + \frac{1}{\alpha}}.$$

For $\alpha_0 = 0$, both $f(\alpha_0)$ and $f'(\alpha_0)$ are undefined. For $\alpha_0 = 1$, the denominator $f'(\alpha_0) = 0$, so α_1 is undefined.

Using the Newton-Raphson method will initial value $\alpha_0 = 0.3$, we obtain

$$\alpha_1 = 0.31663 = 0.317 \text{ (3 d.p.)} \quad \text{and} \quad \alpha_2 = 0.31784 = 0.318 \text{ (3 d.p.)}.$$

Checking, we see that f is continuous of $(0.3175, 0.3184)$ and

$$f(0.3175)f(0.3184) = -8.7 \times 10^{-6} < 0,$$

hence by the Intermediate Value Theorem, $\alpha = (0.3175, 0.3184)$, thus $\alpha = 0.318$ (3 d.p.).

* * * * *

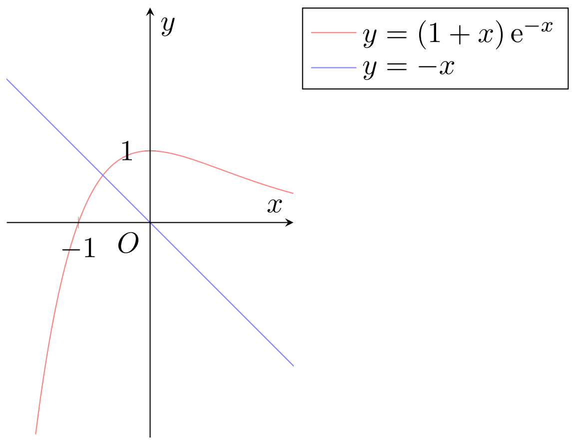
Problem 3. Sketch the graph of $y = (1+x)e^{-x}$, indicating clearly the turning points and asymptotes (if any). State the transformation by which the graph of $y = xe^{1-x}$ may be obtained from the graph of $y = (1+x)e^{-x}$.

By means of a suitable sketch, deduce that $x\left(1 + e^{1-x}\right) = 1$ has exactly one real root α. Show that α lies between 0.3 and 0.4.

Use linear interpolation once to obtain an approximation value, c, for α, giving your answer correct to 4 decimal places.

Using the Newton-Raphson method once with c as the first approximation, obtain a second approximation for α correct to 3 significant figures.

Solution.



The graph of $y = xe^{1-x}$ can be obtained by translating the graph of $y = (1+x)e^{-x}$ one unit in the negative x-direction.

Note that
$$x\left(1 + e^{1-x}\right) = 1 \equiv (1+x)\,e^{-x} = -x.$$
Plotting the graph of $y = -x$, we see that the two graphs intersect at only one point. Thus, $x\left(1 + e^{1-x}\right) = 1$ has only one real root.

Let $f(x) = x\left(1 + e^{1-x}\right) - 1$. Observe that f is continuous on $(0.3, 0.4)$ and
$$f(0.3)f(0.4) = -0.01 < 0,$$
thus by the Intermediate Value Theorem, $\alpha \in (0.3, 0.4)$. Using linear interpolating on $(0.3, 0.4)$,
$$c = \frac{0.3f(0.4) - 0.4f(0.3)}{f(0.4) - f(0.3)} = 0.3427 \text{ (4 d.p.)}.$$
Note that $f'(x) = 1 + e^{1-x}(1-x)$. The Newton-Raphson method employs the recursion
$$\alpha_{n+1} = \alpha_n - \frac{f(\alpha_n)}{f'(\alpha_n)} = \alpha_n - \frac{\alpha_n\left(1 + e^{1-\alpha_n}\right) - 1}{1 + e^{1-\alpha_n}(1-\alpha_n)}.$$
Using the initial condition $\alpha_0 = 0.3427$, we see that $\alpha_1 = 0.3409 = 0.341$ (3 s.f.). Checking, we see that
$$f(0.3405)f(0.3414) = -1.0 \times 10^{-6},$$
thus $\alpha \in (0.3405, 0.3414)$, i.e. $\alpha = 0.341$ (3 s.f.).

* * * * *

Problem 4. In this question, give all your final answers correct to 3 decimal places.

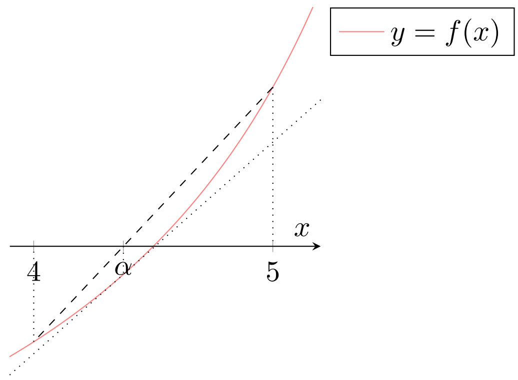
(a) Find, stating your reason, the value of the positive integer n such that
$$n - 1 \leq \sqrt[3]{100} \leq n.$$
Hence, use linear interpolation once only, to find an approximation, α, to the root of the equation $x^3 = 100$. Explain, with the aid of a suitable diagram, whether α is an overestimate or underestimate.

(b) Using the Newton-Raphson method with α as a first approximation, find $\sqrt[3]{100}$. Explain, using the same diagram as in (a), whether this method yields a series of overestimates or underestimates.

Solution.

Part (a). Note that
$$4^3 = 64 < 100 < 125 = 5^3.$$
It follows that $4 < \sqrt[3]{100} < 5$, so $n = 5$.

Let $f(x) = x^3 - 100$. Using linear interpolation on $(4, 5)$, we see that
$$\alpha = \frac{4f(5) - 5f(4)}{f(5) - f(4)} = 4.590 \text{ (3 d.p.)}.$$
Note that over $(4, 5)$,
$$f'(x) = 3x^2 > 0 \quad \text{and} \quad f''(x) = 6x > 0,$$
so $f(x)$ is increasing and convex.



From the above figure, we see that the estimate given by linear interpolation is an underestimate.

Part (b). The Newton-Raphson method uses the recursion
$$\alpha_{n+1} = \alpha_n - \frac{f(\alpha_n)}{f'(a_n)} = \alpha_n - \frac{\alpha_n^3 - 100}{3\alpha_n^2}.$$
With the initial value $\alpha_1 = 4.590$, we have $\alpha_2 = 4.64217 = 4.642$ (3 d.p.). Checking, we see that f is continuous on $(4.6415, 4.6424)$ and
$$f(4.6415)f(4.6424) = -3.0 \times 10^{-4},$$
thus $\alpha \in (4.6415, 4.6424)$ and $\alpha = \sqrt[3]{100} = 4.642$ (3 d.p.).

Since the graph of $y = f(x)$ is convex, it is always above its tangents. Thus, the Newton-Raphson method gives an underestimate.

* * * * *

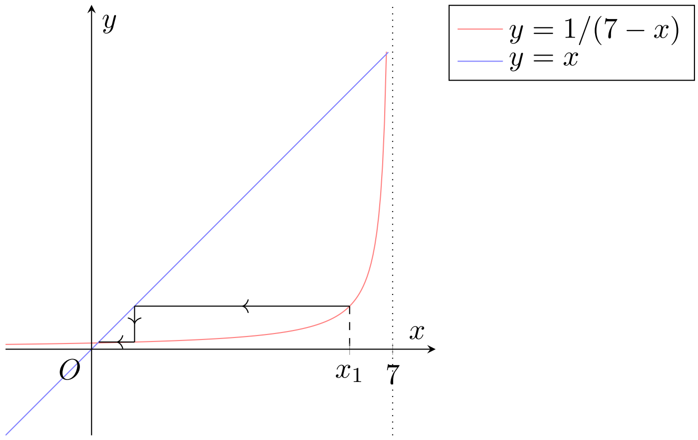
Problem 5. The roots of the quadratic equation $x^2 - 7x + 1 = 0$ are to be calculated by the use of the recurrence relation $x_{r+1} = \frac{1}{7 - x_r}$. Sketch the graphs of $y = x$ and $y = \frac{1}{7-x}$ and hence show

(a) that the equation has 2 roots, which lie between 0 and 7.

(b) if x_1 has a value lying between these roots, then the recurrence relation will always yield an approximation to the smaller root.

Taking $x_1 = 1$, find the smaller root correct to 3 decimal places. Obtain the value of the larger root to the same degree of accuracy.

Solution.



(a) and (b) are obvious from the graph.

Using the given recurrence relation, with initial value $x_1 = 1$, we have

$$x_2 = 0.14394 \quad \text{and} \quad x_3 = 0.14586 = 0.146 \text{ (3 d.p.)}.$$

Checking, we see that $f(x) = x^2 - 7x + 1$ is continuous and

$$f(0.1455)f(0.1464) = -9.0 \times 10^{-6} < 0,$$

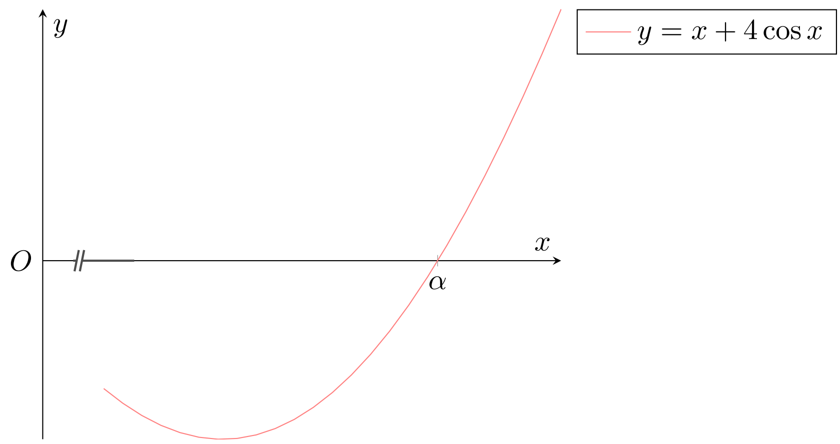
thus $\alpha \in (0.1455, 0.1464)$ and $\alpha = 0.146$ (3 d.p.).

Let β be the other root. By Vieta's formula, $\alpha + \beta = 7$, so $\beta = 7 - 0.14586 = 6.854$ (3 d.p.).

Assignment A2

Problem 1. By considering the graphs of $y = \cos x$ and $y = -\frac{1}{4}x$, or otherwise, show that the equation $x + 4\cos x = 0$ has one negative root and two positive roots.

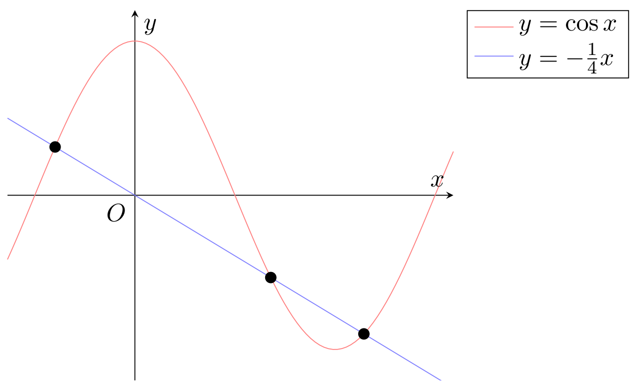
Use linear interpolation, once only, on the interval $[-1.5, 1]$ to find an approximation to the negative root of the equation $x + 4\cos x = 0$ correct to 2 decimal places.



The diagram shows part of the graph of $y = x + 4\cos x$ near the larger positive root, α, of the equation $x + 4\cos x = 0$. Explain why, when using the Newton-Raphson method to find α, an initial approximation which is smaller than α may not be satisfactory.

Use the Newton-Raphson method to find α correct to 2 significant figures. You should demonstrate that your answer has the required accuracy.

Solution.



Note that $x + 4\cos x = 0 \implies \cos x = -\frac{1}{4}x$. Plotting the graphs of $y = \cos x$ and $y = -\frac{1}{4}x$, we see that there is one negative root and two positive roots. Hence, the equation $x + 4\cos x = 0$ has one negative root and two positive roots.

Let $f(x) = x + 4\cos x$. Let β be the negative root of the equation $f(x) = 0$. Using linear interpolation on the interval $[-1.5. - 1]$,

$$\beta = \frac{-1.5f(-1) - (-1)f(-1.5)}{f(-1) - f(1.5)} = -1.24 \text{ (2 d.p.)}.$$

There is a minimum at $x = m$ such that m is between the two positive roots. Hence, when using the Newton-Raphson method, an initial approximation which is smaller than m would result in subsequent approximations being further away from the desired root α. Hence, an initial approximation that is smaller than α may not be satisfactory.

We know from the above graph that $\alpha \in (\pi, 3\pi/2)$. We hence pick $3\pi/2$ as our initial approximation. Using the Newton-Raphson method $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ with $x_1 = 3\pi/2$, we have

r	x_r
1	$\frac{3}{2}\pi$
2	3.7699
3	3.6106
4	3.5955
5	3.5953

Since $f(3.55)f(3.65) = (-0.1)(0.2) < 0$, we have $\alpha \in (3.55, 3.65)$. Hence, $\alpha = 3.6$ (2 s.f.).

* * * * *

Problem 2. Find the coordinates of the stationary points on the graph $y = x^3 + x^2$. Sketch the graph and hence write down the set of values of the constant k for which the equation $x^3 + x^2 = k$ has three distinct real roots.

The positive root of the equation $x^3 + x^2 = 0.1$ is denoted by α.

(a) Find a first approximation to α by linear interpolation on the interval $0 \le x \le 1$.

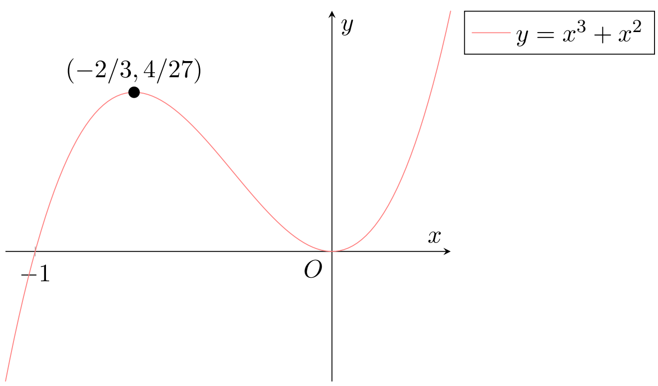
(b) With the aid of a suitable figure, indicate why, in this case, linear interpolation does not give a good approximation to α.

(c) Find an alternative first approximation to α by using the fact that if x is small then x^3 is negligible when compared to x^2.

Solution. For stationary points, $y' = 0$.

$$y' = 0 \implies 3x^2 + 2x = 0 \implies x(3x+2) = 0.$$

Hence, $x = 0$ or $x = -2.3$. When $x = 0$, $y = 0$. When $x = -2/3$, $y = 4/27$. Thus, the coordinates of the stationary points of $y = x^3 + x^2$ are $(0, 0)$ and $(-2/3, 4/27)$.

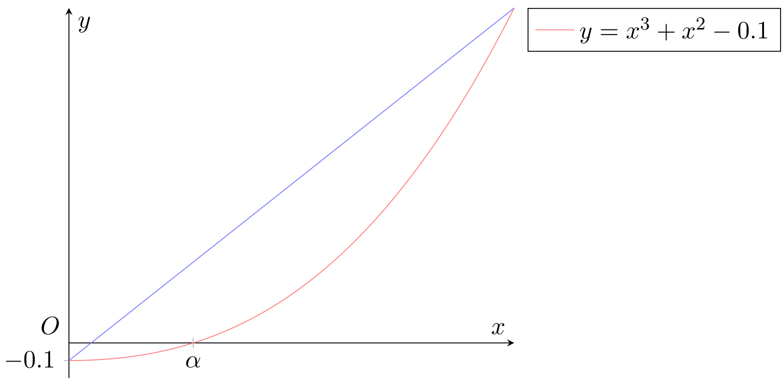


Therefore, $k \in (0, 4/27)$. The solution set of k is thus $\{k \in \mathbb{R} : 0 < k < 4/27\}$.

Part (a). Let $f(x) = x^2 + x^2 - 0.1$. Using linear interpolation on the interval $[0, 1]$,

$$\alpha = \frac{-f(0)}{f(1) - f(0)} = \frac{1}{20}.$$

Part (b).



On the interval $[0, 1]$, the gradient of $y = x^3 + x^2 - 0.1$ changes considerably. Hence, linear interpolation gives an approximation much less than the actual value.

Part (c). For small x, x^3 is negligible when compared to x^2. Consider $g(x) = x^2 - 0.1$. Then the positive root of $g(x) = 0$ is approximately α. Hence, an alternative approximation to α is $\sqrt{0.1} = 0.316$ (3 s.f.).

* * * * *

Problem 3. The equation $2\cos x - x = 0$ has a root α in the interval $[1, 1.2]$. Iterations of the form $x_{n+1} = F(x_n)$ are based on each of the following rearrangements of the equation:

(a) $x = 2\cos x$

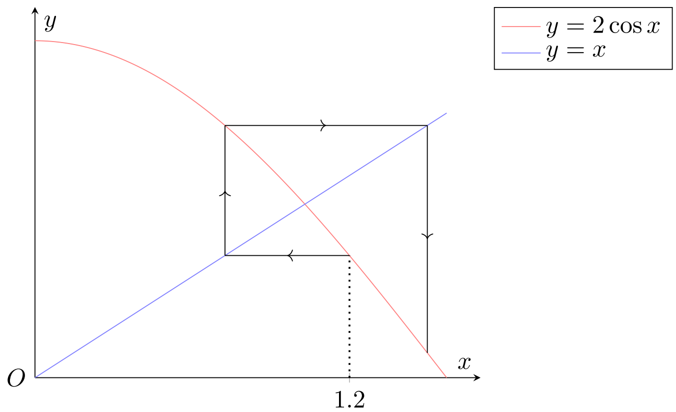
(b) $x = \cos x + \frac{1}{2}x$

(c) $x = \frac{2}{3}(\cos x + x)$

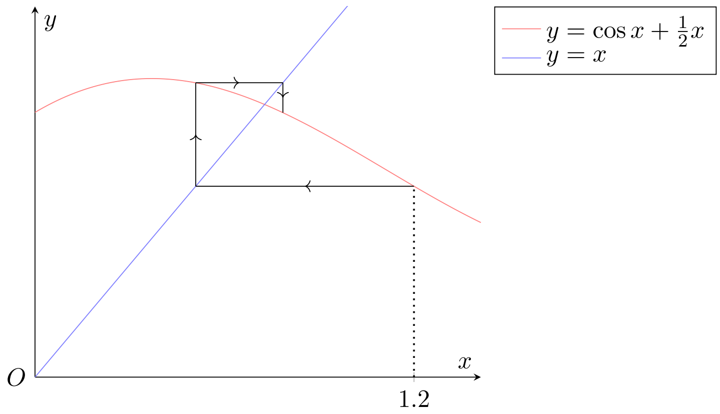
Determine which iteration will converge to α and illustrate your answer by a 'staircase' or 'cobweb' diagram. Use the most appropriate iteration with $x_1 = 1$, to find α to 4 significant figures. You should demonstrate that your answer has the required accuracy.

Solution.

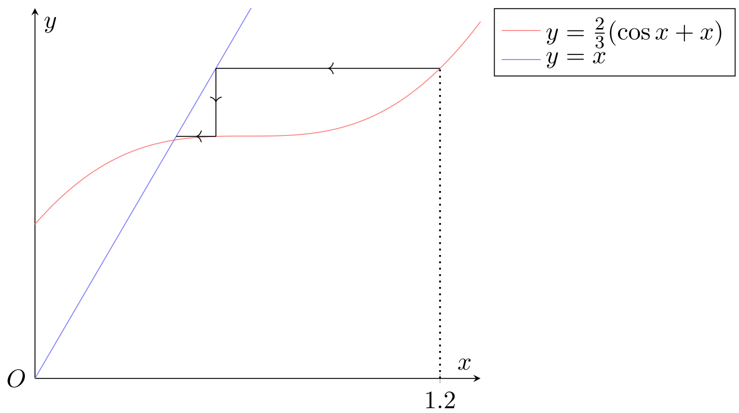
Part (a). Consider $f(x) = 2\cos x$. Then $f'(x) = -2\sin x$. Observe that $\sin x$ is increasing on $[1, 1.2]$. Since $\sin 1 > \frac{1}{2}$, $|f'(x)| > 1$ for all $x \in [1, 1.2]$. Thus, fixed-point iteration fails and will not converge to α.



Part (b). Consider $f(x) = \cos x + \frac{1}{2}x$. Then $f'(x) = -\sin x + \frac{1}{2} - \left(\sin x - \frac{1}{2}\right)$. Since $0 \le \sin x \le 1$ for $x \in \left[0, \frac{\pi}{2}\right]$, and $[1, 1.2] \subset \left[0, \frac{\pi}{2}\right]$, we know $-\frac{1}{2} \le \sin x - \frac{1}{2} \le \frac{1}{2}$ for $x \in [1, 1.2]$. Thus, $0 \le \left|\sin x - \frac{1}{2}\right| \le \frac{1}{2}$ for $x \in [1, 1.2]$. Hence, fixed-point iteration will work and converge to α.



Part (c). Consider $f(x) = \frac{2}{3}(\cos x + x)$. Then $f'(x) = \frac{2}{3}(-\sin x + 1)$. For fixed-point iteration to converge to α, we need $|f'(x)| < 1$ for x near α. It thus suffices to show that $|-\sin x + 1| < \frac{3}{2}$ for all $x \in [1, 1.2]$. Observe that $1 - \sin x$ is strictly decreasing and positive for $x \in \left[0, \frac{\pi}{2}\right]$. Since $1 - \sin 1 < \frac{3}{2}$, and $[1, 1.2] \subset \left[0, \frac{\pi}{2}\right]$, we have that $|-\sin x + 1| < \frac{3}{2}$ for all $x \in [1, 1.2]$. Thus, $|f'(x)| < 1$ for x near α. Hence, fixed-point iteration will work and converge to α.



For $x \in [1, 1.2]$, $\left|\frac{2}{3}(-\sin x + 1)\right| < \left|-\sin x + \frac{1}{2}\right| < 1$. Thus, $x_{n+1} = \frac{2}{3}(\cos x_n + x_n)$ is the most suitable iteration as it will converge to α the quickest. Using $F(x_{n+1}) = \frac{2}{3}(\cos x_n + x_n)$ with $x_1 = 1$,

r	x_r
1	1
2	1.02687
3	1.02958
4	1.02984
5	1.02986

Since $F(1.0295) > 1.0295$ and $F(1.0305) < 1.0305$, we have $\alpha \in (1.0295, 1.0305)$. Hence, $\alpha = 1.030$ (4 s.f.).

A3 Sequences and Series I

Tutorial A3

Problem 1. Determine the behaviour of the following sequences.

(a) $u_n = 3\left(\frac{1}{2}\right)^{n-1}$

(b) $v_n = 2 - n$

(c) $t_n = (-1)^n$

(d) $w_n = 4$

Solution.
Part (a). Decreasing, converges to 0.
Part (b). Decreasing, diverges.
Part (c). Alternating, diverges.
Part (d). Constant, converges to 4.

* * * * *

Problem 2. Find the sum of all even numbers from 20 to 100 inclusive.

Solution. The even numbers from 20 to 100 inclusive form an AP with common difference 2, first term 20 and last term 100. Since we are adding a total of $\frac{100-20}{2} + 1 = 41$ terms, we get a sum of $41\left(\frac{20+100}{2}\right) = 2460$.

* * * * *

Problem 3. A geometric series has first term 3, last term 384 and sum 765. Find the common ratio.

Solution. Let the nth term of the geometric series be ar^{n-1}, where $1 \leq n \leq k$. We hence have $3r^{k-1} = 384$, which gives $r^k = 128r$. Thus,

$$\frac{3(1-r^k)}{1-r} = 765 \implies \frac{3(1-128r)}{1-r} = 765 \implies r = 2.$$

* * * * *

Problem 4.

(a) Find the first four terms of the following sequence $u_{n+1} = \frac{u_n+1}{u_n+2}$, $u_1 = 0$, $n \geq 1$.

(b) Write down the recurrence relation between the terms of these sequences.

  (i) $-1, 2, -4, 8, -16, \ldots$

  (ii) $1, 3, 7, 15, 31, \ldots$

Solution.
Part (a). Using G.C., the first four terms of u_n are 0, $\frac{1}{2}$, $\frac{3}{5}$ and $\frac{8}{13}$.

Part (b).

Part (b)(i). $u_{n+1} = -2u_n$, $u_1 = -1$, $n \geq 1$.

Part (b)(ii). $u_{n+1} = 2u_n + 1$, $u_1 = 1$, $n \geq 1$.

* * * * *

Problem 5. The sum of the first n terms of a series, S_n, is given by $S_n = 2n(n+5)$. Find the nth term and show that the terms are in arithmetic progression.

Solution. We have

$$u_n = S_n - S_{n-1} = 2n(n+5) - 2(n-1)(n+4) = 4n + 8.$$

Observe that $u_n - u_{n-1} = [4n+8] - [4(n-1)+8] = 8$ is a constant. Hence, u_n is in AP.

* * * * *

Problem 6. The sum of the first n terms, S_n, is given by

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

(a) Find an expression for the nth term of the series.

(b) Hence or otherwise, show that it is a geometric series.

(c) State the values of the first term and the common ratio.

(d) Give a reason why the sum of the series converges as n approaches infinity and write down its value.

Solution.

Part (a). Note that

$$u_n = S_n - S_{n-1} = \left[\frac{1}{2} - \left(\frac{1}{2}\right)^{n+1}\right] - \left[\frac{1}{2} - \left(\frac{1}{2}\right)^{n}\right] = \left(\frac{1}{2}\right)^{n+1}.$$

Part (b). Since $\frac{u_{n+1}}{u_n} = \frac{(1/2)^{n+2}}{(1/2)^{n+1}} = \frac{1}{2}$ is constant, u_n is in GP.

Part (c). The first term is $\frac{1}{4}$ and the common ratio is $\frac{1}{2}$.

Part (d). As $n \to \infty$, we clearly have $\left(\frac{1}{2}\right)^{n+1} \to 0$. Hence, $S_\infty = \frac{1}{2}$.

* * * * *

Problem 7. The first term of an arithmetic series is $\ln x$ and the rth term is $\ln\left(xk^{r-1}\right)$, where k is a real constant. Show that the sum of the first n terms of the series is $S_n = \frac{n}{2}\ln\left(x^2k^{n-1}\right)$. If $k = 1$ and $x \neq 1$, find the sum of the series $\mathrm{e}^{S_1} + \mathrm{e}^{S_2} + \mathrm{e}^{S_3} + \ldots + \mathrm{e}^{S_n}$.

Solution. Let u_n be the nth term in the arithmetic series. Then

$$u_n = \ln\left(xk^{r-1}\right) = \ln x + (r-1)\ln k.$$

We thus see that the arithmetic series has first term $\ln x$ and common difference of $\ln k$. Thus,

$$S_n = n\left(\frac{\ln x + (\ln x + (r-1)\ln k)}{2}\right) = \frac{n}{2}\ln\left(x^2k^{r-1}\right).$$

When $k = 1$, we have $S_n = \ln(x^n)$, whence $e^{S_n} = x^n$. Thus,

$$e^{S_1} + e^{S_2} + e^{S_3} + \ldots + e^{S_n} = x + x^2 + x^3 + \ldots x^n = \frac{x(1 - x^{n+1})}{1 - x}.$$

* * * * *

Problem 8. A baker wants to bake a 1-metre tall birthday cake. It comprises 10 cylindrical cakes each of equal height 10 cm. The diameter of the cake at the lowest layer is 30 cm. The diameter of each subsequent layer is 4% less than the diameter of the cake below. Find the volume of this cake in cm^3, giving your answer to the nearest integer.

Solution. Let the diameter of the nth layer be d_n cm. We have $d_{n+1} = 0.96 d_n$ and $d_1 = 30$, whence $d_n = 30 \cdot 0.96^{n-1}$. Let the nth layer have volume v_n cm^3. Then

$$v_n = 10\pi \left(\frac{d_n}{2}\right)^2 = 10\pi \left(\frac{900 \cdot 0.9216^{n-1}}{4}\right) = 2250\pi \cdot 0.9216^{n-1}.$$

The volume of the cake in cm^3 is thus given by

$$2250\pi \left(\frac{1 - 0.9216^{10}}{1 - 0.9216}\right) = 50309.$$

* * * * *

Problem 9. The sum to infinity of a geometric progression is 5 and the sum to infinity of another series is formed by taking the first, fourth, seventh, tenth, ... terms is 4. Find the exact common ratio of the series.

Solution. Let the nth term of the geometric progression be given by ar^{n-1}. Then, we have

$$\frac{a}{1-r} = 5 \implies a = 5(1-r). \tag{1}$$

Note that the first, fourth, seventh, tenth, ... terms forms a new geometric series with common ratio r^3: $a, ar^3, ar^6, ar^9, \ldots$. Thus,

$$\frac{a}{1-r^3} = 4 \implies a = 4(1-r^3). \tag{2}$$

Equating (1) and (2), we have

$$5(1-r) = 4(1-r^3) \implies 4r^3 + 5r + 1 = 0 \implies (r-1)(4r^2 + 4r - 1) = 0.$$

Since $|r| < 1$, we only have $4r^2 + 4r - 1 = 0$, which has solutions $r = \frac{-1+\sqrt{2}}{2}$ or $r = \frac{-1-\sqrt{2}}{2}$. Once again, since $|r| < 1$, we reject $r = \frac{-1-\sqrt{2}}{2}$. Hence, $r = \frac{-1+\sqrt{2}}{2}$.

* * * * *

Problem 10. A geometric series has common ratio r, and an arithmetic series has first term a and common difference d, where a and d are non-zero. The first three terms of the geometric series are equal to the first, fourth and sixth terms respectively of the arithmetic series.

(a) Show that $3r^2 - 5r + 2 = 0$

(b) Deduce that the geometric series is convergent and find, in terms of a, the sum of infinity.

(c) The sum of the first n terms of the arithmetic series is denoted by S. Given that $a > 0$, find the set of possible values of n for which S exceeds $4a$.

Solution.

Part (a). Let the nth term of the geometric series be $G_n = G_1 r^{n-1}$. Let the nth term of the arithmetic series be $A_n = a + (n-1)d$.

Since $G_1 = A_1$, we have $G_1 = a$. We can thus write $G_n = ar^{n-1}$. From $G_2 = A_4$, we have $ar = a + 3d$, which gives $a = \frac{3d}{r-1}$. From $G_3 = A_6$, we have $ar^2 = a + 5d$. Thus,

$$\frac{3d}{r-1} \cdot r^2 = \frac{3d}{r-1} + 5d \implies \frac{3r^2}{r-1} = \frac{3}{r-1} + 5 \implies 3r^2 - 5r + 2 = 0.$$

Part (b). Note that the roots to $3r^2 - 5r + 2 = 0$ are $r = 1$ and $r = 2/3$. Clearly, $r \neq 1$ since $a = 3d/(r-1)$ would be undefined. Hence, $r = 2/3$, whence the geometric series is convergent.

Let S_∞ be the sum to infinity of G_n. Then $S_\infty = a/(1-r) = 3a$.

Part (c). Note that $d = a(r-1)/3 = -\frac{a}{9}$. Hence,

$$S = n\left(\frac{a + [a + (n-1)d]}{2}\right) = n\left(\frac{2a + (n-1)\left(-\frac{a}{9}\right)}{2}\right) = \frac{an}{18}(19 - n).$$

Consider $S > 4a$.

$$S > 4a \implies \frac{n}{18}(19-n) > 4 \implies -n^2 + 19n - 72 > 0.$$

Using G.C., we see that $5.23 < n < 13.8$. Since n is an integer, the set of values that n can take on is $\{n \in \mathbb{Z} : 6 \leq n \leq 13\}$.

* * * * *

Problem 11. Two musical instruments, A and B, consist of metal bars of decreasing lengths.

(a) The first bar of instrument A has length 20 cm and the lengths of the bars form a geometric progression. The 25th bar has length 5 cm. Show that the total length of all the bars must be less than 357 cm, no matter how many bars there are.

Instrument B consists of only 25 bars which are identical to the first 25 bars of instrument A.

(b) Find the total length, L cm, of all the bars of instrument B and the length of the 13th bar.

(c) Unfortunately, the manufacturer misunderstands the instructions and constructs instrument B wrongly, so that the lengths of the bars are in arithmetic progression with a common difference d cm. If the total length of the 25 bars is still L cm and the length of the 25th bar is still 5 cm, find the value of d and the length of the longest bar.

Solution.

Part (a). Let $u_n = u_1 r^{n-1}$ be the length of the nth bar. Since $u_1 = 20$, we have $u_n = 20r^{n-1}$. Since $u_{25} = 5$, we have $r = 4^{-\frac{1}{24}}$. Hence, $u_n = 20 \cdot 4^{-\frac{n-1}{24}}$. Now, consider the sum to infinity of u_n:

$$S_\infty = \frac{u_1}{1-r} = \frac{20}{1 - 4^{-1/24}} = 356.3 < 357.$$

Hence, no matter how many bars there are, the total length of the bars will never exceed 357 cm.

Part (b). We have

$$L = u_1 \left(\frac{1 - r^{25}}{1 - r} \right) = 20 \left(\frac{1 - 4^{-25/24}}{1 - 4^{-1/24}} \right) = 272.26 = 272 \text{ (3 s.f.)}.$$

Note that

$$u_{13} = 20 \cdot \left(4^{-1/24} \right)^{13-1} = 10.$$

The 13th bar is hence 10 cm long.

Part (c). Let $v_n = a + (n-1)d$ be the length of the wrongly-manufactured bars. Since the length of the 25th bar is still 5 cm, we know $v_{25} = a + 24d = 5$. Now, consider the total lengths of the bars, which is still L cm.

$$L = 25 \left(\frac{a+5}{2} \right) = 272.26.$$

Solving, we see that $a = 16.781$. Hence, $d = \frac{5-a}{24} = -0.491$, and the longest bar is $16.8 =$ cm long.

* * * * *

Problem 12. A bank has an account for investors. Interest is added to the account at the end of each year at a fixed rate of 5% of the amount in the account at the beginning of that year. A man a woman both invest money.

(a) The man decides to invest $\$x$ at the beginning of one year and then a further $\$x$ at the beginning of the second and each subsequent year. He also decides that he will not draw any money out of the account, but just leave it, and any interest, to build up.

  (i) How much will there be in the account at the end of 1 year, including the interest?

  (ii) Show that, at the end of n years, when the interest for the last year has been added, he will have a total of $\$21(1.05^n - 1)x$ in his account.

  (iii) After how many complete years will he have, for the first time, at least $\$12x$ in his account?

(b) The woman decides that, to assist her in her everyday expenses, she will withdraw the interest as soon as it has been added. She invests $\$y$ at the beginning of each year. Show that, at the end of n years, she will have received a total of $\$\frac{1}{40}n(n+1)y$ in interest.

Solution.

Part (a).

Part (a)(i). There will be $\$1.05x$ in the account at the end of 1 year.

Part (a)(ii). Let $\$u_n x$ be the amount of money in the account at the end of n years. Then, u_n satisfies the recurrence relation $u_{n+1} = 1.05(1 + u_n)$, with $u_1 = 1.05$. Observe that

$$u_1 = 1.05 \implies u_2 = 1.05 + 1.05^2 \implies u_3 = 1.05 + 1.05^2 + 1.05^3 \implies \cdots.$$

We thus have

$$u_n = 1.05 + 1.05^2 + \cdots + 1.05^n = 1.05 \left(\frac{1 - 1.05^n}{1 - 1.05} \right) = 21\left(1.05^n - 1\right).$$

Hence, there will be $\$21(1.05^n - 1)x$ in the account after n years.

Part (a)(iii). Consider the inequality $u_n \geq 12x$.

$$u_n \geq 12x \implies 21(1.05^n - 1) \geq 12 \implies n \geq 9.26.$$

Since n is an integer, the smallest value of n is 10. Hence, after 10 years, he will have at least \$$12x$ in his account for the first time.

Part (b). After n years, the woman will have \$$ny$ in her account. Hence, the interest she gains at the end of the nth year is $\frac{1}{20}ny$. Thus, the total interest she will gain after n years is

$$\frac{y}{20} + \frac{2y}{20} + \cdots + \frac{ny}{20} = \frac{y}{20}(1 + 2 + \cdots + n) = \frac{y}{20} \cdot \frac{n(n+1)}{2} = \frac{n(n+1)y}{40}.$$

* * * * *

Problem 13. The sum, S_n, of the first n terms of a sequence $U_1, U_2, U_3, \ldots$ is given by

$$S_n = \frac{n}{2}(c - 7n),$$

where c is a constant.

(a) Find U_n in terms of c and n.

(b) Find a recurrence relation of the form $U_{n+1} = f(U_n)$.

Solution.

Part (a). Observe that

$$U_n = S_n - S_{n-1} = \frac{n}{2}(c - 7n) - \frac{n-1}{2}(c - 7(n-1)) = -7n + \frac{7+c}{2}.$$

Part (b). Observe that $U_{n+1} - U_n = -7$. Thus,

$$U_{n+1} = U_n - 7, \quad U_1 = \frac{7+c}{2}, \quad n \geq 1.$$

* * * * *

Problem 14. The positive numbers x_n satisfy the relation

$$x_{n+1} = \sqrt{\frac{9}{2} + \frac{1}{x_n}}$$

for $n = 1, 2, 3, \ldots$.

(a) Given that $n \to \infty$, $x_n \to \theta$, find the exact value of θ.

(b) By considering $x_{n+1}^2 - \theta^2$, or otherwise, show that if $x_n > \theta$, then $0 < x_{n+1} < \theta$.

Solution.

Part (a). Observe that

$$\theta = \lim_{n\to\infty} \sqrt{\frac{9}{2} + \frac{1}{x_n}} = \sqrt{\frac{9}{2} + \frac{1}{\theta}} \implies 2\theta^3 - 9\theta - 2 = 0 \implies (\theta + 2)\left(2\theta^2 - 4\theta - 1\right) = 0.$$

We reject $\theta = -2$ since $\theta > 0$. We thus consider $2\theta^2 - 4\theta - 1 = 0$, which has roots $\theta = 1 + \sqrt{\frac{3}{2}}$ and $\theta = 1 - \sqrt{\frac{3}{2}}$. Once again, we reject $\theta = 1 - \sqrt{\frac{3}{2}}$ since $\theta > 0$. Thus, $\theta = 1 + \sqrt{\frac{3}{2}}$.

Part (b). Suppose $x_n > \theta$. Then

$$x_{n+1}^2 = \frac{9}{2} + \frac{1}{x_n} < \frac{9}{2} + \frac{1}{\theta} = \theta^2 \implies 0 < x_{n+1} < \theta.$$

Self-Practice A3

Problem 1. The sum of the first n terms of a sequence $\{u_n\}$ is given by the formula $S_n = 2n(n-3)$, where $n \in \mathbb{Z}^+$.

(a) Express u_n in terms of n, and show that the sequence $\{u_n\}$ follows an arithmetic progression.

(b) Three terms u_3, u_k and u_{38} of this sequence are consecutive terms in a geometric sequence. Find the value of k.

(c) Explain why the infinite series $\mathrm{e}^{-u_1} + \mathrm{e}^{-u_2} + \mathrm{e}^{-u_3} + \dots$ exists, and determine the value of the infinite sum, leaving your answer in exact form.

Solution.

Part (a). Note that

$$u_n = S_n - S_{n-1} = 2n(n-3) - 2(n-1)(n-1-3) = 4n - 8.$$

Thus,

$$u_n - u_{n-1} = [4n-8] - [4(n-1)-8] = 4.$$

Since $u_n - u_{n-1}$ is a constant, the sequence $\{u_n\}$ follows an arithmetic progression with common difference 4.

Part (b). Note that $u_3 = 4$ and $u_{38} = 144$. Let the common ratio be r. Then

$$u_{38} = r^2 u_3 \implies r^2 = \frac{u_{38}}{u_3} = 36 \implies r = \pm 6.$$

Since $u_k > u_3 > 0$, the common ratio r must be positive. Hence, $r = 6$. Thus,

$$4k - 8 = u_k = ru_3 = 6(4) = 24,$$

whence $k = 8$.

Part (c). Observe that

$$\frac{\mathrm{e}^{-u_n}}{\mathrm{e}^{-u_{n-1}}} = \mathrm{e}^{u_{n-1}-u_n} = \mathrm{e}^{-4}.$$

Hence, $\{\mathrm{e}^{-u_n}\}$ is in geometric progression with common ratio e^{-4}. Since $\left|\mathrm{e}^{-4}\right| < 1$, the sum to infinity exists, and is given by

$$\sum_{n=1}^{\infty} \mathrm{e}^{-u_n} = \mathrm{e}^{-u_1}\left(\frac{1}{1-\mathrm{e}^{-4}}\right) = \frac{\mathrm{e}^4}{1-\mathrm{e}^{-4}}.$$

* * * * *

Problem 2. At the end of December 2010, the amount of water in a large tank was 43 000 litres. The tank was filled with 7000 litres of water at the start of every month. It was observed that 25% of the amount at the start of any month was lost by the end of that month.

(a) Show that at the end of February 2011, the amount of water in the tank was 33 375 litres.

(b) Find the amount of water in the tank, measured in litres, at the end of the nth month after the end of December 2010, expressing your answer in the form $A\left(\frac{3}{4}\right)^n + B$, where A and B are positive integers to be determined.

Solution.

Part (a). Let the amount of water in the tank, measured in litres, at the start of the nth month after the end of December 2010 be u_n. Clearly, $u_0 = 43000$ and

$$u_n = \frac{3}{4}\left(u_{n-1} + 7000\right).$$

Note that

$$u_1 = \frac{3}{4}\left(u_0 + 7000\right) = 37500, \quad u_2 = \frac{3}{4}\left(u_1 + 7000\right) = 33375.$$

Hence, at the end of February 2011, the amount of water in the tank was 33 375 litres.

Part (b). Let k be the constant such that

$$u_n - k = \frac{3}{4}\left(u_{n-1} - k\right).$$

It quickly follows that $k = 21000$. Then

$$u_n - 21000 = \frac{3}{4}\left(u_{n-1} - 21000\right) = \left(\frac{3}{4}\right)^n \left(u_0 - 21000\right).$$

Thus,

$$u_n = 22000\left(\frac{3}{4}\right)^n + 21000,$$

whence $A = 22000$ and $B = 21000$.

* * * * *

Problem 3.

(a) A runner wants to train for the marathon. He runs 8 km during the first day, and increases the distance he runs each subsequent day by 400 m. Find the minimum number of days, n, that he needs to take to complete at least 2000 km.

(b) A sequence of real numbers $\{u_1, u_2, u_3, \ldots\}$, where $u_1 \neq 0$, is defined such that the $(n+1)$th term of the sequence is equal to the sum of the first n terms, where $n \in \mathbb{Z}^+$. Prove that the sequence $\{u_2, u_3, u_4, \ldots\}$ follows a geometric progression. Hence, find $u_1 + u_2 + \cdots + u_{N+1}$ in terms of u_1 and N.

Solution.

Part (a). Let u_n be the distance ran on the nth day, measured in km. Clearly, $\{u_n\}$ is in arithmetic progression with common difference 0.4, and $u_1 = 8$. Thus,

$$u_n = 0.4(n-1) + 8 = 0.4n + 7.6.$$

Let S_n be the total distance ran in n days. We have

$$S_n = \sum_{k=1}^{n} u_k = \sum_{k=1}^{n} (0.4k + 7.6) = 0.4\left(\frac{n(n+1)}{2}\right) + 7.6n.$$

Consider

$$S_n = 0.4\left(\frac{n(n+1)}{2}\right) + 7.6n \geq 2000.$$

Using G.C., we have $n \geq 82.4$ or $n \leq -121.4$. Since n is a positive integer, the least n is 83. Thus, he needs at least 83 days to complete at least 2000 km.

Part (b). Note that $u_2 = u_1$. Observe that

$$S_n - S_{n-1} = u_n = S_{n-1} \implies S_n = 2S_{n-1}.$$

Hence,

$$\frac{u_{n+1}}{u_n} = \frac{S_n}{S_{n-1}} = 2,$$

whence $\{u_2, u_3, u_4, \ldots\}$ is geometric progression with common ratio 2. Thus,

$$u_1 + u_2 + \cdots + u_{N+1} = u_1 + u_2\left(\frac{1-2^N}{1-2}\right) = u_1 + u_1\left(2^N - 1\right) = u_1 2^N.$$

* * * * *

Problem 4.

(a) If the sum of the first n terms of a series is S_n, where $S_n = n - 3n^2$, write down an expression for S_{n-1}. Hence, prove that the series is in an arithmetic series.

(b) Each time a ball falls vertically onto a horizontal surface, it rebounds to two-thirds of the height from which it fell. The ball is initially dropped from a point 12 m above the surface.

Show that the distance the ball has travelled just before it touches the surface for the nth time is $60 - 72\left(\frac{2}{3}\right)^n$.

Hence, find the least number of times the ball has bounced to travel a total distance of more than 52 m.

Solution.

Part (a). Clearly,

$$S_{n-1} = (n-1) - 3(n-1)^2 = -3n^2 + 7n - 4.$$

Hence,

$$u_n = S_n - S_{n-1} = \left(n - 3n^2\right) - \left(-3n^2 + 7n - 4\right) = -6n + 4.$$

Observe that

$$u_n - u_{n-1} = [-6n + 4] - [-6(n-1) - 4] = -6.$$

Hence, $\{u_n\}$ is in arithmetic progression with common ratio -6.

Part (b). Let u_n be the height of the nth "drop" of the ball. We have $u_1 = 12$, and the recurrence relation $u_{n+1} = \frac{2}{3}u_n$. Quite clearly,

$$u_n = \left(\frac{2}{3}\right)^{n-1} u_1 = 18\left(\frac{2}{3}\right)^n.$$

Let D_n be the total distance travelled by the ball just before it touches the surface for the nth time. Observe that the after the initial 12 m, the ball travels up and down before touching the surface again. Hence,

$$D_n = u_1 + 2u_2 + 2u_3 + \cdots + 2u_n = u_1 + \sum_{k=2}^{n} 2u_n.$$

This evaluates as

$$D_n = u_1 + 2\sum_{k=2}^{n} 18\left(\frac{2}{3}\right)^n = 12 + 36 \cdot \left(\frac{2}{3}\right)^2 \left(\frac{1-(2/3)^{n-1}}{1-2/3}\right) = 60 - 72\left(\frac{2}{3}\right)^n.$$

Consider $D_n \geq 52$. Using G.C., we have $n \geq 5.4$. Thus, the ball must bounce at least 6 times.

* * * * *

Problem 5. The sequence $\{2^n, n = 0, 1, 2, \ldots\}$ is grouped into sets such that the rth bracket contains r terms: $\{1\}$, $\left\{2, 2^2\right\}$, $\left\{2^3, 2^4, 2^5\right\}$, $\left\{2^6, 2^7, 2^8, 2^9\right\}$, $\ldots$. Find the total number of terms in the first n brackets. Hence, find the sum of numbers in the first n brackets. Deduce (in any order), in terms of n, the first and the last number in the nth bracket.

Solution. Clearly, the number of terms in the first n brackets is

$$1 + 2 + 3 + \cdots + n = \frac{n(n-1)}{2}.$$

Note that the kth number is given by 2^{k-1}. The sum of number in the first n brackets is hence given by

$$\sum_{k=0}^{n(n+1)/2-1} 2^k = \frac{1 - 2^{n(n+1)/2}}{1-2} = 2^{n(n+1)/2} - 1.$$

The last number in the nth bracket is clearly

$$2^{n(n+1)/2-1}.$$

Note that there are $n(n-1)/2$ terms in the first $(n-1)$ brackets. Thus, the first number in the nth bracket is

$$2^{n(n-1)/2}.$$

Assignment A3

Problem 1. A university student has a goal of saving at least \$1 000 000 (in Singapore dollars). He begins working at the start of the year 2019. In order to achieve his goal, he saves 40% of his annual salary at the end of each year. If his annual salary in the year 2019 is \$40800, and it increases by 5% (of his previous year's salary) every year, find

(a) his annual savings in 2027 (to the nearest dollar),

(b) his total savings at the end of n years.

What is the minimum number of complete years for which he has to work in order to achieve his goal?

Solution. Let $\$u_n$ be his annual salary in the nth year after 2019, with $n \in \mathbb{N}$. Then $u_{n+1} = 1.05 \cdot u_n$, with $u_0 = 40800$. Hence, $u_n = 40800 \cdot 1.05^n$. Let $\$v_n$ be the amount saved in the nth year after 2019. Then $v_n = 0.40 \cdot u_n = 16320 \cdot 1.05^n$.

Part (a). In 2027, $n = 8$. Hence, his annual savings in 2027, in dollars, is given by

$$v_8 = 16320 \cdot 1.05^8 = 24112 \text{ (to the nearest integer).}$$

Part (b). His total savings at the end of n years, in dollars, is given by

$$16320\left(1.05^0 + 1.05^1 + \cdots + 1.05^n\right) = 16320\left(\frac{1 - 1.05^n}{1 - 1.05}\right) = 326400\left(1.05^n - 1\right).$$

Consider $326400\left(1.05^n - 1\right) \geq 1000000$. Using G.C., we see that $n \geq 28.7$. Thus, he needs to work for a minimum of 29 complete years to reach his goal.

* * * * *

Problem 2.

(a) A rope of length 200π cm is cut into pieces to form as many circles as possible, whose radii follow an arithmetic progression with common difference 0.25 cm. Given that the smallest circle has an area of π cm^2, find the area of the largest circle in terms of π.

(b) The sum of the first n terms of a sequence is given by $S_n = \alpha^{-n} - 1$, where α is a non-zero constant, $\alpha \neq 1$.

(i) Show that the sequence is a geometric progression and state its common ratio in terms of α.

(ii) Find the set of values of α for which the sum to infinity of the sequence exists.

(iii) Find the value of the sum to infinity.

Solution.

Part (a). Let the sequence r_n be the radius of the nth smallest circle, in centimetres. Hence, $r_n = \frac{1}{4} + r_{n-1}$. Since the smallest circle has area π cm^2, $r_1 = 1$. Thus, $r_n = 1 + \frac{1}{4}(n-1)$.

Consider the nth partial sum of the circumferences:

$$2\pi r_1 + 2\pi r_2 + \cdots + 2\pi r_n = 2\pi \cdot n\left(\frac{1 + \left[1 + \frac{1}{4}(n-1)\right]}{2}\right) = \frac{\pi(n^2 + 7n)}{4}.$$

Since the rope has length 200π cm, we have the inequality

$$\frac{\pi(n^2+7n)}{4} \leq 200\pi \implies n^2 - 7n - 800 \leq 0 \implies (n+32)(n-25) \leq 0.$$

Hence, $n \leq 25$. Since the rope is cut to form as many circles as possible, $n = 25$. Thus, the largest circle has area $\pi \cdot r_{25}^2 = 49\pi$ cm^2.

Part (b). Let the sequence being summed by $u_1, u_2, \ldots$. Observe that

$$u_n = S_n - S_{n-1} = \left(\alpha^{-n} - 1\right) - \left(\alpha^{-(n-1)} - 1\right) = \alpha^{-n}(1-\alpha).$$

Part (b)(i). Observe that

$$\frac{u_{n+1}}{u_n} = \frac{\alpha^{-(n+1)}(1-\alpha)}{\alpha^{-n}(1-\alpha)} = \alpha^{-1},$$

which is a constant. Thus, u_n is in GP with common ratio α^{-1}.

Part (b)(ii). Consider $S_\infty = \lim_{n\to\infty} S_n = \lim_{n\to\infty}(\alpha^{-n} - 1)$. For S_∞ to exist, we need $\lim_{n\to\infty} \alpha^{-n}$ to exist. Hence, $\left|\alpha^{-1}\right| < 1$, whence $|\alpha| > 1$. Thus, $\alpha < -1$ or $\alpha > 1$. The solution set of α is thus $\{x \in \mathbb{R} : x < -1 \text{ or } x > 1\}$.

Part (b)(iii). Since $\left|\alpha^{-1}\right| < 1$, we know $\lim_{n\to\infty} \alpha^{-n} = 0$. Hence, $S_\infty = -1$.

* * * * *

Problem 3. A sequence $u_1, u_2, u_3, \ldots$ is such that $u_{n+1} = 2u_n + An$, where A is a constant and $n \geq 1$.

(a) Given that $u_1 = 5$ and $u_2 = 15$, find A and u_3.

It is known that the nth term of this sequence is given by

$$u_n = a(2^n) + bn + c,$$

where a, b and c are constants.

(b) Find a, b and c.

Solution.

Part (a). Substituting $n = 1$ into the recurrence relation yields $u_2 = 2u_1 + A$. Thus, $A = u_2 - 2u_1 = 5$. Substituting $n = 2$ into the recurrence relation yields $u_3 = 2u_2 + 2A = 40$.

Part (b). Since $u_1 = 5$, $u_2 = 15$ and $u_3 = 40$, we have the following system

$$\begin{cases} 2a + b + c = 5 \\ 4a + 2b + c = 15 \\ 8a + 3b + c = 40 \end{cases}$$

which has the unique solution $a = \frac{15}{2}$, $b = -5$ and $c = -5$

Problem 4. The graphs of $y = 2^x/3$ and $y = x$ intersect at $x = \alpha$ and $x = \beta$ where $\alpha < \beta$. A sequence of real numbers $x_1, x_2, x_3, \ldots$ satisfies the recurrence relation

$$x_{n+1} = \frac{1}{3} \cdot 2^{x_n}, \qquad n \geq 1.$$

(a) Prove algebraically that, if the sequence converges, then it converges to either α or β.

(b) By using the graphs of $y = \frac{1}{3} \cdot 2^x$ and $y = x$, prove that

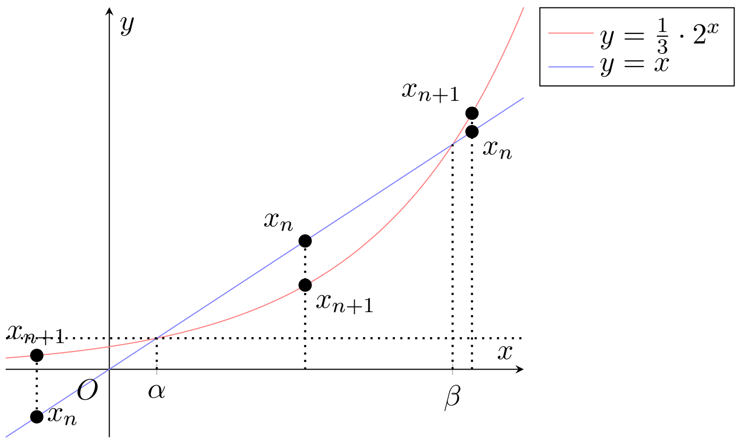
- if $\alpha < x_n < \beta$, then $\alpha < x_{n+1} < x_n$
- if $x_n < \alpha$, then $x_n < x_{n+1} < \alpha$
- if $x_n > \beta$, then $x_n < x_{n+1}$

Describe the behaviour of the sequence for the three cases.

Solution.

Part (a). Let $L = \lim_{n\to\infty} x_n$. Then $L = \frac{1}{3} \cdot 2^L$. Since $y = x$ and $y = \frac{1}{3} \cdot 2^x$ intersect only at $x = \alpha$ and $x = \beta$, then α and β are the only roots of $x = \frac{1}{3} \cdot 2^x$. Since L is also a root of $x = \frac{1}{3} \cdot 2^x$, L must be either α or β.

Part (b).



If $\alpha < x_n < \beta$, then x_n is decreasing and converges to α. If $x_n < \alpha$, then x_n is increasing and converges to α. If $x_n > \beta$, then x_n is increasing and diverges.

A4 Sequences and Series II

Tutorial A4

Problem 1. True or False? Explain your answers briefly.

(a) $\sum_{r=1}^{n}(2r+3) = \sum_{k=1}^{n}(2k+3)$

(b) $\sum_{r=1}^{n}\left(\frac{1}{r}+5\right) = \sum_{r=1}^{n}\frac{1}{r}+5$

(c) $\sum_{r=1}^{n}\frac{1}{r} = 1/\sum_{r=1}^{n} r$

(d) $\sum_{r=1}^{n} c = \sum_{r=0}^{n-1}(c+1)$

Solution.
Part (a). True: A change in index does not affect the sum.
Part (b). False: In general, $\sum_{r=1}^{n} 5$ is not equal to 5.
Part (c). False: In general, $\sum \frac{a}{b} \neq \sum a/\sum b$.
Part (d). False: Since c is a constant, $\sum_{r=1}^{n} c = nc \neq n(c+1) = \sum_{r=0}^{n-1}(c+1)$.

* * * * *

Problem 2. Write the following series in sigma notation twice, with $r = 1$ as the lower limit in the first and $r = 0$ as the lower limit in the second.

(a) $-2+1+4+\ldots+40$

(b) $a^2+a^4+a^6+\ldots+a^{50}$

(c) $\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\ldots+n\text{th term}$

(d) $1-\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\ldots$ to n terms

(e) $\frac{1}{2\cdot 4}+\frac{1}{3\cdot 5}+\frac{1}{4\cdot 6}+\ldots+\frac{1}{28\cdot 30}$

Solution.
Part (a).

$$-2+1+4+\ldots+40 = \sum_{r=1}^{15}(3r-5) = \sum_{r=0}^{14}(3r-2).$$

Part (b).

$$a^2+a^4+a^6+\ldots+a^{50} = \sum_{r=1}^{25} a^{2r} = \sum_{r=0}^{24} a^{2r+2}.$$

Part (c).

$$\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\ldots+n\text{th term} = \sum_{r=1}^{n}\frac{1}{2r+1} = \sum_{r=0}^{n-1}\frac{1}{2r+3}.$$

Part (d).

$$1-\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\ldots \text{ to } n \text{ terms} = \sum_{r=1}^{n}\left(-\frac{1}{2}\right)^{r-1} = \sum_{r=0}^{n-1}\left(-\frac{1}{2}\right)^{r}.$$

Part (e).

$$\frac{1}{2\cdot 4}+\frac{1}{3\cdot 5}+\frac{1}{4\cdot 6}+\ldots+\frac{1}{28\cdot 30} = \sum_{r=1}^{27}\frac{1}{(r+1)(r+3)} = \sum_{r=0}^{26}\frac{1}{(r+2)(r+4)}.$$

* * * * *

Problem 3. Without using the G.C., evaluate the following sums.

(a) $\sum_{r=1}^{50}(2r-7)$

(b) $\sum_{r=1}^{a}(1-a-r)$

(c) $\sum_{r=2}^{n}\left(\ln r+3^r\right)$

(d) $\sum_{r=1}^{\infty}\left(\frac{2^r-1}{3^r}\right)$

Solution.

Part (a).

$$\sum_{r=1}^{50}(2r-7) = 2\sum_{r=1}^{50} r - 7\sum_{r=1}^{50} 1 = 2\left(\frac{50\cdot 51}{2}\right) - 7(50) = 2200.$$

Part (b).

$$\sum_{r=1}^{a}(1-a-r) = (1-a)\sum_{r=1}^{a} 1 - \sum_{r=1}^{a} r = (1-a)a - \frac{a(a+1)}{2} = \frac{a}{2}(1-3a).$$

Part (c).

$$\sum_{r=2}^{n}\left(\ln r+3^r\right) = \sum_{r=2}^{n}\ln r + \sum_{r=2}^{n} 3^r = \ln n! + 3^2\left(\frac{1-3^{n-2+1}}{1-3}\right) = \ln n! + \frac{9}{2}\left(3^{n-1}-1\right).$$

Part (d).

$$\sum_{r=1}^{\infty}\left(\frac{2^r-1}{3^r}\right) = \sum_{r=1}^{\infty}\left(\frac{2}{3}\right)^r - \sum_{r=1}^{\infty}\left(\frac{1}{3}\right)^r = \frac{2/3}{1-2/3} - \frac{1/3}{1-1/3} = \frac{3}{2}.$$

* * * * *

Problem 4. The nth term of a series is $2^{n-2}+3n$. Find the sum of the first N terms.

Solution.

$$\begin{aligned}\sum_{n=1}^{N}\left(2^{n-2}+3r\right) &= \sum_{n=1}^{N} 2^{n-2} + 3\sum_{n=1}^{N} n \\ &= 2^{1-2}\left(\frac{\left(2^N-1\right)}{2-1}\right) + 3\left(\frac{N(N+1)}{2}\right) \\ &= \frac{1}{2}\left(2^N+3N^2+3N-1\right).\end{aligned}$$

Problem 5. The rth term, u_r, of a series is given by $u_r = \left(\frac{1}{3}\right)^{3r-2} + \left(\frac{1}{3}\right)^{3r-1}$. Express $\sum_{r=1}^{n} u_r$ in the form $A\left(1 - \frac{B}{27^n}\right)$, where A and B are constants. Deduce the sum to infinity of the series.

Solution. Observe that

$$u_r = \left(\frac{1}{3}\right)^{3r-2} + \left(\frac{1}{3}\right)^{3r-1} = 12\left(\frac{1}{3}\right)^{3r} = 12\left(\frac{1}{27}\right)^{r}.$$

Hence,

$$\sum_{r=1}^{n} = 12 \cdot \frac{1}{27}\left(\frac{1 - 1/27^n}{1 - 1/27}\right) = \frac{6}{13}\left(1 - \frac{1}{27^n}\right),$$

whence $A = \frac{6}{13}$ and $B = 1$. In the limit as $n \to \infty$, $\frac{1}{27^n} \to 0$. Hence, the sum to infinity is $\frac{6}{13}$.

* * * * *

Problem 6. The rth term, u_r, of a series is given by $u_r = \ln \frac{r}{r+1}$. Find $\sum_{r=1}^{n} u_r$ in terms of n. Comment on whether the series converges.

Solution. Observe that $u_r = \ln \frac{r}{r+1} = \ln r - \ln(r+1)$. Hence,

$$\begin{aligned}\sum_{r=1}^{n} u_r &= \sum_{r=1}^{n} (\ln r - \ln(r+1)) \\ &= [\ln 1 - \ln 2] + [\ln 2 - \ln 3] + \cdots + [\ln n - \ln(n+1)] \\ &= \ln 1 - \ln(n+1) = \ln \frac{1}{n+1}.\end{aligned}$$

As $n \to \infty$, $\ln \frac{1}{n+1} \to \ln 0$. Hence, the series diverges to negative infinity.

* * * * *

Problem 7. Given that $\sum_{r=1}^{n} r^2 = \frac{n}{6}(n+1)(2n+1)$, without using the G.C., find the following sums.

(a) $\sum_{r=0}^{n} [r(r+4) + n]$

(b) $\sum_{r=n+1}^{2n} (2r-1)^2$

(c) $\sum_{r=-15}^{20} r(r-2)$

Solution.

Part (a).

$$\begin{aligned}\sum_{r=0}^{n} [r(r+4) + n] &= \sum_{r=0}^{n} \left(r^2 + 4r + n\right) \\ &= \frac{n}{6}(n+1)(2n+1) + 4\left[\frac{n(n+1)}{2}\right] + n(n+1) \\ &= \frac{n}{6}(n+1)(2n+19).\end{aligned}$$

Part (b).

$$\begin{aligned}\sum_{r=n+1}^{2n}(2r-1)^2 &= \sum_{r=1}^{n}(2(r+n)-1)^2 = \sum_{r=1}^{n}\left(4r^2+4(2n-1)r+(2n-1)^2\right)\\ &= 4\left[\frac{n}{6}(n+1)(2n+1)\right]+4(2n-1)\left[\frac{n(n+1)}{2}\right]+(2n-1)^2n\\ &= \frac{1}{3}n\left(28n^2-1\right)\end{aligned}$$

Part (c).

$$\begin{aligned}\sum_{r=-15}^{20}r(r-2) &= \sum_{r=1}^{36}(r-16)[(r-16)-2] = \sum_{r=1}^{36}\left(r^2-34r+288\right)\\ &= \frac{36}{6}\left[(36+1)(2\cdot 36+1)\right]-34\left[\frac{36\cdot 37}{2}\right]+288(36)\\ &= 3930\end{aligned}$$

* * * * *

Problem 8. Let $S=\sum_{r=0}^{\infty}\frac{(x-2)^r}{3^r}$ where $x\neq 2$. Find the range of values of x such that the series S converges. Given that $x=1$, find

(a) the value of S

(b) S_n, in terms of n, where $S_n=\sum_{r=0}^{n-1}\frac{(x-2)^r}{3^r}$

(c) the least value of n for which $|S_n-S|$ is less than 0.001% of S

Solution. Note that

$$S=\sum_{r=0}^{\infty}\frac{(x-2)^r}{3^r}=\sum_{r=0}^{\infty}\left(\frac{x-2}{3}\right)^r.$$

Hence, for S to converge, we must have $\left|\frac{x-2}{3}\right|<1$, which gives $-1<x<5$, $x\neq 2$.

Part (a). When $x=1$, we get

$$S=\sum_{r=0}^{\infty}\left(-\frac{1}{3}\right)^r=\frac{1}{1-(-\frac{1}{3})}=\frac{3}{4}.$$

Part (b). We have

$$S_n=\sum_{r=0}^{n-1}\left(-\frac{1}{3}\right)^r=\frac{1-(-\frac{1}{3})^n}{1-(-\frac{1}{3})}=\frac{3}{4}\left[1-\left(-\frac{1}{3}\right)^n\right].$$

Part (c). Observe that

$$|S_n-S|<0.001\%S \implies \left|\frac{S_n-S}{S}\right|<\frac{1}{100000} \implies \left|\frac{\frac{3}{4}(1-(-\frac{1}{3})^n)}{\frac{3}{4}}-1\right|<\frac{1}{100000}.$$

Using G.C., the least value of n that satisfies the above inequality is 11.

Problem 9. Given that $\sum_{r=1}^{n} r^2 = \frac{n}{6}(n+1)(2n+1)$,

(a) write down $\sum_{r=1}^{2k} r^2$ in terms of k

(b) find $2^2 + 4^2 + 6^2 + \ldots + (2k)^2$.

Hence, show that $\sum_{r=1}^{k}(2r-1)^2 = \frac{k}{3}(2k+1)(2k-1)$.

Solution.

Part (a).

$$\sum_{r=1}^{2k} r^2 = \frac{2k}{6}(2k+1)(2(2k)+1) = \frac{k}{3}(2k+1)(4k+1).$$

Part (b).

$$2^2 + 4^2 + 6^2 + \ldots + (2k)^2 = \sum_{r=1}^{k}(2r)^2 = \sum_{r=1}^{k} 4r^2 = \frac{2k}{3}(k+1)(2k+1).$$

From parts (a) and (b), we clearly have

$$\sum_{r=1}^{k}(2r-1)^2 = \sum_{r=1}^{2k} r^2 - \sum_{r=1}^{k}(2r)^2 = \frac{k}{3}(2k+1)(4k+1) - \frac{2k}{3}(k+1)(2k+1) = \frac{k}{3}(2k+1)(2k-1).$$

* * * * *

Problem 10. Given that $u_n = e^{nx} - e^{(n+1)x}$, find $\sum_{n=1}^{N} u_n$ in terms of N and x. Hence, determine the set of values of x for which the infinite series $u_1 + u_2 + u_3 + \ldots$ is convergent and give the sum to infinity for cases where this exists.

Solution.

$$\sum_{n=1}^{N} u_n = \left(e^x - e^{2x}\right) + \left(e^{2x} - e^{3x}\right) + \cdots + \left(e^{Nx} + e^{(N+1)x}\right) = e^x - e^{(N+1)x}.$$

For the infinite series to converge, we require $|e^x| < 1$. Hence, $x \in \mathbb{R}_0^-$.

We now consider the sum to infinity.

Case 1. Suppose $x = 0$. Then $e^x = 1$, whence the sum to infinity is clearly 0.

Case 2. Suppose $x < 0$. Then $\lim_{N\to\infty} e^{(N+1)x} \to 0$. Thus, the sum to infinity is e^x.

* * * * *

Problem 11. Given that r is a positive integer and $f(r) = \frac{1}{r^2}$, express $f(r) - f(r+1)$ as a single fraction. Hence, prove that $\sum_{r=1}^{4n}\left(\frac{2r+1}{r^2(r+1)^2}\right) = 1 - \frac{1}{(4n+1)^2}$. Give a reason why the series is convergent and state the sum to infinity. Find $\sum_{r=2}^{4n}\left(\frac{2r-1}{r^2(r-1)^2}\right)$.

Solution.

$$f(r) - f(r+1) = \frac{1}{r^2} - \frac{1}{(r+1)^2} = \frac{(r+1)^2 - r^2}{r^2(r+1)^2} = \frac{2r+1}{r^2(r+1)^2}.$$

$$\begin{aligned}\sum_{r=1}^{4n}\left(\frac{2r+1}{r^2(r+1)^2}\right) &= \sum_{r=1}^{4n}[f(r) - f(r+1)] \\ &= [f(1) - f(2)] + [f(2) - f(3)] + \cdots + [f(4n) - f(4n-1)] \\ &= f(1) - f(4n+1) = 1 - \frac{1}{(4n+1)^2}\end{aligned}$$

As $n \to \infty$, $\frac{1}{(4n+1)^2} \to 0$. Hence, the series converges to 1.

$$\begin{aligned}\sum_{r=2}^{4n}\left(\frac{2r-1}{r^2(r-1)^2}\right) = \sum_{r=1}^{4n-1}\left(\frac{2r+1}{r^2(r+1)^2}\right) &= \sum_{r=1}^{4n-1}[f(r)-f(r+1)]\\ &= [f(1)-f(2)]+[f(2)-f(3)]+\cdots+[f(4n-1)-f(4n)]\\ &= 1-f(4n) = 1-\frac{1}{16n^2}\end{aligned}$$

* * * * *

Problem 12.

(a) Express $\frac{1}{(2x+1)(2x+3)(2x+5)}$ in partial fractions.

(b) Hence, show that $\sum_{r=1}^{n}\frac{1}{(2r+1)(2r+3)(2r+5)} = \frac{1}{60} - \frac{1}{4(2n+3)(2n+5)}$.

(c) Deduce the sum of $\frac{1}{1\cdot 3\cdot 5} + \frac{1}{3\cdot 5\cdot 7} + \frac{1}{3\cdot 5\cdot 7\cdot 9} + \ldots + \frac{1}{41\cdot 43\cdot 45}$.

Solution.

Part (a). Using the cover-up rule, we obtain

$$\frac{1}{(2x+1)(2x+3)(2x+5)} = \frac{1}{8(2x+1)} - \frac{1}{4(2x+3)} + \frac{1}{8(2x+5)}.$$

Part (b).

$$\begin{aligned}\sum_{r=1}^{n}\frac{1}{(2r+1)(2r+3)(2r+5)} &= \sum_{r=1}^{n}\left(\frac{1}{8(2r+1)} - \frac{1}{4(2r+3)} + \frac{1}{8(2r+5)}\right)\\ &= \frac{1}{8}\left[\left(\sum_{r=1}^{n}\frac{1}{2r+1} - \sum_{r=1}^{n}\frac{1}{2r+3}\right) - \left(\sum_{r=1}^{n}\frac{1}{2r+3} - \sum_{r=1}^{n}\frac{1}{2r+5}\right)\right]\end{aligned}$$

Observe that the two terms in brackets clearly telescope, leaving us with

$$\sum_{r=1}^{n}\frac{1}{(2r+1)(2r+3)(2r+5)} = \frac{1}{8}\left[\left(\frac{1}{3} - \frac{1}{2n+3}\right) - \left(\frac{1}{5} - \frac{1}{2n+5}\right)\right],$$

which simplifies to

$$\sum_{r=1}^{n}\frac{1}{(2r+1)(2r+3)(2r+5)} = \frac{1}{60} - \frac{1}{4(2n+3)(2n+5)}$$

as desired.

Part (c).

$$\begin{aligned}&\frac{1}{1\cdot 3\cdot 5} + \frac{1}{3\cdot 5\cdot 7} + \frac{1}{3\cdot 5\cdot 7\cdot 9} + \ldots + \frac{1}{41\cdot 43\cdot 45}\\ &= \frac{1}{1\cdot 3\cdot 5} + \sum_{r=1}^{20}\frac{1}{(2r+1)(2r+3)(2r+5)}\\ &= \frac{1}{15} + \left(\frac{1}{60} - \frac{1}{4(2\cdot 20+3)(2\cdot 20+5)}\right)\\ &= \frac{161}{1935}.\end{aligned}$$

Self-Practice A4

Problem 1. Evaluate $\sum_{r=2}^{n}\left(2^{-r}+2nr+n^2\right)$, giving your answer in terms of n.

Solution. Splitting the sum, we get

$$\sum_{r=2}^{n}\left(2^{-r}+2nr+n^2\right)=\sum_{r=2}^{n}\left(\frac{1}{2}\right)^{r}+2n\sum_{r=2}^{n}r+n^2\sum_{r=2}^{n}1.$$

Hence,

$$\begin{aligned}\sum_{r=2}^{n}\left(2^{-r}+2nr+n^2\right)&=\left(\frac{1}{2}\right)^{2}\left(\frac{1-(1/2)^{n-1}}{1-1/2}\right)+2n\left(\frac{n(n+1)}{2}-1\right)+n^2\left(n-1\right)\\&=\left[\frac{1}{2}-\left(\frac{1}{2}\right)^{n}\right]+n^2\left(n+1\right)-2n+n^2\left(n-1\right)\\&=\frac{1}{2}-\left(\frac{1}{2}\right)^{n}+2n\left(n^2-1\right).\end{aligned}$$

* * * * *

Problem 2. A geometric sequence $\{a_n\}$ has first term a and common ratio r. The sequence of numbers $\{b_n\}$ satisfy the relation $b_n=\ln(a_n)$ for $n\in\mathbb{Z}^+$.

(a) Show that $\{b_n\}$ is an arithmetic sequence and determine the value of the common difference in terms of r.

(b) Find an expression for $\sum_{n=1}^{N+1}b_n$ in terms of a, a_{N+1} and N.

(c) Hence, obtain an expression for $a_1\times a_2\times\cdots\times a_{N+1}$ in terms of a, a_{N+1} and N.

Solution.

Part (a). Note that $a_n=ar^{n-1}$. Hence,

$$b_n=\ln a_n=\ln\left(ar^{n-1}\right)=\ln a+(n-1)\ln r.$$

Hence,

$$b_n-b_{n-1}=[\ln a+n\ln r]-[\ln a+(n-1)\ln r]=\ln r.$$

Thus, $\{b_n\}$ is an arithmetic progression with common difference $\ln r$.

Part (b). Since $\{b_n\}$ is in arithmetic progression, we have

$$\sum_{n=1}^{N+1}b_n=\frac{N+1}{2}\left(b_1+b_{N+1}\right)=\frac{N+1}{2}\left(\ln a_1+\ln a_{N+1}\right)=\frac{N+1}{2}\ln\left(aa_{N+1}\right).$$

Part (c). Since $b_n=\ln a_n$, we can write the sum as

$$\sum_{n=1}^{N+1}b_n=\sum_{n=1}^{N+1}\ln a_n=\ln\prod_{n=1}^{N+1}a_n.$$

Equating this with the above result yields

$$\prod_{n=1}^{N+1}a_n=\left(aa_{N+1}\right)^{(N+1)/2}.$$

Problem 3. It is given that $\sum_{r=1}^{n} r^2 = \frac{1}{6}n(n+1)(2n+1)$.

(a) Show that $\sum_{r=1}^{n}(2r-7)(r+1) = \frac{1}{6}n\left(4n^2-9n-55\right)$.

(b) Find $\sum_{r=1}^{n} 3^{-r}$ in terms of n, and find the least value of n such that

$$\sum_{r=1}^{n}(2r-7)(r+1) > \sum_{r=1}^{n} 3^{-r}.$$

(c) Express $\sum_{r=n+1}^{2n}(2r-7)(r+1)$ in terms of n.

(d) Hence, or otherwise, find the value of

$$43 \times 26 + 45 \times 27 + 47 \times 28 + \cdots + 87 \times 48 + 89 \times 49.$$

Solution.

Part (a). Note that $(2r-7)(r+1) = 2r^2 - 5r - 7$. Hence,

$$\begin{aligned}\sum_{r=1}^{n}(2r-7)(r+1) &= \sum_{r=1}^{n}\left(2r^2-5r-7\right) \\ &= 2\left(\frac{n(n+1)(2n+1)}{6}\right) - 5\left(\frac{n(n+1)}{2}\right) - 7n \\ &= \frac{n\left(4n^2-9n-55\right)}{6}.\end{aligned}$$

Part (b).

$$\sum_{r=1}^{n} 3^{-r} = \sum_{r=1}^{n}\left(\frac{1}{3}\right)^r = \left(\frac{1}{3}\right)\left(\frac{1-(1/3)^{n-1}}{1-1/3}\right) = \frac{1}{2}\left(1-\frac{1}{3^n}\right).$$

The inequality hence becomes

$$\frac{n\left(4n^2-9n-55\right)}{6} > \frac{1}{2}\left(1-\frac{1}{3^n}\right).$$

Using G.C., $n \geq 5.019$. Since n is an integer, the least n that satisfies the inequality is 6.

Part (c). We have

$$\begin{aligned}\sum_{r=n+1}^{2n}(2r-7)(r+1) &= \sum_{r=1}^{2n}(2r-7)(r+1) - \sum_{r=1}^{n}(2r-7)(r+1) \\ &= \frac{2n\left[4(2n)^2-9(2n)-55\right]}{6} - \frac{n\left(4n^2-9n-55\right)}{6} \\ &= \frac{n\left(28n^2-27n-55\right)}{6}.\end{aligned}$$

Part (d). We have

$$\begin{aligned}43 \times 26 + 45 \times 27 + 47 \times 28 + \cdots + 87 \times 48 + 89 \times 49 &= \sum_{r=24+1}^{2(24)}(2r-7)(r+1) \\ &= \frac{(24)\left[28(24)^2-27(24)-55\right]}{6} = 61700.\end{aligned}$$

Problem 4. It is given that $\sum_{r=1}^{n} \frac{2r+1}{r(r+1)(r+2)} = \frac{n(5n+7)}{4(n+1)(n+2)}$.

(a) Show that the series $\sum_{r=1}^{\infty} \frac{2r+1}{r(r+1)(r+2)}$ converges and write down its sum to infinity.

(b) Find $\sum_{r=0}^{n-2} \frac{2r+5}{(r+2)(r+3)(r+4)}$.

Solution.

Part (a). Clearly,

$$\sum_{r=1}^{\infty} \frac{2r+1}{r(r+1)(r+2)} = \lim_{n\to\infty} \frac{n(5n+7)}{4(n+1)(n+2)} = \lim_{n\to\infty} \frac{5+\frac{7}{n}}{4\left(1+\frac{1}{n}\right)\left(1+\frac{2}{n}\right)} = \frac{5}{4}.$$

Thus, the series converges and its sum to infinity is $5/4$.

Part (b). Reindexing $r \mapsto r-2$,

$$\sum_{r=0}^{n-2} \frac{2r+5}{(r+2)(r+3)(r+4)} = \sum_{r=2}^{n} \frac{2r+1}{r(r+1)(r+2)}$$

$$= \sum_{r=1}^{n} \frac{2r+1}{r(r+1)(r+2)} - \frac{2(1)+1}{1(1+1)(1+2)} = \frac{n(5n+7)}{4(n+1)(n+2)} - \frac{1}{2}.$$

Assignment A4

Problem 1. Find $\sum_{r=0}^{n}\left(n^2+1-3r\right)$ in terms of n, giving your answer in factorized form.

Solution.

$$\sum_{r=0}^{n}\left(n^2+1-3r\right)=(n+1)(n^2+1)-3\left[\frac{n(n+1)}{2}\right]=\frac{1}{2}(n+1)\left(2n^2-3n+2\right).$$

* * * * *

Problem 2. Given that $\sum_{k=1}^{n}k!\left(k^2+1\right)=(n+1)!\,n$, find $\sum_{k=1}^{n-1}(k+1)!\left(k^2+2k+2\right)$.

Solution. Reindexing $k+1\mapsto k$,

$$\sum_{k=1}^{n-1}(k+1)!\left(k^2+2k+2\right)=\sum_{k=2}^{n}k!\left(k^2+1\right).$$

Using the given result,

$$\sum_{k=2}^{n}k!\left(k^2+1\right)=\sum_{k=1}^{n}k!\left(k^2+1\right)-1!\left(1^2+1\right)=(n+1)!\,n-2.$$

* * * * *

Problem 3. Given that $\sum_{r=1}^{n}=\frac{1}{6}n(n+1)(2n+1)$, find $\sum_{r=N+1}^{2N}\left(7^{r+1}+3r^2\right)$ in terms of N, simplifying your answer.

Solution. Note that

$$\sum_{r=N+1}^{2N}7^{r+1}=\frac{7^{(N+1)+1}(7^N-1)}{7-1}=\frac{7^{N+2}(7^N-1)}{6}.$$

Next, we split the sum of squares:

$$\sum_{r=N+1}^{2N}3r^2=3\left(\sum_{r=1}^{2N}r^2-\sum_{r=1}^{N}r^2\right).$$

Using the given result,

$$\sum_{r=N+1}^{2N}3r^2=3\left(\frac{(2N)(2N+1)(4N+1)}{6}-\frac{N(N+1)(2N+1)}{6}\right)=\frac{N(2N+1)(7N+1)}{2}.$$

Thus,

$$\sum_{r=N+1}^{2N}\left(7^{r+1}+3r^2\right)=\frac{7^{N+2}(7^N-1)}{6}+\frac{N(2N+1)(7N+1)}{2}.$$

Problem 4. Let $f(r) = \frac{3}{r-1}$.

(a) Show that $f(r+1) - f(r) = -\frac{3}{r(r-1)}$.

(b) Hence, find in terms of N, the sum of the series $S_N = \sum_{r=2}^{N} \frac{1}{r(r-1)}$.

(c) Explain why $\sum_{r=2}^{\infty} \frac{1}{r(r-1)}$ is a convergent series, and find the value of the sum to infinity.

(d) Using the result from part (b), find $\sum_{r=2}^{N} \frac{1}{r(r+1)}$.

Solution.

Part (a).

$$f(r+1) - f(r) = \frac{3}{(r+1)-1} - \frac{3}{r-1} = \frac{3(r-1) - 3r}{r(r-1)} = -\frac{3}{r(r-1)}.$$

Part (b). Observe that

$$S_N = \sum_{r=2}^{N} \frac{1}{r(r-1)} = -\frac{1}{3} \sum_{r=2}^{N} -\frac{3}{r(r-1)} = -\frac{1}{3} \left[\sum_{r=2}^{N} f(r+1) - \sum_{r=2}^{N} f(r) \right],$$

which clearly telescopes. Thus,

$$S_N = -\frac{f(N+1) - f(2)}{3} = -\frac{1}{3} \left(\frac{3}{N+1-1} - \frac{3}{2-1} \right) = 1 - \frac{1}{N}.$$

Part (c).

$$\lim_{N\to\infty} S_N = \lim_{N\to\infty} \left(1 - \frac{1}{N}\right) = 1 - 0 = 1.$$

Since 1 is a constant, $\sum_{r=2}^{\infty} \frac{1}{r(r-1)}$ is a convergent series.

Part (d). Reindexing $r \mapsto r-1$,

$$\sum_{r=2}^{N} \frac{1}{r(r+1)} = \sum_{r=3}^{N+1} \frac{1}{(r-1)r} = \sum_{r=2}^{N} \frac{1}{r(r-1)} - \frac{1}{2(2-1)} + \frac{1}{(N+1)N}.$$

Using the result from part (b),

$$\sum_{r=2}^{N} \frac{1}{r(r+1)} = \left(1 - \frac{1}{N}\right) - \frac{1}{2(2-1)} + \frac{1}{(N+1)N} = \frac{1}{2} - \frac{1}{N+1}.$$

A5 Recurrence Relations

Tutorial A5

Problem 1. Solve these recurrence relations together with the initial conditions.

(a) $u_n = 2u_{n-1}$, for $n \geq 1$, $u_0 = 3$

(b) $u_n = 3u_{n-1} + 7$, for $n \geq 1$, $u_0 = 5$

Solution.
Part (a). $u_n = 2^n \cdot u_0 = 3 \cdot 2^n$.
Part (b). Let k be a constant such that $u_n + k = 3(u_{n-1} + k)$. Then $k = \frac{7}{2}$. Hence,

$$u_n + \frac{7}{2} = 3\left(u_{n-1} + \frac{7}{2}\right) \implies u_n + \frac{7}{2} = 3^n\left(u_0 + \frac{7}{2}\right) \implies u_n = \frac{17}{2} \cdot 3^n - \frac{7}{2}.$$

* * * * *

Problem 2. Solve these recurrence relations together with the initial conditions.

(a) $u_n = 5u_{n-1} - 6u_{n-2}$, for $n \geq 2$, $u_0 = 1$, $u_1 = 0$

(b) $u_n = 4u_{n-2}$, for $n \geq 2$, $u_0 = 0$, $u_1 = 4$

(c) $u_n = 4u_{n-1} - 4u_{n-2}$, for $n \geq 2$, $u_0 = 6$, $u_1 = 8$

(d) $u_n = -6u_{n-1} - 9u_{n-2}$, for $n \geq 2$, $u_0 = 3$, $u_1 = -3$

(e) $u_n = 2u_{n-1} - 2u_{n-2}$, for $n \geq 2$, $u_0 = 2$, $u_1 = 6$

Solution.
Part (a). Note that the characteristic equation of u_n, $x^2 - 5x + 6 = 0$, has roots 2 and 3. Thus,

$$u_n = A \cdot 2^n + B \cdot 3^n.$$

From $u_0 = 1$ and $u_1 = 0$, we have the equations $A + B = 1$ and $2A + 3B = 0$. Solving, we see that $A = 3$ and $B = 2$, whence

$$u_n = 3 \cdot 2^n + 2 \cdot 3^n.$$

Part (b). Note that the characteristic equation of u_n, $x^2 - 4 = 0$, has roots -2 and 2. Thus,

$$u_n = A(-2)^n + B \cdot 2^n.$$

From $u_0 = 0$ and $u_1 = 4$, we get $A + B = 0$ and $-2A + 2B = 4$. Solving, we see that $A = -1$ and $B = 1$, whence

$$u_n = -(-2)^n + 2^n.$$

Part (c). Note that the characteristic equation of u_n, $x^2 - 4x + 4 = 0$, has only one root, 2. Thus,

$$u_n = (A + Bn)2^n.$$

From $u_0 = 6$ and $u_1 = 8$, we obtain $A = 6$ and $A + B = 4$, whence $B = -2$. Thus,

$$u_n = (6 - 2n)2^n.$$

Part (d). Note that the characteristic equation of u_n, $x^2 + 6x + 9 = 0$, has only one root, -3. Thus,

$$u_n = (A + Bn)(-3)^n.$$

From $u_0 = 3$ and $u_1 = -3$, we get $A = 3$ and $A + B = 1$, whence $B = -2$. Thus,

$$u_n = (3 - 2n)2^n.$$

Part (e). Consider the characteristic equation of u_n, $x^2 - 2x + 2 = 0$. By the quadratic formula, this has roots $x = 1 \pm i = \sqrt{2}\exp\left(\pm\frac{i\pi}{4}\right)$. Hence,

$$u_n = A \cdot 2^{\frac{1}{2}n} \cos\left(\frac{n\pi}{4}n\right) + B \cdot 2^{\frac{1}{2}n} \sin\left(\frac{n\pi}{4}\right).$$

From $u_0 = 2$, we obtain $A = 2$. From $u_0 = 6$, we obtain $A + B = 6$, whence $B = 4$. Thus,

$$u_n = 2^{\frac{1}{2}n+1} \cos\left(\frac{n\pi}{4}\right) + 2^{\frac{1}{2}n+2} \sin\left(\frac{n\pi}{4}\right).$$

* * * * *

Problem 3.

(a) A sequence is defined by the formula $b_n = \frac{n!n!}{(2n)!} \cdot 2^n$, where $n \in \mathbb{Z}^+$. Show that the sequence satisfies the recurrence relation $b_{n+1} = \frac{n+1}{2n+1} b_n$.

(b) A sequence is defined recursively by the formula

$$u_{n+1} = 2u_n + 3, \qquad n \in \mathbb{Z}_0^+,\ u_0 = a$$

Show that $u_n = 2^n a + 3\left(2^n - 1\right)$.

Solution.

Part (a).

$$b_{n+1} = \frac{(n+1)!(n+1)!}{(2n+2)!} \cdot 2^{n+1} = \frac{2(n+1)^2}{(2n+1)(2n+2)} \left[\frac{n!n!}{(2n)!} \cdot 2^n\right] = \frac{n+1}{2n+1} b_n.$$

Part (b). Let k be a constant such that $u_{n+1} + k = 2(u_n + k)$. Then $k = 3$. Hence,

$$u_{n+1} + 3 = 2(u_n + 3) \implies u_n + 3 = 2^n(u_0 + 3) \implies u_n = 2^n(a + 3) - 3 = 2^n a + 3\left(2^n - 1\right).$$

Problem 4. The volume of water, in litres, in a storage tank decreases by 10% by the end of each day. However, 90 litres of water is also pumped into the tank at the end of each day. The volume of water in the tank at the end of n days is denoted by x_n and x_0 is the initial volume of water in the tank.

(a) Write down a recurrence relation to represent the above situation.

(b) Show that $x_n = 0.9^n(x_0 - 900) + 900$.

(c) Deduce the amount of water in the tank when n becomes very large.

Solution.

Part (a). $x_{n+1} = 0.9x_n + 90$, $n \in \mathbb{N}$

Part (b). Let k be a constant such that $x_{n+1} + k = 0.9(x_n + k)$. Then $k = -900$. Hence,

$$x_{n+1} - 900 = 0.9(x_n - 900) \implies x_n - 900 = 0.9^n(x_0 - 900) \implies x_n = 0.9^n(x_0 - 900) + 900.$$

Part (c). As $n \to \infty$, $0.9^n \to 0$. Hence, the amount of water in the tank will converge to 900 litres.

* * * * *

Problem 5. A deposit of $100,000 is made to an investment fund at the beginning of a year. On the last day of each year, two dividends are awarded and reinvested into the fund. The first dividend is 20% of the amount in the account during that year. The second dividend is 45% of the amount in the account in the previous year.

(a) Find a recurrence relation $\{P_n\}$ where P_n is the amount at the start of the nth year if no money is ever withdrawn.

(b) How much is in the account after n years if no money is ever withdrawn?

Solution.

Part (a).

$$P_{n+2} = P_{n+1} + 0.2P_{n+1} + 0.45P_n = 1.2P_{n+1} + 0.45P_n.$$

Part (b). Note that the characteristic equation of P_n, $x^2 - 1.2x - 0.45 = 0$, has roots $-\frac{3}{10}$ and $\frac{3}{2}$. Thus,

$$P_n = A\left(-\frac{3}{10}\right)^n + B\left(\frac{3}{2}\right)^n.$$

From $P_0 = 0$ and $P_1 = 100000$, we have $A + B = 0$ and $-\frac{3}{10}A + \frac{3}{2}B = 100000$. Solving, we have $A = -\frac{500000}{9}$ and $B = \frac{500000}{9}$. Thus,

$$P_n = \frac{500000}{9}\left[\left(\frac{3}{2}\right)^n - \left(-\frac{3}{10}\right)^n\right].$$

Hence, there will be $\$\left\{\frac{500000}{9}\left[\left(\frac{3}{2}\right)^n - \left(-\frac{3}{10}\right)^n\right]\right\}$ in the account after n years if no money is ever withdrawn

Problem 6. A pair of rabbits does not breed until they are two months old. After they are two months old, each pair of rabbit produces another pair each month.

(a) Find a recurrence relation $\{f_n\}$ where f_n is the total number of pairs of rabbits, assuming that no rabbits ever die.

(b) What is the number of pairs of rabbits at the end of the nth month, assuming that no rabbits ever die?

Solution.

Part (a). $f_{n+2} = f_{n+1} + f_n$, $n \geq 2$, $f_0 = 0, f_1 = 1$

Part (b). Consider the characteristic equation of f_n, $x^2 - x - 1 = 0$. By the quadratic formula, the roots of the characteristic equation are $\frac{1+\sqrt{5}}{2}$ and $\frac{1-\sqrt{5}}{2}$. Hence,

$$f_n = A\left(\frac{1+\sqrt{5}}{2}\right)^n + B\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

From $f_0 = 0$, we get $A + B = 0$. From $f_1 = 1$, we get $A\left(\frac{1+\sqrt{5}}{2}\right) + B\left(\frac{1-\sqrt{5}}{2}\right) = 1$. Solving, we get $A = \frac{1}{\sqrt{5}}$ and $B = -\frac{1}{\sqrt{5}}$. Hence,

$$f_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

* * * * *

Problem 7. For $n \in \{2^j : j \in \mathbb{Z}, j \geq 1\}$, it is given that $T_n = 3T_{n/2} + 17$, where $T_1 = 4$. By considering the substitution $n = 2^i$ and another suitable substitution, show that the recurrence relation can be expressed in the form

$$t_i = 3t_{i-1} + 17, \qquad i \in \mathbb{Z}^+$$

Hence, find an expression for T_n in terms of n.

Solution. Let $n = 2^i \iff i = \log_2 n$. The given recurrence relation transforms to

$$T_{2^i} = 3T_{2^{i-1}} + 17, T_{2^0} = 4.$$

Let $t_i = T_{2i}$. Then

$$t_i = 3t_{i-1} + 17, t_0 = 4.$$

Let k be a constant such that $t_i + k = 3(t_{i-1} + k)$. Then $k = \frac{17}{2}$. We thus obtain a formula for t_i:

$$t_i + \frac{17}{2} = 3\left(t_{i-1} + \frac{17}{2}\right) \implies t_i + \frac{17}{2} = 3^i\left(t_0 + \frac{17}{2}\right) \implies t_i = \frac{25}{2}\cdot 3^i - \frac{17}{2}.$$

Thus,

$$T_{2i} = \frac{25}{2}\cdot 3^i - \frac{17}{2} \implies T_n = \frac{25}{2}\cdot 3^{\log_2 n} - \frac{17}{2}.$$

Problem 8. Consider the sequence $\{a_n\}$ given by the recurrence relation

$$a_{n+1} = 2a_n + 5^n, \qquad n \geq 1.$$

(a) Given that $a_n = k\,(5^n)$ satisfies the recurrent relation, find the value of the constant k.

(b) Hence, by considering the sequence $\{b_n\}$ where $b_n = a_n - k(5^n)$, find the particular solution to the recurrence relation for which $a_1 = 2$.

Solution.

Part (a).

$$a_{n+1} = 2a_n + 5^n \implies k\left(5^{n+1}\right) = 2 \cdot k\,(5^n) + 5^n \implies 5k = 2k + 1 \implies k = \frac{1}{3}.$$

Part (b).

$$b_n = a_n - \frac{5^n}{3} = \left(2a_{n-1} - 5^{n-1}\right) - \frac{5^n}{3} = 2a_{n-1} - \frac{2}{3} \cdot 5^{n-1} = 2\left(a_{n-1} - \frac{5^{n-1}}{3}\right) = 2b_{n-1}.$$

Hence, $b_n = b_1 \cdot 2^{n-1}$. Note that $b_1 = a_1 - \frac{5}{3} = \frac{1}{3}$. Thus, $b_n = \frac{2^{n-1}}{3}$, which gives

$$b_n = a_n - \frac{5^n}{3} = \frac{2^{n-1}}{3} \implies a_n = \frac{2^n + 2 \cdot 5^n}{6}.$$

* * * * *

Problem 9. The sequence $\{X_n\}$ is given by

$$\sqrt{X_{n+2}} = \frac{X_{n+1}}{X_n^2}, \qquad n \geq 1.$$

By applying the natural logarithm to the recurrence relation, use a suitable substitution to find the general solution of the sequence, expressing your answer in trigonometric form.

Solution. Taking the natural logarithm of the recurrence relation and simplifying, we get

$$\ln X_{n+2} = 2 \ln X_{n+1} - 4 \ln X_n.$$

Let $L_n = \ln X_n \iff X_n = \exp(L_n)$. Then,

$$L_{n+2} = 2L_{n+1} - 4L_n.$$

Consider the characteristic equation of L_n, $x^2 - 2x + 4 = 0$. By the quadratic formula, this has roots $1 \pm \sqrt{3}\mathrm{i} = 2\exp\left(\pm\frac{i\pi}{3}\right)$. Thus, we can express L_n as

$$L_n = A \cdot 2^n \cos\frac{n\pi}{3} + B \cdot 2^n \sin\frac{n\pi}{3} = 2^n\left(A\cos\frac{n\pi}{3} + B\sin\frac{n\pi}{3}\right).$$

Thus, X_n has the general solution

$$X_n = \exp\left(2^n\left(A\cos\frac{n\pi}{3} + B\sin\frac{n\pi}{3}\right)\right).$$

Problem 10. The sequence $\{X_n\}$ is given by $X_1 = 2$, $X_2 = 15$ and

$$X_{n+2} = 5\left(1 + \frac{1}{n+2}\right) X_{n+1} - 6\left(1 + \frac{2}{n+1}\right) X_n, \qquad n \geq 1.$$

By dividing the recurrence relation throughout by $n + 3$, use a suitable substitution to determine X_n as a function of n.

Solution. Dividing the recurrence relation by $n + 3$, we obtain

$$\frac{X_{n+2}}{n+3} = 5\left(\frac{1}{n+3} + \frac{1}{(n+2)(n+3)}\right) X_{n+1} - 6\left(\frac{1}{n+3} + \frac{2}{(n+1)(n+3)}\right) X_n.$$

Note that $\frac{1}{(n+2)(n+3)} = \frac{1}{n+2} - \frac{1}{n+3}$ and $\frac{2}{(n+1)(n+3)} = \frac{1}{n+1} - \frac{1}{n+3}$. Thus,

$$\frac{X_{n+2}}{n+3} = 5\left(\frac{X_{n+1}}{n+2}\right) - 6\left(\frac{X_n}{n+1}\right).$$

Let $Y_n = \frac{n+1}{X_n} \iff X_n = (n+1)Y_n$. Then,

$$Y_{n+2} = 5Y_{n+1} - 6Y_n.$$

Note that the characteristic equation of Y_n, $x^2 - 5x + 6 = 0$, has roots 2 and 3. Hence,

$$Y_n = A \cdot 2^n + B \cdot 3^n \implies X_n = (n+1)\left(A \cdot 2^n + B \cdot 3^n\right).$$

From $X_1 = 2$ and $X_2 = 15$, we have $2A + 3B = 1$ and $4A + 9B = 5$. Solving, we obtain $A = -1$ and $B = 1$. Thus,

$$X_n = (n+1)\left(3^n - 2^n\right).$$

* * * * *

Problem 11. A logistics company set up an online platform providing delivery services to users on a monthly paid subscription basis. The company's sales manager models the number of subscribers that the company has at the end of each month. She notes that approximately 10% of the existing subscribers leave each month, and that there will be a constant number k of new subscribers in each subsequent month after the first.

Let T_n, $n \geq 1$, denote the number of subscribers the company has at the end of the nth month after the online platform was set up.

(a) Express T_{n+1} in terms of T_n.

The company has 250 subscribers at the end of the first month.

(b) Find T_n in terms of n and k.

(c) Find the least number of subscribers the company needs to attract in each subsequent month after the first if it aims to have at least 350 subscribers by the end of the 12th month.

Let $k = 50$ for the rest of the question.

The monthly running cost of the company is assumed to be fixed at \$4,000. The monthly subscription fee is \$10 per user which is charged at the end of each month.

(d) Given that the mth month is the first month in which the company's revenue up to and including that month is able to cover its cost up to and including that month, find the value of m.

(e) Using your answer to part (b), determine the long-term behaviour of the number of subscribers that the company has. Hence, explain whether this behaviour is appropriate in terms of long-term prospects for the company's success.

Solution.

Part (a). $T_{n+1} = 0.9T_n + k$

Part (b). Let m be a constant such that $T_{n+1} + m = 0.9\,(T_n + m)$. Then $m = -10k$. Hence,

$$T_{n+1} - 10k = 0.9\,(T_n - 10k) \implies T_n - 10k = 0.9^{n-1}\,(T_0 - 10k)\,.$$

Since $T_0 = 250$, we get

$$T_n = 0.9^{n-1}\,(250 - 10k) + 10k.$$

Part (c). Consider $T_{12} \geq 350$.

$$T_{12} \geq 350 \implies 0.9^{12-1}\,(250 - 10k) + 10k \geq 350.$$

Using G.C., $k \geq 39.6$. Hence, the company needs to attract at least 40 subscribers in each subsequent month.

Part (d). Since $k = 50$, $T_n = -250 \cdot 0.9^{n-1} + 500$. Let $\$S_m$ be the total revenue for the first m months.

$$\begin{aligned} S_m &= 10\sum_{n=1}^{m} T_n = 10\sum_{n=1}^{m}\left(-250 \cdot 0.9^{n-1} + 500\right) \\ &= 10\left[-250\left(\frac{1-0.9^m}{1-0.9}\right) + 500m\right] = 25000\,(0.9^m - 1) + 5000m. \end{aligned}$$

Note that the total cost for the first m months is $\$4000m$. Hence, the total profit for the first m months is given by $\$(S_m - 4000m)$. Hence, we consider $S_m - 4000m \geq 0$:

$$S_m - 4000m \geq 0 \implies 25000\,(0.9^m - 1) + 1000m \geq 0.$$

Using G.C., we obtain $m \geq 22.7$, whence the least value of m is 23.

Part (e). As $n \to \infty$, $0.9^{n-1} \to 0$. Hence, $T_n \to 500$. Hence, as n becomes very large, the profit per month approaches $500 \cdot 10 - 4000 = 1000$ dollars. Thus, this behaviour is appropriate as the business will remain profitable in the long run.

Self-Practice A5

Problem 1. Tom wants to buy a new Aphone11. To save up for his purchase, Tom takes up a part-time job that pays him \$400 per month which will be credited into his bank account on the 25th of each month, starting from January 2012. On the first day of every month of 2012, he withdraws half of the total amount of money from his bank account for food and transportation. Assuming that Tom has \$250 in this bank account on 31 December 2011,

(a) write down a recurrence relation for u_n, where u_n denotes the amount in his bank account on the last day of the nth month after December 2011, and

(b) show that $u_n = 800 - 550\,(0.5^n)$.

Given that a new Aphone 11 costs \$850,

(c) explain why Tom is unable to buy the Aphone11, and

(d) find the maximum percentage of the total amount of money in the bank that Tom should spend on transport and food every month in order to be able to buy the Aphone11 on the last day of December 2012.

Solution.

Part (a). We have

$$u_n = \frac{1}{2}u_{n-1} + 400, \quad u_0 = 250.$$

Part (b). Note that the complementary solution is

$$u_n^{(c)} = C\left(\frac{1}{2}\right)^n,$$

where C is an arbitrary constant. Let the particular solution be $u_n^{(p)} = k$. Then

$$k = \frac{1}{2}k + 400 \implies k = 800.$$

Hence,

$$u_n = u_n^{(c)} + u_n^{(p)} = C\left(\frac{1}{2}\right)^n + 800.$$

Using the condition $u_0 = 250$, we get

$$250 = C + 800 \implies C = -500,$$

whence

$$u_n = 800 - 500\left(\frac{1}{2}\right)^n.$$

Part (c). Clearly, $-500(1/2)^n < 0$ for all $n > 0$. Hence,

$$u_n = 800 - 500\left(\frac{1}{2}\right)^n < 800 < 850.$$

Thus, Tom is unable to buy the Aphone11.

Part (d). Let the desired percentage be $p\%$. Then

$$u_n = \left(1 - \frac{p}{100}\right) u_{n-1} + 400.$$

Let the particular solution be $u_n^{(p)} = k$. Then

$$k = \left(1 - \frac{p}{100}\right) k + 400 \implies k = \frac{40000}{p}.$$

We thus want

$$\frac{40000}{p} \geq 850 \implies p \leq \frac{800}{17} = 47.059.$$

Hence, the maximum percentage is 47%.

* * * * *

Problem 2. A sequence of real numbers $u_1, u_2, u_3, \ldots$ satisfies the recurrence relation

$$u_n = 2u_{n-1} + 1, \quad n \geq 1.$$

Given that $u_1 = 2$, show that $u_n = 2^n + 2^{n-1} - 1$. Hence, determine the behaviour of the sequence.

Solution. Note that the complementary solution is

$$u_n^{(c)} = C2^n,$$

where C is an arbitrary constant. Let the particular solution be $u_n^{(p)} = k$. Then

$$k = 2k + 1 \implies k = -1.$$

Hence,

$$u_n = u_n^{(c)} + u_n^{(p)} = C2^n - 1.$$

Using the condition $u_1 = 2$, we get

$$2 = 2C - 1 \implies C = \frac{3}{2},$$

whence

$$u_n = \frac{3}{2} \cdot 2^n - 1 = (2+1)2^{n-1} - 1 = 2^n + 2^{n-1} - 1.$$

Clearly, u_n is increasing and diverges to infinity.

* * * * *

Problem 3. Solve these recurrence relations together with the initial conditions.

(a) $u_n = 7u_{n-1} - 10u_{n-2}$ for $n \geq 2$, $u_0 = 2$, $u_1 = 1$.

(b) $u_n = \frac{1}{4}u_{n-2}$ for $n \geq 2$, $u_0 = 1$, $u_1 = 0$.

(c) $u_n = -4u_{n-1} - 4u_{n-2}$ for $n \geq 2$, $u_0 = 0$, $u_1 = 1$.

(d) $u_{n+2} = -4u_{n+1} + 5u_n$ for $n \geq 0$, $u_0 = 2$, $u_1 = 8$.

Solution.

Part (a). Consider the characteristic equation $x^2 - 7x + 10 = 0$, which has distinct roots $x = 2$ and $x = 5$. Hence,

$$u_n = A\left(2^n\right) + B\left(5^n\right).$$

Using the conditions $u_0 = 2$ and $u_1 = 1$, we get the system

$$\begin{cases} A + B = 2 \\ 2A + 5B = 1 \end{cases},$$

whence $A = 3$ and $B = -1$. Thus,

$$u_n = 3\left(2^n\right) - 5^n.$$

Part (b). Consider the characteristic equation $x^2 = 1/4$, which has distinct roots $x = \pm 1/2$. Hence,

$$u_n = A\left(\frac{1}{2}\right)^n + B\left(-\frac{1}{2}\right)^n = \frac{1}{2^n}\left[A + (-1)^n B\right].$$

Using the conditions $u_0 = 1$ and $u_1 = 0$, we get the system

$$\begin{cases} A - B = 1 \\ A + B = 0 \end{cases},$$

whence $A = 1/2$ and $B = -1/2$. Thus,

$$u_n = \frac{1}{2^n}\left[\frac{1}{2} + (-1)^n\left(-\frac{1}{2}\right)\right] = \frac{1 + (-1)^{n-1}}{2^{n+1}}.$$

Part (c). Consider the characteristic equation $x^2 - 4x + 4 = 0$, which has the unique root $x = -2$. Hence,

$$u_n = (A + Bn)(-2)^n.$$

Using the conditions $u_0 = 0$ and $u_1 = 1$, we get the system

$$\begin{cases} A = 0 \\ 2A - 2B = 1 \end{cases},$$

whence $A = 0$ and $B = -1/2$. Thus,

$$u_n = \left(0 - \frac{n}{2}\right)(-2)^n = n(-2)^{n-1}.$$

Part (d). Consider the characteristic equation $x^2 + 4x - 5 = 0$, which has distinct roots $x = -5$ and $x = 1$. Hence,

$$u_n = A(-5)^n + B(1)^n = A(-5)^n + B.$$

Using the conditions $u_0 = 2$ and $u_1 = 8$, we get the system

$$\begin{cases} A + B = 2 \\ 5A + B = 8 \end{cases},$$

whence $A = -1$ and $B = 3$. Thus,

$$u_n = 3 - (-5)^n.$$

Problem 4 (🌶). Find the unit digit of the number $(3+\sqrt{5})^{2016}+(3-\sqrt{5})^{2016}$.

Solution. Let u_n be a sequence such that

$$u_n = \left(3+\sqrt{5}\right)^n + \left(3-\sqrt{5}\right)^n.$$

We aim to find a recurrence relation for u_n. First, observe that $3+\sqrt{5}$ and $3-\sqrt{5}$ are roots to the characteristic polynomial $P(x)$ of u_n:

$$P(x) = \left[x - \left(3+\sqrt{5}\right)\right]\left[x - \left(3-\sqrt{5}\right)\right] = x^2 - 6x + 4.$$

Thus, u_n satisfies the recurrence relation

$$u_n = 6u_{n-1} - 4u_{n-2}.$$

Since we are interested in the unit digit of u_{2016}, we consider $u_n \pmod{10}$:

$$u_n = 6u_{n-1} - 4u_{n-2} \equiv 6u_{n-1} + 6u_{n-2} = 6\left(u_{n-1}+u_{n-2}\right) \pmod{10}.$$

Since $u_0 = 2$ and $u_1 = 6$, we construct the following table:

n	$u_n \pmod{10}$
0	2
1	6
2	8
3	4
4	2
5	6

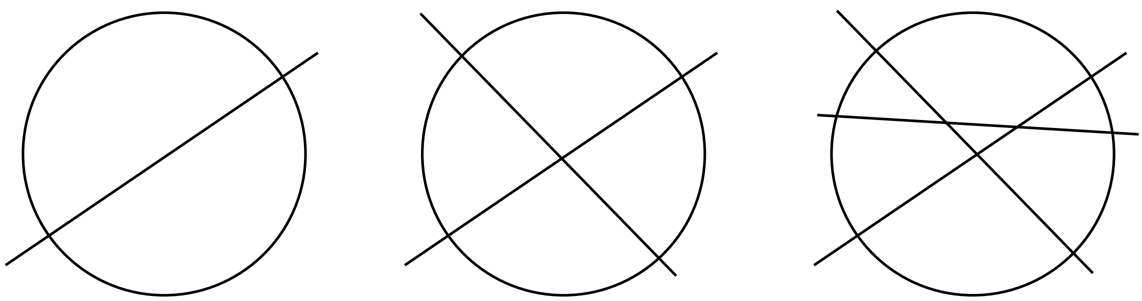
Observe that the pattern repeats every four terms: 2, 6, 8, 4, 2, 6, 8, 4, 2, ... Thus,

$$u_n \pmod{10} \equiv \begin{cases} 2, & n \equiv 0 \pmod{4} \\ 6, & n \equiv 1 \pmod{4} \\ 8, & n \equiv 2 \pmod{4} \\ 4, & n \equiv 3 \pmod{4} \end{cases}.$$

Since $2016 \equiv 0 \pmod{4}$, it follows that the unit digit of u_{2016} is 2.

* * * * *

Problem 5 (🌶). A person attempts to cut a circular pizza into as many pieces as possible with a given number of straight cuts. In order to have as many slices as possible with each cut, no three cuts are concurrent, no two cuts are parallel, and the intersection of any two cuts should lie in the interior of the pizza.



$n=1$ $\qquad$ $n=2$ $\qquad$ $n=3$

Find the maximum number of slices of a circular pizza that a person can obtain by making n straight cuts with a knife.

Solution. Let u_n be the maximum number of slices obtainable from n cuts. From the above diagrams, we see that the nth slice can add at most n new slices. Hence,

$$u_n = u_{n-1} + n.$$

We can rewrite this as

$$u_n - u_{n-1} = n.$$

Summing over $k = 2, 3, \ldots, n$,

$$u_n - a_1 = \sum_{k=1}^{n} (u_k - u_{k-1}) = \sum_{k=2}^{n} k = \frac{n^2+n}{2} - 1.$$

Since $a_1 = 2$, we have

$$u_n = \frac{n^2+n}{2} + 1 = \frac{n^2+n+2}{2}.$$

* * * * *

Problem 6 (🌶). Solve the simultaneous recurrence relations:

$$a_n = 3a_{n-1} + 2b_{n-1}, \quad b_n = a_{n-1} + 2b_{n-1}$$

with $a_0 = 1$ and $b_0 = 2$.

Solution. Adding the two equations together, we see that $\{a_n + b_n\}$ is in geometric progression:

$$a_n + b_n = 4\left(a_{n-1} + b_{n-1}\right) = 4^n\left(a_0 + b_0\right) = 3 \cdot 4^n.$$

Substituting this into the first equation, we get

$$a_n - a_{n-1} = 2\left(a_{n-1} + b_{n-1}\right) = 6 \cdot 4^{n-1}.$$

Summing over $k = 1, 2, \ldots, n$,

$$a_n - a_0 = \sum_{k=1}^{n} (a_k - a_{k-1}) = \sum_{k=1}^{n} 6 \cdot 4^{k-1} = 6\left(\frac{1-4^n}{1-4}\right) = 2\left(4^n - 1\right).$$

Thus,

$$a_n = a_0 + 2\left(4^n - 1\right) = 2^{2n+1} - 1$$

and

$$b_n = 3 \cdot 4^n - a_n = 3 \cdot 2^{2n} - \left(2^{2n+1} - 1\right) = 2^{2n} + 1.$$

Assignment A5

Problem 1. In an auction at a charity gala dinner, a group of wealthy businessmen are competing with each other to be the highest bidder. Each time one of them makes a bid amount, another counter-bids by 50% more, less a service charge of ten dollars (e.g. If A bids \$1000, then B will bid \$1490). Let u_n be the amount at the nth bid and u_1 be the initial amount.

(a) Write down a recurrence relation that describes the bidding process.

(b) Show that $u_n = \$(1.5^{n-1}(u_1 - 20) + 20)$.

(c) The target amount to be raised is \$1 234 567 and the bidding stops when the bid amount meets or crosses this target amount. Given that $u_1 = 111$,

  (i) state the least number of bids required to meet this amount.

  (ii) find the winning bid amount, correct to the nearest thousand dollars.

Solution.

Part (a). $u_{n+1} = 1.5u_n - 10$.

Part (b). Let k be the constant such that $u_{n+1} + k = 1.5(u_n + k)$. Then $k = -20$. Hence, $u_{n+1} - 20 = 1.5(u_n - 20)$.

$$u_{n+1} - 20 = 1.5(u_n - 20) \implies u_n - 20 = 1.5^{n-1}(u_1 - 20) \implies u_n = 1.5^{n-1}(u_1 - 20) + 20.$$

Part (c).

Part (c)(i). Let m be the least integer such that $u_m \geq 1234567$. Consider $u_m \geq 1234567$:

$$u_m \geq 1234567 \implies 1.5^{m-1}(111 - 20) + 20 \geq 1234567.$$

Using G.C., $m \geq 24.5$. Hence, it takes at least 25 bids to meet this amount.

Part (c)(ii). Since $u_{25} = 1.5^{25-1}(111 - 20) = 1532000$ (to the nearest thousand), the winning bid is \$1 532 000.

* * * * *

Problem 2. Solve these recurrence relations together with the initial conditions.

(a) $u_{n+2} = -u_n + 2u_{n+1}$, for $n \geq 0$, $u_0 = 5$, $u_1 = -1$.

(b) $4u_n = 4u_{n-1} + u_{n-2}$, for $n \geq 2$, $u_0 = a$, $u_1 = b$, $a, b \in \mathbb{R}$.

Solution.

Part (a). Observe that the characteristic equation of u_n, $x^2 - 2x + 1 = 0$, has only one root, namely $x = 1$. Thus,

$$u_n = (A + Bn) \cdot 1^n = A + Bn.$$

Thus, u_n is in AP. Since $u_0 = 5$ and $u_1 = -1$, it follows that

$$u_n = 5 - 6n.$$

Part (b). Rewriting the given recurrence relation, we have $u_n = u_{n-1} + \frac{1}{4}u_{n-2}$. Thus, the characteristic equation is $x^2 - x - \frac{1}{4} = 0$, which has roots $\frac{1}{2}(1 \pm \sqrt{2})$. Thus,

$$u_n = A\left(\frac{1+\sqrt{2}}{2}\right)^n + B\left(\frac{1-\sqrt{2}}{2}\right)^n.$$

Since $u_0 = a$, we obviously have $A+B = a$. Since $u_1 = b$, we get $A\left(\frac{1+\sqrt{2}}{2}\right)+B\left(\frac{1-\sqrt{2}}{2}\right) = b$. Solving, we get

$$A = \frac{\sqrt{2}-1}{2\sqrt{2}}a + \frac{1}{\sqrt{2}}b, \quad B = \frac{\sqrt{2}+1}{2\sqrt{2}}a - \frac{1}{\sqrt{2}}b.$$

Thus,

$$u_n = \left(\frac{\sqrt{2}-1}{2\sqrt{2}}a + \frac{1}{\sqrt{2}}b\right)\left(\frac{1+\sqrt{2}}{2}\right)^n + \left(\frac{\sqrt{2}+1}{2\sqrt{2}}a - \frac{1}{\sqrt{2}}b\right)\left(\frac{1-\sqrt{2}}{2}\right)^n.$$

* * * * *

Problem 3. A passcode is generated using the digits 1 to 5, with repetitions allowed. The passcodes are classified into two types. A Type A passcode has an even number of the digit 1, while a Type B passcode has an odd number of the digit 1. For example, a Type A passcode is 1231, and a Type B passcode is 1541213. Let a_n and b_n denote the number of n-digit Type A and Type B passcodes respectively.

(a) State the values of a_1 and a_2.

(b) By considering the relationship between a_n and b_n, show that

$$a_n = xa_{n-1} + y^{n-1}, \qquad n \geq 2$$

where x and y are constants to be determined.

(c) Using the substitution $c_n = za_n + y^n$, where z is a constant to be determined, find a first order linear recurrence relation for c_n. Hence, find the general term formula for a_n.

Solution.

Part (a). $a_1 = 4$, $a_2 = 17$.

Part (b). Let P be an n-digit passcode with Type T, where T is either A or B. Let Type T' be the other type.

By concatenating a digit from 1 to 5 to P, five $(n+1)$-digit passcodes can be created. Let P' denote a new passcode that is created via this process. If the digit 1 is concatenated, then P' is of Type T'. If the digit 1 is not concatenated, then P' is of Type T. There are 4 choices for such a case. This hence gives the recurrence relations

$$\begin{cases} a_n = 4a_{n-1} + b_{n-1} \\ b_n = 4b_{n-1} + a_{n-1} \end{cases}$$

Adding the two equations, we see that $a_n + b_n = 5(a_{n-1} + b_{n-1})$. Thus,

$$a_n + b_n = 5^{n-1}(a_1 + b_1) = 5^{n-1}(4+1) = 5^n.$$

Hence,

$$a_n = 4a_{n-1} + b_{n-1} = 3a_{n-1} + a_{n-1} + b_{n-1} = 3a_{n-1} + 5^{n-1},$$

whence $x = 3$ and $y = 5$.

Part (c). Observe that

$$\begin{aligned}c_n &= za_n + 5^n = z\left(3a_{n-1} + 5^{n-1}\right) + 5^n = 3\left(za_{n-1} + 5^{n-1}\right) + (2+z)5^{n-1} \\ &= 3c_{n-1} + (2+z)5^{n-1}.\end{aligned}$$

Let $z = -2$. Then,

$$c_n = 3c_{n-1} = 3^{n-1}c_1 = 3^{n-1}\left(-2a_1 + 5\right) = -3^n.$$

Note that $a_n = \frac{1}{z}\left(c_n - y^n\right)$. Thus,

$$a_n = \frac{-3^n - 5^n}{-2} = \frac{3^n + 5^n}{2}.$$

A6 Polar Coordinates

Tutorial A6

Problem 1.

(a) Find the rectangular coordinates of the following points.

(i) $(3, -\frac{\pi}{4})$

(ii) $(1, \pi)$

(iii) $(\frac{1}{2}, \frac{3}{2}\pi)$

(b) Find the polar coordinates of the following points.

(i) $(3, 3)$

(ii) $(-1, -\sqrt{3})$

(iii) $(2, 0)$

(iv) $(4, 2)$

Solution.

Part (a).

Part (a)(i). Note that $r = 3$ and $\theta = -\frac{\pi}{4}$. This gives

$$x = r\cos\theta = \frac{3}{\sqrt{2}}, \quad y = r\sin\theta = -\frac{3}{\sqrt{2}}.$$

Hence, the rectangular coordinate of the point is $(3/\sqrt{2}, -3\sqrt{2})$.

Part (a)(ii). Note that $r = 1$ and $\theta = \pi$. This gives

$$x = r\cos\theta = -1, \quad y = r\sin\theta = 0.$$

Hence, the rectangular coordinate of the point is $(-1, 0)$.

Part (a)(iii). Note that $r = \frac{1}{2}$ and $\theta = \frac{3}{2}\pi$. This gives

$$x = \rho\cos\theta = 0, \quad y = r\sin\theta = -\frac{1}{2}.$$

Hence, the rectangular coordinate of the point is $(0, -1/2)$.

Part (b).

Part (b)(i). Note that $x = 3$ and $y = -3$. This gives

$$r^2 = x^2 + y^2 \implies r = 3\sqrt{2}, \quad \tan\theta = \frac{y}{x} \implies \theta = -\frac{\pi}{4}.$$

Hence, the polar coordinate of the point is $(3\sqrt{2}, -\pi/4)$.

Part (b)(ii). Note that $x = -1$ and $y = -\sqrt{3}$. This gives

$$r^2 = x^2 + y^2 \implies r = 2, \quad \tan\theta = \frac{y}{x} \implies \theta = \frac{\pi}{3}.$$

Hence, the polar coordinate of the point is $(2, \pi/3)$.

Part (b)(iii). Note that $x = 2$ and $y = 0$. This gives

$$r^2 = x^2 + y^2 \implies r = 2, \quad \tan\theta = \frac{y}{x} \implies \theta = 0.$$

Hence, the polar coordinate of the point is $(2, 0)$.

Part (b)(iv). Note that $x = 4$ and $y = 2$. This gives

$$r^2 = x^2 + y^2 \implies r = 2\sqrt{5}, \quad \tan\theta = \frac{y}{x} \implies \theta = \arctan\frac{1}{2}.$$

Hence, the polar coordinate of the point is $(2\sqrt{5}, \arctan(1/2))$.

* * * * *

Problem 2. Rewrite the following equations in polar form.

(a) $2x^2 + 3y^2 = 4$

(b) $y = 2x^2$

Solution.

Part (a).

$$2x^2 + 3y^2 = 2(r\cos\theta)^2 + 3(r\sin\theta)^2 = 4 \implies r^2 = \frac{4}{2\cos^2\theta + 3\sin^2\theta} = \frac{4}{2 + \sin^2\theta}.$$

Part (b).

$$y = 2x^2 \implies \frac{y}{x} = 2x \implies \tan\theta = 2r\cos\theta \implies r = \frac{1}{2}\tan\theta\sec\theta.$$

* * * * *

Problem 3. Rewrite the following equations in rectangular form.

(a) $r = \frac{1}{1-2\cos\theta}$

(b) $r = \sin\theta$

Solution.

Part (a).

$$r = \frac{1}{1 - 2\cos\theta} \implies r - 2r\cos\theta = 1 \implies r = 2x + 1 \implies r^2 = 4x^2 + 4x + 1$$
$$\implies x^2 + y^2 = 4x^2 + 4x + 1 \implies y^2 = 3x^2 + 4x + 1.$$

Part (b).

$$r = \sin\theta \implies r^2 = r\sin\theta \implies x^2 + y^2 = y.$$

Problem 4.

(a) Show that the curve with polar equation $r = 3a\cos\theta$, where a is a positive constant, is a circle. Write down its centre and radius.

(b) By finding the Cartesian equation, sketch the curve whose polar equation is $r = a\sec\left(\theta - \frac{\pi}{4}\right)$, where a is a positive constant.

Solution.

Part (a).

$$r = 3a\cos\theta \implies r^2 = 3ar\cos\theta \implies x^2 + y^2 = 3ax \implies x^2 - 3ax + y^2 = 0.$$

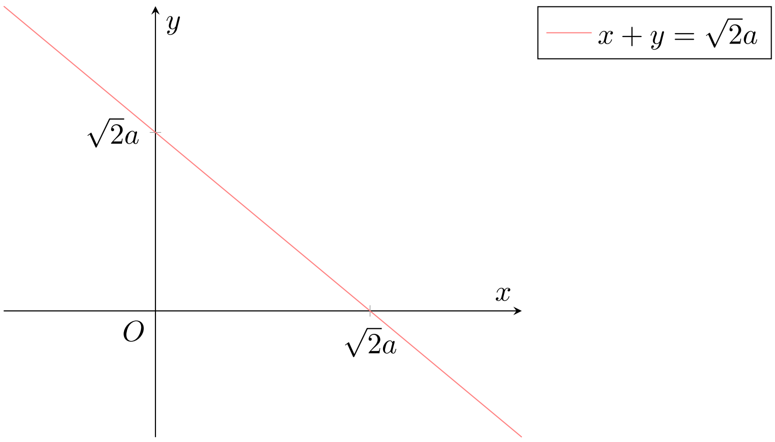
Completing the square, we get

$$\left(x - \frac{3a}{2}\right)^2 + y^2 \left(\frac{3a}{2}\right)^2.$$

Thus, the circle has centre $(3a/2, 0)$ and radius $3a/2$.

Part (b).

$$r = a\sec\left(\theta - \frac{\pi}{4}\right) \implies r\cos\left(\theta - \frac{\pi}{4}\right) = a \implies r\left(\cos\theta + \sin\theta\right) = \sqrt{2}a \implies x + y = \sqrt{2}a.$$



Problem 5. Sketch the following polar curves, where r is non-negative and $0 \le \theta \le 2\pi$.

(a) $r = 2$

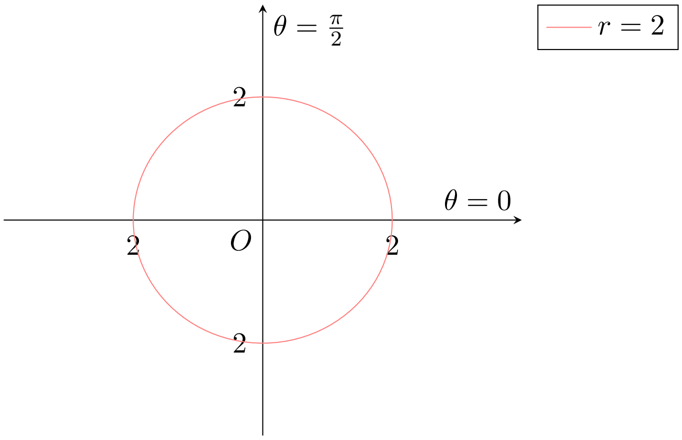
(b) $\theta = \frac{\pi}{4}$

(c) $r = \frac{1}{2}\theta$

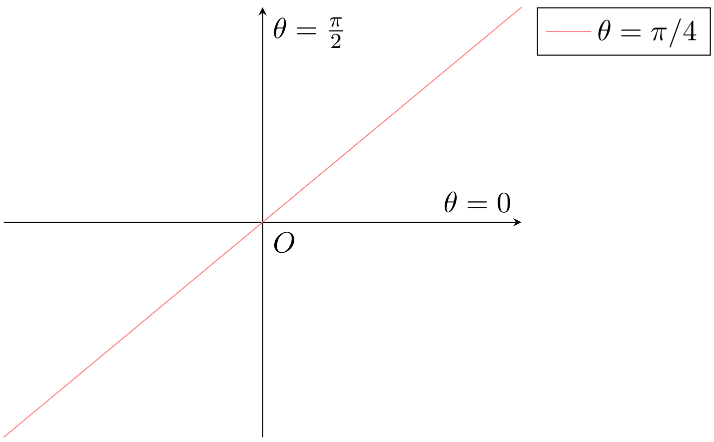
(d) $r = 2\csc\theta$

Solution.

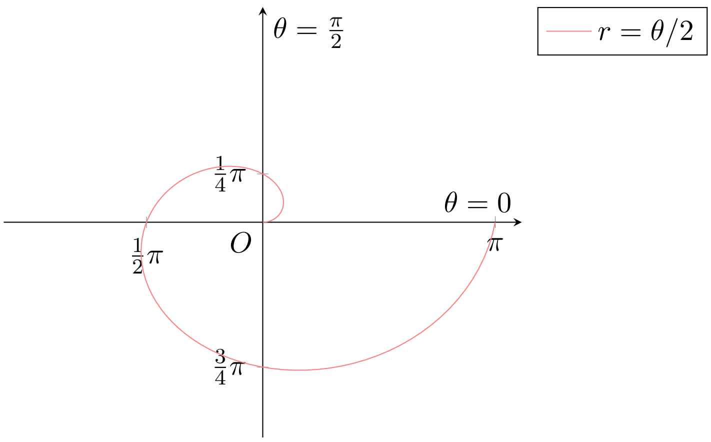
Part (a).



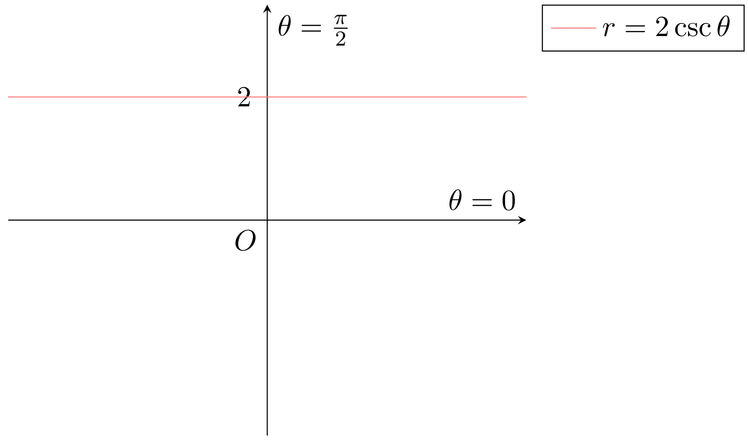
Part (b).



Part (c).

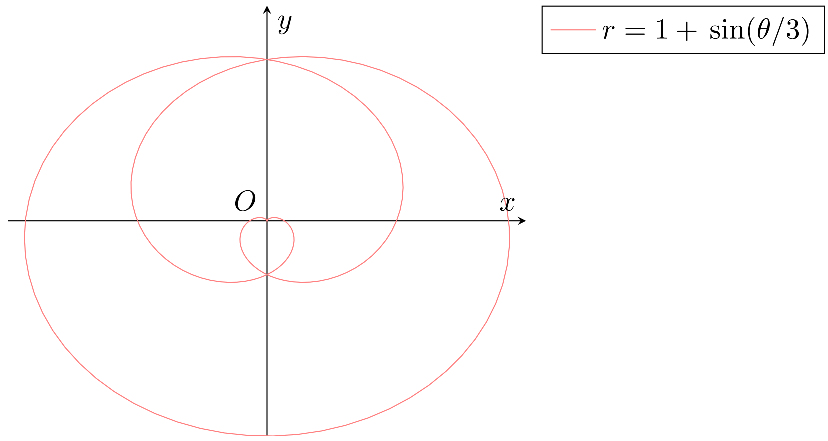


Part (d).



* * * * *

Problem 6. A sketch of the curve $r = 1 + \sin\frac{\theta}{3}$ is shown. Copy the diagram and indicate the x- and y-intercepts.



Solution. Observe that the curve is symmetric about the y-axis. Also observe that $\frac{\theta}{3} \in [0, 2\pi)$, hence we take $\theta \in [0, 6\pi)$.

For x-intercepts, $y = r\sin\theta = 0 \implies \theta = n\pi$, where $n \in \mathbb{Z}$. Due to the symmetry of the curve, we consider only $n = 0, 2, 4$.

Case 1. $n = 0 \implies r = 1 + \sin\frac{0}{3}\pi = 1$.

Case 2. $n = 2 \implies r = 1 + \sin\frac{2}{3}\pi = 1 + \frac{\sqrt{3}}{2}$.

Case 3. $n = 4 \implies r = 1 + \sin\frac{4}{3}\pi = 1 - \frac{\sqrt{3}}{2}$.

Hence, the curve intersects the x-axis at $x = 1, 1 + \frac{\sqrt{3}}{2}, 1 - \frac{\sqrt{3}}{2}$. Correspondingly, the curve also intersects the x-axis at $x = -1, -1 - \frac{\sqrt{3}}{2}, -1 + \frac{\sqrt{3}}{2}$.

For y-intercepts, $x = r\cos\theta = 0 \implies \theta = (n + \frac{1}{2})\pi$, where $n \in \mathbb{Z}$. Due to the restriction on θ, we consider $n \in [0, 5)$.

Case 1. $n = 0, r = 1 + \sin\frac{1/2}{3}\pi = \frac{3}{2}$.

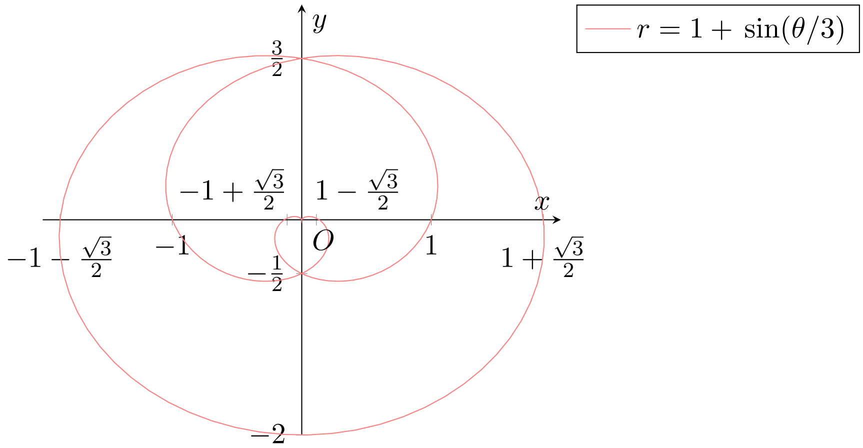
Case 2. $n = 1, r = 1 + \sin\frac{3/2}{3}\pi = 2$.

Case 3. $n = 2, r = 1 + \sin\frac{5/2}{3}\pi = \frac{3}{2}$.

Case 4. $n = 3, r = 1 + \sin\frac{7/2}{3}\pi = \frac{1}{2}$.

Case 5. $n = 4, r = 1 + \sin\frac{9/2}{3}\pi = 0$.

Hence, the curve intersects the y-axis at $y = -2, -\frac{1}{2}, \frac{3}{2}$.



* * * * *

Problem 7.

(a) A graph has polar equation $r = \frac{2}{\cos\theta\sin\alpha - \sin\theta\cos\alpha}$, where α is a constant. Express the equation in Cartesian form. Hence, sketch the graph in the case $\alpha = \frac{\pi}{4}$, giving the Cartesian coordinates of the intersection with the axes.

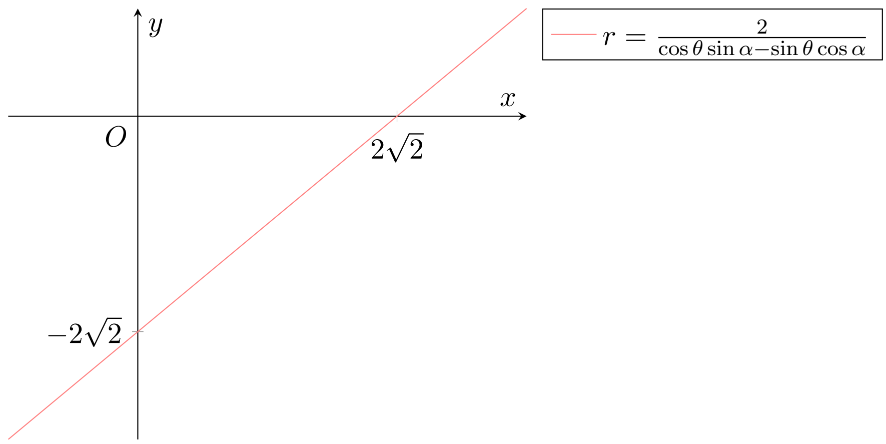
(b) A graph has Cartesian equation $(x^2 + y^2)^2 = 4x^2$. Express the equation in polar form. Hence, or otherwise, sketch the graph.

Solution.

Part (a).

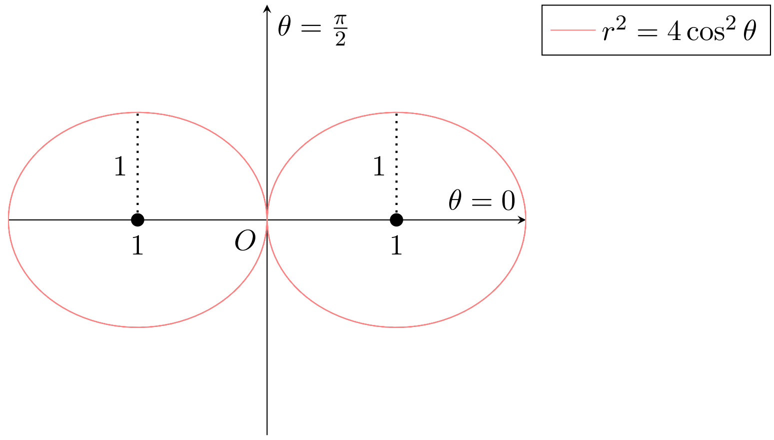
$$r = \frac{2}{\cos\theta\sin\alpha - \sin\theta\cos\alpha} \implies r\cos\theta\sin\alpha - r\sin\theta\cos\alpha = x\sin\alpha - y\cos\alpha = 2$$
$$\implies y = x\tan\alpha - 2\sec\alpha.$$

When $\alpha = \frac{\pi}{4}$, we have $y = x - 2\sqrt{2}$.



Part (b).

$$(x^2+y^2)^2 = 4x^2 \implies \left(r^2\right)^2 = 4(r\cos\theta)^2 \implies r^4 = 4r^2\cos^2\theta \implies r^2 = 4\cos^2\theta.$$

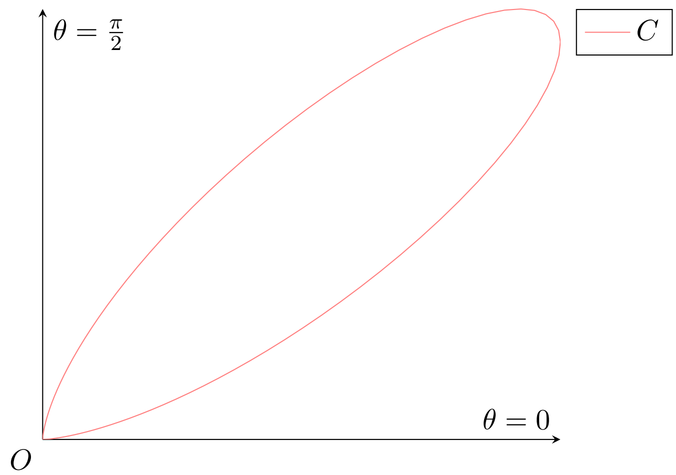


* * * * *

Problem 8. Find the polar equation of the curve C with equation $x^5 + y^5 = 5bx^2y^2$, where b is a positive constant. Sketch the part of the curve C where $0 \leq \theta \leq \pi/2$.

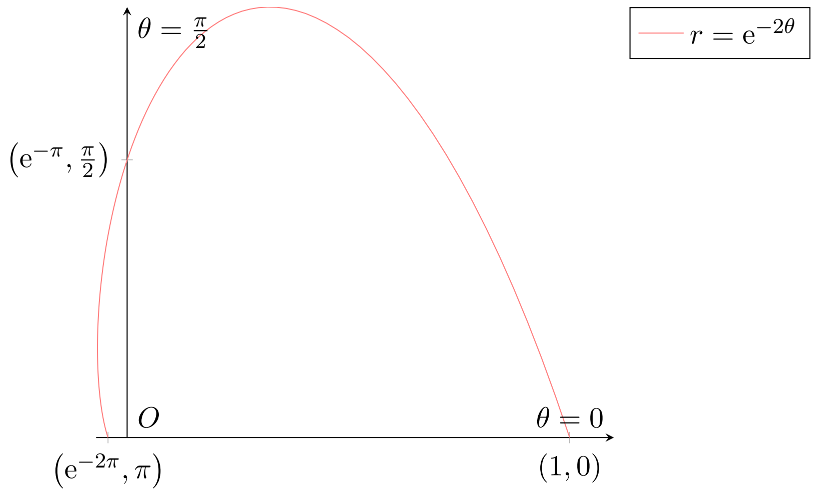
Solution.

$$x^5 + y^5 = 5bx^2y^2 \implies (r\cos\theta)^5 + (r\sin\theta)^5 = 5b(r\cos\theta)^2(r\sin\theta)^2$$
$$\implies r\left(\cos^5\theta + \sin^5\theta\right) = 5b\cos^2\theta\sin^2\theta \implies r = \frac{5b\cos^2\theta\sin^2\theta}{\cos^5\theta + \sin^5\theta}.$$



Problem 9. The equation of a curve, in polar coordinates, is $r = \mathrm{e}^{-2\theta}$, for $0 \le \theta \le \pi$. Sketch the curve, indicating clearly the polar coordinates of any axial intercepts.

Solution.



* * * * *

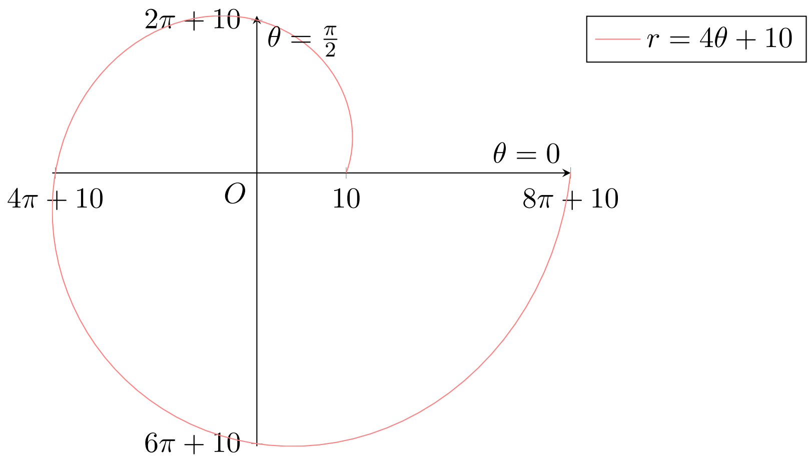
Problem 10. Suppose that a long thin rod with one end fixed at the pole of a polar coordinate system rotates counter-clockwise at the constant rate of 0.5 rad/sec. At time $t = 0$, a bug on the rod is 10 mm from the pole and is moving outward along the rod at a constant speed of 2 mm/sec. Find an equation of the form $r = f(\theta)$ for the part of motion of the bug, assuming that $\theta = 0$ when $t = 0$. Sketch the path of the bug on the polar coordinate system for $0 \le t \le 4\pi$.

Solution. Let $\theta(t)$ and $r(t)$ be functions of time, with $\theta(0) = 0$ and $r(0) = 10$. We know that $\mathrm{d}\theta/\mathrm{d}t = 0.5$ and $\mathrm{d}r/\mathrm{d}t = 2$. Hence,

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}t} \cdot \frac{\mathrm{d}t}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}t} \cdot \left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^{-1} = 2 \cdot (0.5)^{-1} = 4.$$

Thus, $r = 4\theta + r(0) = 4\theta + 10$.

Since $\mathrm{d}\theta/\mathrm{d}t = 0.5$ and $\theta(0) = 0$, we have $\theta = 0.5t$. Hence, $0 \le t \le 4\pi \implies 0 \le \theta \le 2\pi$.



Problem 11. The equation, in polar coordinates, of a curve C is $r = ae^{\frac{1}{2}\theta}$, $0 \le \theta \le 2\pi$, where a is a positive constant. Write down, in terms of θ, the Cartesian coordinates, x and y, of a general point P on the curve. Show that the gradient at P is given by $\frac{dy}{dx} = \frac{\tan\theta+2}{1-2\tan\theta}$.

Hence, show that the tangent at P is inclined to $\overrightarrow{OP}$ at a constant angle α, where $\tan\alpha = 2$. Sketch the curve C.

Solution. Note that $x = r\cos\theta$ and $y = r\sin\theta$, whence $x = ae^{\frac{1}{2}\theta}\cos\theta$ and $y = ae^{\frac{1}{2}\theta}\sin\theta$. Hence, $P\left(ae^{\frac{1}{2}\theta}\cos\theta, ae^{\frac{1}{2}\theta}\sin\theta\right)$.

Observe that $\frac{dr}{d\theta} = \frac{1}{2}ae^{\frac{1}{2}\theta} = \frac{1}{2}r$. Hence,

$$\frac{dy}{dx} = \frac{\frac{dr}{d\theta}\sin\theta + r\cos\theta}{\frac{dr}{d\theta}\cos\theta - r\sin\theta} = \frac{\frac{1}{2}r\sin\theta + r\cos\theta}{\frac{1}{2}r\cos\theta - r\sin\theta} = \frac{\sin\theta + 2\cos\theta}{\cos\theta - 2\sin\theta} = \frac{\tan\theta + 2}{1 - 2\tan\theta}.$$

Let $\mathbf{t} = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$ represent the direction of the tangent line. Then

$$\mathbf{t} = \begin{pmatrix} 1 \\ dy/dx \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{\tan\theta+2}{1-2\tan\theta} \end{pmatrix} = \frac{1}{1-2\tan\theta}\begin{pmatrix} 1-2\tan\theta \\ \tan\theta+2 \end{pmatrix}$$

and

$$\overrightarrow{OP} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ae^{\frac{1}{2}\theta}\cos\theta \\ ae^{\frac{1}{2}\theta}\sin\theta \end{pmatrix} = ae^{\frac{1}{2}\theta}\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}.$$

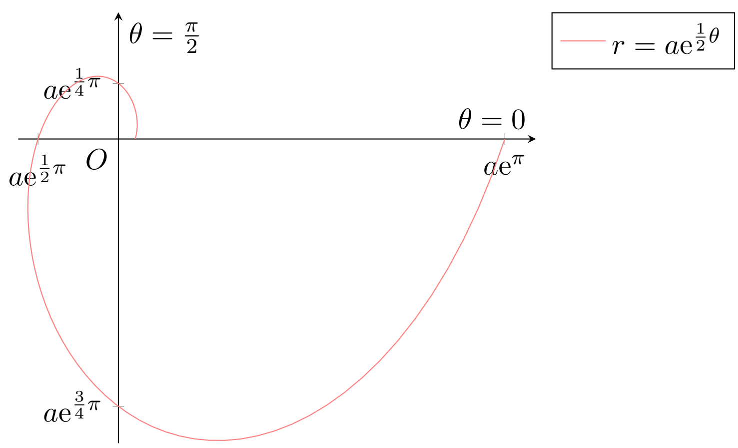
By the definition of the dot-product, we have $\mathbf{t}\cdot\overrightarrow{OP} = |\mathbf{t}|\left|\overrightarrow{OP}\right|\cos\alpha$, whence

$$\begin{aligned}\cos\alpha = \frac{\mathbf{t}\cdot\overrightarrow{OP}}{|\mathbf{t}|\left|\overrightarrow{OP}\right|} &= \frac{(1-2\tan\theta)\cos\theta + (\tan\theta+2)\sin\theta}{\sqrt{(1-2\tan\theta)^2 + (\tan\theta+2)^2}\cdot\sqrt{\cos^2\theta + \sin^2\theta}} \\ &= \frac{\cos\theta + \tan\theta\sin\theta}{\sqrt{5\tan^2\theta + 5}} = \frac{\cos^2\theta + \sin^2\theta}{\sqrt{5\sin^2\theta + 5\cos^2\theta}} = \frac{1}{\sqrt{5}}.\end{aligned}$$

Thus, $\alpha = \arccos\frac{1}{\sqrt{5}}$. Since $\tan(\arccos x) = \frac{\sqrt{1-x^2}}{x}$,

$$\tan\alpha = \tan\left(\arccos\frac{1}{\sqrt{5}}\right) = \frac{\sqrt{1-\left(1/\sqrt{5}\right)^2}}{1/\sqrt{5}} = 2.$$

Hence, the tangent at P is inclined to $\overrightarrow{OP}$ at a constant angle α, where $\tan\alpha = 2$.



Problem 12. The polar equation of a curve is given by $r = \mathrm{e}^{\theta}$ where $0 \leq \theta \leq \frac{\pi}{2}$. Cartesian axes are taken at the pole O. Express x and y in terms of θ and hence find the Cartesian equation of the tangent at $\left(\mathrm{e}^{\frac{\pi}{2}}, \frac{\pi}{2}\right)$.

Solution. Recall that $x = r\cos\theta$ and $y = r\sin\theta$, whence $x = \mathrm{e}^{\theta}\cos\theta$ and $y = \mathrm{e}^{\theta}\sin\theta$. Thus, $\frac{\mathrm{d}x}{\mathrm{d}\theta} = \mathrm{e}^{\theta}(\cos\theta - \sin\theta)$, and $\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^{\theta}(\cos\theta + \sin\theta)$. Hence,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = \frac{\mathrm{e}^{\theta}(\cos\theta + \sin\theta)}{\mathrm{e}^{\theta}(\cos\theta - \sin\theta)} = \frac{\cos\theta + \sin\theta}{\cos\theta - \sin\theta}.$$

At $\left(\mathrm{e}^{\frac{\pi}{2}}, \frac{\pi}{2}\right)$, we clearly have $x = 0$ and $y = \mathrm{e}^{\pi/2}$. Also, $\mathrm{d}y/\mathrm{d}x = -1$. By the point-slope formula, the equation of the tangent line at $\left(\mathrm{e}^{\frac{\pi}{2}}, \frac{\pi}{2}\right)$ is given by $y = -x + \mathrm{e}^{\frac{\pi}{2}}$.

* * * * *

Problem 13. A curve C has polar equation $r = a\cot\theta$, $0 < \theta \leq \pi$, where a is a positive constant.

(a) Show that $y = a$ is an asymptote of C.

(b) Find the tangent at the pole.

Hence, sketch C and find the Cartesian equation of C in the form $y^2(x^2 + y^2) = bx^2$, where b is a constant to be determined.

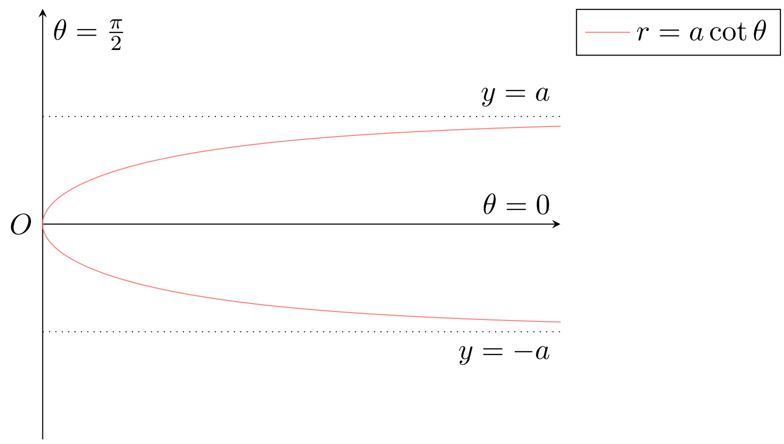
Solution.

Part (a). Note that

$$r = a\cot\theta \implies y = r\sin\theta = a\cos\theta.$$

As $\theta \to 0$, $r \to \infty$. Hence, there is an asymptote at $\theta = 0$. Since $\cos\theta = 1$ when $\theta = 0$, the line $y = a\cos\theta = a$ is an asymptote of C.

Part (b). For tangents at the pole, $r = 0 \implies \cot\theta = 0 \implies \theta = \frac{\pi}{2}$.



Note that

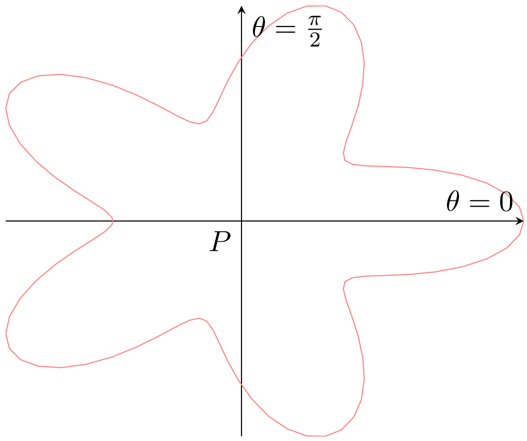
$$r = a\cot\theta = a\left(\frac{r\cos\theta}{r\sin\theta}\right) = a\left(\frac{x}{y}\right).$$

Thus,

$$x^2 + y^2 = r^2 = a^2\left(\frac{x^2}{y^2}\right) \implies y^2\left(x^2 + y^2\right) = a^2x^2,$$

whence $b = a^2$.

Problem 14.



Relative to the pole P and the initial line $\theta = 0$, the polar equation of the curve shown is either

i. $r = a + b \sin n\theta$, or

ii. $r = a + b \cos n\theta$

where a, b and n are positive constants. State, with a reason, whether the equation is (i) or (ii) and state the value of n.

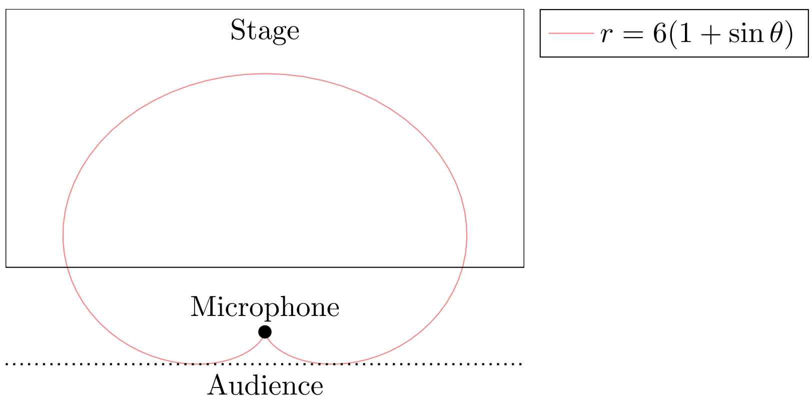
The maximum value of r is $\frac{11}{2}$ and the minimum value of r is $\frac{5}{2}$. Find the values of a and b.

Solution. Since the curve is symmetrical about the horizontal half-line $\theta = 0$, the polar equation of the curve is a function of $\cos n\theta$ only. Hence, the polar equation of the curve is $r = a + b \cos n\theta$, with $n = 5$.

Observe that the maximum value of r is achieved when $\cos 5\theta = 1$, whence $r = a + b$. Thus, $a + b = \frac{11}{2}$. Also observe that the minimum value of r is achieved when $\cos 5\theta = -1$, whence $r = a - b$. Thus, $a - b = \frac{5}{2}$. Solving, we get $a = 4$ and $b = \frac{3}{2}$.

* * * * *

Problem 15.



Sound engineers often use a microphone with a cardioid acoustic pickup pattern to record live performances because it reduces pickup from the audience. Suppose a cardioid microphone is placed 3 metres from the front of the stage, and the boundary of the optimal pickup region is given by the cardioid with polar equation

$$r = 6(1 + \sin\theta)$$

where r is measured in metres and the microphone is at the pole.

Find the minimum distance from the front of the stage the first row of the audience can be seated such that the microphone does not pick up noise from the audience.

Solution. Note that $r = 6(1+\sin\theta) = 6(1+\frac{y}{r})$, whence $r^2 = 6r + 6y$. Thus,

$$r^2 - 6r - 6y = 0 \implies r = 3 \pm \sqrt{9+6y} \implies 9 + 6y = (r-3)^2.$$

Since $9 + 6y = (r-3)^2 \geq 0$, we have $y \geq -1.5$. Thus, the furthest distance the audience has to be from the stage is $|-1.5| + 3 = 4.5$ m.

* * * * *

Problem 16. To design a flower pendant, a designer starts off with a curve C_1, given by the Cartesian equation

$$\left(x^2+y^2\right)^2 = a^2\left(3x^2-y^2\right)$$

where a is a positive constant.

(a) Show that a corresponding polar equation of C_1 is $r^2 = a^2(1+2\cos 2\theta)$.

(b) Find the equations of the tangents to C_1 at the pole.

Another curve C_2 is obtained by rotating C_1 anti-clockwise about the origin by $\frac{\pi}{3}$ radians.

(c) State a polar equation of C_2.

(d) Sketch C_1 and C_2 on the same diagram, stating clearly the exact polar coordinates of the points of intersection of the curves with the axes. Find also the exact polar coordinates of the points of intersection with C_1 and C_2.

The curve C_3 is obtained by reflecting C_2 in the line $\theta = \frac{\pi}{2}$.

(e) State a polar equation of C_3.

(f) The designer wishes to enclose the 3 curves inside a circle given by the polar equation $r = r_1$. State the minimum value of r_1 in terms of a.

Solution.

Part (a). Observe that $\left(x^2+y^2\right)^2 = r^4$ and $3x^2 - y^2 = r^2\left(3\cos^2\theta - \sin^2\theta\right)$. Hence,

$$\left(x^2+y^2\right)^2 = a^2\left(3x^2-y^2\right) \implies r^2 = a^2\left(3\cos^2\theta - \sin^2\theta\right).$$

Note that

$$3\cos^2\theta - \sin^2\theta = 1 + 2\cos^2\theta - 2\sin^2\theta = 1 + 2\cos 2\theta.$$

Thus,

$$r^2 = a^2\left(1+2\cos 2\theta\right).$$

Part (b). For tangents at the pole,

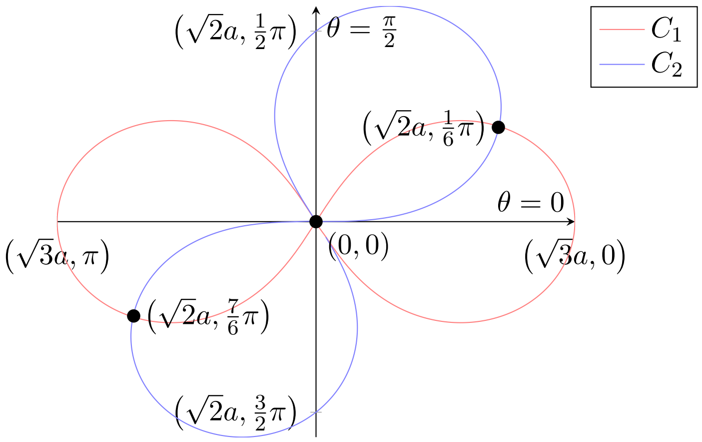
$$r = 0 \implies 1 + 2\cos 2\theta = 0 \implies \cos 2\theta = -\frac{1}{2}.$$

Since $0 \leq 2\theta \leq 2\pi$, we have $\theta = \pi/3, 2\pi/3$. For full lines, we also have $\theta = 4\pi/3$ and $\theta = 5\pi/3$.

Part (c).

$$r^2 = a^2\left[1+2\cos\left(2\left(\theta-\frac{\pi}{3}\right)\right)\right] = a^2\left[1+2\cos\left(2\theta-\frac{2}{3}\pi\right)\right].$$

Part (d).



Consider the horizontal intercepts of C_1. When $\theta = 0$, $r = \sqrt{3}a$. Hence, by symmetry, C_1 intercepts the horizontal axis at $\left(\sqrt{3}a, 0\right)$ and $\left(\sqrt{3}a, \pi\right)$.

Consider the vertical intercepts of C_2. When $\theta = \pi/2$, $r = \sqrt{2}a$. Hence, by symmetry, C_2 intercepts the vertical axis at $\left(\sqrt{2}a, \pi/2\right)$ and $\left(\sqrt{2}a, 3\pi/2\right)$.

Now consider the intersections between C_1 and C_2. By symmetry, it is obvious that the points of intersections must lie along the half-lines $\pi/6$ and $7\pi/6$, or along the half-lines $4\pi/6$ and $10\pi/6$. By symmetry, we consider only the half-lines $\pi/6$ and $4\pi/6$.

Case 1: $\theta = \pi/6$. Substituting $\theta = \pi/6$ into the equation of C_1, we obtain $r = \sqrt{2}a$. Hence, C_1 and C_2 intersect at $\left(\sqrt{2}a, \pi/6\right)$ and, by symmetry, at $\left(\sqrt{2}a, 7\pi/6\right)$.

Case 2.$\theta = 4\pi/6$ Substituting $\theta = 4\pi/6$ into the equation of C_1, we obtain $r = 0$. Hence, C_1 and C_2 intersect at $(0,0)$.

Part (e). Reflecting about the line $\theta = \pi/2$ is equivalent to applying the map $\theta \mapsto \theta + \pi/3$ to C_1. Hence,

$$r^2 = a^2\left[1+2\cos\left(2\left(\theta+\frac{1}{3}\pi\right)\right)\right] = a^2\left[1+2\cos\left(2\theta+\frac{2}{3}\pi\right)\right].$$

Part (f). $r_1 = \sqrt{3}a$.

Self-Practice A6

Problem 1. A curve C has equation, in polar coordinates, $r = a\sqrt{\left(4+\sin^2\theta\right)\cos\theta}$, $-\frac{1}{2}\pi \le \theta \le \frac{1}{2}\pi$, where a is a positive constant.

(a) Show that $\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\left(4+\sin^2\theta\right)\cos\theta\right] = -\left(2+3\sin^2\theta\right)\sin\theta$. Hence, state, with a reason, whether r increases or decreases as θ increases, for $0 < \theta \le \frac{1}{2}\pi$.

(b) Sketch the curve C.

(c) Find the Cartesian equation of C in the form $\left(x^2+y^2\right)^m = a^2x\left(bx^2+cy^2\right)$, giving the numerical values of m, b and c.

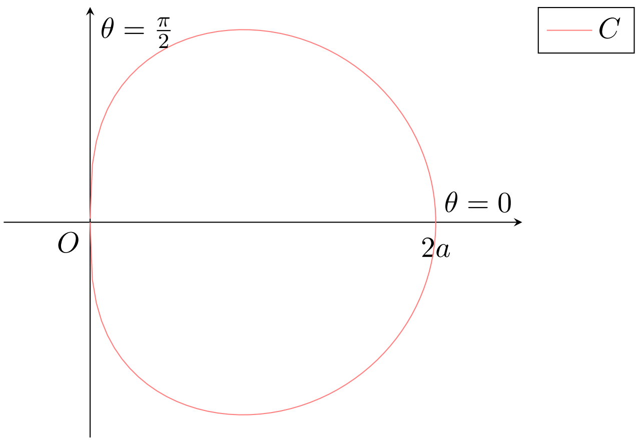
Solution.

Part (a).

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\left(4+\sin^2\theta\right)\cos\theta\right] &= -\left(4+\sin^2\theta\right)\sin\theta + 2\sin\theta\cos^2\theta \\ &= -\sin\theta\left(\sin^2\theta - 2\cos^2\theta + 4\right) \\ &= -\sin\theta\left[\sin^2\theta - 2\left(1-\sin^2\theta\right) + 4\right] \\ &= -\sin\theta\left(3\sin^2\theta + 2\right).\end{aligned}$$

For $t \in (0, \pi/2]$, we have $\sin\theta > 0$ and $3\sin^2\theta + 2 > 0$. Hence, r is decreasing.

Part (b).



Part (c). Squaring, we have

$$r^2 = a^2\left(4+\sin^2\theta\right)\cos\theta.$$

Recall that $x = r\cos\theta$ and $y = r\sin\theta$, so

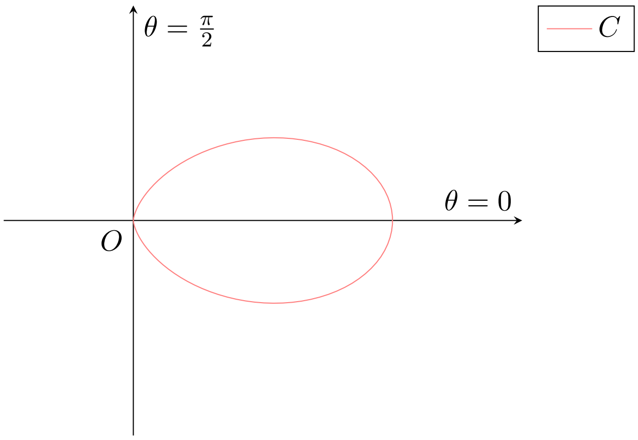
$$r^2 = a^2\left[4+\left(\frac{y}{r}\right)^2\right]\left(\frac{x}{r}\right) \implies r^5 = a^2x\left(4r^2+y^2\right).$$

Since $x^2+y^2 = r^2$, we get

$$\left(x^2+y^2\right)^{5/2} = a^2x\left(4x^2+5y^2\right),$$

whence $m = 5/2$, $b = 4$ and $c = 5$.

Problem 2. The diagram shows a sketch of the curve C with polar equation $r = a\cos^2\theta$, where a is a positive constant and $-\frac{1}{2}\pi \le \theta \le \frac{1}{2}\pi$.



(a) Explain briefly about how you can tell from this form of the equation that C is symmetrical about the line $\theta = 0$ and that the tangent to C at the pole O is perpendicular to the line $\theta = 0$.

(b) Show that the equation of C in Cartesian coordinates may be expressed in the form $y^2 = a^{2/3}x^{4/3} - x^2$.

Solution.

Part (a). Observe that

$$a\cos^2\theta = a\cos^2(-\theta)\,.$$

Hence, C is invariant under the transformation $\theta \mapsto -\theta$, whence it is symmetrical about the line $\theta = 0$.

For tangents to the pole, we have $r = 0$. Since $a > 0$, we require $\cos\theta = 0$, whence $\theta = \pm\pi/2$, which are clearly perpendicular to the line $\theta = 0$.

Part (b). We have

$$r = a\cos^2\theta = a\left(\frac{x}{r}\right)^2 \implies r^3 = ax^2.$$

Hence,

$$x^2 + y^2 = r^2 = \left(ax^2\right)^{2/3} \implies y^2 = a^{2/3}x^{4/3} - x^2.$$

* * * * *

Problem 3. The equation of curve C is given in polar coordinates by $r = 1 + \sin 2\theta$, $0 \le \theta \le 2\pi$.

(a) Prove that C is symmetric about the pole.

(b) Sketch C and any tangents to C at the pole. Label any points of intersection with the axes, and show clearly the symmetries and curvature near the pole.

(c) Determine whether each loop of C is a circle. Justify your answer.

(d) Show that the Cartesian equation of C is $\left(x^2 + y^2\right)^3 = (x + y)^4$.

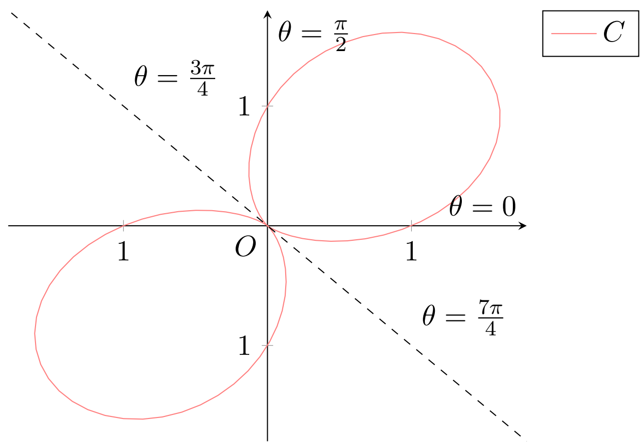
Solution.

Part (a). Observe that

$$1 + \sin 2\theta = 1 + \sin(2\theta + 2\pi) = 1 + \sin(2(\theta + \pi)).$$

Hence, C is invariant under the transformation $\theta \mapsto \theta + \pi$, whence C is symmetric about the pole.

Part (b).



Part (c). Consider the top-right loop. r attains a maximum of 2 when $\theta = \pi/4$. Suppose the loop is a circle (with radius 1). Then the centre should be $(1, \pi/4)$, which is $(1/\sqrt{2}, 1\sqrt{2})$ in Cartesian coordinates. The distance between $(1/\sqrt{2}, 1/\sqrt{2})$ and $(1, 0)$ is given by

$$\sqrt{\left(\frac{1}{\sqrt{2}} - 1\right)^2 + \left(\frac{1}{\sqrt{2}} - 0\right)^2} = \sqrt{2 - \sqrt{2}} \neq 1.$$

Hence, the loop is not a circle.

Part (d). We have

$$r = 1 + \sin 2\theta = 1 + 2\cos\theta \sin\theta = 1 + 2\left(\frac{x}{r}\right)\left(\frac{y}{r}\right).$$

Thus,

$$r^3 = r^2 + 2xy \implies \left(x^2 + y^2\right)^{3/2} = x^2 + y^2 + 2xy = (x + y)^2.$$

Squaring both sides yields the desired equation:

$$\left(x^2 + y^2\right)^3 = (x + y)^4.$$

* * * * *

Problem 4 (🌶). Prove that at all points of intersection of the polar curves with equations $r = a(1 + \cos\theta)$ and $r = b(1 - \cos\theta)$, the tangent lines are perpendicular.

Solution. Consider the gradient of C_1. Firstly, we have

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}\theta}\cos\theta - r\sin\theta = -a\sin\theta - 2a\sin\theta\cos\theta = -a\left(\sin\theta + \sin 2\theta\right).$$

Next, we have

$$\frac{\mathrm{d}y}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}t}\sin\theta + r\cos\theta = a\cos\theta - a\cos^2\theta + a\sin^2\theta = a\left(\cos\theta - \cos 2\theta\right).$$

Thus,
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = -\left(\frac{\cos\theta - \cos 2\theta}{\sin\theta + \sin 2\theta}\right).$$

Consider the gradient of C_2. Firstly, we have
$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}\theta}\cos\theta - r\sin\theta = -b\sin\theta + 2b\cos\theta\sin\theta = b\left(\sin 2\theta - \sin\theta\right).$$

Next, we have
$$\frac{\mathrm{d}y}{\mathrm{d}\theta} = \frac{\mathrm{d}r}{\mathrm{d}t}\sin\theta + r\cos\theta = b\cos\theta - b\cos^2\theta + b\sin^2\theta = b\left(\cos\theta - \cos 2\theta\right).$$

Thus,
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = \frac{\cos\theta - \cos 2\theta}{\sin 2\theta - \sin\theta}.$$

Consider the product of the gradients:
$$-\left(\frac{\cos\theta - \cos 2\theta}{\sin\theta + \sin 2\theta}\right)\left(\frac{\cos\theta - \cos 2t}{\sin 2\theta - \sin\theta}\right) = -\frac{\cos^2\theta - \cos^2 2\theta}{\sin^2 2\theta - \sin^2\theta}.$$

Observe that
$$\cos^2\theta - \cos^2 2\theta = \cos^2\theta - \left(2\cos^2\theta - 1\right)^2 = -4\cos^4\theta + 5\cos^2\theta - 1.$$

Also observe that
$$\begin{aligned}\sin^2 2\theta - \sin^2\theta &= 4\sin^2\theta\cos^2\theta - \sin^2\theta\\ &= 4\left(1-\cos^2\theta\right)\cos^2\theta - \left(1-\cos^2\theta\right)\\ &= -4\cos^4\theta + 5\cos^2\theta - 1.\end{aligned}$$

Hence, the product of the gradients is
$$-\frac{\cos^2\theta - \cos^2 2\theta}{\sin^2 2\theta - \sin^2\theta} = -\frac{-4\cos^4\theta + 5\cos^2\theta - 1}{-4\cos^4\theta + 5\cos^2\theta - 1} = -1.$$

Thus, for any given θ, the tangents of C_1 and C_2 are perpendicular. This immediately implies that the tangent lines at all intersection points of C_1 and C_2 are perpendicular.

Assignment A6

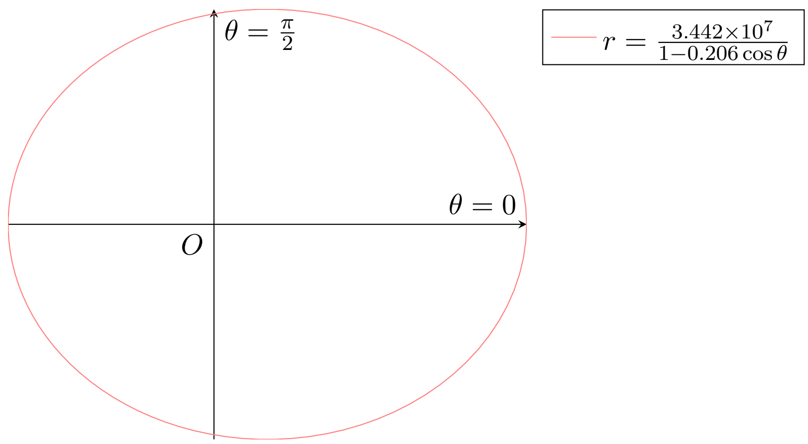
Problem 1. The planet Mercury travels around the sun in an elliptical orbit given approximately by

$$r = \frac{3.442 \times 10^7}{1 - 0.206\cos\theta},$$

where r is measured in miles and the sun is at the pole.

Sketch the orbit and find the distance from Mercury to the sun at the aphelion (the greatest distance from the sun) and at the perihelion (the shortest distance from the sun).

Solution.



Observe that r attains a maximum when $\cos\theta$ is also at its maximum. Since the maximum value of $\cos\theta$ is 1,

$$r = \frac{3.442 \times 10^7}{1 - 0.206(1)} = 4.34 \times 10^7 \text{ (3 s.f.)}.$$

Hence, the distance from Mercury to the sun at the aphelion is 4.34×10^7 miles.

Observe that r attains a minimum when $\cos\theta$ is also at its minimum. Since the minimum value of $\cos\theta$ is -1,

$$r = \frac{3.442 \times 10^7}{1 - 0.206(-1)} = 2.85 \times 10^7 \text{ (3 s.f.)}.$$

Hence, the distance from Mercury to the sun at the perihelion is 2.85×10^7 miles.

* * * * *

Problem 2. A variable point P has polar coordinates (r, θ), and fixed points A and B have polar coordinates $(1, 0)$ and $(1, \pi)$ respectively. Given that P moves so that the product $PA \cdot PB = 2$, show that

$$r^2 = \cos 2\theta + \sqrt{3 + \cos^2 2\theta}.$$

(a) Given that $r \geq 0$ and $0 \leq \theta \leq 2\pi$, find the maximum and minimum values of r, and the values of θ at which they occur.

(b) Verify that the path taken by P is symmetric about the lines $\theta = 0$ and $\theta = \dfrac{\pi}{2}$, giving your reasons.

Solution. Note that A and B have Cartesian coordinates $(1, 0)$ and $(-1, 0)$ respectively. Let $P(x, y)$. Then

$$PA^2 = (x-1)^2 + y^2, \qquad PB^2 = (x+1)^2 + y^2.$$

Hence,

$$PA \cdot PB = \left((x-1)^2 + y^2\right)\left((x+1)^2 + y^2\right) = \left(x^2 + y^2\right)^2 - 2\left(x^2 - y^2\right) + 1.$$

Since $x^2 - y^2 = r^2\left(\cos^2\theta - \sin^2\theta\right) = r^2\cos 2\theta$, the polar equation of the locus of P is

$$r^4 - 2r^2\cos 2\theta + 1 = (PA \cdot PB)^2 = 4 \implies r^4 - 2r^2\cos 2\theta - 3 = 0.$$

By the quadratic formula, we have

$$r^2 = \frac{2\cos 2\theta \pm \sqrt{4\cos^2 2\theta + 12}}{2} = \cos 2\theta \pm \sqrt{\cos^2 2\theta + 3}.$$

Since $\sqrt{\cos^2 2\theta + 3} > \cos 2\theta$ and $r^2 \geq 0$, we reject the negative case. Thus,

$$r^2 = \cos 2\theta + \sqrt{3 + \cos^2 2\theta}.$$

Part (a). Differentiating with respect to θ, we obtain

$$2r\frac{\mathrm{d}r}{\mathrm{d}\theta} = -2\sin 2\theta\left(1 + \frac{1}{2\sqrt{3 + \cos^2 2\theta}}\right).$$

For stationary points, $\mathrm{d}r/\mathrm{d}\theta = 0$. Since $1 + 1/2\sqrt{3 + \cos^2 2\theta} > 0$, we must have $\sin 2\theta = 0$, whence $\theta = 0, \pi/2, \pi, 3\pi/2$. By symmetry, we only consider $\theta = 0$ and $\theta = \pi/2$.

Case 1. When $\theta = 0$, we have $r^2 = 3$, whence $r = \sqrt{3}$.

Case 2. When $\theta = \pi/2$, we have $r^2 = 1$, whence $r = 1$.

Thus, $\max r = \sqrt{3}$ and occurs when $\theta = 0, \pi$, while $\min r = 1$ and occurs when $\theta = \pi/2, 3\pi/2$.

Part (b). Recall that the path taken by P is given by

$$\left((x-1)^2 + y^2\right)\left((x+1)^2 + y^2\right) = 4.$$

Observe that the above equation is invariant under the transformations $x \mapsto -x$ and $y \mapsto -y$. Hence, the path is symmetric about both the x- and y-axes, i.e. the lines $\theta = 0$ and $\theta = \pi/2$.

* * * * *

Problem 3.

(a) Explain why the curve with equation $x^3 + 2xy^2 - a^2y = 0$ where a is a positive constant lies entirely in the region $|x| \leq 2^{-\frac{3}{4}}a$.

(b) Show that the polar equation of this curve is $r^2 = \dfrac{a^2\tan\theta}{2 - \cos^2\theta}$.

(c) Sketch the curve.

Solution.

Part (a). Consider the discriminant Δ of $x^3 + 2xy^2 - a^2y = 0$ with respect to y:

$$\Delta = \left(-a^2\right)^2 - 4(2x) = a^4 - 8x^4.$$

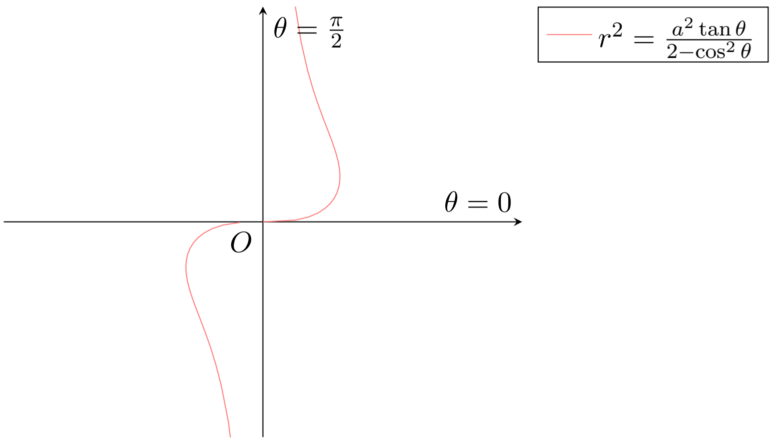
For points on the curve, we clearly have $\Delta \geq 0$. Thus,

$$a^3 - 8x^4 \geq 0 \implies x^4 \leq 2^{-3}a^4 \implies |x| \leq 2^{-3/4}a.$$

Part (b).

$$x^3 + 2xy^2 - a^2 y = 0 \implies 2\left(x^2 + y^2\right) - x^2 - a^2 \frac{y}{x} = 0 \implies 2r^2 - r^2 \cos^2\theta - a^2 \tan\theta = 0$$
$$\implies r^2 = \frac{a^2 \tan\theta}{2 - \cos^2\theta}.$$

Part (c).



* * * * *

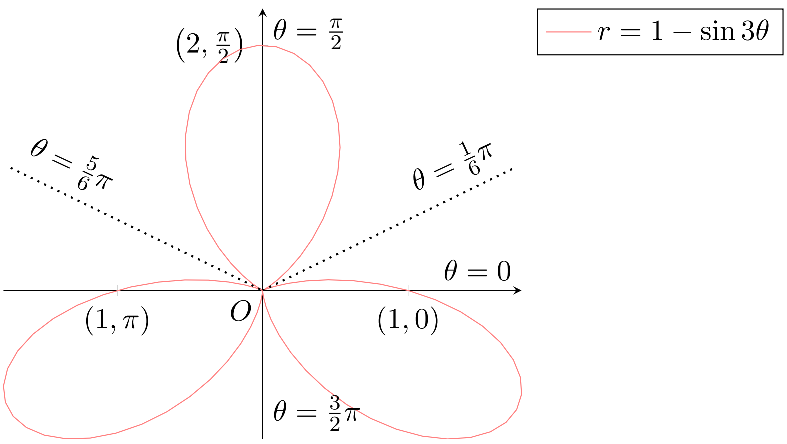
Problem 4. The curve C has polar equation $r = 1 - \sin 3\theta$, where $0 \le \theta \le 2\pi$.

(a) Sketch the curve C, showing the tangents at the pole and the intersections with the axes.

(b) Find the gradient of the curve at the point where $\theta = \dfrac{\pi}{3}$, giving your answer in the form $a + b\sqrt{3}$, where a and b are constants to be determined.

Solution.

Part (a).



When $\theta = 0$ or $\theta = \pi$, we have $r = 1$. Thus, C intersects the horizontal axis at $(1, 0)$ and $(1, \pi)$. When $\theta = \pi/2$, we have $r = 2$. Thus, C intersects the vertical axis at $(2, \pi/2)$. When $\theta = 3\pi/2$, we have $r = 0$. Thus, C passes through the pole.

For tangents at the pole, $r = 0 \implies \sin 3\theta = 1 \implies \theta = \pi/6, 5\pi/6, 3\pi/2$.

Part (b). Note that $\mathrm{d}r/\mathrm{d}\theta = -3\cos 3\theta$ evaluates to 3 when $\theta = \pi/3$. Thus,

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{\theta=\frac{\pi}{3}} = \left.\frac{\frac{\mathrm{d}r}{\mathrm{d}\theta}\sin\theta + r\cos\theta}{\frac{\mathrm{d}r}{\mathrm{d}\theta}\cos\theta - r\sin\theta}\right|_{\theta=\frac{\pi}{3}} = \frac{3\sqrt{3}+1}{3-\sqrt{3}} = \frac{12+10\sqrt{3}}{6} = 2 + \frac{5}{3}\sqrt{3}.$$

Hence, when $\theta = \pi/3$, the gradient of the curve is $2 + 5\sqrt{3}/2$.

A7 Vectors I - Basic Properties and Vector Algebra

Tutorial A7

Problem 1. The vector $\mathbf{v}$ is defined by $3\mathbf{i} - 4\mathbf{j} + \mathbf{k}$. Find the unit vector in the direction of $\mathbf{v}$ and hence find a vector of magnitude 25 which is parallel to $\mathbf{v}$.

Solution.

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{1}{\sqrt{3^2 + (-4)^2 + 1^2}} \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{26}} \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}, \quad 25\hat{\mathbf{v}} = \frac{25}{\sqrt{26}} \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}.$$

* * * * *

Problem 2. With respect to an origin O, the position vectors of the points A, B, C and D are $4\mathbf{i} + 7\mathbf{j}$, $\mathbf{i} + 3\mathbf{j}$, $2\mathbf{i} + 4\mathbf{j}$ and $3\mathbf{i} + d\mathbf{j}$ respectively.

(a) Find the vectors $\overrightarrow{BA}$ and $\overrightarrow{BC}$.

(b) Find the value of d if B, C and D are collinear. State the ratio $\frac{BC}{BD}$.

Solution.

Part (a). Note that

$$\overrightarrow{BA} = \overrightarrow{OA} - \overrightarrow{OB} = \begin{pmatrix} 4 \\ 7 \end{pmatrix} - \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad \overrightarrow{BC} = \overrightarrow{OC} - \overrightarrow{OB} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

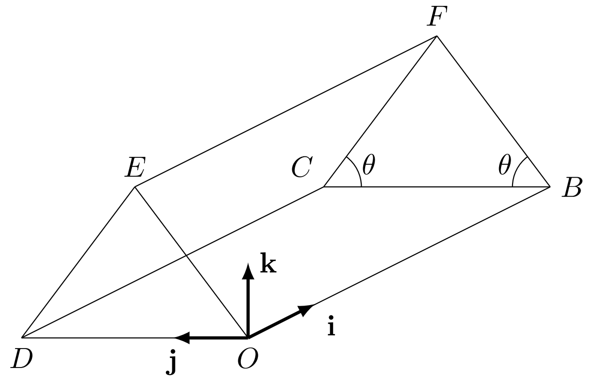
Part (b). If B, C and D are collinear, then $\overrightarrow{BC} = \lambda\overrightarrow{CD}$ for some $\lambda \in \mathbb{R}$.

$$\overrightarrow{BC} = \lambda\overrightarrow{CD} \implies \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \lambda\left(\overrightarrow{OD} - \overrightarrow{OC}\right) = \lambda\left[\begin{pmatrix} 3 \\ d \end{pmatrix} - \begin{pmatrix} 2 \\ 4 \end{pmatrix}\right] = \begin{pmatrix} \lambda \\ \lambda(d-4) \end{pmatrix}.$$

Hence, $\lambda = 1$ and $\lambda(d-4) = 1$, whence $d = 5$. Also, $\overrightarrow{BC} = \overrightarrow{CD}$. Thus,

$$\frac{BC}{BD} = \frac{BC}{BC + CD} = \frac{BC}{BC + BC} = \frac{1}{2}.$$

Problem 3. The diagram shows a roof, with horizontal rectangular base $OBCD$, where $OB = 10$ m and $BC = 6$ m. The triangular planes ODE and BCF are vertical and the ridge EF is horizontal to the base. The planes $OBFE$ and $DCFE$ are each inclined at an angle θ to the horizontal, where $\tan\theta = 4/3$. The point O is taken as the origin and vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, each of length 1 m, are taken along OB, OD and vertically upwards from O respectively.



Find the position vectors of the points B, C, D, E and F.

Solution. Note that $\overrightarrow{OB} = 10\mathbf{i}$ and $\overrightarrow{BC} = 6\mathbf{j}$. Thus, $\overrightarrow{OC} = \overrightarrow{OB} + \overrightarrow{BC} = 10\mathbf{i} + 6\mathbf{j}$. Also, note that $\triangle ODE \cong \triangle BCF$. Hence, $\overrightarrow{OD} = \overrightarrow{BC} = 6\mathbf{j}$. Note that $\triangle ODE$ is isosceles. Let G be the mid-point of OD. Since $\tan\theta = 4/3$, we have

$$\frac{EG}{DG} = \frac{4}{3} \implies EG = \frac{4}{3}DG = \frac{2}{3}OD = \frac{2}{3}\cdot 6 = 4 \implies \overrightarrow{GE} = 4\mathbf{k}.$$

Hence,

$$\overrightarrow{OE} = \overrightarrow{OG} + \overrightarrow{GE} = \frac{1}{2}\overrightarrow{OD} + \overrightarrow{GE} = 3\mathbf{j} + 4\mathbf{k}.$$

Hence,

$$\overrightarrow{OF} = \overrightarrow{OB} + \overrightarrow{BF} = \overrightarrow{OB} + \overrightarrow{OE} = 10\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}.$$

Thus,

$$\overrightarrow{OB} = 10\mathbf{i}, \quad \overrightarrow{OC} = 10\mathbf{i} + 6\mathbf{j}, \quad \overrightarrow{OD} = 6\mathbf{j}, \quad \overrightarrow{OE} = 3\mathbf{j} + 4\mathbf{k}, \quad \overrightarrow{OF} = 10\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}.$$

* * * * *

Problem 4. Find $\mathbf{u}\cdot\mathbf{v}$, $\mathbf{u}\times\mathbf{v}$ and the angle between $\mathbf{u}$ and $\mathbf{v}$ given that

(a) $\mathbf{u} = \mathbf{i} - \mathbf{j} + \mathbf{k}$, $\mathbf{v} = 3\mathbf{i} + 2\mathbf{j} + 7\mathbf{k}$

(b) $\mathbf{u} = 2\mathbf{i} - 3\mathbf{k}$, $\mathbf{v} = -\mathbf{i} + 7\mathbf{j} + 2\mathbf{k}$

Solution.

Part (a). We have $\mathbf{u} = (1,\ -1,\ 1)^\mathsf{T}$ and $\mathbf{v} = (3,\ 2,\ 7)^\mathsf{T}$. Hence,

$$\mathbf{u}\cdot\mathbf{v} = (1)(3) + (-1)(2) + (1)(7) = 8, \quad \mathbf{u}\times\mathbf{v} = \begin{pmatrix} (-1)(7) - (2)(1) \\ (1)(3) - (7)(1) \\ (1)(2) - (3)(-1) \end{pmatrix} = \begin{pmatrix} -9 \\ -4 \\ 5 \end{pmatrix}.$$

Let the angle between $\mathbf{u}$ and $\mathbf{v}$ be θ.

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} = \frac{8}{\sqrt{3}\sqrt{62}} \implies \theta = 54.1^\circ \text{ (1 d.p.)}.$$

Part (b). We have $\mathbf{u} = (2,\ 0,\ -3)^\mathsf{T}$ and $\mathbf{v} = (-1,\ 7,\ 2)^\mathsf{T}$. Hence,

$$\mathbf{u} \cdot \mathbf{v} = (2)(-1) + (0)(7) + (-3)(2) = -8, \quad \mathbf{u} \cdot \mathbf{v} = \begin{pmatrix} (0)(2) - (7)(-3) \\ (-3)(-1) - (2)(2) \\ (2)(7) - (-1)(0) \end{pmatrix} = \begin{pmatrix} 21 \\ -1 \\ 14 \end{pmatrix}.$$

Let the angle between $\mathbf{u}$ and $\mathbf{v}$ be θ.

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} = \frac{-8}{\sqrt{13}\sqrt{54}} \implies \theta = 107.6^\circ \text{ (1 d.p.)}.$$

* * * * *

Problem 5. Find $\mathbf{u} \cdot \mathbf{v}$ and $|\mathbf{u} \times \mathbf{v}|$ given that $\mathbf{u} = 2\mathbf{a} - \mathbf{b}$, $\mathbf{v} = -\mathbf{a} + 3\mathbf{b}$, where $|\mathbf{a}| = 2$, $|\mathbf{b}| = 1$ and the angle between $\mathbf{a}$ and $\mathbf{b}$ is 60°.

Solution.

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= (2\mathbf{a} - \mathbf{b}) \cdot (-\mathbf{a} + 3\mathbf{b}) = -2\mathbf{a} \cdot \mathbf{a} + 6\mathbf{a} \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{a} - 3\mathbf{b} \cdot \mathbf{b} \\ &= -2\,|\mathbf{a}|^2 - 3\,|\mathbf{b}|^2 + 7\,|\mathbf{a}|\,|\mathbf{b}| \cos\theta = -2(2)^2 - 3(1)^2 + 7(2)(1)\cos 60^\circ = -4. \\ |\mathbf{u} \times \mathbf{v}| &= |(2\mathbf{a} - \mathbf{b}) \times (-\mathbf{a} + 3\mathbf{b})| = |-2\mathbf{a} \times \mathbf{a} + 6\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{a} - 3\mathbf{b} \times \mathbf{b}| \\ &= |5\mathbf{a} \times \mathbf{b}| = 5\,|\mathbf{a}|\,|\mathbf{b}| \sin\theta = 5(2)(1)\sin 60^\circ = 5\sqrt{3}. \end{aligned}$$

* * * * *

Problem 6. If $\mathbf{a} = \mathbf{i} + 4\mathbf{j} - \mathbf{k}$, $\mathbf{b} = \mathbf{i} - \mathbf{j} + 3\mathbf{k}$ and $\mathbf{c} = 2\mathbf{i} + \mathbf{j}$, find

(a) a unit vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$,

(b) a vector perpendicular to both $(3\mathbf{b} - 5\mathbf{c})$ and $(7\mathbf{b} + \mathbf{c})$.

Solution.

Part (a). Note that $\mathbf{a} \times \mathbf{b} = (11,\ -4,\ -5)^\mathsf{T}$. Hence, $\widehat{\mathbf{a} \times \mathbf{b}} = \frac{1}{\sqrt{162}} (11,\ -4,\ -5)^\mathsf{T}$.

Part (b). Observe that $(3\mathbf{b} - 5\mathbf{c}) \times (7\mathbf{b} + \mathbf{c}) = \lambda \mathbf{b} \times \mathbf{c}$ for some $\lambda \in \mathbb{R}$. It hence suffices to find $\mathbf{b} \times \mathbf{c}$, which works out to be $(-3,\ 6,\ 3)^\mathsf{T}$.

* * * * *

Problem 7. The position vectors of the points A, B and C are given by $\mathbf{a} = 2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}$, $\mathbf{b} = 5\mathbf{i} - \mathbf{j} + 2\mathbf{k}$, $\mathbf{c} = 11\mathbf{i} + \lambda\mathbf{j} + 14\mathbf{k}$ respectively. Find

(a) a unit vector parallel to $\overrightarrow{AB}$;

(b) the position vector of the point D such that $ABCD$ is a parallelogram, leaving your answer in terms of λ;

(c) the value of λ if A, B and C are collinear;

(d) the position vector of the point P on AB is $AP : PB = 2 : 1$.

Solution.

Part (a).

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a} = \begin{pmatrix} 5 \\ -1 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} = \begin{pmatrix} 3 \\ -4 \\ 6 \end{pmatrix}.$$

Note that $\left|\overrightarrow{AB}\right| = \sqrt{61}$. Hence, the required vector is $\frac{1}{\sqrt{61}} (3, -4, 6)^\mathsf{T}$.

Part (b). Since $ABCD$ is a parallelogram, we have that $\overrightarrow{AD} = \overrightarrow{BC}$. Thus,

$$\overrightarrow{OD} - \mathbf{a} = \mathbf{c} - \mathbf{b} \implies \overrightarrow{OD} = \mathbf{a} - \mathbf{b} + \mathbf{c} = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} - \begin{pmatrix} 5 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 11 \\ \lambda \\ 14 \end{pmatrix} = \begin{pmatrix} 8 \\ \lambda + 4 \\ 8 \end{pmatrix}.$$

Part (c). Given that A, B and C are collinear, we have $\overrightarrow{AB} = k\overrightarrow{BC}$ for some $k \in \mathbb{R}$. Hence,

$$\begin{pmatrix} 3 \\ -4 \\ 6 \end{pmatrix} = k\,(\mathbf{c} - \mathbf{b}) = k \left[\begin{pmatrix} 11 \\ \lambda \\ 14 \end{pmatrix} - \begin{pmatrix} 5 \\ -1 \\ 2 \end{pmatrix} \right] = k \begin{pmatrix} 6 \\ \lambda + 1 \\ 12 \end{pmatrix}.$$

We hence see that $k = 1/2$, whence $\lambda = -9$.

Part (d). By the ratio theorem,

$$\overrightarrow{OP} = \frac{\mathbf{a} + 2\mathbf{b}}{2+1} = \frac{1}{3} \left[\begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} + 2 \begin{pmatrix} 5 \\ -1 \\ 2 \end{pmatrix} \right] = \frac{1}{3} \begin{pmatrix} 12 \\ 1 \\ 0 \end{pmatrix}.$$

* * * * *

Problem 8. $ABCD$ is a square, and M and N are the midpoints of BC and CD respectively. Express $\overrightarrow{AC}$ in terms of $\mathbf{p}$ and $\mathbf{q}$, where $\overrightarrow{AM} = \mathbf{p}$ and $\overrightarrow{AN} = \mathbf{q}$.

Solution. Let $ABCD$ be a square with side length $2k$ with A at the origin. Then $\mathbf{p} = \overrightarrow{AM} = (2k, -k)^\mathsf{T}$ and $\mathbf{q} = \overrightarrow{AN} = (k, -2k)^\mathsf{T}$. Hence, $\mathbf{p} + \mathbf{q} = (3k, -3k)^\mathsf{T}$. Thus, $\overrightarrow{AC} = (2k, -2k)^\mathsf{T} = \frac{2}{3}(3k, -3k)^\mathsf{T} = \frac{2}{3}(\mathbf{p} + \mathbf{q})$.

* * * * *

Problem 9. The points A, B have position vectors $\mathbf{a}$, $\mathbf{b}$ respectively, referred to an origin O, where $\mathbf{a}$ and $\mathbf{b}$ are not parallel to each other. The point C lies on AB between A and B and is such that $\frac{AC}{CB} = 2$, and D is the mid-point of OC. The line AD produced meets OB at E.

Find, in terms of $\mathbf{a}$ and $\mathbf{b}$,

(a) the position vector of C (referred to O),

(b) the vector $\overrightarrow{AD}$. Find the values of $\frac{OE}{EB}$ and $\frac{AE}{ED}$.

Solution.

Part (a). By the ratio theorem,

$$\overrightarrow{OC} = \frac{\mathbf{a} + 2\mathbf{b}}{2+1} = \frac{1}{3}\mathbf{a} + \frac{2}{3}\mathbf{b}.$$

Part (b). Since D is the midpoint of OC, we have $\overrightarrow{OD} = \frac{1}{6}\mathbf{a} + \frac{1}{3}\mathbf{b}$. Hence,

$$\overrightarrow{AD} = \overrightarrow{OD} - \overrightarrow{OA} = \left(\frac{1}{6}\mathbf{a} + \frac{1}{3}\mathbf{b}\right) - \mathbf{a} = -\frac{5}{6}\mathbf{a} + \frac{1}{3}\mathbf{b}.$$

Using Menelaus' theorem on $\triangle BCO$,

$$\frac{BA}{AC}\frac{CD}{DO}\frac{OE}{EB} = 1 \implies \frac{OE}{EB} = \frac{2}{3}.$$

Using Menelaus' theorem on $\triangle BEA$,

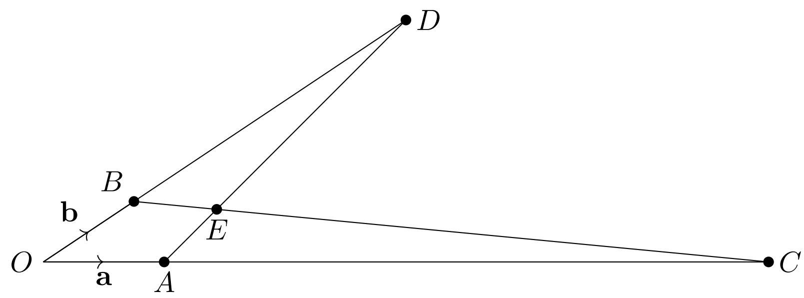
$$\frac{BO}{OE}\frac{ED}{DA}\frac{AC}{CB} = 1 \implies \frac{ED}{AD} = \frac{1}{5} \implies \frac{AE}{ED} = \frac{AD + DE}{ED} = 6.$$

* * * * *

Problem 10.

(a) The angle between the vectors $(3\mathbf{i} - 2\mathbf{j})$ and $(6\mathbf{i} + d\mathbf{j} - \sqrt{7}\mathbf{k})$ is $\arccos \frac{6}{13}$. Show that $2d^2 - 117d + 333 = 0$.

(b) With reference to the origin O, the points A, B, C and D are such that $\overrightarrow{OA} = \mathbf{a}$, $\overrightarrow{OB} = \mathbf{b}$, $\overrightarrow{AC} = 5\mathbf{a}$, $\overrightarrow{BD} = 3\mathbf{b}$. The lines AD and BC cross at E.



(i) Find $\overrightarrow{OE}$ in terms of $\mathbf{a}$ and $\mathbf{b}$.

(ii) The point F divides the line CD in the ratio $5 : 3$. Show that O, E and F are collinear, and find $OE : EF$.

Solution.

Part (a). Let $\mathbf{a} = (3, -2, 0)^\mathsf{T}$ and $\mathbf{b} = \left(6, d, -\sqrt{7}\right)^\mathsf{T}$. Note that $\mathbf{a} \cdot \mathbf{b} = 18 - 2d$. Let θ be the angle between $\mathbf{a}$ and $\mathbf{b}$.

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} \implies \frac{6}{13} = \frac{18 - 2d}{\sqrt{43 + d^2}\sqrt{13}} \implies \frac{9}{13} = \frac{(9-d)^2}{43 + d^2}$$
$$\implies 9(43 + d^2) = 13(d^2 - 18d + 81) \implies 2d^2 - 117d + 333 = 0.$$

Part (b).

Part (b)(i). By Menelaus' theorem,

$$\frac{OC}{CA}\frac{AE}{ED}\frac{DB}{BO} = 1 \implies \frac{AE}{ED} = \frac{5}{18} \implies \overrightarrow{AE} = \frac{5}{23}\overrightarrow{AD} \implies \overrightarrow{OE} = \overrightarrow{OA} + \frac{5}{23}\overrightarrow{AD}.$$

Since $\overrightarrow{AD} = \overrightarrow{OD} - \overrightarrow{OA} = 4\mathbf{b} - \mathbf{a}$. Thus,

$$\overrightarrow{OE} = \mathbf{a} + \frac{5}{23}(4\mathbf{b} - \mathbf{a}) = \frac{18}{23}\mathbf{a} + \frac{20}{23}\mathbf{b}.$$

Part (b)(ii). By the ratio theorem,

$$\overrightarrow{OF} = \frac{3\mathbf{c} + 5\mathbf{d}}{5+3} = \frac{23}{8}\left(\frac{18}{23}\mathbf{a} + \frac{20}{23}\mathbf{b}\right) = \frac{23}{8}\overrightarrow{OE}.$$

Thus, $OE : OF = 8 : 23$.

* * * * *

Problem 11. Relative to the origin O, two points A and B have position vectors given by $\mathbf{a} = 14\mathbf{i} + 14\mathbf{j} + 14\mathbf{k}$ and $\mathbf{b} = 11\mathbf{i} - 13\mathbf{j} + 2\mathbf{k}$ respectively.

(a) The point P divides the line AB in the ratio $2 : 1$. Find the coordinates of P.

(b) Show that AB and OP are perpendicular.

(c) The vector $\mathbf{c}$ is a unit vector in the direction of $\overrightarrow{OP}$. Write $\mathbf{c}$ as a column vector and give the geometrical meaning of $|\mathbf{a} \cdot \mathbf{c}|$.

(d) Find $\mathbf{a} \times \mathbf{p}$, where $\mathbf{p}$ is the vector $\overrightarrow{OP}$, and give the geometrical meaning of $|\mathbf{a} \times \mathbf{p}|$. Hence, write down the area of triangle OAP.

Solution.

Part (a). We have $\mathbf{a} = (14,\ 14,\ 14)^\mathsf{T} = 14\,(1,\ 1,\ 1)^\mathsf{T}$ and $\mathbf{b} = (11,\ -13,\ 2)^\mathsf{T}$. By the ratio theorem,

$$\overrightarrow{OP} = \frac{\mathbf{a} + 2\mathbf{b}}{2+1} = \frac{1}{3}\left[\begin{pmatrix}14\\14\\14\end{pmatrix} + 2\begin{pmatrix}11\\-13\\2\end{pmatrix}\right] = \begin{pmatrix}12\\-4\\6\end{pmatrix} = 2\begin{pmatrix}6\\-2\\3\end{pmatrix}.$$

Hence, $P(12, -4, 6)$

Part (b). Consider $\overrightarrow{AB} \cdot \overrightarrow{OP}$.

$$\overrightarrow{AB} \cdot \overrightarrow{OP} = \left[\begin{pmatrix}11\\-13\\2\end{pmatrix} - \begin{pmatrix}14\\14\\14\end{pmatrix}\right] \cdot \begin{pmatrix}12\\-4\\6\end{pmatrix} = -3\begin{pmatrix}1\\9\\4\end{pmatrix} \cdot 2\begin{pmatrix}6\\-2\\3\end{pmatrix} = 0.$$

Since $\overrightarrow{AB} \cdot \overrightarrow{OP} = 0$, AB and OP must be perpendicular.

Part (c). We have

$$\mathbf{c} = \frac{\overrightarrow{OP}}{\left|\overrightarrow{OP}\right|} = \frac{1}{\sqrt{6^2 + (-2)^2 + 3^2}}\begin{pmatrix}6\\-2\\3\end{pmatrix} = \frac{1}{7}\begin{pmatrix}6\\-2\\3\end{pmatrix}.$$

$|\mathbf{a} \cdot \mathbf{c}|$ is the length of the projection of $\mathbf{a}$ on $\overrightarrow{OP}$.

Part (d). We have

$$\mathbf{a} \times \mathbf{p} = 14\begin{pmatrix}1\\1\\1\end{pmatrix} \times 2\begin{pmatrix}6\\-2\\3\end{pmatrix} = 28\begin{pmatrix}1 \cdot 3 - (-2) \cdot 1\\ 1 \cdot 6 - 3 \cdot 1\\ 1 \cdot -2 - 6 \cdot 1\end{pmatrix} = 28\begin{pmatrix}5\\3\\-8\end{pmatrix}.$$

$|\mathbf{a} \times \mathbf{p}|$ is twice the area of $\triangle OAP$.

$$[\triangle OAP] = \frac{1}{2}\left|\mathbf{a} \times \mathbf{p}\right| = 14\sqrt{98} = 98\sqrt{2} \text{ units}^2.$$

Problem 12. The points A, B and C have position vectors given by $\mathbf{i}-\mathbf{j}+\mathbf{k}$, $\mathbf{j}-\mathbf{k}$ and $2\mathbf{i}-\mathbf{j}-\mathbf{k}$ respectively.

(a) Find the area of the triangle ABC. Hence, find the sine of the angle BAC.

(b) Find a vector perpendicular to the plane ABC.

(c) Find the projection vector of $\overrightarrow{AC}$ onto $\overrightarrow{AB}$.

(d) Find the distance of C to AB.

Solution.

Part (a). We have $\overrightarrow{OA} = (1, -1, 1)^\mathsf{T}$, $\overrightarrow{OB} = (0, 1, -1)^\mathsf{T}$ and $\overrightarrow{OC} = (2, -1, -1)^\mathsf{T}$. Note that $\overrightarrow{AB} = (-1, 2, -2)^\mathsf{T}$ and $\overrightarrow{AC} = (1, 0, -2)^\mathsf{T}$. Thus,

$$[\triangle ABC] = \frac{1}{2}\left|\overrightarrow{AB}\times\overrightarrow{AC}\right| = \frac{1}{2}\left|\begin{pmatrix}-4\\-4\\-2\end{pmatrix}\right| = \frac{1}{2}\cdot 6 = 3 \text{ units}^2.$$

We have

$$\sin BAC = \frac{\left|\overrightarrow{AB}\times\overrightarrow{AC}\right|}{\left|\overrightarrow{AB}\right|\left|\overrightarrow{AC}\right|} = \frac{6}{3\sqrt{5}} = \frac{2\sqrt{5}}{5}.$$

Part (b). $(2, 2, 1)^\mathsf{T}$ is parallel to $\overrightarrow{AB}\times\overrightarrow{AC}$ and is hence perpendicular to the plane ABC.

Part (c). The projection vector of $\overrightarrow{AC}$ onto $\overrightarrow{AB}$ is given by

$$\left(\overrightarrow{AC}\cdot\frac{\overrightarrow{AB}}{\left|\overrightarrow{AB}\right|}\right)\frac{\overrightarrow{AB}}{\left|\overrightarrow{AB}\right|} = \frac{1}{3}\begin{pmatrix}-1\\2\\-2\end{pmatrix}.$$

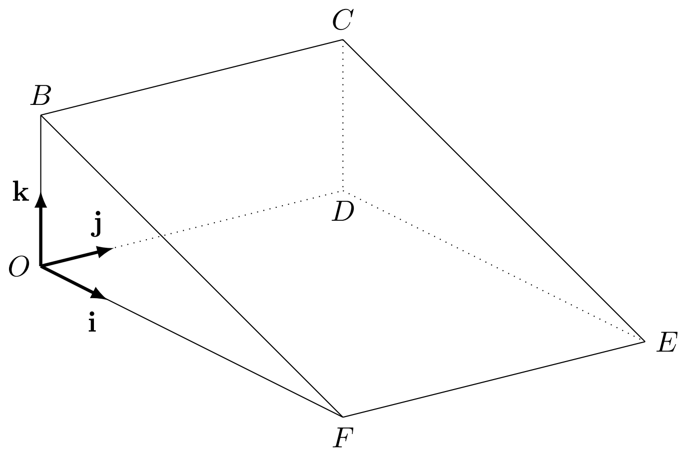
Part (d). Observe that

$$\left|\overrightarrow{AC}\times\frac{\overrightarrow{AB}}{\left|\overrightarrow{AB}\right|}\right| = \frac{1}{3}\left|\overrightarrow{AB}\times\overrightarrow{AC}\right| = 2.$$

Hence, the perpendicular distance between C and AB is 2 units.

* * * * *

Problem 13.



The diagram shows a vehicle ramp $OBCDEF$ with horizontal rectangular base $ODEF$ and vertical rectangular face $OBCD$. Taking the point O as the origin, the perpendicular unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are parallel to the edges OF, OD and OB respectively. The lengths of OF, OD and OB are $2h$ units, 3 units and h units respectively.

(a) Show that $\overrightarrow{OC} = 3\mathbf{j} + h\mathbf{k}$.

(b) The point P divides the segment CF in the ratio $2:1$. Find $\overrightarrow{OP}$ in terms of h.

For parts (c) and (d), let $h = 1$.

(c) Find the length of projection of $\overrightarrow{OP}$ onto $\overrightarrow{OC}$.

(d) Using the scalar product, find the angle that the rectangular face $BCEF$ makes with the horizontal base.

Solution.

Part (a). We have
$$\overrightarrow{OC} = \overrightarrow{OD} + \overrightarrow{DC} = \overrightarrow{OD} + \overrightarrow{OB} = 3\mathbf{j} + h\mathbf{k}.$$

Part (b). By the ratio theorem,
$$\overrightarrow{OP} = \frac{\overrightarrow{OC} + 2\overrightarrow{OF}}{2+1} = \frac{1}{3}\left[\begin{pmatrix}0\\3\\h\end{pmatrix} + 2\begin{pmatrix}2h\\0\\0\end{pmatrix}\right] = \frac{1}{3}\begin{pmatrix}4h\\3\\h\end{pmatrix}.$$

Part (c). The length of projection of $\overrightarrow{OP}$ onto $\overrightarrow{OC}$ is given by
$$\left|\overrightarrow{OP} \cdot \frac{\overrightarrow{OC}}{\left|\overrightarrow{OC}\right|}\right| = \frac{1}{3\sqrt{10}}\left|\begin{pmatrix}4\\3\\1\end{pmatrix} \cdot \begin{pmatrix}0\\3\\1\end{pmatrix}\right| = \frac{\sqrt{10}}{3} \text{ units.}$$

Part (d). Note that $\overrightarrow{OF} = (2,\, 0,\, 0)^\mathsf{T}$ and $\overrightarrow{BF} = \overrightarrow{OF} - \overrightarrow{OB} = (2,\, 0,\, -1)^\mathsf{T}$. Let θ be the angle the rectangular face $BCEF$ makes with the horizontal base.
$$\cos\theta = \frac{\overrightarrow{OF} \cdot \overrightarrow{BF}}{\left|\overrightarrow{OF}\right|\left|\overrightarrow{BF}\right|} = \frac{4}{2\sqrt{5}} \implies \theta = 26.6^\circ \text{ (1 d.p.).}$$

* * * * *

Problem 14. The position vectors of the points A and B relative to the origin O are $\overrightarrow{OA} = \mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$ and $\overrightarrow{OB} = 2\mathbf{i} - 3\mathbf{j} + 6\mathbf{k}$ respectively. The point P on AB is such that $AP : PB = \lambda : 1 - \lambda$. Show that $\overrightarrow{OP} = (1+\lambda)\mathbf{i} + (2-5\lambda)\mathbf{j} + (-2+8)\mathbf{k}$ where λ is a real parameter.

(a) Find the value of λ for which OP is perpendicular to AB.

(b) Find the value of λ for which angles $\angle AOP$ and $\angle POB$ are equal.

Solution. By the ratio theorem,
$$\overrightarrow{OP} = \frac{\lambda\overrightarrow{OB} + (1-\lambda)\overrightarrow{OA}}{\lambda + (1-\lambda)} = \lambda\begin{pmatrix}2\\-3\\6\end{pmatrix} + (1-\lambda)\begin{pmatrix}1\\2\\-2\end{pmatrix} = \begin{pmatrix}1+\lambda\\2-5\lambda\\-2+8\lambda\end{pmatrix}.$$

Part (a). Note that $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = (1, -5, 8)^{\mathsf{T}}$. For OP to be perpendicular to AB, we must have $\overrightarrow{OP} \cdot \overrightarrow{AB} = 0$.

$$\overrightarrow{OP} \cdot \overrightarrow{AB} = 0 \implies \begin{pmatrix} 1+\lambda \\ 2-5\lambda \\ -2+8\lambda \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -5 \\ 8 \end{pmatrix} = 0 \implies -25 + 90\lambda = 0 \implies \lambda = \frac{5}{18}.$$

Part (b). Suppose $\angle AOP = \angle POB$. Then $\cos\angle AOP = \cos\angle POB$. Thus,

$$\frac{\overrightarrow{OP} \cdot \overrightarrow{OA}}{\left|\overrightarrow{OP}\right| \left|\overrightarrow{OA}\right|} = \frac{\overrightarrow{OP} \cdot \overrightarrow{OB}}{\left|\overrightarrow{OP}\right| \left|\overrightarrow{OB}\right|} \implies \overrightarrow{OP} \cdot \left(\frac{1}{3}\overrightarrow{OA} - \frac{1}{7}\overrightarrow{OB}\right) = 0 \implies \overrightarrow{OP} \cdot \left(7\overrightarrow{OA} - 3\overrightarrow{OB}\right) = 0.$$

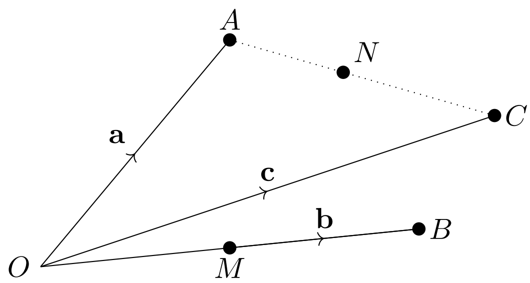
This gives

$$\begin{pmatrix} 1+\lambda \\ 2-5\lambda \\ -2+8\lambda \end{pmatrix} \cdot \left[7\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} - 3\begin{pmatrix} 2 \\ -3 \\ 6 \end{pmatrix}\right] = \begin{pmatrix} 1+\lambda \\ 2-5\lambda \\ -2+8\lambda \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 23 \\ -32 \end{pmatrix} = 0.$$

Taking the dot product and simplifying, we see that $111 - 370\lambda = 0$, whence $\lambda = \frac{3}{10}$.

* * * * *

Problem 15.



The origin O and the points A, B and C lie in the same plane, where $\overrightarrow{OA} = \mathbf{a}$, $\overrightarrow{OB} = \mathbf{b}$ and $\overrightarrow{OC} = \mathbf{c}$,

(a) Explain why $\mathbf{c}$ can be expressed as $\mathbf{c} = \lambda\mathbf{a} + \mu\mathbf{b}$, for constants λ and μ.

The point N is on AC such that $AN : NC = 3 : 4$.

(b) Write down the position vector of N in terms of $\mathbf{a}$ and $\mathbf{c}$.

(c) It is given that the area of triangle ONC is equal to the area of triangle OMC, where M is the mid-point of OB. By finding the areas of these triangles in terms of $\mathbf{a}$ and $\mathbf{b}$, find λ in terms of μ in the case where λ and μ are both positive.

Solution.

Part (a). Since $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are co-planar and $\mathbf{a}$ is not parallel to $\mathbf{b}$, $\mathbf{c}$ can be written as a linear combination of $\mathbf{a}$ and $\mathbf{b}$.

Part (b). By the ratio theorem,

$$\overrightarrow{ON} = \frac{4\mathbf{a} + 3\mathbf{c}}{3+4} = \frac{4}{7}\mathbf{a} + \frac{3}{7}\mathbf{c}.$$

Part (c). Let $\mathbf{c} = \lambda\mathbf{a} + \mu\mathbf{b}$. The area of $\triangle ONC$ is given by

$$[\triangle ONC] = \frac{1}{2}\left|\overrightarrow{ON} \times \overrightarrow{OC}\right| = \frac{1}{2}\left|\left[\frac{4}{7}\mathbf{a} + \frac{3}{7}(\lambda\mathbf{a} + \mu\mathbf{b})\right] \times (\lambda\mathbf{a} + \mu\mathbf{b})\right| = \frac{2\mu}{7}\left|\mathbf{a} \times \mathbf{b}\right|.$$

Meanwhile, the area of $\triangle OMC$ is given by

$$[\triangle OMC] = \frac{1}{2}\left|\overrightarrow{OM} \times \overrightarrow{OC}\right| = \frac{1}{2}\left|\frac{1}{2}\mathbf{b} \times (\lambda\mathbf{a} + \mu\mathbf{b})\right| = \frac{\lambda}{4}\left|\mathbf{a} \times \mathbf{b}\right|.$$

Since the two areas are equal,

$$[\triangle ONC] = [\triangle OMC] \implies \frac{2\mu}{7}\left|\mathbf{a} \times \mathbf{b}\right| = \frac{\lambda}{4}\left|\mathbf{a} \times \mathbf{b}\right| \implies \lambda = \frac{8}{7}\mu.$$

Self-Practice A7

Problem 1. The position vector of points A, B and C relative to an origin O are $\mathbf{a}$, $\mathbf{b}$ and $k\mathbf{a}$ respectively. The point P lies on AB and is such that $AP = 2PB$. The point Q lies on BC such that $CQ = 6QB$. Find, in terms of $\mathbf{a}$ and $\mathbf{b}$, the position vector of P and Q. Given that OPQ is a straight line, find

(a) the value of k,

(b) the ratio of $OP : PQ$.

The position vector of a point R is $\frac{7}{3}\mathbf{a}$. Show that PR is parallel to BC.

Solution. By the ratio theorem,

$$\overrightarrow{OP} = \frac{\mathbf{a}+2\mathbf{b}}{1+2} = \frac{1}{3}\mathbf{a} + \frac{2}{3}\mathbf{b}$$

and

$$\overrightarrow{OQ} = \frac{k\mathbf{a}+6\mathbf{b}}{6+1} = \frac{k}{7}\mathbf{a} + \frac{6}{7}\mathbf{b}.$$

Part (a). Since OPQ is a straight line, there exists some $\lambda \in \mathbb{R}$ such that

$$\overrightarrow{OQ} = \lambda\overrightarrow{OP} \implies \frac{k}{7}\mathbf{a} + \frac{6}{7}\mathbf{b} = \frac{\lambda}{3}\mathbf{a} + \frac{2\lambda}{3}\mathbf{b}.$$

Comparing coefficients of $\mathbf{b}$ terms, we have $\lambda = 9/7$, whence

$$\frac{k}{7} = \frac{9/7}{3} \implies k = 3.$$

Part (b). Note that $\overrightarrow{OQ} = \frac{9}{7}\overrightarrow{OP}$. Hence, $OP : PQ = 2 : 7$.

Note that

$$\overrightarrow{PR} = \frac{7}{3}\mathbf{a} - \left(\frac{1}{3}\mathbf{a} + \frac{2}{3}\mathbf{b}\right) = 2\mathbf{a} - \frac{2}{3}\mathbf{b}.$$

Hence,

$$\overrightarrow{BC} = 3\mathbf{a} - \mathbf{b} = \frac{3}{2}\left(2\mathbf{a} - \frac{2}{3}\mathbf{b}\right) = \frac{3}{2}\overrightarrow{PR}.$$

Hence, $PR \parallel BC$.

* * * * *

Problem 2. The position vectors of the points P and R, relative to an origin O, are $\mathbf{p}$ and $\mathbf{r}$ respectively, where $\mathbf{p}$ and $\mathbf{r}$ are not parallel to each other. Q is a point such that $\overrightarrow{OQ} = 2\overrightarrow{OP}$ and S is a point such that $\overrightarrow{OS} = 2\overrightarrow{OR}$. T is the midpoint of QS.

Find, in terms of $\mathbf{p}$ and $\mathbf{r}$,

(a) $\overrightarrow{PR}$,

(b) $\overrightarrow{QT}$,

(c) $\overrightarrow{TR}$.

What shape is the quadrilateral $PRTQ$? Name another quadrilateral that has the same shape as $PRTQ$.

Solution. By the midpoint theorem,

$$\overrightarrow{OT} = \frac{\overrightarrow{OQ} + \overrightarrow{OS}}{2} = \mathbf{p} + \mathbf{r}.$$

Part (a).

$$\overrightarrow{PR} = \mathbf{r} - \mathbf{p}.$$

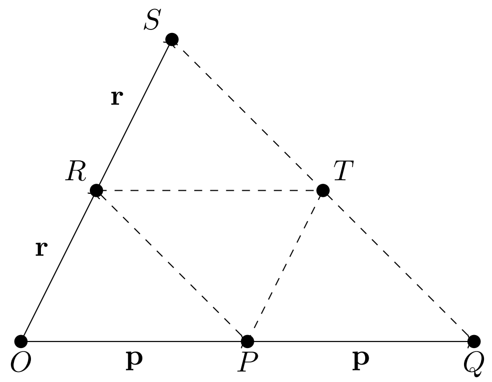
Part (b).

$$\overrightarrow{QT} = (\mathbf{p} + \mathbf{r}) - (2\mathbf{p}) = \mathbf{r} - \mathbf{p}.$$

Part (c).

$$\overrightarrow{TR} = \mathbf{r} - (\mathbf{r} + \mathbf{p}) = -\mathbf{p}.$$

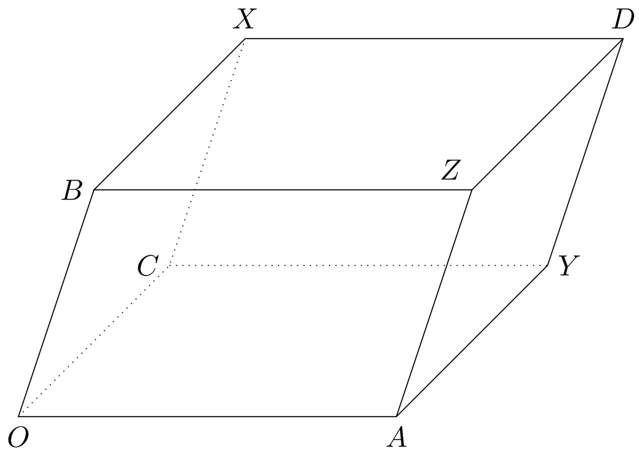
Consider the following diagram:



Clearly, $PRTQ$ is a parallelogram. Likewise, $ORTP$ is also a parallelogram.

* * * * *

Problem 3. The position vectors of points A, B, C are given by $\overrightarrow{OA} = 5\mathbf{i}$, $\overrightarrow{OB} = \mathbf{i} + 3\mathbf{k}$, $\overrightarrow{OC} = \mathbf{i} + 4\mathbf{j}$. A parallelepiped has OA, OB and OC as three edges, and the remaining edges are X, Y, Z and D as shown in the diagram.



(a) Write down the position vectors of X, Y, Z and D in terms of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, and calculate the length of OD.

(b) Calculate the size of angle OZY.

(c) The point P divides CZ in the ratio $\lambda : 1$. Write down the position vector of P, and evaluate λ if $\overrightarrow{OP}$ is perpendicular to $\overrightarrow{CZ}$.

Solution.

Part (a). We have

$$\begin{aligned}\overrightarrow{OX} &= \phantom{\overrightarrow{OA}+{}} \overrightarrow{OB} + \overrightarrow{OC} &&= 2\mathbf{i} + 4\mathbf{j} + 3\mathbf{k},\\ \overrightarrow{OY} &= \overrightarrow{OA} \phantom{+\overrightarrow{OB}} + \overrightarrow{OC} &&= 6\mathbf{i} + 4\mathbf{j},\\ \overrightarrow{OZ} &= \overrightarrow{OA} + \overrightarrow{OB} &&= 6\mathbf{i} + 3\mathbf{k},\\ \overrightarrow{OD} &= \overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC} &&= 7\mathbf{i} + 4\mathbf{j} + 3\mathbf{k}.\end{aligned}$$

Part (b). Note that $\overrightarrow{ZY} = (0,\, 4,\, -3)^\mathsf{T}$. Hence,

$$\cos \angle OZY = \frac{\overrightarrow{OZ} \cdot \overrightarrow{ZY}}{\left|\overrightarrow{OZ}\right| \left|\overrightarrow{ZY}\right|} = \frac{9}{\sqrt{45}\sqrt{25}} \implies \angle OZY = 74.4^\circ \text{ (1 d.p.)}.$$

Part (c). By the ratio theorem,

$$\overrightarrow{OP} = \frac{\overrightarrow{OC} + \lambda \overrightarrow{OZ}}{1+\lambda} = \frac{1}{1+\lambda}\left[3\lambda \begin{pmatrix}2\\0\\1\end{pmatrix} + \begin{pmatrix}1\\4\\0\end{pmatrix}\right].$$

Note that $\overrightarrow{CZ} = (5,\, -4,\, 3)^\mathsf{T}$. Since $\overrightarrow{OP} \perp \overrightarrow{CZ}$, we have

$$\overrightarrow{OP} \cdot \overrightarrow{CZ} = 0.$$

Hence,

$$3\lambda \begin{pmatrix}2\\0\\1\end{pmatrix} \cdot \begin{pmatrix}5\\-4\\3\end{pmatrix} + \begin{pmatrix}1\\4\\0\end{pmatrix} \cdot \begin{pmatrix}5\\-4\\3\end{pmatrix} = 39\lambda - 11 = 0,$$

whence $\lambda = 11/39$.

* * * * *

Problem 4. The vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are such that $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{c} = 0$ and $\mathbf{a} \cdot \mathbf{c} = 2$. Given that $|\mathbf{a}| = 1$, $|\mathbf{b}| = 2$, $|\mathbf{c}| = 3$, find

(a) $|\mathbf{a} - \mathbf{b}|$;

(b) $|\mathbf{a} - \mathbf{b} - \mathbf{c}|$.

Solution.

Part (a). Observe that

$$|\mathbf{a} - \mathbf{b}|^2 = (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) = \mathbf{a} \cdot \mathbf{a} - \mathbf{a} \cdot \mathbf{b} - \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b}.$$

Since $\mathbf{a} \cdot \mathbf{b} = 0$, we get

$$|\mathbf{a} - \mathbf{b}|^2 = \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} = |\mathbf{a}|^2 + |\mathbf{b}|^2 = 1^1 + 2^2 = 5.$$

Thus, $|\mathbf{a} - \mathbf{b}| = \sqrt{5}$.

Part (b). Observe that

$$|\mathbf{a} - \mathbf{b} - \mathbf{c}|^2 = (\mathbf{a} - \mathbf{b} - \mathbf{c}) \cdot (\mathbf{a} - \mathbf{b} - \mathbf{c}) = (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) - 2\mathbf{c} \cdot (\mathbf{a} - \mathbf{b}) + \mathbf{c} \cdot \mathbf{c}.$$

Since $\mathbf{a} \cdot \mathbf{c} = 2$ and $\mathbf{b} \cdot \mathbf{c} = 0$, we have

$$|\mathbf{a} - \mathbf{b} - \mathbf{c}|^2 = |\mathbf{a} - \mathbf{b}|^2 - 2(2) + |\mathbf{c}|^2 = 5 - 2(2) + 3^2 = 10.$$

Thus, $|\mathbf{a} - \mathbf{b} - \mathbf{c}| = \sqrt{10}$.

* * * * *

Problem 5. The position vectors of the points M and N are given by

$$\overrightarrow{OM} = \lambda\mathbf{i} + (2\lambda - 1)\mathbf{j} + \mathbf{k}, \qquad \overrightarrow{ON} = (1-\lambda)\mathbf{i} + 3\lambda\mathbf{j} - 2\mathbf{k},$$

where λ is a scalar. Find the values of λ for which $\overrightarrow{OM}$ and $\overrightarrow{ON}$ are perpendicular. When $\lambda = 1$, find the size of $\angle MNO$ to the nearest degree.

Solution. Since $\overrightarrow{OM} \perp \overrightarrow{ON}$, we have

$$\overrightarrow{OM} \cdot \overrightarrow{ON} = \begin{pmatrix} \lambda \\ 2\lambda - 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1-\lambda \\ 3\lambda \\ -2 \end{pmatrix} = 5\lambda^2 - 2\lambda - 2 = 0.$$

Solving the quadratic, we get

$$\lambda = \frac{1 \pm \sqrt{11}}{5}.$$

When $\lambda = 1$, we have

$$\overrightarrow{OM} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \overrightarrow{ON} = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}, \quad \overrightarrow{MN} = \begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix}.$$

Hence,

$$\cos\angle MNO = \frac{\overrightarrow{ON} \cdot \overrightarrow{MN}}{\left|\overrightarrow{ON}\right|\left|\overrightarrow{MN}\right|} = \frac{12}{\sqrt{13}\sqrt{14}} \implies \angle MNO = 27^\circ.$$

* * * * *

Problem 6. The points A, B, C and D have position vectors $\mathbf{i}-2\mathbf{j}+5\mathbf{k}$, $\mathbf{i}+3\mathbf{j}$, $10\mathbf{i}+\mathbf{j}+2\mathbf{k}$ and $-2\mathbf{i}+4\mathbf{j}+5\mathbf{k}$ respectively, with respect to an origin O. The point P on AB is such that $AP : PB = \lambda : 1-\lambda$ and point Q on CD is such that $CQ : QD = \mu : 1-\mu$. Find $\overrightarrow{OP}$ and $\overrightarrow{OQ}$ in terms of λ and μ respectively.

Given that PQ is perpendicular to both AB and CD, show that $\overrightarrow{PQ} = \mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$.

Solution. By the ratio theorem,

$$\overrightarrow{OP} = \frac{(1-\lambda)\overrightarrow{OA} + \lambda\overrightarrow{OB}}{(1-\lambda)+\lambda} = \overrightarrow{OA} + \lambda\left(\overrightarrow{OB} - \overrightarrow{OA}\right) = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} + 5\lambda\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

and

$$\overrightarrow{OQ} = \frac{(1-\mu)\overrightarrow{OC} + \overrightarrow{OD}}{(1-\mu)+\mu} = \overrightarrow{OC} + \mu\left(\overrightarrow{OD} - \overrightarrow{OC}\right) = \begin{pmatrix} 10 \\ 1 \\ 2 \end{pmatrix} + 3\mu\begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix}.$$

Note that

$$\overrightarrow{PQ} = \left[\begin{pmatrix} 10 \\ 1 \\ 2 \end{pmatrix} + 3\mu\begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix}\right] - \left[\begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} + 5\lambda\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right] = 3\begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} + 3\mu\begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix} - 5\lambda\begin{pmatrix} 0 \\ 1 \\ 01 \end{pmatrix}.$$

Since PQ is perpendicular to AB, we have

$$\overrightarrow{PQ} \cdot \overrightarrow{AB} = \left[3\begin{pmatrix}3\\1\\-1\end{pmatrix} + 3\mu\begin{pmatrix}-4\\1\\1\end{pmatrix} - 5\lambda\begin{pmatrix}0\\1\\01\end{pmatrix}\right] \cdot \left[5\begin{pmatrix}0\\1\\-1\end{pmatrix}\right] = 5(6-10\lambda) = 0.$$

Thus, $\lambda = 3/5$.

Since PQ is perpendicular to CD, we have

$$\overrightarrow{PQ} \cdot \overrightarrow{CD} = \left[3\begin{pmatrix}3\\1\\-1\end{pmatrix} + 3\mu\begin{pmatrix}-4\\1\\1\end{pmatrix} - 5\lambda\begin{pmatrix}0\\1\\01\end{pmatrix}\right] \cdot \left[3\begin{pmatrix}-4\\1\\1\end{pmatrix}\right] = 3(-36+54\mu) = 0.$$

Thus, $\mu = 2/3$.

Hence,

$$\overrightarrow{PQ} = 3\begin{pmatrix}3\\1\\-1\end{pmatrix} + 3\left(\frac{2}{3}\right)\begin{pmatrix}-4\\1\\1\end{pmatrix} - 5\left(\frac{3}{5}\right)\begin{pmatrix}0\\1\\-1\end{pmatrix} = \begin{pmatrix}1\\2\\2\end{pmatrix}.$$

* * * * *

Problem 7. The position vectors of the vertices A, B and C of a triangle are $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ respectively. If O is the origin and not within the triangle, show that the area of triangle OAB is $\frac{1}{2}|\mathbf{a}\times\mathbf{b}|$, and deduce and expression for the area of the triangle ABC.

Hence, or otherwise, show that the perpendicular distance from B to AC is

$$\frac{|\mathbf{a}\times\mathbf{b}+\mathbf{b}\times\mathbf{c}+\mathbf{c}\times\mathbf{a}|}{|\mathbf{c}-\mathbf{a}|}.$$

Solution. Let $\theta = \angle AOB$ be the angle between $\mathbf{a}$ and $\mathbf{b}$. Clearly,

$$[\triangle OAB] = \frac{1}{2}(OA)(OB)\sin\theta = \frac{1}{2}|\mathbf{a}\times\mathbf{b}|.$$

Note that $AB = |\mathbf{b}-\mathbf{a}|$ and $AC = |\mathbf{c}-\mathbf{a}|$. Hence,

$$[\triangle ABC] = \frac{1}{2}|(\mathbf{b}-\mathbf{a})\times(\mathbf{c}-\mathbf{a})|.$$

Expanding, we get

$$[\triangle ABC] = \frac{1}{2}|\mathbf{b}\times\mathbf{c}-\mathbf{b}\times\mathbf{a}-\mathbf{a}\times\mathbf{c}| = \frac{1}{2}|\mathbf{a}\times\mathbf{b}+\mathbf{b}\times\mathbf{c}+\mathbf{c}\times\mathbf{a}|.$$

Let the perpendicular distance from B to AC be h. Then

$$[\triangle ABC] = \frac{1}{2}h(AC) = \frac{1}{2}h|\mathbf{c}-\mathbf{a}|.$$

Hence,

$$h = \frac{2[\triangle ABC]}{|\mathbf{c}-\mathbf{a}|} = \frac{|\mathbf{a}\times\mathbf{b}+\mathbf{b}\times\mathbf{c}+\mathbf{c}\times\mathbf{a}|}{|\mathbf{c}-\mathbf{a}|}.$$

* * * * *

Problem 8 (🌶). The points A, B and C lie on a circle with centre O and diameter AC. It is given that $\overrightarrow{OA} = \mathbf{a}$ and $\overrightarrow{OB} = \mathbf{b}$.

(a) Find $\overrightarrow{BC}$ in terms of $\mathbf{a}$ and $\mathbf{b}$. Hence, show that AB is perpendicular to BC.

(b) Given that $\angle AOB = 30^\circ$, find $\overrightarrow{OF}$ where F is the foot of perpendicular of B to AC. Hence, find $\overrightarrow{OB'}$, where B' is the reflection of B in the line AC.

Solution.

Part (a). Since A, B and C lie on the same circle, $|\mathbf{a}| = |\mathbf{b}| = |\mathbf{c}|$. Since AC is the diameter of the circle, $\mathbf{c}$ is in the opposite direction as $\mathbf{a}$. Hence, $\mathbf{c} = -\mathbf{a}$. Thus,

$$\overrightarrow{BC} = \overrightarrow{OC} - \overrightarrow{OB} = -\mathbf{a} - \mathbf{b}.$$

Also note that

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \mathbf{b} - \mathbf{a}.$$

Consider $\overrightarrow{AB} \cdot \overrightarrow{BC}$:

$$\overrightarrow{AB} \cdot \overrightarrow{BC} = (\mathbf{b} - \mathbf{a}) \cdot -(\mathbf{a} + \mathbf{b}) = -(\mathbf{b} \cdot \mathbf{b} - \mathbf{a} \cdot \mathbf{a}) = -\left(|\mathbf{b}|^2 - |\mathbf{a}|^2\right) = 0.$$

Thus, AB is perpendicular to BC.

Part (b). Observe that

$$\frac{\sqrt{3}}{2} = \cos \angle AOB = \frac{OF}{OB} = \frac{\left|\overrightarrow{OF}\right|}{|\mathbf{a}|} \implies \left|\overrightarrow{OF}\right| = \frac{\sqrt{3}}{2}\,|\mathbf{a}|\,.$$

Since $\overrightarrow{OF}$ is in the same direction as $\overrightarrow{OA}$, we have

$$\overrightarrow{OF} = \frac{\sqrt{3}}{2}\mathbf{a}.$$

Note that

$$\overrightarrow{BF} = \frac{\sqrt{3}}{2}\mathbf{a} - \mathbf{b}.$$

By the midpoint theorem,

$$\overrightarrow{OF} = \frac{\overrightarrow{OB} + \overrightarrow{OB'}}{2} \implies \overrightarrow{OB'} = 2\overrightarrow{OF} - \overrightarrow{OB} = \sqrt{3}\mathbf{a} - \mathbf{b}.$$

Assignment A7

Problem 1. The points A and B have position vectors relative to the origin O, denoted by $\mathbf{a}$ and $\mathbf{b}$ respectively, where $\mathbf{a}$ and $\mathbf{b}$ are non-parallel vectors. The point P lies on AB such that $AP : PB = \lambda : 1$. The point Q lies on OP extended such that $OP = 2PQ$ and $\overrightarrow{BQ} = \overrightarrow{OA} + \mu\overrightarrow{OB}$. Find the values of the real constants λ and μ.

Solution. By the ratio theorem,

$$\overrightarrow{OP} = \frac{\mathbf{a} + \lambda\mathbf{b}}{1+\lambda} \implies \overrightarrow{OQ} = \frac{3}{2}\overrightarrow{OP} = \frac{3}{2} \cdot \frac{\mathbf{a} + \lambda\mathbf{b}}{1+\lambda}.$$

However, we also have

$$\overrightarrow{OQ} = \overrightarrow{OB} + \overrightarrow{BQ} = \mathbf{a} + (1+\mu)\mathbf{b}.$$

This gives the equality

$$\frac{3}{2} \cdot \frac{\mathbf{a} + \lambda\mathbf{b}}{1+\lambda} = \mathbf{a} + (1+\mu)\mathbf{b}.$$

Since $\mathbf{a}$ and $\mathbf{b}$ are non-parallel, we can compare the $\mathbf{a}$- and $\mathbf{b}$-components of both vectors separately. This gives us

$$\frac{3}{2} \cdot \frac{1}{1+\lambda} = 1, \quad \frac{3}{2} \cdot \frac{\lambda}{1+\lambda} = 1 + \mu,$$

which has the unique solution $\lambda = 1/2$ and $\mu = -1/2$.

* * * * *

Problem 2. Given that $\mathbf{a} = \mathbf{i} + \mathbf{j}$, $\mathbf{b} = 4\mathbf{i} - 2\mathbf{j} + 6\mathbf{k}$ and $\mathbf{p} = \lambda\mathbf{a} + (1-\lambda)\mathbf{b}$ where $\lambda \in \mathbb{R}$, find the possible value(s) of λ for which the angle between $\mathbf{p}$ and $\mathbf{k}$ is 45°.

Solution. Observe that

$$\mathbf{p} = \lambda\mathbf{a} + (1-\lambda)\mathbf{b} = \lambda \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (1-\lambda) \begin{pmatrix} 4 \\ -2 \\ 6 \end{pmatrix} = \begin{pmatrix} 4 - 3\lambda \\ -2 + 3\lambda \\ 6 - 6\lambda \end{pmatrix}.$$

Thus,

$$|\mathbf{p}|^2 = (4-3\lambda)^2 + (-2+3\lambda)^2 + (6-6\lambda)^2 = 54\lambda^2 - 108\lambda + 56.$$

Since the angle between $\mathbf{p}$ and $\mathbf{k}$ is 45°,

$$\cos 45° = \frac{\mathbf{p} \cdot \mathbf{k}}{|\mathbf{p}|\,|\mathbf{k}|} \implies \frac{1}{\sqrt{2}} = \frac{6 - 6\lambda}{|\mathbf{p}|} \implies \frac{|\mathbf{p}|^2}{2} = (6-6\lambda)^2.$$

We thus obtain the quadratic equation

$$\frac{54\lambda^2 - 108\lambda + 56}{2} = 36\lambda^2 - 72\lambda + 36 \implies 9\lambda^2 - 18\lambda + 8 = 0,$$

which has solutions $\lambda = 2/3$ and $\lambda = 4/3$. However, we must reject $\lambda = 4/3$ since $6 - 6\lambda = |\mathbf{p}|/\sqrt{2} > 0 \implies \lambda < 1$. Thus, $\lambda = 2/3$.

Problem 3.

(a) $\mathbf{a}$ and $\mathbf{b}$ are non-zero vectors such that $\mathbf{a} = (\mathbf{a} \cdot \mathbf{b})\mathbf{b}$. State the relation between the directions of $\mathbf{a}$ and $\mathbf{b}$, and find $|\mathbf{b}|$.

(b) $\mathbf{a}$ is a non-zero vector such that $\mathbf{a} = \sqrt{3}$ and $\mathbf{b}$ is a unit vector. Given that $\mathbf{a}$ and $\mathbf{b}$ are non-parallel and the angle between them is $5\pi/6$, find the exact value of the length of projection of $\mathbf{a}$ on $\mathbf{b}$. By considering $(2\mathbf{a} + \mathbf{b}) \cdot (2\mathbf{a} + \mathbf{b})$, or otherwise, find the exact value of $|2\mathbf{a} + \mathbf{b}|$.

Solution.

Part (a). $\mathbf{a}$ and $\mathbf{b}$ either have the same or opposite direction. Let $\mathbf{b} = \lambda\mathbf{a}$ for some $\lambda \in \mathbb{R}$.

$$\mathbf{a} = (\mathbf{a} \cdot \mathbf{b})\mathbf{b} = (\mathbf{a} \cdot \lambda\mathbf{a})\lambda\mathbf{a} = \lambda^2 |\mathbf{a}|^2 \mathbf{a} \implies \lambda^2 |\mathbf{a}|^2 = 1 \implies |\mathbf{b}| = |\lambda |\mathbf{a}|| = 1.$$

Part (b). Note that $|\mathbf{a} \cdot \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \cos(5\pi/6) = -3/2$. Hence, the length of projection of $\mathbf{a}$ on $\mathbf{b}$ is $\left|\mathbf{a} \cdot \hat{\mathbf{b}}\right| = 3/2$ units.

Observe that

$$|2\mathbf{a} + \mathbf{b}|^2 = (2\mathbf{a} + \mathbf{b}) \cdot (2\mathbf{a} + \mathbf{b}) = 4|\mathbf{a}|^2 + 4(\mathbf{a} \cdot \mathbf{b}) + |\mathbf{b}|^2 = 7.$$

Thus, $|2\mathbf{a} + \mathbf{b}| = \sqrt{7}$.

* * * * *

Problem 4. The points A, B, C, D have position vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ given by $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$, $\mathbf{b} = \mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$, $\mathbf{c} = 3\mathbf{i} + 2\mathbf{j} + \mathbf{k}$, $\mathbf{d} = 4\mathbf{i} - \mathbf{j} - \mathbf{k}$, respectively. The point P lies on AB produced such that $AP = 2AB$, and the point Q is the mid-point of AC.

(a) Show that PQ is perpendicular to AQ.

(b) Find the area of the triangle APQ.

(c) Find a vector perpendicular to the plane ABC.

(d) Find the cosine of the angle between $\overrightarrow{AD}$ and $\overrightarrow{BD}$.

Solution. Note that $\overrightarrow{AB} = (0,\ 0,\ -1)^\mathsf{T}$, $\overrightarrow{AC} = (2,\ 0,\ -2)^\mathsf{T}$ and $\overrightarrow{AD} = (3,\ -3,\ -4)^\mathsf{T}$.

Part (a). Note that

$$\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{AP} = \overrightarrow{OA} + 2\overrightarrow{AB} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

and

$$\overrightarrow{OQ} = \overrightarrow{OA} + \frac{1}{2}\overrightarrow{AC} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

Thus,

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \overrightarrow{AQ} = \overrightarrow{OQ} - \overrightarrow{OA} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Since $\overrightarrow{PQ} \cdot \overrightarrow{AQ} = 0$, the two vectors are perpendicular, whence $PQ \perp AQ$.

Part (b). Note that $\overrightarrow{AP} = (0,\, 0,\, -2)^{\mathsf{T}}$. Hence,

$$[\triangle APQ] = \frac{1}{2}\left|\overrightarrow{AP} \times \overrightarrow{AQ}\right| = \frac{1}{2}\left|\begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right| = 1 \text{ units}^2.$$

Part (c). The vector $\overrightarrow{AB} \times \overrightarrow{AC} = (0,\, -2,\, 0)^{\mathsf{T}}$ is perpendicular to the plane ABC.

Part (d). Let the angle between $\overrightarrow{AD}$ and $\overrightarrow{BD}$ be θ. Note that $\overrightarrow{BD} = -3\,(-1,\, 1,\, 1)^{\mathsf{T}}$. Hence,

$$\cos\theta = \frac{\overrightarrow{AD} \cdot \overrightarrow{BD}}{\left|\overrightarrow{AD}\right| \left|\overrightarrow{BD}\right|} = \frac{30}{\sqrt{34} \cdot 3\sqrt{3}} = \frac{10}{\sqrt{102}}.$$

A8 Vectors II - Lines

Tutorial A8

Problem 1. For each of the following, write down a vector equivalent of the line l and convert it to parametric and Cartesian forms.

(a) l passes through the point with position vector $-\mathbf{i}+\mathbf{k}$ and is parallel to the vector $\mathbf{i}+\mathbf{j}$.

(b) l passes through the points $P(1,-1,3)$ and $Q(2,1,-2)$.

(c) l passes through the origin and is parallel to the line $m : \mathbf{r} = (1,\ -1,\ 3)^{\mathsf{T}} + \lambda\,(1,\ 2,\ 3)^{\mathsf{T}}$, where $\lambda \in \mathbb{R}$.

(d) l is the x-axis.

(e) l passes through the point $C(4,-1,2)$ and is parallel to the z-axis.

Solution.

Part (a).

Form	Expression
Vector	$\mathbf{r} = (-1,\ 0,\ 1)^{\mathsf{T}} + \lambda\,(1,\ 1,\ 0)^{\mathsf{T}},\ \lambda \in \mathbb{R}$
Parametric	$x = \lambda - 1,\ y = \lambda,\ z = 1$
Cartesian	$x + 1 = y,\ z = 1$

Part (b).

Form	Expression
Vector	$\mathbf{r} = (1,\ -1,\ 3)^{\mathsf{T}} + \lambda\,(1,\ 2,\ -5)^{\mathsf{T}},\ \lambda \in \mathbb{R}$
Parametric	$x = \lambda + 1,\ y = 2\lambda - 1,\ z = -5\lambda + 3$
Cartesian	$x - 1 = \frac{y+1}{2} = \frac{3-z}{5}$

Part (c).

Form	Expression
Vector	$\mathbf{r} = \lambda\,(1,\ 2,\ 3)^{\mathsf{T}},\ \lambda \in \mathbb{R}$
Parametric	$x = \lambda,\ y = 2\lambda,\ z = 3\lambda$
Cartesian	$x = \frac{y}{2} = \frac{z}{3}$

Part (d).

Form	Expression
Vector	$\mathbf{r} = \lambda\,(1,\ 0,\ 0)^{\mathsf{T}},\ \lambda \in \mathbb{R}$
Parametric	$x = \lambda,\ y = 0,\ z = 0$
Cartesian	$x \in \mathbb{R},\ y = 0,\ z = 0$

Part (e).

Form	Expression
Vector	$\mathbf{r} = (4,\ -1,\ 2)^\mathsf{T} + \lambda\,(0,\ 0,\ 1)^\mathsf{T},\ \lambda \in \mathbb{R}$
Parametric	$x = 4,\ y = -1,\ z = \lambda + 2$
Cartesian	$x = 4,\ y = -1,\ z \in \mathbb{R}$

* * * * *

Problem 2. For each of the following, determine if l_1 and l_2 are parallel, intersecting or skew. In the case of intersecting lines, find the position vector of the point of intersection. In addition, find the acute angle between the lines l_1 and l_2.

(a) $l_1 : x - 1 = -y = z - 2$ and $l_2 : \frac{x-2}{2} = -\frac{y+1}{2} = \frac{z-4}{2}$

(b) $l_1 : \mathbf{r} = (1,\ 0,\ 0)^\mathsf{T} + \alpha\,(4,\ -2,\ -3)^\mathsf{T},\ \alpha \in \mathbb{R}$ and $l_2 : \mathbf{r} = (0,\ 10,\ 1)^\mathsf{T} + \beta\,(3,\ 8,\ 1)^\mathsf{T}$

(c) $l_1 : \mathbf{r} = (\mathbf{i} - 5\mathbf{k}) + \lambda(\mathbf{i} - \mathbf{j} + \mathbf{k}),\ \lambda \in \mathbb{R}$ and $l_2 : \mathbf{r} = (\mathbf{i} - \mathbf{j} + \mathbf{k}) + \mu(5\mathbf{i} - 4\mathbf{j} - \mathbf{k}),\ \mu \in \mathbb{R}$

Solution.

Part (a). Note that l_1 and l_2 have vector form

$$l_1 : \mathbf{r} = \begin{pmatrix}1\\0\\2\end{pmatrix} + \lambda\begin{pmatrix}1\\-1\\1\end{pmatrix},\ \lambda \in \mathbb{R} \text{ and } l_2 : \mathbf{r} = \begin{pmatrix}2\\1\\4\end{pmatrix} + \mu\begin{pmatrix}2\\-2\\2\end{pmatrix},\ \mu \in \mathbb{R}.$$

Since $(2,\ -2,\ 2)^\mathsf{T} = 2\,(1,\ -1,\ 1)^\mathsf{T}$, l_1 and l_2 are parallel ($\theta = 0$). Since $(1,\ 0,\ 2)^\mathsf{T} \neq (2,\ 1,\ 4)^\mathsf{T} + \mu\,(2,\ -2,\ 2)^\mathsf{T}$ for all real μ, we have that l_1 and l_2 are distinct.

Part (b). Since $(4,\ -2,\ 3)^\mathsf{T} \neq \beta\,(3,\ 8,\ 1)^\mathsf{T}$ for all real β, it follows that l_1 and l_2 are not parallel.

Consider $l_1 = l_2$.

$$l_1 = l_2 \implies \begin{pmatrix}1\\0\\0\end{pmatrix} + \alpha\begin{pmatrix}4\\-2\\-3\end{pmatrix} = \begin{pmatrix}0\\10\\1\end{pmatrix} + \beta\begin{pmatrix}3\\8\\1\end{pmatrix} \implies \alpha\begin{pmatrix}4\\-2\\-3\end{pmatrix} - \beta\begin{pmatrix}3\\8\\1\end{pmatrix} = \begin{pmatrix}-1\\10\\1\end{pmatrix}.$$

This gives the following system:

$$\begin{cases} 4\alpha - 3\beta = -1 \\ -2\alpha - 8\beta = 10 \\ -3\alpha - \ \beta = 1 \end{cases}$$

There are no solutions to the above system. Hence, l_1 and l_2 do not intersect and are thus skew.

Let θ be the acute angle between l_1 and l_2.

$$\cos\theta = \frac{\left|(4,\ -2,\ -3)^\mathsf{T} \cdot (3,\ 8,\ 1)^\mathsf{T}\right|}{\left|(4,\ -2,\ -3)^\mathsf{T}\right|\left|(3,\ 8,\ 1)^\mathsf{T}\right|} = \frac{7}{\sqrt{2146}} \implies \theta = 81.3^\circ \text{ (1 d.p.)}.$$

Part (c). Note that l_1 and l_2 have vector form

$$l_1 : \mathbf{r} = \begin{pmatrix}1\\0\\-5\end{pmatrix} + \lambda\begin{pmatrix}1\\-1\\1\end{pmatrix} \text{ and } l_2 : \mathbf{r} = \begin{pmatrix}1\\-1\\1\end{pmatrix} + \mu\begin{pmatrix}5\\-4\\-1\end{pmatrix},\ \lambda, \mu \in \mathbb{R}.$$

Since $(1,\ -1,\ 1)^{\mathsf{T}} \neq \mu\,(5,\ -4,\ -1)^{\mathsf{T}}$ for all real μ, it follows that l_1 and l_2 are not parallel. Consider $l_1 = l_2$.

$$l_1 = l_2 \implies \begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 5 \\ -4 \\ -1 \end{pmatrix} \implies \lambda \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - \mu \begin{pmatrix} 5 \\ -4 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 6 \end{pmatrix}.$$

This gives the following system:

$$\begin{cases} -5\mu + \lambda = 0 \\ 4\mu - \lambda = -1 \\ \mu + \lambda = 6 \end{cases}$$

The above system has the unique solution $\lambda = 5$ and $\mu = 1$. Hence, l_1 and l_2 intersect at $(1,\ 0,\ -5)^{\mathsf{T}} + 5\,(1,\ -1,\ 1)^{\mathsf{T}} = (6,\ -5,\ 0)^{\mathsf{T}}$.

Let θ be the acute angle between l_1 and l_2.

$$\cos\theta = \frac{\left|(1,\ -1,\ 1)^{\mathsf{T}} \cdot (5,\ -4,\ -1)^{\mathsf{T}}\right|}{\left|(1,\ -1,\ 1)^{\mathsf{T}}\right| \left|(5,\ -4,\ -1)^{\mathsf{T}}\right|} = \frac{8}{3\sqrt{14}} \implies \theta = 44.5^\circ \text{ (1 d.p.).}$$

* * * * *

Problem 3.

(a) Find the shortest distance from the point $(1, 2, 3)$ to the line with equation $\mathbf{r} = 3\mathbf{i} + 2\mathbf{j} + 4\mathbf{k} + \lambda(\mathbf{i} + 2\mathbf{j} + 2\mathbf{k})$, $\lambda \in \mathbb{R}$.

(b) Find the length of projection of $4\mathbf{i} - 5\mathbf{j} + 6\mathbf{k}$ onto the line with equation $\frac{x+5}{4} = \frac{y-5}{3} = 10 - 2z$.

(c) Find the projection of $4\mathbf{i} - 5\mathbf{j} + 6\mathbf{k}$ onto the line with equation $\frac{x+5}{4} = \frac{y-5}{3} = 10 - 2z$.

Solution.

Part (a). Let $\overrightarrow{OP} = (1,\ 2,\ 3)^{\mathsf{T}}$ and $\overrightarrow{OA} = (3,\ 2,\ 4)^{\mathsf{T}}$. Note that $\overrightarrow{AP} = (-2,\ 0,\ -1)^{\mathsf{T}}$. The shortest distance between P and the line is thus

$$\text{Shortest distance} = \frac{\left|(-2,\ 0,\ -1)^{\mathsf{T}} \times (1,\ 2,\ 2)^{\mathsf{T}}\right|}{\left|(1,\ 2,\ 2)^{\mathsf{T}}\right|} = \frac{\left|(2,\ -3,\ -4)^{\mathsf{T}}\right|}{3} = \frac{\sqrt{29}}{3} \text{ units.}$$

Part (b). Note that the line has vector form

$$\mathbf{r} = \begin{pmatrix} -5 \\ 5 \\ 5 \end{pmatrix} + \lambda' \begin{pmatrix} 4 \\ 3 \\ -1/2 \end{pmatrix} = \begin{pmatrix} -5 \\ 5 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 8 \\ 6 \\ -1 \end{pmatrix},\ \lambda \in \mathbb{R}.$$

The length of projection of $(4,\ -5,\ 6)^{\mathsf{T}}$ onto the line is thus given by

$$\text{Length of projection} = \frac{\left|(4,\ -5,\ 6)^{\mathsf{T}} \cdot (8,\ 6,\ -1)^{\mathsf{T}}\right|}{\left|(8,\ 6,\ -1)^{\mathsf{T}}\right|} = \frac{4}{\sqrt{101}} \text{ units.}$$

Part (c).

$$\text{Projection} = \left[\frac{(4,\ -5,\ 6)^{\mathsf{T}} \cdot (8,\ 6,\ -1)^{\mathsf{T}}}{\left|(8,\ 6,\ -1)^{\mathsf{T}}\right|}\right] \cdot \frac{(8,\ 6,\ -1)^{\mathsf{T}}}{\left|(8,\ 6,\ -1)^{\mathsf{T}}\right|} = \frac{-4}{101} \begin{pmatrix} 8 \\ 6 \\ -1 \end{pmatrix}$$

Problem 4. The points P and Q have coordinates $(0, -1, -1)$ and $(3, 0, 1)$ respectively, and the equations of the lines l_1 and l_2 are given by

$$l_1 : \mathbf{r} = \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \lambda \in \mathbb{R} \text{ and } l_2 : \mathbf{r} = \begin{pmatrix} -3 \\ 3 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \mu \in \mathbb{R}.$$

(a) Show that P lies on l_1 but not on l_2.

(b) Determine if l_2 passes through Q.

(c) Find the coordinates of the foot of the perpendicular from P to l_2. Hence, or otherwise, find the perpendicular distance from P to l_2.

(d) Find the length of projection of $\overrightarrow{PQ}$ onto l_2.

Solution. We have that $\overrightarrow{OP} = (0, -1, -1)^\mathsf{T}$ and $\overrightarrow{OQ} = (3, 0, 1)^\mathsf{T}$.

Part (a). When $\lambda = -2$, we have $(0, 1, -3)^\mathsf{T} - 2\,(0, 1, -1)^\mathsf{T} = (0, -1, -1)^\mathsf{T} = \overrightarrow{OP}$. Hence, P lies on l_1.

Observe that all points on l_2 have a z-coordinate of 1. Since P has a z-coordinate of -1, P does not lie on l_2.

Part (b). When $\mu = 3$, we have $(-3, 3, 1)^\mathsf{T} + 3\,(2, -1, 0)^\mathsf{T} = (3, 0, 1)^\mathsf{T} = \overrightarrow{OQ}$. Hence, l_2 passes through Q.

Part (c). Let the foot of the perpendicular from P to l_2 be F. Since F is on l_2, we have that $\overrightarrow{OF} = (-3, 3, 1)^\mathsf{T} + \mu\,(2, -1, 0)^\mathsf{T}$ for some real μ. We also have that $\overrightarrow{PF} \cdot (2, -1, 0)^\mathsf{T} = 0$. Note that

$$\overrightarrow{PF} = \overrightarrow{OF} - \overrightarrow{OP} = \begin{pmatrix} -3 \\ 3 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 + 2\mu \\ 4 - \mu \\ 2 \end{pmatrix}.$$

Hence,

$$\overrightarrow{PF} \cdot \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = 0 \implies \begin{pmatrix} -3 + 2\mu \\ 4 - \mu \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = 0 \implies -10 + 5\mu = 0 \implies \mu = 2.$$

Hence, $\overrightarrow{OF} = (-3, 3, 1)^\mathsf{T} + 2\,(3, -1, 0)^\mathsf{T} = (1, 1, 1)^\mathsf{T}$. Thus, $F(1, 1, 1)$. The perpendicular distance from P to l_2 is thus $\left|\overrightarrow{PF}\right| = \left|(1, 2, 2)^\mathsf{T}\right| = 3$ units.

Part (d). Note that $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$. The length of projection of $\overrightarrow{PQ}$ onto l_2 is thus given by

$$\text{Length of projection} = \frac{\left|(3, 1, 2)^\mathsf{T} \cdot (2, -1, 0)^\mathsf{T}\right|}{\left|(2, -1, 0)^\mathsf{T}\right|} = \frac{5}{\sqrt{5}} = \sqrt{5} \text{ units.}$$

* * * * *

Problem 5. The lines l_1 and l_2 have equations

$$\mathbf{r} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \text{ and } \mathbf{r} = \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} + t \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

respectively. Find the position vectors of the points P on l_1 and Q on l_2 such that O, P and Q are collinear, where O is the origin.

Solution. We have that $\overrightarrow{OP} = (0,\, 1,\, 2)^\mathsf{T} + s\,(1,\, 0,\, 3)^\mathsf{T}$ and $\overrightarrow{OQ} = (-2,\, 3,\, 1)^\mathsf{T} + t\,(2,\, 1,\, 0)^\mathsf{T}$ for some $s, t \in \mathbb{R}$. For O, P and Q to be collinear, we need $\overrightarrow{OP} = \lambda\overrightarrow{OQ}$ for some $\lambda \in \mathbb{R}$:

$$\begin{pmatrix}0\\1\\2\end{pmatrix} + s\begin{pmatrix}1\\0\\3\end{pmatrix} = \lambda\left[\begin{pmatrix}-2\\3\\1\end{pmatrix} + t\begin{pmatrix}2\\1\\0\end{pmatrix}\right] \implies \begin{pmatrix}s\\1\\2+3s\end{pmatrix} = \lambda\begin{pmatrix}-2+2t\\3+t\\1\end{pmatrix}.$$

This gives us the system:

$$\begin{cases} s = \lambda(-2+2t) \\ 1 = \lambda(3+t) \\ 2+3s = \lambda \end{cases}$$

Substituting the third equation into the first two gives the reduced system:

$$\begin{cases} s = (2+3s)(-2+2t) \\ 1 = (2+3s)(3+t) \end{cases}$$

Subtracting twice of the second equation from the first yields $s - 2 = -8(2+3s)$, whence $s = -14/25$. It quickly follows that $t = 1/8$. Hence,

$$\overrightarrow{OP} = \begin{pmatrix}0\\1\\2\end{pmatrix} - \frac{14}{25}\begin{pmatrix}1\\0\\3\end{pmatrix} = \frac{1}{25}\begin{pmatrix}-14\\25\\8\end{pmatrix}, \quad \overrightarrow{OQ} = \begin{pmatrix}-2\\3\\1\end{pmatrix} + \frac{1}{8}\begin{pmatrix}2\\1\\0\end{pmatrix} = \frac{1}{8}\begin{pmatrix}-14\\25\\8\end{pmatrix}.$$

* * * * *

Problem 6. Relative to the origin O, the points A, B and C have position vectors $5\mathbf{i} + 4\mathbf{j} + 10\mathbf{k}$, $-4\mathbf{i} + 4\mathbf{j} - 2\mathbf{k}$ and $-5\mathbf{i} + 9\mathbf{j} + 5\mathbf{k}$ respectively.

(a) Find the Cartesian equation of the line AB.

(b) Find the length of projection of $\overrightarrow{AC}$ onto the line AB. Hence, find the perpendicular distance from C to the line AB.

(c) Find the position vector of the foot N of the perpendicular from C to the line AB.

(d) The point D is such that it is a reflection of point C about the line AB. Find the position vector of D.

Solution. We have that $\overrightarrow{OA} = (5,\, 4,\, 10)^\mathsf{T}$, $\overrightarrow{OB} = (-4,\, 4,\, -2)^\mathsf{T}$ and $\overrightarrow{OC} = (-5,\, 9,\, 5)^\mathsf{T}$.

Part (a). Note that $\overrightarrow{AB} = (-9,\, 0,\, -12)^\mathsf{T} = -3\,(3,\, 0,\, 4)^\mathsf{T}$. The line AB hence has the vector form

$$\mathbf{r} = \begin{pmatrix}5\\4\\10\end{pmatrix} + \lambda\begin{pmatrix}3\\0\\4\end{pmatrix},\ \lambda \in \mathbb{R}$$

and Cartesian form $\frac{x-5}{3} = \frac{z-10}{4}$, $y = 4$.

Part (b). Note that $\overrightarrow{AC} = (-10,\, 5,\, -5)^\mathsf{T} = -5\,(2,\, -1,\, 1)^\mathsf{T}$. Hence, the length of projection of $\overrightarrow{AC}$ onto the line AB is given by

$$\text{Length of projection} = \frac{\left|\overrightarrow{AC} \cdot \overrightarrow{AB}\right|}{\left|\overrightarrow{AB}\right|} = \frac{1}{15}\left|5\begin{pmatrix}2\\-1\\1\end{pmatrix} \cdot 3\begin{pmatrix}3\\0\\4\end{pmatrix}\right| = 10 \text{ units.}$$

Since $\left|\overrightarrow{AC}\right| = 5\sqrt{6}$, the perpendicular distance from C to the line AB is $\sqrt{\left(5\sqrt{6}\right)^2 - 10^2} = 5\sqrt{2}$ units.

Part (c). Let $\overrightarrow{AN} = \lambda\,(-9,\ 0,\ -12)^\mathsf{T}$ for some $\lambda \in \mathbb{R}$ such that $\left|\overrightarrow{AN}\right| = 10$.

$$\left|\overrightarrow{AN}\right| = 10 \implies 15\lambda = 10 \implies \lambda = \frac{2}{3}.$$

Hence, $\overrightarrow{AN} = \frac{2}{3}\,(-9,\ 0,\ -12)^\mathsf{T} = (-6,\ 0,\ -8)^\mathsf{T}$. Thus, $\overrightarrow{ON} = \overrightarrow{OA} + \overrightarrow{AN} = (-1,\ 4,\ 2)^\mathsf{T}$.

Part (d). Note that $\overrightarrow{NC} = \overrightarrow{OC} - \overrightarrow{ON} = (-4,\ 5,\ 3)^\mathsf{T}$. Since D is the reflection of C about AB, we have that $\overrightarrow{ND} = -\overrightarrow{NC}$. Thus,

$$\overrightarrow{OD} = \overrightarrow{ON} + \overrightarrow{ND} = \overrightarrow{ON} - \overrightarrow{NC} = \begin{pmatrix}-1\\4\\2\end{pmatrix} - \begin{pmatrix}-4\\5\\3\end{pmatrix} = \begin{pmatrix}3\\-1\\-1\end{pmatrix}.$$

* * * * *

Problem 7. The points A and B have coordinates $(0, 9, c)$ and $(d, 5, -2)$ respectively, where c and d are constants. The line l has equation $\frac{x+3}{-1} = \frac{y-1}{4} = \frac{z-5}{3}$.

(a) Given that $d = 22/7$ and the line AB intersects l, find the value of c. Find also the coordinates of the foot of the perpendicular from A to l.

(b) Given instead that the lines AB and l are parallel, state the value of c and d and find the shortest distance between the lines AB and l.

Solution. We have that $\overrightarrow{OA} = (0,\ 9,\ c)^\mathsf{T}$ and $\overrightarrow{OB} = (d,\ 5,\ -2)^\mathsf{T}$. We also have that the line l is given by the vector $\mathbf{r} = (-3,\ 1,\ 5)^\mathsf{T} + \lambda\,(-1,\ 4,\ 3)^\mathsf{T}$ for $\lambda \in \mathbb{R}$.

Note that $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = (d,\ -4,\ -2-c)^\mathsf{T}$. Hence, the line AB is given by the vector $\mathbf{r}_{AB} = (d,\ 5,\ -2)^\mathsf{T} + \mu\,(d,\ -4,\ -2-c)^\mathsf{T}$ for $\mu \in \mathbb{R}$.

Part (a). Consider the direction vectors of AB and l. Since $(22/7,\ -4,\ -2-c)^\mathsf{T} \neq \lambda\,(-1,\ 4,\ 3)^\mathsf{T}$ for all real λ and c, the lines AB and l are not parallel. Hence, AB and l intersect at only one point. Thus, there must be a unique solution to $\mathbf{r} = \mathbf{r}_{AB}$.

$$\begin{aligned}\mathbf{r} = \mathbf{r}_{AB} &\implies \begin{pmatrix}-3\\1\\5\end{pmatrix} + \lambda\begin{pmatrix}-1\\4\\3\end{pmatrix} = \begin{pmatrix}22/7\\5\\-2\end{pmatrix} + \mu\begin{pmatrix}22/7\\-4\\-2-c\end{pmatrix}\\ &\implies \lambda\begin{pmatrix}-7\\28\\21\end{pmatrix} - \mu\begin{pmatrix}22\\-28\\-14-7c\end{pmatrix} = \begin{pmatrix}43\\28\\-49\end{pmatrix}\end{aligned}$$

This gives the following system:

$$\begin{cases}-\lambda - 22\mu = 43\\ 4\lambda + 28\mu = 28\\ 3\lambda + (14+7c)\,\mu = -49\end{cases}$$

Solving the first two equations gives $\lambda = 91/3$ and $\mu = -10/3$. It follows from the third equation that $c = 4$.

Let F be the foot of the perpendicular from A to l. We have that $\overrightarrow{OF} = (-3,\ 1,\ 5)^\mathsf{T} + \lambda\,(-1,\ 4,\ 3)^\mathsf{T}$ for some $\lambda \in \mathbb{R}$. We also have that $\overrightarrow{AF} \cdot (-1,\ 4,\ 3)^\mathsf{T} = 0$. Note that

$$\overrightarrow{AF} = \overrightarrow{OF} - \overrightarrow{OA} = \begin{pmatrix}-3-\lambda\\-8+4\lambda\\1+3\lambda\end{pmatrix}.$$

Hence,

$$\overrightarrow{AF} \cdot \begin{pmatrix} -1 \\ 4 \\ 3 \end{pmatrix} = 0 \implies \begin{pmatrix} -3-\lambda \\ -8+4\lambda \\ 1+3\lambda \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 4 \\ 3 \end{pmatrix} = 0 \implies -26+26\lambda = 0 \implies \lambda = 1.$$

Hence, $\overrightarrow{OF} = (-3,\ 1,\ 5)^\mathsf{T} + (-1,\ 4,\ 3)^\mathsf{T} = (-4,\ 5,\ 8)^\mathsf{T}$. The foot of the perpendicular from A to l hence has coordinates $(-4, 5, 8)$.

Part (b). Given that AB is parallel to l, one of their direction vectors must be a scalar multiple of the other. Hence, for some real λ, $(-1,\ 4,\ 3)^\mathsf{T} = \lambda\,(d,\ -4,\ -2-c)^\mathsf{T}$. It is obvious that $\lambda = -1$, whence $c = 1$ and $d = 1$.

Note that the direction vector of l and AB is $(-1,\ 4,\ 3)^\mathsf{T}$. Also note that l passes through $(-3, 1, 5)$ and AB passes through $(1, 5, -2)$. Since $(1,\ 5,\ -2)^\mathsf{T} - (-3,\ 1,\ 5)^\mathsf{T} = (4,\ 4,\ -7)^\mathsf{T}$, the shortest distance between AB and l is

$$\frac{\left|(-1,\ 4,\ 3)^\mathsf{T} \times (4,\ 4,\ -7)^\mathsf{T}\right|}{\left|(-1,\ 4,\ 3)^\mathsf{T}\right|} = \frac{1}{\sqrt{26}} \left| \begin{pmatrix} -40 \\ -5 \\ -20 \end{pmatrix} \right| = \frac{45}{\sqrt{26}} \text{ units.}$$

* * * * *

Problem 8. The equation of the line L is $\mathbf{r} = (1,\ 3,\ 7)^\mathsf{T} + t\,(2,\ -1,\ 5)^\mathsf{T}$, $t \in \mathbb{R}$. The points A and B have position vectors $(9,\ 3,\ 26)^\mathsf{T}$ and $(13,\ 9,\ \alpha)^\mathsf{T}$ respectively. The line L intersects the line through A and B at P.

(a) Find α and the acute angle between line L and AB.

The point C has position vector $(2,\ 5,\ 1)^\mathsf{T}$ and the foot of the perpendicular from C to L is Q.

(b) Find the position vector of Q. Hence, find the shortest distance from C to L.

(c) Find the position vector of the point of reflection of the point C about the line L. Hence, find the reflection of the line passing through C and the point $(1, 3, 7)$ about the line L.

Solution.

Part (a). Note that $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = (4,\ 6,\ \alpha - 26)^\mathsf{T}$. The line AB is thus given by $\mathbf{r}_{AB} = (9,\ 3,\ 26)^\mathsf{T} + u\,(4,\ 6,\ \alpha - 26)^\mathsf{T}$ for $u \in \mathbb{R}$. Note that AB is not parallel to L. Hence, $\overrightarrow{OP}$ is the only solution to the equation $\mathbf{r} = \mathbf{r}_{AB}$.

$$\begin{pmatrix} 1 \\ 3 \\ 7 \end{pmatrix} + t \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \\ 26 \end{pmatrix} + u \begin{pmatrix} 4 \\ 6 \\ \alpha - 26 \end{pmatrix} \implies t \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix} - u \begin{pmatrix} 4 \\ 6 \\ \alpha - 26 \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \\ 19 \end{pmatrix}.$$

This gives the following system:

$$\begin{cases} 2t - 4u = 8 \\ -t - 6u = 0 \\ 5t - (\alpha - 26)\,u = 19 \end{cases}$$

Solving the first two equations gives $t = 3$ and $u = -\frac{1}{2}$. It follows from the third equation that $\alpha = 34$.

Let the acute angle between L and AB be θ.

$$\cos\theta = \frac{\left|(2,\,-1,\,5)^\mathsf{T}\cdot(4,\,6,\,8)^\mathsf{T}\right|}{\left|(2,\,-1,\,5)^\mathsf{T}\right|\left|(4,\,6,\,8)^\mathsf{T}\right|} = \frac{42}{\sqrt{30}\sqrt{116}} \implies \theta = 44.6^\circ \text{ (1 d.p.)}.$$

Part (b). Since Q is on L, we have that $\overrightarrow{OQ} = (1,\,3,\,7)^\mathsf{T} + t\,(2,\,-1,\,5)^\mathsf{T}$ for some real t. Further, since $\overrightarrow{CQ} \perp L$, we have that $\overrightarrow{CQ}\cdot(2,\,-1,\,5)^\mathsf{T} = 0$. Note that

$$\overrightarrow{CQ} = \overrightarrow{OQ} - \overrightarrow{OC} = \begin{pmatrix}-1+2t\\-2-t\\6+5t\end{pmatrix}.$$

Thus,

$$\overrightarrow{CQ}\cdot\begin{pmatrix}2\\-1\\5\end{pmatrix} = 0 \implies \begin{pmatrix}-1+2t\\-2-t\\6+5t\end{pmatrix}\cdot\begin{pmatrix}2\\-1\\5\end{pmatrix} = 0 \implies 30+30t=0 \implies t=1.$$

Hence, $\overrightarrow{OQ} = (1,\,3,\,7)^\mathsf{T} + (2,\,-1,\,5)^\mathsf{T} = (-1,\,4,\,2)^\mathsf{T}$. The shortest distance from C to L is thus

$$\left|\overrightarrow{CQ}\right| = \left|\begin{pmatrix}-1\\4\\2\end{pmatrix} - \begin{pmatrix}2\\5\\1\end{pmatrix}\right| = \left|\begin{pmatrix}-3\\-1\\1\end{pmatrix}\right| = \sqrt{11} \text{ units}.$$

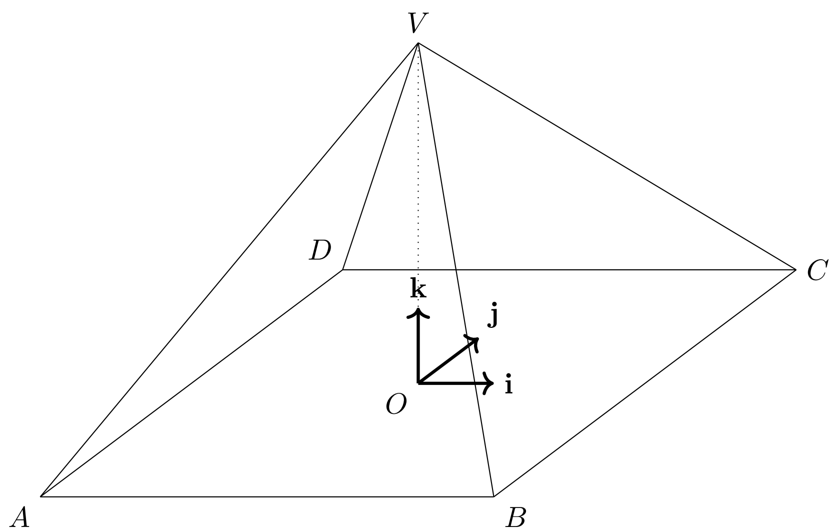
Part (c). Let C' be the reflection of C about L. Note that

$$\overrightarrow{OC'} = \overrightarrow{OQ} - \overrightarrow{QC} = \overrightarrow{OQ} + \overrightarrow{CQ} = \begin{pmatrix}-1\\4\\2\end{pmatrix} + \begin{pmatrix}-3\\-1\\1\end{pmatrix} = \begin{pmatrix}-4\\3\\3\end{pmatrix}.$$

Note that $(1,3,7)$ is on L and is hence invariant under a reflection about L. Let the reflection about L of the line passing through C and $(1,3,7)$ be L'. Since $(-4,\,3,\,3)^\mathsf{T} - (1,\,3,\,7)^\mathsf{T} = (-5,\,0,\,-4)^\mathsf{T} \parallel (5,\,0,\,4)^\mathsf{T}$, L' hence has direction vector $(5,\,0,\,4)^\mathsf{T}$. Thus, L' is given by $\mathbf{r}' = (1,\,3,\,7)^\mathsf{T} + \lambda\,(5,\,0,\,4)^\mathsf{T}$ for $\lambda\in\mathbb{R}$.

* * * * *

Problem 9.



In the diagram, O is the origin of the square base $ABCD$ of a right pyramid with vertex V. The perpendicular unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are parallel to AB, AD and OV respectively. The length of AB is 4 units and the length of OV is $2h$ units. P, Q, M and N are the mid-points of AB, BC, CV and VA respectively. The point O is taken as the origin for position vectors.

Show that the equation of the line PM may be expressed as $\mathbf{r} = (0,\ -2,\ 0)^\mathsf{T} + t\,(1,\ 3,\ h)^\mathsf{T}$, where t is a parameter.

(a) Find an equation for the line QN.

(b) Show that the lines PM and QN intersect and that the position vector $\overrightarrow{OX}$ of their point of intersection is $\mathbf{r} = \frac{1}{2}\,(1,\ -1,\ h)^\mathsf{T}$.

(c) Given that OX is perpendicular to VB, find the value of h and calculate the acute angle between PM and QN, giving your answer correct to the nearest $0.1°$.

Solution. We are given that $\overrightarrow{OP} = (0,\ -2,\ 0)^\mathsf{T}$, $\overrightarrow{OC} = (2,\ 2,\ 0)^\mathsf{T}$ and $\overrightarrow{OV} = (0,\ 0,\ 2h)^\mathsf{T}$. Hence, $\overrightarrow{CV} = \overrightarrow{OV} - \overrightarrow{OC} = (-2,\ -2,\ 2h)^\mathsf{T}$. Thus, $\overrightarrow{CM} = \frac{1}{2}\overrightarrow{CV} = (-1,\ -1,\ h)^\mathsf{T}$. Since $\overrightarrow{OM} = \overrightarrow{OC} + \overrightarrow{CM} = (1,\ 1,\ h)^\mathsf{T}$, we have that $\overrightarrow{PM} = \overrightarrow{OM} - \overrightarrow{OP} = (1,\ 3,\ h)^\mathsf{T}$. Thus, PM is given by

$$\mathbf{r} = \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 3 \\ h \end{pmatrix},\ t \in \mathbb{R}.$$

Part (a). Since $\overrightarrow{OM} = (1,\ 1,\ h)^\mathsf{T}$, by symmetry, $\overrightarrow{ON} = (-1,\ -1,\ h)^\mathsf{T}$. Given that $\overrightarrow{OQ} = (2,\ 0,\ 0)^\mathsf{T}$, we have that $\overrightarrow{QN} = \overrightarrow{ON} - \overrightarrow{OQ} = (-3,\ -1,\ h)^\mathsf{T}$. Thus, QN is given by

$$\mathbf{r} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + u\begin{pmatrix} -3 \\ -1 \\ h \end{pmatrix},\ u \in \mathbb{R}.$$

Part (b). Consider $PM = QN$.

$$PM = QN \implies \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 3 \\ h \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + u\begin{pmatrix} -3 \\ -1 \\ h \end{pmatrix} \implies t\begin{pmatrix} 1 \\ 3 \\ h \end{pmatrix} - u\begin{pmatrix} -3 \\ -1 \\ h \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}.$$

This gives the following system:

$$\begin{cases} t + 3u = 2 \\ 3t + u = 2 \\ ht - hu = 0 \end{cases}$$

From the first two equations, we see that $t = \frac{1}{2}$ and $u = \frac{1}{2}$, which is consistent with the third equation. Hence, $\overrightarrow{OX} = (0,\ -2,\ 0)^\mathsf{T} + \frac{1}{2}\,(1,\ 3,\ h)^\mathsf{T} = \frac{1}{2}\,(1,\ -1,\ h)^\mathsf{T}$.

Part (c). Note that $\overrightarrow{OB} = (2,\ -2,\ 6)^\mathsf{T}$, whence $\overrightarrow{VB} = \overrightarrow{OB} - \overrightarrow{OV} = (2,\ -2,\ -2h)^\mathsf{T}$. Since OX is perpendicular to VB, we have that $\overrightarrow{OX} \cdot \overrightarrow{VB} = 0$.

$$\overrightarrow{OX} \cdot \overrightarrow{VB} = 0 \implies \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ h \end{pmatrix} \cdot 2\begin{pmatrix} 1 \\ -1 \\ -h \end{pmatrix} = 0 \implies h^2 = 2.$$

We hence have that $h = \sqrt{2}$. Note that we reject $h = -\sqrt{2}$ since $h > 0$.

Let the acute angle between PM and QN be θ.

$$\cos\theta = \frac{\left|\overrightarrow{PM} \cdot \overrightarrow{QN}\right|}{\left|\overrightarrow{PM}\right|\left|\overrightarrow{QN}\right|} = \frac{1}{\sqrt{12}\sqrt{12}}\left|\begin{pmatrix} 1 \\ 3 \\ \sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} -3 \\ -1 \\ \sqrt{2} \end{pmatrix}\right| = \frac{1}{3} \implies \theta = 70.5°\ (1\ \text{d.p.}).$$

Self-Practice A8

Problem 1. The points A and B have positions vectors $(8, 3, 2)^{\mathsf{T}}$ and $(-2, 3, 4)^{\mathsf{T}}$ respectively.

(a) Show that $AB = 2\sqrt{26}$.

(b) Find the Cartesian equation for the line AB.

(c) The line l has equation $\mathbf{r} = (-2, 3, 4)^{\mathsf{T}} + t\,(2, 6, 5)^{\mathsf{T}}$. Find the length of the projection of AB onto l.

(d) Calculate the acute angle between AB and l, giving your answer correct to the nearest degree.

(e) Find the position vector of the foot N of the perpendicular from A to l. Hence, find the position vector of the image of A in the line l.

Solution.

Part (a). Note that

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \begin{pmatrix} -2 \\ 3 \\ 4 \end{pmatrix} - \begin{pmatrix} 8 \\ 3 \\ 2 \end{pmatrix} = 2\begin{pmatrix} -5 \\ 0 \\ 1 \end{pmatrix}.$$

Hence,

$$AB = \left|\overrightarrow{AB}\right| = 2\sqrt{(-5)^2 + 0^2 + 1^2} = 2\sqrt{26} \text{ units.}$$

Part (b). The vector equation of the line AB is

$$\mathbf{r} = \begin{pmatrix} 8 \\ 3 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} -5 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Hence, the Cartesian equation is

$$\frac{x-8}{-5} = z - 2, \; y = 3.$$

Part (c). The length of projection of AB onto l is given by

$$\frac{\left|2\,(-5, 0, 1)^{\mathsf{T}} \cdot (2, 6, 5)^{\mathsf{T}}\right|}{\left|(2, 6, 5)^{\mathsf{T}}\right|} = \frac{10}{\sqrt{65}} \text{ units.}$$

Part (d). Let the acute angle be θ.

$$\cos\theta = \frac{\left|(-5, 0, 1)^{\mathsf{T}} \cdot (2, 6, 5)^{\mathsf{T}}\right|}{\left|(-5, 0, 1)^{\mathsf{T}}\right|\left|(2, 6, 5)^{\mathsf{T}}\right|} = \frac{5}{\sqrt{65}\sqrt{26}} \implies \theta = 83^\circ.$$

Part (e). Since N is on l, there exists some $t \in \mathbb{R}$ such that

$$\overrightarrow{ON} = \frac{-2}{3}4 + t\begin{pmatrix} 2 \\ 6 \\ 5 \end{pmatrix}.$$

Hence,

$$\overrightarrow{AN} = \left[\begin{pmatrix}-2\\3\\4\end{pmatrix} + t\begin{pmatrix}2\\6\\5\end{pmatrix}\right] - \begin{pmatrix}8\\3\\2\end{pmatrix} = 2\begin{pmatrix}-5\\0\\1\end{pmatrix} + t\begin{pmatrix}2\\6\\5\end{pmatrix}.$$

Since AN is perpendicular to l, we have

$$\overrightarrow{AN} \cdot \begin{pmatrix}2\\6\\5\end{pmatrix} = \left[2\begin{pmatrix}-5\\0\\1\end{pmatrix} + t\begin{pmatrix}2\\6\\5\end{pmatrix}\right] \cdot \begin{pmatrix}2\\6\\5\end{pmatrix} = -10 + 65t = 0.$$

Hence, $t = 2/13$, whence

$$\overrightarrow{ON} = \begin{pmatrix}-2\\3\\4\end{pmatrix} + \frac{2}{13}\begin{pmatrix}2\\6\\5\end{pmatrix} = \frac{1}{13}\begin{pmatrix}-22\\51\\62\end{pmatrix}.$$

Let the image of A in l be A'. By the midpoint theorem,

$$\overrightarrow{ON} = \frac{\overrightarrow{OA} + \overrightarrow{OA'}}{2}.$$

Hence,

$$\overrightarrow{OA'} = 2\overrightarrow{ON} - \overrightarrow{OA} = \frac{2}{13}\begin{pmatrix}-22\\51\\62\end{pmatrix} - \begin{pmatrix}8\\3\\2\end{pmatrix} = \frac{1}{13}\begin{pmatrix}-148\\63\\98\end{pmatrix}.$$

* * * * *

Problem 2. The position vectors of the points A and B are $\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$ and $2\mathbf{i} + 3\mathbf{j} + p\mathbf{k}$ respectively, where p is a constant. The point C is such that $OABC$ is a rectangle, where O is the origin.

(a) Show that $p = 2$.

(b) Write down the position vector of C.

(c) Find a vector equation of the line BC.

The equation of line l is given by $\frac{x-1}{3} = \frac{y-1}{3}$, $z = 1$.

(e) Show that the lines BC and l are skew.

Solution.

Part (a). Note that

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \begin{pmatrix}2\\3\\p\end{pmatrix} - \begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}1\\1\\p-3\end{pmatrix}.$$

Since $OABC$ is a rectangle, $\overrightarrow{OA} \perp \overrightarrow{AB}$. Hence,

$$\overrightarrow{OA} \cdot \overrightarrow{AB} = \begin{pmatrix}1\\2\\3\end{pmatrix} \cdot \begin{pmatrix}1\\1\\p-3\end{pmatrix} = 3 + 3(p-3) = 0 \implies p = 2.$$

Part (b). Since $OABC$ is a rectangle,

$$\overrightarrow{OC} = \overrightarrow{AB} = (1,\ 1,\ -1)^\mathsf{T}.$$

Part (c). Since $OABC$ is a rectangle,

$$\overrightarrow{BC} = \overrightarrow{OA} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Thus, the vector equation of line BC is

$$l_{BC} : \mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Part (d). Note that the vector equation of l is

$$\mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}.$$

Consider $l \cap l_{BC}$:

$$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix} \implies \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \mu \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$$

This gives the system

$$\begin{cases} \lambda - 3\mu = 3 \\ 2\lambda - 3\mu = 3 \\ 3\lambda = 2 \end{cases},$$

which has no solution. Since the direction vectors of l and l_{BC} are not parallel (i.e. $(1,\ 2,\ 3)^\mathsf{T} \nparallel (3,\ 3,\ 0)^\mathsf{T}$), the two lines are skew.

* * * * *

Problem 3. The lines l_1 and l_2 have equations $\mathbf{r} = (3,\ 1,\ 2)^\mathsf{T} + \lambda\,(b,\ 1,\ -1)^\mathsf{T}$, where $b > 1$, and $\mathbf{r} = (4,\ 0,\ 1)^\mathsf{T} + \mu\,(-1,\ -1,\ 1)^\mathsf{T}$ respectively.

(a) Given that the acute angle between l_1 and l_2 is 30°, find the value of b, giving your answer correct to 2 decimal places.

For the rest of the question, use $b = 3$.

(b) Find the coordinates of the points A and B where l_1 and l_2 meet the xy-plane respectively.

(c) The point C has position vector $2\mathbf{i} + 7\mathbf{j} + 3\mathbf{k}$. Find whether C is closer to l_1 or l_2.

Solution.

Part (a).

$$\frac{\sqrt{3}}{2} = \cos 30^\circ = \frac{\left|(b,\ 1,\ -1)^\mathsf{T} \cdot (-1,\ -1,\ 1)^\mathsf{T}\right|}{\left|(b,\ 1,\ -1)^\mathsf{T}\right| \left|(-1,\ -1,\ 1)^\mathsf{T}\right|} = \frac{|-b-2|}{\sqrt{b^2+2}\sqrt{3}}.$$

Since $b > 1$, we clearly have $|-b-2| = b+2$. Thus,

$$\frac{b+2}{\sqrt{b^2+2}} = \frac{3}{2}.$$

Using G.C., we have $b = 0.13$ or $b = 3.07$. Since $b > 1$, we take $b = 3.07$.

Part (b). Note that the xy-plane has equation $z = 0$. Consider the intersection between l_1 and the xy-plane. Clearly, we need $\lambda = 2$, whence

$$\overrightarrow{OA} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + 2 \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \\ 0 \end{pmatrix},$$

and $A(9, 3, 0)$.

Consider the intersection between l_2 and the xy-plane. Clearly, we need $\mu = -1$, whence

$$\overrightarrow{OB} = \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix},$$

and $B(5, 1, 0)$.

Part (c). The perpendicular distance between C and l_1 is given by

$$\frac{\left|\left[(2,\,7,\,3)^\mathsf{T} - (3,\,1,\,2)^\mathsf{T}\right] \times (3,\,1,\,-1)^\mathsf{T}\right|}{\left|(3,\,1,\,-1)^\mathsf{T}\right|} = \frac{\left|(-7,\,2,\,19)^\mathsf{T}\right|}{\sqrt{11}} = \frac{\sqrt{414}}{\sqrt{11}} = 6.13 \text{ units.}$$

The perpendicular distance between C and l_2 is given by

$$\frac{\left|\left[(2,\,7,\,3)^\mathsf{T} - (4,\,0,\,1)^\mathsf{T}\right]\right| \times (-1,\,-1,\,1)^\mathsf{T}}{\left|(-1,\,-1,\,1)^\mathsf{T}\right|} = \frac{\left|(9,\,0,\,9)^\mathsf{T}\right|}{\sqrt{3}} = \frac{\sqrt{162}}{\sqrt{3}} = 7.35 \text{ units.}$$

Thus, C is closer to l_1.

* * * * *

Problem 4. Relative to an origin O, points C and D have position vectors $(7,\,3,\,2)^\mathsf{T}$ and $(10,\,a,\,b)^\mathsf{T}$ respectively, where a and b are constants.

(a) The straight line through C and D has equation $\mathbf{r} = (7,\,3,\,2)^\mathsf{T} + t\,(1,\,3,\,0)^\mathsf{T}$, $t \in \mathbb{R}$. Find the values of a and b.

(b) Find the position vector of the point P on the line CD such that $\overrightarrow{OP}$ is perpendicular to $\overrightarrow{CD}$.

(c) Find the position vector of the point Q on the line CD such that the angle between $\overrightarrow{OQ}$ and $\overrightarrow{OC}$ is equal to the angle between $\overrightarrow{OQ}$ and $\overrightarrow{OD}$.

Solution.

Part (a). Note that

$$\overrightarrow{CD} = \begin{pmatrix} 10 \\ a \\ b \end{pmatrix} - \begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ a-3 \\ b-2 \end{pmatrix}.$$

Since $\overrightarrow{CD}$ is parallel to $(1,\ 3,\ 0)^{\mathsf{T}}$, we have

$$\begin{pmatrix} 3 \\ a-3 \\ b-2 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix},$$

whence $a = 12$ and $b = 2$.

Part (b). Since P is on CD, there exists some $t \in \mathbb{R}$ such that

$$\overrightarrow{OP} = \begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix} + t\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}.$$

Since $\overrightarrow{OP}$ is perpendicular to $\overrightarrow{CD}$, we have

$$\overrightarrow{OP} \cdot \overrightarrow{CD} = \left[\begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix} + t\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}\right] \cdot 3\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} = 16 + 10t = 0,$$

whence $t = -8/5$ and

$$\overrightarrow{OP} = \begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix} - \frac{8}{5}\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 27 \\ -9 \\ 10 \end{pmatrix}.$$

Part (c). By the angle bisector theorem,

$$\frac{OC}{CQ} = \frac{OD}{DQ} \implies CQ : QD = OC : OD.$$

Since

$$OC = \left|\begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix}\right| = \sqrt{62} \quad \text{and} \quad OD = \left|\begin{pmatrix} 10 \\ 12 \\ 2 \end{pmatrix}\right| = \sqrt{248},$$

we have

$$CQ : QD = \sqrt{62} : \sqrt{248} = 1 : 2.$$

By the ratio theorem,

$$\overrightarrow{OQ} = \frac{\overrightarrow{OD} + 2\overrightarrow{OC}}{1+2} = \frac{1}{3}\left[\begin{pmatrix} 10 \\ 12 \\ 2 \end{pmatrix} + 2\begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix}\right] = \begin{pmatrix} 8 \\ 6 \\ 2 \end{pmatrix}.$$

* * * * *

Problem 5. Relative to an origin O, points A and B have position vectors $(3,\ 4,\ 1)^{\mathsf{T}}$ and $(-1,\ 2,\ 0)^{\mathsf{T}}$ respectively. The line l has vector equation $\mathbf{r} = (6,\ a,\ 0)^{\mathsf{T}} + t\,(1,\ 3,\ a)^{\mathsf{T}}$, where t is a real parameter and a is a constant. The line m passes through the point A and is parallel to the line OB.

(a) Find the position vector of the point P on m such that OP is perpendicular to m.

(b) Show that the two lines l and m have no common point.

(c) If the acute angle between the line l and the z-axis is $60°$, find the exact values of the constant a.

Solution.

Part (a). Note that the line m has vector equation

$$m : \mathbf{r} = \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} + s \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \quad s \in \mathbb{R}.$$

Since P is on m, there exists some $s \in \mathbb{R}$ such that

$$\overrightarrow{OP} = \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} + s \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}.$$

Since $\overrightarrow{OP}$ is perpendicular to m, we have

$$\overrightarrow{OP} \cdot \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \left[\begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} + s \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right] \cdot \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = 5 + 5s = 0,$$

whence $s = -1$ and

$$\overrightarrow{OP} = \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} - \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}.$$

Part (b). Consider $l \cap m$:

$$\begin{pmatrix} 6 \\ a \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 3 \\ a \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} + s \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}.$$

Comparing z-coordinates, we have

$$ta = 1 \implies t = \frac{1}{a}.$$

Substituting this into the equation, we get

$$\begin{pmatrix} 6 \\ a \\ 0 \end{pmatrix} + \frac{1}{a} \begin{pmatrix} 1 \\ 3 \\ a \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} + s \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}.$$

This yields the system

$$6 + \frac{1}{a} = 3 - s$$

$$a + \frac{3}{a} = 4 + 2s$$

Adding the second equation to twice the first yields

$$2\left(6 + \frac{1}{a}\right) + \left(a + \frac{3}{a}\right) = 2(3 - s) + (4 + 2s) \implies a + \frac{5}{a} + 2 = 0.$$

Multiplying through by a gives the quadratic

$$a^2 + 2a + 5 = (a + 1)^2 + 4 = 0,$$

which clearly has no real solution. Hence, $l \cap m$ has no solution, whence the two lines do not have any common point

Part (c). Note that the z-axis is parallel to the vector $(0,\ 0,\ 1)^\mathsf{T}$. Thus,

$$\frac{1}{2} = \cos 60^\circ = \frac{\left|(1,\ 3,\ a)^\mathsf{T} \cdot (0,\ 0,\ 1)^\mathsf{T}\right|}{\left|(1,\ 3,\ a)^\mathsf{T}\right| \left|(0,\ 0,\ 1)^\mathsf{T}\right|} = \frac{|a|}{\sqrt{10+a^2}\sqrt{1}}.$$

Squaring, we get

$$\frac{1}{4} = \frac{a^2}{10+a^2} \implies 10 + a^2 = 4a^2 \implies a^2 = \frac{10}{3} \implies a = \pm\sqrt{\frac{10}{3}}.$$

* * * * *

Problem 6. The lines l_1 and l_2 have vector equations

$$\mathbf{r} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

respectively, where λ and μ are real parameters.

(a) Find the acute angle between the two lines l_1 and l_2, giving your answer to the nearest 0.1°.

(b) Show that l_1 passes through the point P with position vector $(1,\ -4,\ 2)^\mathsf{T}$. Hence, show that the distance between point P and any point on the line l_2 is given by $\sqrt{6\mu^2 - 12\mu + 20}$. Deduce the shortest distance between point P and the line l_2.

Solution.

Part (a). Let the acute angle be θ. Then

$$\cos\theta = \frac{\left|(0,\ 2,\ 1)^\mathsf{T} \cdot (1,\ -2,\ 1)^\mathsf{T}\right|}{\left|(0,\ 2,\ 1)^\mathsf{T}\right| (1,\ -2,\ 1)^\mathsf{T}} = \frac{3}{\sqrt{5}\sqrt{6}} \implies \theta = 56.8^\circ.$$

Part (b). Take $\lambda = -1$. Then

$$\mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \\ 2 \end{pmatrix}.$$

Hence, l_1 passes through $P(1, -4, 2)$.

Note that l_2 has vector equation

$$\mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+\mu \\ -2\mu \\ 4+\mu \end{pmatrix}.$$

Hence,

$$\mathbf{r} - \overrightarrow{OP} = \begin{pmatrix} 1+\mu \\ -2\mu \\ 4+\mu \end{pmatrix} - \begin{pmatrix} 1 \\ -4 \\ 2 \end{pmatrix} = \begin{pmatrix} \mu \\ 4-2\mu \\ 2+\mu \end{pmatrix}.$$

Thus, the distance between P and any point on l_2 is given by

$$\begin{aligned} \sqrt{\mu^2 + (4-2\mu)^2 + (2+\mu)^2} &= \sqrt{\mu^2 + (4\mu^2 - 16\mu + 16) + (\mu^2 + 4\mu + 4)} \\ &= \sqrt{6\mu^2 - 12\mu + 20} \text{ units.} \end{aligned}$$

Since $6\mu^2 - 12\mu + 20 = 6(\mu+1)^2 + 13$, the shortest distance is $\sqrt{14}$ units.

* * * * *

Problem 7 (🌶). The coordinates of the points A, B and C are given by $A(0,2,4)$, $B(4,6,11)$ and $C(8,1,0)$.

(a) Show that the triangle with vertices A, B and C is an isosceles right-angled triangle.

(b) Find the position vector of point D in the same plane as A, B and C such that BCD is an equilateral triangle.

Solution.

Part (a). Observe that

$$\overrightarrow{AB} = \begin{pmatrix}4\\6\\11\end{pmatrix} - \begin{pmatrix}0\\2\\4\end{pmatrix} = \begin{pmatrix}4\\4\\7\end{pmatrix} \implies AB = \sqrt{4^2+4^2+7^2} = 9$$

and

$$\overrightarrow{CA} = \begin{pmatrix}0\\2\\4\end{pmatrix} - \begin{pmatrix}8\\1\\0\end{pmatrix} = \begin{pmatrix}-8\\1\\4\end{pmatrix} \implies AC = \sqrt{(-8)^2+1^2+4^2} = 9.$$

Since $AB = AC$, triangle ABC is isosceles.

Consider $\overrightarrow{AB} \cdot \overrightarrow{CA}$:

$$\overrightarrow{AB} \cdot \overrightarrow{CA} = \begin{pmatrix}4\\4\\7\end{pmatrix} \cdot \begin{pmatrix}-8\\1\\4\end{pmatrix} = -32 + 4 + 28 = 0.$$

Thus, $\overrightarrow{AB} \perp \overrightarrow{CA}$, whence triangle ABC is a right-angled triangle.

Hence, triangle ABC is an isosceles right-angled triangle.

Part (b). Let N be the foot of perpendicular of A on BC. Since $\triangle ABC$ is isosceles, with $AB = AC$, by symmetry, N is the midpoint of BC:

$$\overrightarrow{ON} = \frac{\overrightarrow{OB} + \overrightarrow{OC}}{2} = \frac{1}{2}\begin{pmatrix}12\\7\\11\end{pmatrix}.$$

Consider point D. Since $\triangle BCD$ is equilateral, it must also be isosceles, with $DB = DC$. Hence, D lies on AN (extended). Also, we have $ND/BC = \sin 60° = \sqrt{3}/2$.

Since

$$\overrightarrow{AN} = \frac{1}{2}\begin{pmatrix}12\\7\\11\end{pmatrix} - \begin{pmatrix}0\\2\\4\end{pmatrix} = \frac{3}{2}\begin{pmatrix}4\\1\\1\end{pmatrix},$$

the line AN has vector equation

$$\mathbf{r} = \begin{pmatrix}0\\2\\4\end{pmatrix} + \lambda\begin{pmatrix}4\\1\\1\end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Hence, there exists some $\lambda \in \mathbb{R}$ such that

$$\overrightarrow{OD} = \begin{pmatrix}0\\2\\4\end{pmatrix} + \lambda\begin{pmatrix}4\\1\\1\end{pmatrix}.$$

Thus,

$$\overrightarrow{ND} = \left[\begin{pmatrix}0\\2\\4\end{pmatrix} + \lambda\begin{pmatrix}4\\1\\1\end{pmatrix}\right] - \frac{1}{2}\begin{pmatrix}12\\7\\11\end{pmatrix} = \left(\lambda - \frac{3}{2}\right)\begin{pmatrix}4\\1\\1\end{pmatrix}.$$

Note that $\overrightarrow{BC} = (4,\ -5,\ -11)^\mathsf{T}$. Hence,

$$\frac{ND}{BC} = \frac{|\lambda - 3/2|\sqrt{4^2+1^2+1^2}}{\sqrt{4^2+(-5)^2+(-11)^2}} = \frac{|\lambda - 3/2|\sqrt{18}}{\sqrt{162}} = \frac{\sqrt{3}}{2}.$$

Rearranging, we get

$$\left|\lambda - \frac{3}{2}\right| = \frac{\sqrt{3}\sqrt{162}}{2\sqrt{18}} = \frac{3\sqrt{3}}{2} \implies \lambda = \frac{3 \pm 3\sqrt{3}}{2}.$$

Thus,

$$\overrightarrow{OD} = \begin{pmatrix}0\\2\\4\end{pmatrix} + \frac{3 \pm 3\sqrt{3}}{2}\begin{pmatrix}4\\1\\1\end{pmatrix}.$$

* * * * *

Problem 8 (🌶). The equations of the lines l_1 and l_2 are given by

$$l_1 : \mathbf{r} = \begin{pmatrix}1\\0\\0\end{pmatrix} + \lambda\begin{pmatrix}1\\2\\3\end{pmatrix},\ \lambda \in \mathbb{R} \quad \text{and} \quad l_2 : \mathbf{r} = \begin{pmatrix}1\\0\\0\end{pmatrix} + \mu\begin{pmatrix}1\\0\\-1\end{pmatrix},\ \mu \in \mathbb{R}.$$

(a) The point P with coordinates $(2, 2, 3)$ lies on the line l_1. Find the reflection of P in the line l_2.

(b) The line l_3 is the reflection of the line l_1 in the line l_2. Find an equation for the line l_3.

(c) The line l_4 is such that it is parallel to l_1 and its distance between the two lines is $\sqrt{13/14}$. Find two possible vector equations of l_4.

Solution.

Part (a). Let N be the foot of perpendicular of P on l_2. Since N lies on l_2, there exists some $\mu \in \mathbb{R}$ such that

$$\overrightarrow{ON} = \begin{pmatrix}1\\0\\0\end{pmatrix} + \mu\begin{pmatrix}1\\0\\-1\end{pmatrix}.$$

Thus,

$$\overrightarrow{PN} = \left[\begin{pmatrix}1\\0\\0\end{pmatrix} + \mu\begin{pmatrix}1\\0\\-1\end{pmatrix}\right] - \begin{pmatrix}2\\2\\3\end{pmatrix} = -\begin{pmatrix}1\\2\\3\end{pmatrix} + \mu\begin{pmatrix}1\\0\\-1\end{pmatrix}.$$

Since PN is perpendicular to l_2,

$$\overrightarrow{PN} \cdot \begin{pmatrix}1\\0\\-1\end{pmatrix} = \left[-\begin{pmatrix}1\\2\\3\end{pmatrix} + \mu\begin{pmatrix}1\\0\\-1\end{pmatrix}\right] \cdot \begin{pmatrix}1\\0\\-1\end{pmatrix} = 2 + 2\mu = 0,$$

whence $\mu = -1$ and

$$\overrightarrow{ON} = \begin{pmatrix}1\\0\\0\end{pmatrix} - \begin{pmatrix}1\\0\\-1\end{pmatrix} = \begin{pmatrix}0\\0\\1\end{pmatrix}.$$

Let P be the reflection of P in l_2. By the midpoint theorem,

$$\overrightarrow{ON} = \frac{\overrightarrow{OP}+\overrightarrow{OP'}}{2} \implies \overrightarrow{OP'} = 2\overrightarrow{ON} - \overrightarrow{OP} = 2\begin{pmatrix}0\\0\\1\end{pmatrix} - \begin{pmatrix}2\\2\\3\end{pmatrix} = -\begin{pmatrix}2\\2\\1\end{pmatrix}.$$

Part (b). Note that l_1 and l_2 have a common point $(1, 0, 0)$. Under reflection, this point is an invariant. Hence, l_3 must also contain the point $(1, 0, 0)$. Additionally, l_3 must contain P', the reflection of P in l_2. Since

$$-\begin{pmatrix}2\\2\\1\end{pmatrix} - \begin{pmatrix}1\\0\\0\end{pmatrix} = -\begin{pmatrix}3\\2\\1\end{pmatrix} \parallel \begin{pmatrix}3\\2\\1\end{pmatrix},$$

l_3 has vector equation

$$\lambda_3 : \mathbf{r} = \begin{pmatrix}1\\0\\0\end{pmatrix} + \nu\begin{pmatrix}3\\2\\1\end{pmatrix}, \quad \nu \in \mathbb{R}.$$

Part (c). Clearly, l_4 is given by

$$l_4 : \mathbf{r} = \begin{pmatrix}a\\b\\c\end{pmatrix} \xi \begin{pmatrix}1\\2\\3\end{pmatrix}, \quad \xi \in \mathbb{R}.$$

The perpendicular distance between l_1 and l_4 is hence given by

$$\frac{\left|\left[(a,\, b,\, c)^\mathsf{T} - (1,\, 0,\, 0)^\mathsf{T}\, 0\right] \times (1,\, 2,\, 3)^\mathsf{T}\right|}{\left|(1,\, 2,\, 3)^\mathsf{T}\right|} = \frac{\left|(3b-2c,\, c-3a+3,\, 2a-2-b)^\mathsf{T}\right|}{\sqrt{14}} = \frac{\sqrt{13}}{\sqrt{14}}.$$

Hence,

$$\left|\begin{pmatrix}3b-2c\\c-3a+3\\2a-2-b\end{pmatrix}\right| = \sqrt{13}.$$

This immediately gives

$$(3b-2c)^2 + (c-3a+3)^2 + (2a-2-b)^2 = 13.$$

Taking $a = 0$, $b = 0$, this reduces to

$$(-2c)^2 + (c+3)^2 + (-2)^2 = 13 \implies 5c^2 + 6c = 0 \implies c = 0 \text{ or } -\frac{6}{5}.$$

Thus,

$$l_4 : \mathbf{r} = \begin{pmatrix}0\\0\\0\end{pmatrix} + \xi\begin{pmatrix}1\\2\\3\end{pmatrix}, \quad \xi \in \mathbb{R}$$

or

$$l_4 : \mathbf{r} = \begin{pmatrix}0\\0\\-6/5\end{pmatrix} + \xi\begin{pmatrix}1\\2\\3\end{pmatrix}, \quad \xi \in \mathbb{R}.$$

Assignment A8

Problem 1. Find the position vector of the foot of the perpendicular from the point with position vector $\mathbf{c}$ to the line with equation $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, $\lambda \in \mathbb{R}$. Leave your answers in terms of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$.

Solution. Let the foot of the perpendicular be F. We have that $\overrightarrow{OF} = \mathbf{a} + \lambda\mathbf{b}$ for some real λ, and $\overrightarrow{CF} \cdot \mathbf{b} = 0$. Note that $\overrightarrow{CF} = \overrightarrow{OF} - \overrightarrow{OC} = \mathbf{a} + \lambda\mathbf{b} - \mathbf{c}$. Thus,

$$\overrightarrow{CF} \cdot \mathbf{b} = 0 \implies (\mathbf{a} + \lambda\mathbf{b} - \mathbf{c}) \cdot \mathbf{b} = 0 \implies \lambda\left|\mathbf{b}\right|^2 + (\mathbf{a} - \mathbf{c}) \cdot \mathbf{b} = 0 \implies \lambda = \frac{(\mathbf{c} - \mathbf{a}) \cdot \mathbf{b}}{\left|\mathbf{b}\right|^2}.$$

Thus,

$$\overrightarrow{OF} = \mathbf{a} + \left(\frac{(\mathbf{c} - \mathbf{a}) \cdot \mathbf{b}}{\left|\mathbf{b}\right|^2}\right)\mathbf{b}.$$

* * * * *

Problem 2. The point O is the origin, and points A, B, C have position vectors given by $\overrightarrow{OA} = 6\mathbf{i}$, $\overrightarrow{OB} = 3\mathbf{j}$, $\overrightarrow{OC} = 4\mathbf{k}$. The point P is on the line AB between A and B, and is such that $AP = 2PB$. The point Q has position vector given by $\overrightarrow{OQ} = q\mathbf{i}$, where q is a scalar.

(a) Express, in terms of $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, the vector $\overrightarrow{CP}$.

(b) Show that the line BQ has equation $\mathbf{r} = 3\mathbf{j} + t(q\mathbf{i} - 3\mathbf{j})$, where t is a parameter. Give an equation of the line CP in a similar form.

(c) Find the value of q for which the lines CP and BQ are perpendicular.

(d) Find the sine of the acute angle between the lines CP and BQ in terms of q.

Solution. We have that $\overrightarrow{OA} = (6,\ 0,\ 0)^\mathsf{T}$, $\overrightarrow{OB} = (0,\ 3,\ 0)^\mathsf{T}$ and $\overrightarrow{OC} = (0,\ 0,\ 4)^\mathsf{T}$.

Part (a). By the ratio theorem,

$$\overrightarrow{OP} = \frac{2\,\overrightarrow{OB} + \overrightarrow{OA}}{1+2} = \frac{1}{3}\left[2\begin{pmatrix}0\\3\\0\end{pmatrix} + \begin{pmatrix}6\\0\\0\end{pmatrix}\right] = \begin{pmatrix}2\\2\\0\end{pmatrix} \implies \overrightarrow{CP} = \overrightarrow{OP} - \overrightarrow{OC} = \begin{pmatrix}2\\2\\-4\end{pmatrix}.$$

Hence, $\overrightarrow{CP} = 2\mathbf{i} + 2\mathbf{j} - 4\mathbf{k}$.

Part (b). Note that $\overrightarrow{BQ} = \overrightarrow{OQ} - \overrightarrow{OB} = (q,\ -3,\ 0)^\mathsf{T}$. Thus, BQ is given by

$$\mathbf{r} = \begin{pmatrix}0\\3\\0\end{pmatrix} + t\begin{pmatrix}q\\-3\\0\end{pmatrix},\ t \in \mathbb{R} \iff \mathbf{r} = 3\mathbf{j} + t(q\mathbf{i} - 3\mathbf{j}),\ t \in \mathbb{R}.$$

Note that $\overrightarrow{CP} = (2,\ 2,\ -4)^\mathsf{T} = 2\,(1,\ 1,\ -2)^\mathsf{T}$. Hence, CP is given by

$$\mathbf{r} = \begin{pmatrix}0\\0\\4\end{pmatrix} + u\begin{pmatrix}1\\1\\-2\end{pmatrix},\ u \in \mathbb{R} \iff \mathbf{r} = 4\mathbf{k} + u(\mathbf{i} + \mathbf{j} - 2\mathbf{k}),\ u \in \mathbb{R}.$$

Part (c). Since CP is perpendicular to BQ, we have $\overrightarrow{CP} \cdot \overrightarrow{BQ} = 0$. Thus,

$$\overrightarrow{CP} \cdot \overrightarrow{BQ} = 0 \implies 2\begin{pmatrix}1\\1\\-2\end{pmatrix} \cdot \begin{pmatrix}q\\-3\\0\end{pmatrix} = 0 \implies q - 3 + 0 = 0 \implies q = 3.$$

Part (d). Let θ be the acute angle between CP and BQ.

$$\sin\theta = \frac{\left|(1,\,1,\,-2)^\mathsf{T} \times (q,\,-3,\,0)^\mathsf{T}\right|}{\left|(1,\,1,\,-2)^\mathsf{T}\right|\left|(q,\,-3,\,0)^\mathsf{T}\right|} = \frac{\left|(-6,\,2q,\,3-q)^\mathsf{T}\right|}{\sqrt{6}\sqrt{q^2+9}} = \sqrt{\frac{5q^2-6q+45}{6q^2+54}}.$$

* * * * *

Problem 3. Line l_1 passes through the point A with position vector $3\mathbf{i}-2\mathbf{k}$ and is parallel to $-2\mathbf{i}+4\mathbf{j}-\mathbf{j}$. Line l_2 has Cartesian equation given by $\frac{x-1}{2} = y = z+3$.

(a) Show that the two lines intersect and find the coordinates of their point of intersection.

(b) Find the acute angle between the two lines l_1 and l_2. Hence, or otherwise, find the shortest distance from point A to line l_2.

(c) Find the position vector of the foot N of the perpendicular from A to the line l_2. The point B lies on the line AN produced and is such that N is the mid-point of AB. Find the position vector of B.

Solution. We have

$$l_1 : \mathbf{r} = \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix},\ \lambda \in \mathbb{R}, \quad l_2 : \mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix},\ \mu \in \mathbb{R}.$$

Part (a). Consider $l_1 = l_2$.

$$l_1 = l_2 \implies \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \implies \mu \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \lambda \begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

This gives the following system:

$$\begin{cases} 2\lambda + 2\mu = 2 \\ -4\lambda + \mu = 0 \\ \lambda + \mu = 1 \end{cases}$$

which has the unique solution $\mu = 4/5$ and $\lambda = 1/5$. Thus, the intersection point P has position vector $(3,\,0,\,-2)^\mathsf{T} + \frac{1}{5}(-2,\,4,\,-1)^\mathsf{T} = \frac{1}{5}(13,\,4,\,-11)^\mathsf{T}$ and thus has coordinates $(13/5, 4/5, -11/5)$.

Part (b). Let θ be the acute angle between l_1 and l_2.

$$\cos\theta = \frac{\left|(-2,\,4,\,-1)^\mathsf{T} \cdot (2,\,1,\,1)^\mathsf{T}\right|}{\left|(-2,\,4,\,-1)^\mathsf{T}\right|\left|(2,\,1,\,1)^\mathsf{T}\right|} = \frac{1}{\sqrt{126}} \implies \theta = 84.9^\circ \text{ (1 d.p.)}.$$

Note that

$$AP = \sqrt{\left(\frac{17}{5}-3\right)^2 + \left(-\frac{4}{5}-0\right)^2 + \left(-\frac{9}{5}-(-2)\right)^2} = \sqrt{\frac{21}{25}} = \frac{\sqrt{21}}{5}.$$

Since $\sin\theta = \frac{AN}{AP}$, we have that $AN = AP\sin\theta$. Note that

$$\sin\theta = \sin\arccos\frac{1}{\sqrt{126}} = \frac{\sqrt{\left(\sqrt{126}\right)^2 - 1}}{\sqrt{126}} = \frac{\sqrt{125}}{\sqrt{126}} = \frac{5\sqrt{5}}{\sqrt{6}\sqrt{21}}.$$

Thus,

$$AN = \frac{\sqrt{21}}{5} \cdot \frac{5\sqrt{5}}{\sqrt{6}\sqrt{21}} = \sqrt{\frac{5}{6}}.$$

The shortest distance between A and l_2 is hence $\sqrt{\frac{5}{6}}$ units.

Part (c). Since N is on l_2, we have that $\overrightarrow{ON} = (1,\, 0,\, -3)^\mathsf{T} + \mu\,(2,\, 1,\, 1)^\mathsf{T}$ for some real μ. Additionally, since $\overrightarrow{AN} \perp l_2$, we have $\overrightarrow{AN} \cdot (2,\, 1,\, 1)^\mathsf{T} = 0$. Note that

$$\overrightarrow{AN} = \overrightarrow{ON} - \overrightarrow{OA} = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} -2+2\mu \\ \mu \\ -1+\mu \end{pmatrix}.$$

Thus,

$$\overrightarrow{AN} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = 0 \implies \begin{pmatrix} -2+2\mu \\ \mu \\ -1+\mu \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = 0 \implies -5+6\mu = 0 \implies \mu = \frac{5}{6}.$$

Hence,

$$\overrightarrow{ON} = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} + \frac{5}{6} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 16 \\ 5 \\ -13 \end{pmatrix}.$$

Note that $\overrightarrow{ON} = \frac{\overrightarrow{OA}+\overrightarrow{OB}}{2}$. Hence,

$$\overrightarrow{OB} = 2\overrightarrow{ON} - \overrightarrow{OA} = \frac{2}{6} \begin{pmatrix} 16 \\ 5 \\ -13 \end{pmatrix} - \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 7 \\ 5 \\ -7 \end{pmatrix}.$$

A9 Vectors III - Planes

Tutorial A9

Problem 1. A student claims that a unique plane can always be defined based on the given information. True or False? (Whenever a line is mentioned, assume the vector equation is known.)

	Statement	**T/F**
(a)	Any 2 vectors parallel to the plane and a point lying on the plane.	False
(b)	Any 3 distinct points lying on the plane.	False
(c)	A vector perpendicular to the plane and a point lying on the plane.	True
(d)	A line l perpendicular to the plane and a particular point on l lying on the plane.	True
(e)	A line l lying on the plane.	False
(f)	A line l and a point not on l, both lying on the plane.	True
(g)	A pair of distinct, intersecting lines, both lying on the plane.	True
(h)	A pair of distinct, parallel lines, both lying on the plane.	True
(i)	A pair of skew lines both parallel to the plane.	False
(j)	2 intersecting lines both parallel to the plane.	False

* * * * *

Problem 2. Find the equations of the following planes in parametric, scalar product and Cartesian form:

(a) The plane passes through the point with position vector $7\mathbf{i} + 2\mathbf{j} - 3\mathbf{j}$ and is parallel to $\mathbf{i} + 3\mathbf{j}$ and $4\mathbf{j} - 2\mathbf{k}$.

(b) The plane passes through the points $A(2, 0, 1)$, $B(1, -1, 2)$ and $C(1, 3, 1)$.

(c) The plane passes through the point with position vector $7\mathbf{i}$ and is parallel to the plane $\mathbf{r} = (2 - p + q)\mathbf{i} + (p + 3q)\mathbf{j} + (-2 - 3q)\mathbf{k}$, $p, q \in \mathbb{R}$.

(d) The plane contains the line $l : \mathbf{r} = (-2\mathbf{i} + 5\mathbf{j} - 3\mathbf{k}) + \lambda(2\mathbf{i} + \mathbf{j} + 2\mathbf{k})$, $\lambda \in \mathbb{R}$ and is perpendicular to the plane $\pi : \mathbf{r} \cdot (7\mathbf{i} + 4\mathbf{j} + 5\mathbf{k}) = 2$.

Solution.

Part (a). Parametric. Note that $(0, 4, -2)^\mathsf{T} \parallel (0, 2, -1)^\mathsf{T}$. Hence, the plane has parametric form

$$\mathbf{r} = \begin{pmatrix} 7 \\ 2 \\ -3 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}, \ \lambda, \mu \in \mathbb{R}.$$

Scalar Product. Note that $\mathbf{n} = (1, 3, 0)^\mathsf{T} \times (0, 2, -1)^\mathsf{T} = (-3, 1, 2)^\mathsf{T} \implies d = (7, 2, -3)^\mathsf{T} \cdot (-3, 1, 2)^\mathsf{T} = -25$. Thus, the plane has scalar product form

$$\mathbf{r} \cdot \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} = -2.$$

Cartesian. Let $\mathbf{r} = (x,\ y,\ z)^\mathsf{T}$. From the scalar product form, we have

$$-3x + y + 2z = -25.$$

Part (b). Parametric. Since the plane passes through the points A, B and C, it is parallel to both $\overrightarrow{AB} = -(1,\ 1,\ -1)^\mathsf{T}$ and $\overrightarrow{AC} = (-1,\ 3,\ 0)^\mathsf{T}$. Hence, the plane has parametric form

$$\mathbf{r} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix},\ \lambda, \mu \in \mathbb{R}.$$

Scalar Product. Note that $\mathbf{n} = (1,\ 1,\ -1)^\mathsf{T} \times (-1,\ 3,\ 0)^\mathsf{T} = (3,\ 1,\ 4)^\mathsf{T} \implies d = (2,\ 0,\ 1)^\mathsf{T} \cdot (3,\ 1,\ 4)^\mathsf{T} = 10$. Thus, the plane has scalar product form

$$\mathbf{r} \cdot \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} = 10.$$

Cartesian. Let $\mathbf{r} = (x,\ y,\ z)^\mathsf{T}$. From the scalar product form, we have

$$3x + y + 4z = 10.$$

Part (c). Parametric. Note that the plane is parallel to $\mathbf{r} = (2,\ 0,\ -1)^\mathsf{T} + p\,(-1,\ 1,\ 0)^\mathsf{T} + q\,(1,\ 3,\ -3)^\mathsf{T}$ and passes through $(7, 0, 0)$. Hence, the plane has parametric form

$$\mathbf{r} = \begin{pmatrix} 7 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 3 \\ -3 \end{pmatrix},\ \lambda, \mu \in \mathbb{R}.$$

Scalar Product. Note that $(-1,\ 1,\ 0)^\mathsf{T} \times (1,\ 3,\ -3)^\mathsf{T} = (-3,\ -3,\ -4)^\mathsf{T} \parallel (3,\ 3,\ 4)^\mathsf{T}$. We hence take $\mathbf{n} = (3,\ 3,\ 4)^\mathsf{T}$, whence $d = (7,\ 0,\ 0)^\mathsf{T} \cdot (3,\ 3,\ 4)^\mathsf{T} = 21$. Thus, the plane has scalar product form

$$\mathbf{r} \cdot \begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix} = 21.$$

Cartesian. Let $\mathbf{r} = (x,\ y,\ z)^\mathsf{T}$. From the scalar product form, we have

$$3x + 3y + 4z = 21.$$

Part (d). Parametric. Since the plane contains the line with equation $\mathbf{r} = (-2,\ 5,\ -3)^\mathsf{T} + \lambda\,(2,\ 1,\ 2)^\mathsf{T}$, $\lambda \in \mathbb{R}$, the plane passes through $(-2, 5, -3)$ and is parallel to the vector $(2,\ 1,\ 2)^\mathsf{T}$. Furthermore, since the plane is perpendicular to the plane with normal $(7,\ 4,\ 5)^\mathsf{T}$, it must be parallel to said vector. Thus, the plane has the following parametric form:

$$\mathbf{r} = \begin{pmatrix} -2 \\ 5 \\ -3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 7 \\ 4 \\ 5 \end{pmatrix},\ \lambda, \mu \in \mathbb{R}.$$

Scalar Product. Note that $\mathbf{n} = (2,\ 1,\ 2)^\mathsf{T} \times (7,\ 4,\ 5)^\mathsf{T} = (-3,\ 4,\ 1)^\mathsf{T} \implies d = (-2,\ 5,\ -3)^\mathsf{T} \cdot (-3,\ 4,\ 1)^\mathsf{T} = 23$. Thus, the plane has scalar product form

$$\mathbf{r} \cdot \begin{pmatrix} -3 \\ 4 \\ 1 \end{pmatrix} = 23.$$

Cartesian. Let $\mathbf{r} = (x,\, y,\, z)^\mathsf{T}$. From the scalar product form, we have

$$-3x + 4y + z = 23.$$

* * * * *

Problem 3. The line l passes through the points A and B with coordinates $(1, 2, 4)$ and $(-2, 3, 1)$ respectively. The plane p has equation $3x - y + 2z = 17$. Find

(a) the coordinates of the point of intersection of l and p,

(b) the acute angle between l and p,

(c) the perpendicular distance from A to p, and

(d) the position vector of the foot of the perpendicular from B to p.

The line m passes through the point C with position vector $6\mathbf{i} + \mathbf{j}$ and is parallel to $2\mathbf{j} + \mathbf{k}$.

(e) Determine whether m lies in p.

Solution. Note that $\overrightarrow{OA} = (1,\, 2,\, 4)^\mathsf{T}$ and $\overrightarrow{OB} = (-2,\, 3,\, 1)^\mathsf{T}$, whence $\overrightarrow{AB} = -\,(3,\, -1,\, 3)^\mathsf{T}$. Thus, the line l has vector equation

$$\mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix},\ \lambda \in \mathbb{R}.$$

Also note that the equation of the plane p can be written as

$$\mathbf{r} \cdot \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} = 17.$$

Part (a). Let the point of intersection of l and p be P. Consider $l = p$.

$$l = p \implies \left[\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix}\right] \cdot \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} = 17 \implies 9 + 16\lambda = 17 \implies \lambda = \frac{1}{2}.$$

Thus, $\overrightarrow{OP} = (1,\, 2,\, 4)^\mathsf{T} + \frac{1}{2}\,(3,\, -1,\, 3)^\mathsf{T} = (5/2,\, 3/2,\, 11/2)^\mathsf{T}$, whence $P(5/2, 3/2, 11/2)$.

Part (b). Let θ be the acute angle between l and p.

$$\sin\theta = \frac{\left|(3,\, -1,\, 3)^\mathsf{T} \cdot (3,\, -1,\, 2)^\mathsf{T}\right|}{\left|(3,\, -1,\, 3)^\mathsf{T}\right| \left|(3,\, -1,\, 2)^\mathsf{T}\right|} = \frac{16}{\sqrt{266}} \implies \theta = 78.8^\circ \text{ (1 d.p.)}.$$

Part (c). Note that $\overrightarrow{AP} = \frac{1}{2}\,(3,\, -1,\, 3)^\mathsf{T}$. The perpendicular distance from A to p is hence

$$\left|\overrightarrow{AP} \cdot \hat{\mathbf{n}}\right| = \frac{\left|\frac{1}{2}\,(3,\, -1,\, 3)^\mathsf{T} \cdot (3,\, -1,\, 2)^\mathsf{T}\right|}{\left|(3,\, -1,\, 2)^\mathsf{T}\right|} = \frac{8}{\sqrt{14}} \text{ units.}$$

Part (d). Let F be the foot of the perpendicular from B to p. Since F is on p, we have $\overrightarrow{OF} \cdot (3, -1, 2)^\mathsf{T} = 17$. Furthermore, since BF is perpendicular to p, we have $\overrightarrow{BF} = \lambda\mathbf{n} = \lambda\,(3, -1, 2)^\mathsf{T}$ for some $\lambda \in \mathbb{R}$. We hence have $\overrightarrow{OF} = \overrightarrow{OB} + \overrightarrow{BF} = (-2, 3, 1)^\mathsf{T} + \lambda\,(3, -1, 2)^\mathsf{T}$. Thus,

$$\left[\begin{pmatrix}-2\\3\\1\end{pmatrix} + \lambda\begin{pmatrix}3\\-1\\2\end{pmatrix}\right] \cdot \begin{pmatrix}3\\-1\\2\end{pmatrix} = 17 \implies -7 + 14\lambda = 17 \implies \lambda = \frac{12}{7}.$$

Hence, $\overrightarrow{OF} = (-2, 3, 1)^\mathsf{T} + \frac{12}{7}\,(3, -1, 2)^\mathsf{T} = \frac{1}{7}\,(22, 9, 31)^\mathsf{T}$.

Part (e). Note that m has the vector equation

$$\mathbf{r}_m = \begin{pmatrix}6\\1\\0\end{pmatrix} + \lambda\begin{pmatrix}0\\2\\1\end{pmatrix},\ \lambda \in \mathbb{R}.$$

Consider $\mathbf{r}_m \cdot \mathbf{n}$:

$$\mathbf{r}_m \cdot \mathbf{n} = \left[\begin{pmatrix}6\\1\\0\end{pmatrix} + \lambda\begin{pmatrix}0\\2\\1\end{pmatrix}\right] \cdot \begin{pmatrix}3\\-1\\2\end{pmatrix} = 17.$$

Since $\mathbf{r}_m \cdot \mathbf{n} = 17$ for all $\lambda \in \mathbb{R}$, it follows that m lies in p.

* * * * *

Problem 4. A plane contains distinct points P, Q, R and S, of which no 3 points are collinear. What can be said about the relationship between the vectors $\overrightarrow{PQ}$, $\overrightarrow{PR}$ and $\overrightarrow{PS}$?

Solution. Each of the three vectors can be expressed as a unique linear combination of the other two.

* * * * *

Problem 5.

(a) Interpret geometrically the vector equation $\mathbf{r} = \mathbf{a} + t\mathbf{b}$ where $\mathbf{a}$ and $\mathbf{b}$ are constant vectors and t is a parameter.

(b) Interpret geometrically the vector equation $\mathbf{r} \cdot \mathbf{n} = d$, where $\mathbf{n}$ is a constant unit vector and d is a constant scalar, stating what d represents.

(c) Given that $\mathbf{b} \cdot \mathbf{n} \neq 0$, solve the equations $\mathbf{r} = \mathbf{a} + t\mathbf{b}$ and $\mathbf{r} \cdot \mathbf{n} = d$ to find $\mathbf{r}$ in terms of $\mathbf{a}$, $\mathbf{b}$, $\mathbf{n}$ and d. Interpret the solution geometrically.

Solution.

Part (a). The vector equation $\mathbf{r} = \mathbf{a} + t\mathbf{b}$ represents a line with direction vector $\mathbf{b}$ that passes through the point with position vector $\mathbf{a}$.

Part (b). The vector equation $\mathbf{r} \cdot \mathbf{n} = d$ represents a plane perpendicular to $\mathbf{n}$ that has a perpendicular distance of d units from the origin. Here, a negative value of d corresponds to a plane d units from the origin in the opposite direction of $\mathbf{n}$.

Part (c).

$$\mathbf{r} \cdot \mathbf{n} = d \implies (\mathbf{a} + t\mathbf{b}) \cdot \mathbf{n} = d \implies \mathbf{a} \cdot \mathbf{n} + t\mathbf{b} \cdot \mathbf{n} = d$$
$$\implies t = \frac{d - \mathbf{a} \cdot \mathbf{n}}{\mathbf{b} \cdot \mathbf{n}} \implies \mathbf{r} = \mathbf{a} + \frac{d - \mathbf{a} \cdot \mathbf{n}}{\mathbf{b} \cdot \mathbf{n}}\,\mathbf{b}.$$

$\mathbf{a} + \frac{d - \mathbf{a} \cdot \mathbf{n}}{\mathbf{b} \cdot \mathbf{n}}\,\mathbf{b}$ is the position vector of the point of intersection of the line and plane.

Problem 6. The planes p_1 and p_2 have equations $\mathbf{r} \cdot (2,\,-2,\,1)^\mathsf{T} = 1$ and $\mathbf{r} \cdot (-6,\,3,\,2)^\mathsf{T} = -1$ respectively, and meet in the line l.

(a) Find the acute angle between p_1 and p_2.

(b) Find a vector equation for l.

(c) The point $A(4, 3, c)$ is equidistant from the planes p_1 and p_2. Calculate the two possible values of c.

Solution.

Part (a). Let θ the acute angle between p_1 and p_2.

$$\cos\theta = \frac{\left|(2,\,-2,\,1)^\mathsf{T} \cdot (-6,\,3,\,2)^\mathsf{T}\right|}{\left|(2,\,-2,\,1)^\mathsf{T}\right| \left|(-6,\,3,\,2)^\mathsf{T}\right|} = \frac{16}{21} \implies \theta = 40.4^\circ \text{ (1 d.p.)}.$$

Part (b). Observe that p_1 has the Cartesian equation $2x - 2y + z = 1$ and p_2 has the Cartesian equation $-6x + 3y + 2z = -1$. Consider $p_1 = p_2$. Solving both Cartesian equations simultaneously gives the solution

$$x = -\frac{1}{6} + \frac{7}{6}t, \quad y = -\frac{2}{3} + \frac{5}{3}t, \quad z = t$$

for all $t \in \mathbb{R}$. The line l thus has vector equation

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = -\frac{1}{6}\begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix} + t\begin{pmatrix} 7 \\ 10 \\ 6 \end{pmatrix}, \; t \in \mathbb{R}.$$

Part (c). Let Q be the point with position vector $-\frac{1}{6}(1,\,4,\,0)^\mathsf{T}$. Then $\overrightarrow{AQ} = -\frac{1}{6}(25,\,22,\,6c)^\mathsf{T}$. Since Q lies on l, it lies on both p_1 and p_2. Since A is equidistant to p_1 and p_2, the perpendicular distances from A to p_1 and p_2 are equal.

The perpendicular distance from A to p_1 is given by:

$$\frac{\left|\overrightarrow{AQ} \cdot (2,\,-2,\,1)^\mathsf{T}\right|}{\left|(2,\,-2,\,1)^\mathsf{T}\right|} = \frac{1}{3}\left|-\frac{1}{6}\begin{pmatrix} 25 \\ 22 \\ 6c \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}\right| = \frac{1}{3}\left|1 + c\right|.$$

Meanwhile, the perpendicular distance from A to p_2 is given by:

$$\frac{\left|\overrightarrow{AQ} \cdot (-6,\,3,\,2)^\mathsf{T}\right|}{\left|(-6,\,3,\,2)^\mathsf{T}\right|} = \frac{1}{7}\left|-\frac{1}{6}\begin{pmatrix} 25 \\ 22 \\ 6c \end{pmatrix} \cdot \begin{pmatrix} -6 \\ 3 \\ 2 \end{pmatrix}\right| = \frac{1}{7}\left|-14 + 2c\right|.$$

Equating the two gives

$$\frac{1}{3}\left|1 + c\right| = \frac{1}{7}\left|-14 + 2c\right| \implies \left|7 + 7c\right| = \left|-42 + 6c\right|.$$

This splits into the following two cases:

Case 1. $(7 + 7c)(-42 + 6c) > 0 \implies 7 + 7c = -42 + 6c \implies c = -49.$

Case 2. $(7 + 7c)(-42 + 6c) < 0 \implies 7 + 7c = -(-42 + 6c) \implies c = -35/13.$

Problem 7. A plane Π has equation $\mathbf{r} \cdot (2\mathbf{i} + 3\mathbf{j}) = -6$.

(a) Find, in vector form, an equation for the line passing through the point P with position vector $2\mathbf{i} + \mathbf{j} + 4\mathbf{k}$ and normal to the plane Π.

(b) Find the position vector of the foot Q of the perpendicular from P to the plane Π and hence find the position vector of the image of P after the reflection in the plane Π.

(c) Find the sine of the acute angle between OQ and the plane Π.

The plane Π' has equation $\mathbf{r} \cdot (\mathbf{i} + \mathbf{j} + \mathbf{k}) = 5$.

(d) Find the position vector of the point A where the planes Π, Π' and the plane with equation $\mathbf{r} \cdot \mathbf{i} = 0$ meet.

(e) Hence, or otherwise, find also the vector equation of the line of intersection of planes Π and Π'.

Solution.

Part (a). Let l be the required line. Since l is normal to Π, it is parallel to the normal vector of Π, $(2,\ 3,\ 0)^\mathsf{T}$. Thus, l has vector equation

$$l : \mathbf{r} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix},\ \lambda \in \mathbb{R}.$$

Part (b). Since Q is on Π, $\overrightarrow{OQ} \cdot (2,\ 3,\ 0)^\mathsf{T} = -6$. Furthermore, observe that Q is also on the line l. Thus, $\overrightarrow{OQ} = (2,\ 1,\ 4)^\mathsf{T} + \lambda\,(2,\ 3,\ 0)^\mathsf{T}$ for some $\lambda \in \mathbb{R}$. Hence,

$$\overrightarrow{OQ} \cdot \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = -6 \implies \left[\begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}\right] \cdot \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = -6 \implies 7 + 13\lambda = -6 \implies \lambda = -1.$$

Thus, $\overrightarrow{OQ} = (2,\ 1,\ 4)^\mathsf{T} - (2,\ 3,\ 0)^\mathsf{T} = (0,\ -2,\ 4)^\mathsf{T}$.

Let the reflection of P in Π be P'. Then

$$\overrightarrow{PQ} = \overrightarrow{QP'} \implies \overrightarrow{OQ} - \overrightarrow{OP} = \overrightarrow{OP'} - \overrightarrow{OQ} \implies \overrightarrow{OP'} = 2\overrightarrow{OQ} - \overrightarrow{OP}.$$

Hence, $\overrightarrow{OP'} = 2\,(0,\ -2,\ 4)^\mathsf{T} - (2,\ 1,\ 4)^\mathsf{T} = (-2,\ -5,\ 4)^\mathsf{T}$.

Part (c). Let θ be the acute angle between OQ and Π.

$$\sin\theta = \frac{\left|(0,\ -2,\ 4)^\mathsf{T} \cdot (2,\ 3,\ 0)^\mathsf{T}\right|}{\left|(0,\ -2,\ 4)^\mathsf{T}\right|\left|(2,\ 3,\ 0)^\mathsf{T}\right|} = \frac{3}{\sqrt{65}}.$$

Part (d). Let $\overrightarrow{OA} = (x,\ y,\ z)^\mathsf{T}$. We thus have the following system:

$$\begin{cases} (x,\ y,\ z)^\mathsf{T} \cdot (2,\ 3,\ 0)^\mathsf{T} = -6 & \implies 2x + 3y = -6 \\ (x,\ y,\ z)^\mathsf{T} \cdot (1,\ 1,\ 1)^\mathsf{T} = 5 & \implies x + y + z = 5 \\ (x,\ y,\ z)^\mathsf{T} \cdot (1,\ 0,\ 0)^\mathsf{T} = 0 & \implies x = 0 \end{cases}$$

Solving, we obtain $x = 0$, $y = -2$ and $z = 7$, whence $\overrightarrow{OA} = (0,\ -2,\ 7)^\mathsf{T}$.

Part (e). Let the line of intersection of Π and Π' be l'. Observe that A is on Π and Π' and thus lies on l'. Hence,

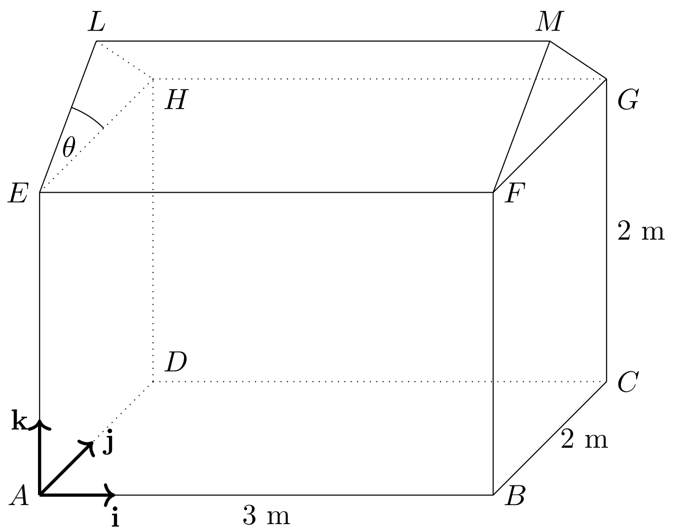
$$l' : \mathbf{r} = \begin{pmatrix} 0 \\ -2 \\ 7 \end{pmatrix} + \lambda \mathbf{b},\ \lambda \in \mathbb{R}.$$

Since l' lies on both Π and Π', $\mathbf{b}$ is perpendicular to the normals of both planes, i.e. $(2,\ 3,\ 0)^\mathsf{T}$ and $(1,\ 1,\ 1)^\mathsf{T}$. Thus, $\mathbf{b} = (2,\ 3,\ 0)^\mathsf{T} \times (1,\ 1,\ 1)^\mathsf{T} = (3,\ -2,\ -1)^\mathsf{T}$ and

$$l' : \mathbf{r} = \begin{pmatrix} 0 \\ -2 \\ 7 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ -2 \\ -1 \end{pmatrix},\ \lambda \in \mathbb{R}.$$

* * * * *

Problem 8.



The diagram shows a garden shed with horizontal base $ABCD$, where $AB = 3$ m and $BC = 2$ m. There are two vertical rectangular walls $ABFE$ and $DCGH$, where $AE = BF = CG = DH = 2$ m. The roof consists of two rectangular planes $EFML$ and $HGML$, which are inclined at an angle θ to the horizontal such that $\tan\theta = \frac{3}{4}$.

The point A is taken as the origin and the vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, each of length 1 m, are taken along AB, AD and AE respectively.

(a) Verify that the plane with equation $\mathbf{r} \cdot (22\mathbf{i} + 33\mathbf{j} - 12\mathbf{k}) = 66$ passes through B, D and M.

(b) Find the perpendicular distance, in metres, from A to the plane BDM.

(c) Find a vector equation of the straight line EM.

(d) Show that the perpendicular distance from C to the straight line EM is 2.91 m, correct to 3 significant figures.

Solution.

Part (a). We have $\overrightarrow{AB} = (3,\ 0,\ 0)^\mathsf{T}$, $\overrightarrow{BF} = \overrightarrow{AE} = (0,\ 0,\ 2)^\mathsf{T}$ and $\overrightarrow{FG} = \overrightarrow{AD} = (0,\ 2,\ 0)^\mathsf{T}$. Let T be the midpoint of FG. We have $\overrightarrow{FT} = (0,\ 1,\ 0)^\mathsf{T}$ and $TM/FT = \tan\theta = 3/4$, whence $\overrightarrow{TM} = (0,\ 0,\ 3/4)^\mathsf{T}$. Hence,

$$\overrightarrow{AM} = \overrightarrow{AB} + \overrightarrow{BF} + \overrightarrow{FT} + \overrightarrow{TM} = \begin{pmatrix}3\\0\\0\end{pmatrix} + \begin{pmatrix}0\\0\\2\end{pmatrix} + \begin{pmatrix}0\\1\\0\end{pmatrix} + \begin{pmatrix}0\\0\\3/4\end{pmatrix} = \frac{1}{4}\begin{pmatrix}12\\4\\11\end{pmatrix}.$$

Consider $\overrightarrow{AB} \cdot (22,\ 33,\ -12)^\mathsf{T}$, $\overrightarrow{AD} \cdot (22,\ 33,\ -12)^\mathsf{T}$ and $\overrightarrow{AM} \cdot (22,\ 33,\ -12)^\mathsf{T}$.

$$\begin{aligned}
\overrightarrow{AB} \cdot \begin{pmatrix}22\\33\\-12\end{pmatrix} &= \begin{pmatrix}3\\0\\0\end{pmatrix} \cdot \begin{pmatrix}22\\33\\-12\end{pmatrix} &= 66\\
\overrightarrow{AD} \cdot \begin{pmatrix}22\\33\\-12\end{pmatrix} &= \begin{pmatrix}0\\2\\0\end{pmatrix} \cdot \begin{pmatrix}22\\33\\-12\end{pmatrix} &= 66\\
\overrightarrow{AM} \cdot \begin{pmatrix}22\\33\\-12\end{pmatrix} &= \frac{1}{4}\begin{pmatrix}12\\4\\11\end{pmatrix} \cdot \begin{pmatrix}22\\33\\-12\end{pmatrix} &= 66
\end{aligned}$$

Since $\overrightarrow{AB}$, $\overrightarrow{AD}$ and $\overrightarrow{AM}$ satisfy the equation $\mathbf{r} \cdot (22,\ 33,\ -12)^\mathsf{T} = 66$, they all lie on the plane with said equation.

Part (b). The perpendicular distance from A to the plane BDM is given by

$$\text{Perpendicular distance} = \left|\overrightarrow{AB} \cdot \hat{\mathbf{n}}\right| = \frac{\left|(3,\ 0,\ 0)^\mathsf{T} \cdot (22,\ 33,\ -12)^\mathsf{T}\right|}{\left|(22,\ 33,\ -12)^\mathsf{T}\right|} = \frac{66}{\sqrt{1717}} \text{ m}.$$

Part (c). Observe that $\overrightarrow{EM} = \overrightarrow{AM} - \overrightarrow{AE} = \frac{1}{4}(12,\ 4,\ 3)^\mathsf{T}$. Hence, the line EM has vector equation

$$\mathbf{r} = \begin{pmatrix}0\\0\\2\end{pmatrix} + \lambda\begin{pmatrix}12\\4\\3\end{pmatrix},\ \lambda \in \mathbb{R}.$$

Part (d). Note that $\overrightarrow{EC} = \overrightarrow{AC} - \overrightarrow{AE} = (3,\ 2,\ -2)^\mathsf{T}$. The perpendicular distance from C to the line EM is hence given by

$$\frac{\left|\overrightarrow{EC} \times (12,\ 4,\ 3)^\mathsf{T}\right|}{\left|(12,\ 4,\ 3)^\mathsf{T}\right|} = \frac{1}{13}\left|\begin{pmatrix}3\\2\\-2\end{pmatrix} \times \begin{pmatrix}12\\4\\3\end{pmatrix}\right| = \frac{1}{13}\left|\begin{pmatrix}14\\-33\\-12\end{pmatrix}\right| = \frac{\sqrt{1429}}{13} = 2.91 \text{ m (3 s.f.)}.$$

* * * * *

Problem 9. The planes π_1 and π_2 have equations

$$x + y - z = 0 \text{ and } 2x - 4y + z + 12 = 0$$

respectively. The point P has coordinates $(3, 8, 2)$ and O is the origin.

(a) Verify that the vector $\mathbf{i} + \mathbf{j} + 2\mathbf{k}$ is parallel to both π_1 and π_2.

(b) Find the equation of the plane which passes through P and is perpendicular to both π_1 and π_2.

(c) Verify that $(0, 4, 4)$ is a point common to both π_1 and π_2, and hence or otherwise, find the equation of the line of intersection of π_1 and π_2, giving your answer in the form $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, $\lambda \in \mathbb{R}$.

(d) Find the coordinates of the point in which the line OP meets π_2.

(e) Find the length of projection of OP on π_1.

Solution. Note that π_1 and π_2 have vector equations $\mathbf{r}\cdot(1,\ 1,\ -1)^\mathsf{T} = 0$ and $\mathbf{r}\cdot(2,\ -4,\ 1)^\mathsf{T} = -12$ respectively.

Part (a). Observe that $(1,\ 1,\ 2)^\mathsf{T} \cdot (1,\ 1,\ -1)^\mathsf{T} = (1,\ 1,\ 2)^\mathsf{T} \cdot (2,\ -4,\ 1)^\mathsf{T} = 0$. Thus, the vector $(1,\ 1,\ 2)^\mathsf{T}$ is perpendicular to the normal vectors of both π_1 and π_2 and is hence parallel to them.

Part (b). Let the required plane be π_3. Since π_3 is perpendicular to both π_1 and π_2, its normal vector is parallel to both planes. Thus, $\mathbf{n} = (1,\ 1,\ 2)^\mathsf{T} \implies d = (3,\ 8,\ 2)^\mathsf{T} \cdot (1,\ 1,\ 2)^\mathsf{T} = 15$. π_3 hence has the vector equation

$$\mathbf{r} \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = 15.$$

Part (c). Since $(0,\ 4,\ 4)^\mathsf{T} \cdot (1,\ 1,\ -1)^\mathsf{T} = 0$ and $(0,\ 4,\ 4)^\mathsf{T} \cdot (2,\ -4,\ 1)^\mathsf{T} = -12$, $(0, 4, 4)$ satisfies the vector equation of both π_1 and π_2 and thus lies on both planes.

Let l be the line of intersection of π_1 and π_2. Since $(0, 4, 4)$ is a point common to both planes, l passes through it. Furthermore, since l lies on both π_1 and π_2, it is perpendicular to the normal vector of both planes and hence has direction vector $(1,\ 1,\ -1)^\mathsf{T} \times (2,\ -4,\ 1)^\mathsf{T} = -3\,(1,\ 1,\ 2)^\mathsf{T}$. Thus, l can be expressed as

$$l : \mathbf{r} = \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix},\ \lambda \in \mathbb{R}.$$

Part (d). Note that the line OP, denoted l_{OP} has equation

$$l_{OP} : \mathbf{r} = \mu \begin{pmatrix} 3 \\ 8 \\ 2 \end{pmatrix},\ \mu \in \mathbb{R}.$$

Consider the intersection between l_{OP} and π_2.

$$\mu \begin{pmatrix} 3 \\ 8 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -4 \\ 1 \end{pmatrix} = -12 \implies -24\mu = -12 \implies \mu = \frac{1}{2}.$$

Hence, OP meets π_2 at $(3/2, 4, 1)$.

Part (e). The length of projection of OP on π_1 is given by

$$\frac{\overrightarrow{OP} \times (1,\ 1,\ -1)^\mathsf{T}}{\left|(1,\ 1,\ -1)^\mathsf{T}\right|} = \frac{1}{\sqrt{3}} \left| \begin{pmatrix} 3 \\ 8 \\ 2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right| = \frac{1}{\sqrt{3}} \left| \begin{pmatrix} -10 \\ 5 \\ -5 \end{pmatrix} \right| = \frac{5\sqrt{6}}{\sqrt{3}} = 5\sqrt{2} \text{ units}.$$

Problem 10. The line l_1 passes through the point A, whose position vector is $3\mathbf{i}-5\mathbf{j}-4\mathbf{k}$, and is parallel to the vector $3\mathbf{i}+4\mathbf{j}+2\mathbf{k}$. The line l_2 passes through the point B, whose position vector is $2\mathbf{i}+3\mathbf{j}+5\mathbf{k}$, and is parallel to the vector $\mathbf{i}-\mathbf{j}-4\mathbf{k}$. The point P on l_1 and Q on l_2 are such that PQ is perpendicular to both l_1 and l_2. The plane Π contains PQ and l_1.

(a) Find a vector parallel to PQ.

(b) Find the equation of Π in the forms $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b} + \mu\mathbf{c}$, $\lambda, \mu \in \mathbb{R}$ and $\mathbf{r} \cdot \mathbf{n} = D$.

(c) Find the perpendicular distance from B to Π.

(d) Find the acute angle between Π and l_2.

(e) Find the position vectors of P and Q.

Solution.

Part (a). Note that l_1 and l_2 have vector equations

$$\mathbf{r} = \begin{pmatrix} 3 \\ -5 \\ -4 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}, \ \lambda \in \mathbb{R} \text{ and } \mathbf{r} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \\ -4 \end{pmatrix}, \ \mu \in \mathbb{R}$$

respectively. Since PQ is perpendicular to both l_1 and l_2, it is parallel to $(3,\ 4,\ 2)^\mathsf{T} \times (1,\ -1,\ -4)^\mathsf{T} = (-14,\ 14,\ -7)^\mathsf{T} = -7\,(2,\ -2,\ 1)^\mathsf{T}$.

Part (b). Since Π contains PQ and l_1, it is parallel to $(2,\ -2,\ 1)^\mathsf{T}$ and $(3,\ 4,\ 2)^\mathsf{T}$. Also note that Π contains $(3,\ -5,\ -4)^\mathsf{T}$. Thus,

$$\Pi : \mathbf{r} = \begin{pmatrix} 3 \\ -5 \\ -4 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}, \ \lambda, \mu \in \mathbb{R}.$$

Note that $(2,\ -2,\ 1)^\mathsf{T} \times (3,\ 4,\ 2)^\mathsf{T} = (-8,\ -1,\ 14)^\mathsf{T} \parallel (8,\ 1,\ -14)^\mathsf{T}$. We hence take $\mathbf{n} = (8,\ 1,\ -14)^\mathsf{T}$, whence $d = (3,\ -5,\ -4)^\mathsf{T} \cdot (8,\ 1,\ -14)^\mathsf{T} = 75$. Thus, $|Pi$ is also given by

$$\Pi : \mathbf{r} \cdot \begin{pmatrix} 8 \\ 1 \\ -14 \end{pmatrix} = 75.$$

Part (c). Note that $\overrightarrow{AB} = (-1,\ 8,\ 9)^\mathsf{T}$. Hence, the perpendicular distance from B to Π is given by

$$\frac{\left|(-1,\ 8,\ 9)^\mathsf{T} \cdot (8,\ 1,\ -14)^\mathsf{T}\right|}{\left|(8,\ 1,\ -14)^\mathsf{T}\right|} = \frac{126}{\sqrt{261}} \text{ units.}$$

Part (d). Let θ be the acute angle between Π and l_2.

$$\sin\theta = \frac{\left|(1,\ -1,\ -4)^\mathsf{T} \cdot (8,\ 1,\ -14)^\mathsf{T}\right|}{\left|(1,\ -1,\ -4)^\mathsf{T}\right| \left|(8,\ 1,\ -14)^\mathsf{T}\right|} = \frac{7}{\sqrt{58}} \implies \theta = 66.8^\circ \text{ (1 d.p.).}$$

Part (e). Since P is on l_1, we have $\overrightarrow{OP} = (3,\ -5,\ -4)^\mathsf{T} + \lambda\,(3,\ 4,\ 2)^\mathsf{T}$ for some $\lambda \in \mathbb{R}$. Similarly, since Q is on l_2, we have $\overrightarrow{OQ} = (2,\ 3,\ 5)^\mathsf{T} + \mu\,(1,\ -1,\ 4)^\mathsf{T}$ for some $\mu \in \mathbb{R}$. Thus,

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \begin{pmatrix} -1 \\ 8 \\ 9 \end{pmatrix} - \lambda \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \\ -4 \end{pmatrix}.$$

Recall that PQ is parallel to $(2,\ -2,\ 1)^\mathsf{T}$. Hence, $\overrightarrow{PQ}$ can be expressed as $\nu\,(2,\ -2,\ 1)^\mathsf{T}$ for some $\nu \in \mathbb{R}$. Equating the two expressions for $\overrightarrow{PQ}$, we obtain

$$\begin{pmatrix}-1\\8\\9\end{pmatrix} - \lambda\begin{pmatrix}3\\4\\2\end{pmatrix} + \mu\begin{pmatrix}1\\-1\\-4\end{pmatrix} = \nu\begin{pmatrix}2\\-2\\1\end{pmatrix} \implies \lambda\begin{pmatrix}3\\4\\2\end{pmatrix} + \mu\begin{pmatrix}-1\\1\\4\end{pmatrix} + \nu\begin{pmatrix}2\\-2\\1\end{pmatrix} = \begin{pmatrix}-1\\8\\9\end{pmatrix}.$$

This gives the following system:

$$\begin{cases} 3\lambda - \mu + 2\nu = -1 \\ 4\lambda + \mu - 2\nu = 8 \\ 2\lambda + 4\mu + \nu = 9 \end{cases}$$

which has the unique solution $\lambda = 1$, $\mu = 2$ and $\nu = -1$. Thus,

$$\overrightarrow{OP} = \begin{pmatrix}3\\-5\\-4\end{pmatrix} + \begin{pmatrix}3\\4\\2\end{pmatrix} = \begin{pmatrix}6\\-1\\-2\end{pmatrix}, \quad \overrightarrow{OQ} = \begin{pmatrix}2\\3\\5\end{pmatrix} + 2\begin{pmatrix}1\\-1\\-4\end{pmatrix} = \begin{pmatrix}4\\1\\-3\end{pmatrix}.$$

* * * * *

Problem 11. The equations of three planes p_1, p_2 and p_3 are

$$\begin{aligned} 2x - 5y + 3z &= 3 \\ 3x + 2y - 5z &= -5 \\ 5x + \lambda y + 17z &= \mu \end{aligned}$$

respectively, where λ and μ are constants. The planes p_1 and p_2 intersect in a line l.

(a) Find a vector equation of l.

(b) Given that all three planes meet in the line l, find λ and μ.

(c) Given instead that the three planes have no point in common, what can be said about the values of λ and μ?

(d) Find the Cartesian equation of the plane which contains l and the point $(1, -1, 3)$.

Solution.

Part (a). Consider the intersection of p_1 and p_2:

$$\begin{cases} 2x - 5y + 3z = 3 \\ 3x + 2y - 5z = -5 \end{cases}$$

The above system has solution

$$x = -1 + t, \quad y = -1 + t, \quad z = t$$

for all $t \in \mathbb{R}$. Thus, the line l has vector equation

$$l : \mathbf{r} = \begin{pmatrix}-1\\-1\\0\end{pmatrix} + t\begin{pmatrix}1\\1\\1\end{pmatrix}, \quad t \in \mathbb{R}.$$

Part (b). Since all three planes meet in the line l, l must satisfy the equation of p_3. Substituting the above solution to the given equation, we have

$$5(-1+t)+\lambda(-1+t)+17t=\mu \implies (22+\lambda)t-(5+\lambda+\mu)=0.$$

Comparing the coefficients of t and the constant terms, we have the following system:

$$\begin{cases} \lambda \qquad\quad + 22 = 0 \\ \lambda + \mu + \ \ 5 = 0 \end{cases}$$

which has the unique solution $\lambda = -22$ and $\mu = 17$.

Part (c). If the three planes have no point in common, we have

$$(22+\lambda)t-(5+\lambda+\mu)\neq 0$$

for all $t \in \mathbb{R}$. To satisfy this relation, we need $22+\lambda=0$ and $5+\lambda+\mu\neq 0$, whence $\lambda=-22$ and $\mu \neq 17$.

Part (d). Note that $(-1,\,-1,\,0)^{\mathsf{T}}$ lies on l and is thus contained on the required plane. Observe that $(-1,\,-1,\,0)^{\mathsf{T}}-(1,\,-1,\,3)^{\mathsf{T}}=(-2,\,0,\,-3)^{\mathsf{T}}$. Thus, the required plane is parallel to $(1,\,1,\,1)^{\mathsf{T}}$ and $(-2,\,0,\,-3)^{\mathsf{T}}$ and hence has vector equation

$$\mathbf{r}=\begin{pmatrix}-1\\-1\\0\end{pmatrix}+\lambda\begin{pmatrix}1\\1\\1\end{pmatrix}+\mu\begin{pmatrix}-2\\0\\-3\end{pmatrix}, \quad \lambda,\mu\in\mathbb{R}.$$

Observe that $\mathbf{n}=(1,\,1,\,1)^{\mathsf{T}}\times(-2,\,0,\,3)^{\mathsf{T}}=(-3,\,1,\,2)^{\mathsf{T}}$, whence $d=(-1,\,-1,\,0)^{\mathsf{T}}\cdot(-3,\,1,\,2)^{\mathsf{T}}=2$. The required plane thus has the equation

$$\mathbf{r}\cdot\begin{pmatrix}-3\\1\\2\end{pmatrix}=2.$$

Let $\mathbf{r}=(x,\,y,\,z)^{\mathsf{T}}$. It follows that the plane has Cartesian equation

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

* * * * *

Problem 12. The planes p_1 and p_2, which meet in line l, have equations $x-2y+2z=0$ and $2x-2y+z=0$ respectively.

(a) Find an equation of l in Cartesian form.

The plane p_3 has equation $(x-2y+2z)+c(2x-2y+z)=d$.

(b) Given that $d=0$, show that all 3 planes meet in the line l for any constant c.

(c) Given instead that the 3 planes have no point in common, what can be said about the value of d?

Solution.

Part (a). Consider the intersection of p_1 and p_2. This gives the system

$$\begin{cases} x - 2y + 2z = 0 \\ 2x - 2y + z = 0 \end{cases}$$

which has solution $x = t$, $y = \frac{3}{2}t$ and $z = t$. Thus, l has Cartesian equation

$$x = \frac{2}{3}y = z.$$

Part (b). When $d = 0$, p_3 has equation

$$(x - 2y + 2z) + c(2x - 2y + z) = 0.$$

Observe that the line l satisfies the equations $x - 2y + 2z = 0$ and $2x - 2y + z = 0$. Hence, l also satisfies the equation that gives p_3 for all c. Thus, p_3 contains l, implying that all 3 planes meet in the line l.

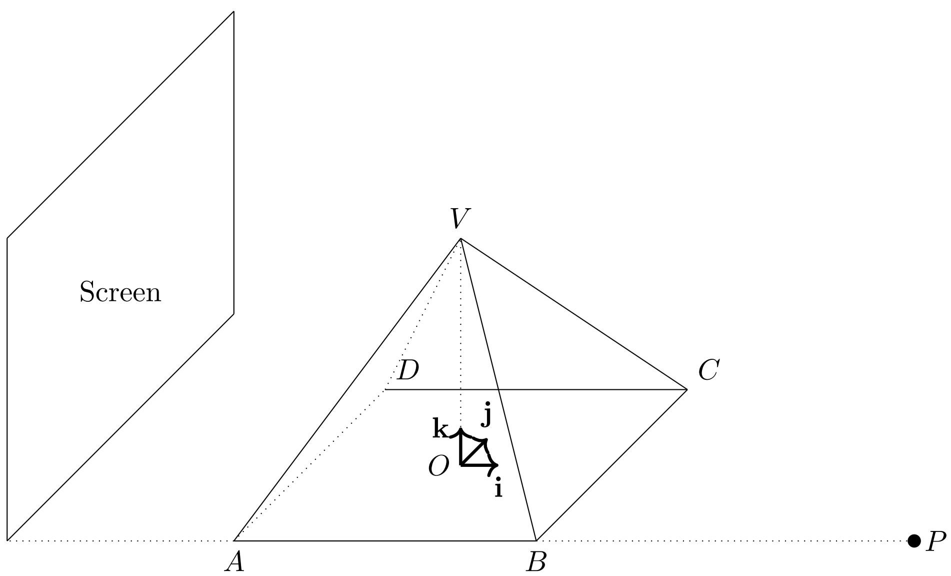
Part (c). If the 3 planes have no point in common, then l does not have any point in common with p_3. That is, all points on l satisfy the relation

$$(x - 2y + 2z) + c(2x - 2y + z) \neq d.$$

Since $x - 2y + 2z = 0$ and $2x - 2y + z = 0$ for all points on l, the LHS simplifies to 0. Thus, to satisfy the above relation, we require $d \neq 0$.

* * * * *

Problem 13.



A right opaque pyramid with square base $ABCD$ and vertex V is placed at ground level for a shadow display, as shown in the diagram. O is the centre of the square base $ABCD$, and the perpendicular unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are in the directions of $\overrightarrow{AB}$, $\overrightarrow{AD}$ and $\overrightarrow{OV}$ respectively. The length of AB is 8 units and the length of OV is $2h$ units.

A point light source for this shadow display is placed at the point $P(20, -4, 0)$ and a screen of height 35 units is placed with its base on the ground such that the screen lies on a plane with vector equation $\mathbf{r} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \alpha$, where $\alpha < -4$.

(a) Find a vector equation of the line depicting the path of the light ray from P to V in terms of h.

(b) Find an inequality between α and h so that the shadow of the pyramid cast on the screen will not exceed the height of the screen.

The point light source is now replaced by a parallel light source whose light rays are perpendicular to the screen. It is also given that $h = 10$.

(c) Find the exact length of the shadow cast by the edge VB on the screen.

A mirror is placed on the plane VBC to create a special effect during the display.

(d) Find a vector equation of the plane VBC and hence find the angle of inclination made by the mirror with the ground.

Solution.

Part (a). Note that $\overrightarrow{OV} = (0,\ 0,\ 2h)^\mathsf{T}$ and $\overrightarrow{OP} = (20,\ -4,\ 0)^\mathsf{T}$, whence $\overrightarrow{PV} = (-20,\ 4,\ 2h)^\mathsf{T} = 2\,(-10,\ 2,\ h)^\mathsf{T}$. Thus, the line from P to V, denoted l_{PV}, has the vector equation

$$l_{PV} : \mathbf{r} = \begin{pmatrix} 20 \\ -4 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -10 \\ 2 \\ h \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Part (b). Let the point of intersection between l_{PV} and the screen be I.

$$\left[\begin{pmatrix} 20 \\ -4 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -10 \\ 2 \\ h \end{pmatrix}\right] \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \alpha \implies 20 - 10\lambda = \alpha \implies \lambda = \frac{20-\alpha}{10}.$$

Hence, $\overrightarrow{OI} = (20,\ -4,\ 0)^\mathsf{T} + \frac{20-\alpha}{10}\,(-10,\ 2,\ h)^\mathsf{T}$. To prevent the shadow from exceeding the screen, we require the $\mathbf{k}$-component of $\overrightarrow{OI}$ to be less than the height of the screen, i.e. 35 units. This gives the inequality $\frac{20-\alpha}{10} \cdot h \leq 35$, whence we obtain

$$h \leq \frac{350}{20-\alpha}.$$

Part (c). Since the light rays emitted by the light source are now perpendicular to the screen, the image of some point with coordinates (a, b, c) on the screen is given by (α, b, c). Thus, the image of $B(4, -4, 0)$ and $V(0, 0, 20)$ on the screen have coordinates $(\alpha, -4, 0)$ and $(\alpha, 0, 20)$. The length of the shadow cast by VB is thus

$$\sqrt{(\alpha-\alpha)^2 + (-4-0)^2 + (0-20)^2} = 4\sqrt{26} \text{ units}.$$

Part (d). Note that $\overrightarrow{BV} = 4\,(-1,\ 1,\ 5)^\mathsf{T}$ and $\overrightarrow{BC} = 8\,(0,\ 1,\ 0)^\mathsf{T}$. Hence, the plane VBC is parallel to $(-1,\ 1,\ 5)^\mathsf{T}$ and $(0,\ 1,\ 0)^\mathsf{T}$. Note that $(-1,\ 1,\ 5)^\mathsf{T} \times (0,\ 1,\ 0)^\mathsf{T} = -\,(5,\ 0,\ 1)^\mathsf{T}$. Thus, $\mathbf{n} = (5,\ 0,\ 1)^\mathsf{T}$, whence $d = (0,\ 0,\ 20)^\mathsf{T} \cdot (5,\ 0,\ 1)^\mathsf{T} = 20$. Thus, the plane VBC has the vector equation

$$\mathbf{r} \cdot \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix} = 20.$$

Observe that the ground is given by the vector equation $\mathbf{r} \cdot (0,\ 0,\ 1)^\mathsf{T} = 0$. Let θ be the angle of inclination made by the mirror with the ground.

$$\cos\theta = \frac{(5,\ 0,\ 1)^\mathsf{T} \cdot (0,\ 0,\ 1)^\mathsf{T}}{\left|(5,\ 0,\ 1)^\mathsf{T}\right| \left|(0,\ 0,\ 1)^\mathsf{T}\right|} = \frac{1}{\sqrt{26}} \implies \theta = 78.7^\circ \text{ (1 d.p.)}.$$

Self-Practice A9

Problem 1. The position vectors of the vertices of A, B and C of a triangle are $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ respectively.

If O is the origin, show that the area of triangle OAB is $\frac{1}{2}|\mathbf{a}\times\mathbf{b}|$ and deduce an expression for the area of the triangle ABC.

Hence, or otherwise, show that the perpendicular distance from B to AC is

$$\frac{|\mathbf{a}\times\mathbf{b}+\mathbf{b}\times\mathbf{c}+\mathbf{c}\times\mathbf{a}|}{|\mathbf{c}-\mathbf{a}|}.$$

Solution. Let θ be the angle between OA and OB. Then

$$[\triangle OAB]=\frac{1}{2}(OA)(OB)\sin\theta=\frac{1}{2}\left|\overrightarrow{OA}\right|\left|\overrightarrow{OB}\right|\sin\theta=\frac{1}{2}|\mathbf{a}\times\mathbf{b}|.$$

Similarly, let φ be the angle between AB and AC. Then

$$\begin{aligned}[\triangle ABC]&=\frac{1}{2}(AB)(AC)\sin\varphi=\frac{1}{2}\left|\overrightarrow{AB}\times\overrightarrow{AC}\right|=\frac{1}{2}|(\mathbf{b}-\mathbf{a})\times(\mathbf{c}-\mathbf{a})|\\&=\frac{1}{2}|\mathbf{b}\times\mathbf{c}-\mathbf{a}\times\mathbf{c}-\mathbf{b}\times\mathbf{a}+\mathbf{a}\times\mathbf{a}|=\frac{1}{2}|\mathbf{b}\times\mathbf{c}-\mathbf{a}\times\mathbf{c}-\mathbf{b}\times\mathbf{a}|\\&=\frac{1}{2}|\mathbf{a}\times\mathbf{b}+\mathbf{b}\times\mathbf{c}+\mathbf{c}\times\mathbf{a}|.\end{aligned}$$

Let h be the perpendicular distance from B to AC. Then

$$[\triangle ABC]=\frac{1}{2}(AC)(h)\implies h=\frac{2[\triangle ABC]}{\left|\overrightarrow{AC}\right|}=\frac{|\mathbf{a}\times\mathbf{b}+\mathbf{b}\times\mathbf{c}+\mathbf{c}\times\mathbf{a}|}{|\mathbf{c}-\mathbf{a}|}.$$

* * * * *

Problem 2. Points A, B, C and D have position vectors, relative to the origin O, given by $\overrightarrow{OA}=\mathbf{i}+2\mathbf{j}-\mathbf{k}$, $\overrightarrow{OB}=-\mathbf{i}+2\mathbf{j}+c\mathbf{k}$, $\overrightarrow{OC}=2\mathbf{i}+\mathbf{j}+4\mathbf{k}$ and $\overrightarrow{OD}=\mathbf{i}+\mathbf{j}+\mathbf{k}$, where c is a constant. It is given that OA and OB are perpendicular.

(a) Find the value of c.

(b) Show that OA is normal to the plane OBC.

(c) Find an equation of the plane through D and parallel to OBC.

Also, find the position vector of the point of intersection of this plane and the line AC. Find the acute angle between the plane OBC and the plane through D normal to OD.

Solution.

Part (a). Since OA and OB are perpendicular, we have

$$\overrightarrow{OA}\cdot\overrightarrow{OB}=\begin{pmatrix}1\\2\\-1\end{pmatrix}\cdot\begin{pmatrix}-1\\2\\c\end{pmatrix}=3-c=0\implies c=3.$$

Part (b). The normal vector of the plane OBC is given by

$$\overrightarrow{OB}\times\overrightarrow{OC}=\begin{pmatrix}-1\\2\\3\end{pmatrix}\times\begin{pmatrix}2\\1\\4\end{pmatrix}=5\begin{pmatrix}1\\2\\-1\end{pmatrix}=5\overrightarrow{OA},$$

hence $\overrightarrow{OA}$ is normal to the plane OBC.

Part (c). Note that

$$\overrightarrow{OD} \cdot \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = 2.$$

Thus, the equation of the plane through D and parallel to OBC is given by

$$\Pi : \quad \mathbf{r} \cdot \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = 2.$$

Note that the line AC has vector equation

$$\mathbf{r} = \overrightarrow{OA} + \lambda \overrightarrow{AC} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \lambda \left[\begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \right] = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

When this line intersects Π, we have

$$\left[\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix} \right] \cdot \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = 6 - 6\lambda = 2 \implies \lambda = \frac{2}{3}.$$

Thus, the point of intersection has position vector

$$\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \frac{2}{3} \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 5 \\ 4 \\ 7 \end{pmatrix}.$$

Let θ be the acute angle between the plane OBC and the plane through D normal to OD. Then

$$\cos\theta = \frac{\left| (1,\, 1,\, 1)^\mathsf{T} \cdot (1,\, 2,\, -1)^\mathsf{T} \right|}{\left| (1,\, 1,\, 1)^\mathsf{T} \right| \left| (1,\, 2,\, -1)^\mathsf{T} \right|} = \frac{2}{\sqrt{3}\sqrt{6}} \implies \theta = 61.9^\circ \text{ (1 d.p.)}.$$

* * * * *

Problem 3. The equations of the line l_1 and the plane Π_1 are as follows:

$$l_1 : \mathbf{r} = \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \lambda \in \mathbb{R},$$

$$\Pi_1 : xa + z = 5a + 4, \quad a \in \mathbb{R}^+.$$

(a) If the angle between l_1 and Π_1 is $\pi/6$, show that $a = 1$.

Using the value of a in (a),

(b) Verify that l_1 and Π_1 intersect at the point $A(5, -1, 4)$.

(c) Given that $C(7, -3, 4)$, find the length of projection of $\overrightarrow{AC}$ on Π_1.

(d) Find the position vector of N, the foot of perpendicular of C to Π_1.

(e) Point C' is obtained by reflecting C about Π_1. Determine the vector equation of the line that passes through A and C'.

Solution. Note that Π_1 has vector equation

$$\Pi_1: \quad \mathbf{r} \cdot \begin{pmatrix} a \\ 0 \\ 1 \end{pmatrix} = 5a + 4.$$

Part (a). Since the angle between l_1 and Π_1 is $\pi/6$, we have

$$\frac{1}{2} = \sin\frac{\pi}{6} = \frac{\left|(1,\, -1,\, 0)^\mathsf{T} \cdot (a,\, 0,\, 1)^\mathsf{T}\right|}{\left|(1,\, -1,\, 0)^\mathsf{T}\right| \cdot \left|(a,\, 0,\, 1)^\mathsf{T}\right|} = \frac{a}{\sqrt{2}\sqrt{a^2+1}},$$

which yields $a = 1$.

Part (b). $(5, -1, 4)$ is clearly on l_1. Since

$$\begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 9 = 5(1) + 4,$$

it follows that $(5, -1, 4)$ is also on Π_1. Thus, l_1 and Π_1 intersect at $(5, -1, 4)$.

Part (c). Note that

$$\overrightarrow{AC} = \overrightarrow{OC} - \overrightarrow{OA} = \begin{pmatrix} 7 \\ -3 \\ 4 \end{pmatrix} - \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}.$$

The length of projection of $\overrightarrow{AC}$ on Π_1 is hence given by

$$\frac{\left|(2,\, -2,\, 0)^\mathsf{T} \times (1,\, 0,\, 1)^\mathsf{T}\right|}{\left|(1,\, 0,\, 1)^\mathsf{T}\right|} = \frac{\left|(-2,\, -2,\, 2)^\mathsf{T}\right|}{\left|(1,\, 0,\, 1)^\mathsf{T}\right|} = \sqrt{6}.$$

Part (d). Observe that $\overrightarrow{CN}$ is parallel to the normal vector of Π_1, so

$$\overrightarrow{ON} = \overrightarrow{OC} + \overrightarrow{CN} = \begin{pmatrix} 7 \\ -3 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

for some $\mu \in \mathbb{R}$. Since N lies on Π_1, we have

$$\left[\begin{pmatrix} 7 \\ -3 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right] \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 11 + 2\mu = 9 \implies \mu == -1.$$

Thus, the position vector of N is

$$\overrightarrow{ON} = \begin{pmatrix} 7 \\ -3 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \\ -3 \end{pmatrix}.$$

Part (e). By the midpoint theorem,

$$\overrightarrow{ON} = \frac{\overrightarrow{OC} + \overrightarrow{OC'}}{2} \implies \overrightarrow{OC'} = 2\overrightarrow{ON} - \overrightarrow{OC} = 2\begin{pmatrix} 6 \\ -3 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \\ 2 \end{pmatrix}.$$

Thus,

$$\overrightarrow{AC'} = \overrightarrow{OC'} - \overrightarrow{OA} = \begin{pmatrix} 5 \\ -3 \\ 2 \end{pmatrix} - \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix} = -2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

hence the vector equation of the line AC' is given by

$$\mathbf{r} = \begin{pmatrix} 5 \\ -1 \\ 4 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \nu \in \mathbb{R}.$$

* * * * *

Problem 4. The equation of the plane Π_1 is $x + y - 2z = 3$.

(a) Find the vector equation of the line l_1, which lies in both the plane Π_1 and the yz plane.

(b) Another plane Π_2 contains the line l_2 with equation $x = 1$, $\frac{y+1}{2} = z$ and is perpendicular to Π_1. Find the equation of the plane Π_2 in scalar product form. Determine whether l_1 lies on Π_2.

Solution. Note that the vector equations of Π_1 and the yz plane are

$$\Pi_1 : \ \mathbf{r} \cdot \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = 3 \quad \text{and} \quad \mathbf{r} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 0$$

respectively.

Part (a). Note that

$$\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix} = -\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}.$$

Since $(0, 1, -1)$ lies on both Π_1 and the yz plane, it follows that the vector equation of l_1 is

$$l_1 : \quad \mathbf{r} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Part (b). Let the normal vector of Π_2 be $(x,\, y,\, z)^{\mathsf{T}}$, so it has vector equation

$$\Pi_2 : \quad \mathbf{r} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = d$$

for some constant d.

The vector equation of l_2 is

$$l_2 : \quad \mathbf{r} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad \mu \in \mathbb{R}.$$

Since l_2 lies on Π_2, for all $\mu \in \mathbb{R}$, we must have

$$\left[\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right] \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = d.$$

This simplifies to

$$(x-y)+\mu\,(2y+z)=d,$$

whence we conclude that $2y+z=0$ and $x-y=d$. The vector equation of Π_2 hence updates as

$$\Pi_2: \quad \mathbf{r}\cdot\begin{pmatrix}x\\y\\-2y\end{pmatrix}=x-y.$$

Since Π_1 and Π_2, we have that

$$0=\cos\frac{\pi}{2}=\frac{\left|(1,\,1,\,-2)^\mathsf{T}\cdot(x,\,y,\,-2y)^\mathsf{T}\right|}{\left|(1,\,1,\,-1)^\mathsf{T}\right|\left|(x,\,y,\,-2y)^\mathsf{T}\right|} \implies x+5y=0.$$

Thus, the normal vector is

$$\begin{pmatrix}x\\y\\-2y\end{pmatrix}=\begin{pmatrix}-5y\\y\\-2y\end{pmatrix}=y\begin{pmatrix}-5\\1\\-2\end{pmatrix}.$$

Taking $y=1$, we get $x=-5$, so $d=x-y=-6$. Thus, the vector equation of Π_2 is

$$\Pi_2: \quad \mathbf{r}\cdot\begin{pmatrix}-5\\1\\-2\end{pmatrix}=-6.$$

Note that l_1 is parallel to l_2. Since l_2 lies on Π_2, this implies that l_1 is parallel to Π_2. Since

$$\begin{pmatrix}0\\1\\-1\end{pmatrix}\cdot\begin{pmatrix}-5\\1\\-2\end{pmatrix}=3\neq-6,$$

it follows that $(0,1,-1)$ does not lie on Π_2, thus l_1 does not lie on Π_2.

* * * * *

Problem 5. The lines l_1 and l_2 intersect at the point P with position vector $\mathbf{i}+5\mathbf{j}+12\mathbf{k}$. The equations of l_1 and l_2 are $\mathbf{r}=(1+3\lambda)\mathbf{i}+(5+2\lambda)\mathbf{j}+(12-2\lambda)\mathbf{k}$ and $\mathbf{r}=(1+8\mu)\mathbf{i}+(5+11\mu)\mathbf{j}+(12+6\mu)\mathbf{k}$ respectively, where λ and μ are real parameters.

(a) Find an equation of the plane Π_1, which contains l_1 and l_2 in the form $\mathbf{r}\cdot\mathbf{n}=d$.

Π_2 and Π_3 are two planes with equations $2x+az=b$ and $x-3y-z=7$ respectively, where a and b are constants.

(b) Find the line of intersection between Π_1 and Π_3.

(c) (i) Find the condition satisfied by a if the three planes Π_1, Π_2 and Π_3 intersect at one unique point.

(ii) Given that all three planes meet in a line l, find a and b.

(iii) Given instead that the three planes have no point in common, what can be said about the values of a and b?

Solution. Rewriting, we see that the equations of l_1 and l_2 are

$$l_1: \quad \mathbf{r} = \begin{pmatrix} 1 \\ 5 \\ 12 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 2 \\ -2 \end{pmatrix}, \quad \lambda \in \mathbb{R},$$

$$l_2: \quad \mathbf{r} = \begin{pmatrix} 1 \\ 5 \\ 12 \end{pmatrix} + \mu \begin{pmatrix} 8 \\ 11 \\ 6 \end{pmatrix}, \quad \mu \in \mathbb{R}.$$

Part (a). Note that

$$\begin{pmatrix} 3 \\ 2 \\ -2 \end{pmatrix} \times \begin{pmatrix} 8 \\ 11 \\ 6 \end{pmatrix} = \begin{pmatrix} 34 \\ -34 \\ 17 \end{pmatrix} = 17 \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}.$$

Thus, the equation of Π_1 is

$$\Pi_1: \quad \mathbf{r} \cdot \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 12 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} = 4.$$

Part (b). Note that Π_3 has vector equation

$$\Pi_3: \quad \mathbf{r} \cdot \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix} = 7.$$

By inspection, we see that $(-1/2, -3/2, 0)$ lies on both Π_1 and Π_3. Since

$$\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ -4 \end{pmatrix},$$

the vector equation of the line of intersection l is

$$\mathbf{r} = -\frac{1}{2} \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 5 \\ 3 \\ -4 \end{pmatrix}, \quad \nu \in \mathbb{R}.$$

Part (c). Note that Π_2 has vector equation

$$\Pi_2: \quad \mathbf{r} \cdot \begin{pmatrix} 2 \\ 0 \\ a \end{pmatrix} = b.$$

Part (c)(i). If the three planes intersect at a common point, it must be that l intersects Π_2 at a single point. Consider now the intersection between l and Π_2:

$$\left[-\frac{1}{2} \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 5 \\ 3 \\ -4 \end{pmatrix} \right] \cdot \begin{pmatrix} 2 \\ 0 \\ a \end{pmatrix} = -1 + \nu (10 - 4a) = b.$$

In order for this equation to have a unique solution, we must be able to write

$$\nu = \frac{b+1}{10-4a},$$

i.e. $10 - 4a \neq 0$. Thus, so long as $a \neq 5/2$, the three planes will intersect at a unique point.

Part (c)(ii). If the planes intersect at a common line, then l must lie on Π_2. Thus,

$$-1+\nu(10-4a)=b$$

must hold true for all $\nu\in\mathbb{R}$. This can only happen when $10-4a=0$ and $b=-1$. Hence, the three planes meet in a line when $a=5/2$ and $b=-1$.

Part (c)(iii). The complement of $(a\neq 5/2)$ or $(a=5/2$ and $b=-1)$ is $(a=5/2$ and $b\neq -1)$, which corresponds to the case where the three planes neither meet in a point nor in a line, i.e. they have no common point.

* * * * *

Problem 6. The point A and B have position vectors $3\mathbf{i}+\mathbf{j}$ and $3\mathbf{i}+3\mathbf{j}$ respectively. The line l_1 and the planes Π_1 and Π_2 have equations as follows:

$$l_1:\mathbf{r}=\overrightarrow{OA}+\alpha\begin{pmatrix}2\\1\\-1\end{pmatrix},\quad \Pi_1: x+2z=3,\quad \Pi_2:\mathbf{r}=\lambda\begin{pmatrix}1\\1\\0\end{pmatrix}+\mu\begin{pmatrix}0\\1\\1\end{pmatrix},$$

where α, λ and $\mu\in\mathbb{R}$.

It is given that the planes Π_1 and Π_2 intersect in the line l_2 and B lies on l_2.

(a) Find a vector equation of the line l_2 and show that the line l_2 is parallel to the line l_1. Hence, find the shortest distance between the lines l_1 and l_2.

(b) The plane Π_3 is parallel to the plane Π_2 and is equidistant to both point A and the plane Π_2. Show that the equation of the plane Π_3 is given by $\mathbf{r}\cdot(\mathbf{i}-\mathbf{j}+\mathbf{k})=1$. Find the position vector of the foot of perpendicular from the point A to the plane Π_3.

Solution. Note that

$$\begin{pmatrix}1\\1\\0\end{pmatrix}\times\begin{pmatrix}0\\1\\1\end{pmatrix}=\begin{pmatrix}1\\-1\\1\end{pmatrix},$$

hence Π_2 has vector equation

$$\Pi_2:\quad \mathbf{r}\cdot\begin{pmatrix}1\\-1\\1\end{pmatrix}=0.$$

Part (a). Note that

$$\begin{pmatrix}1\\0\\2\end{pmatrix}\times\begin{pmatrix}1\\-1\\1\end{pmatrix}=\begin{pmatrix}2\\1\\-1\end{pmatrix}.$$

Thus, the equation of l_2 is

$$l_2:\quad \mathbf{r}=\begin{pmatrix}3\\3\\0\end{pmatrix}+t\begin{pmatrix}2\\1\\-1\end{pmatrix},\quad t\in\mathbb{R}.$$

Since l_1 and l_2 have the same direction vector, they are parallel. The shortest distance between them is

$$\frac{\left|\overrightarrow{AB}\times(2,\,1,\,-1)^\mathsf{T}\right|}{\left|(2,\,1,\,-1)^\mathsf{T}\right|}=\frac{\left|(0,\,2,\,0)^\mathsf{T}\times(2,\,1,\,-1)^\mathsf{T}\right|}{\left|(2,\,1,\,-1)^\mathsf{T}\right|}=\frac{\left|(-2,\,0,\,-4)^\mathsf{T}\right|}{\left|(2,\,1,\,-1)^\mathsf{T}\right|}=\frac{\sqrt{20}}{\sqrt{6}}=\sqrt{\frac{10}{3}}\text{ units.}$$

Part (b). Let A' be the reflection of A in Π_3. Let M be foot of perpendicular from A to Π_3, so that it is the midpoint of AA'. By the midpoint theorem,

$$\overrightarrow{OM} = \frac{\overrightarrow{OA} + \overrightarrow{OA'}}{2}.$$

Since Π_3 is parallel to Π_2, it is normal to $(1, -1, 1)^\mathsf{T}$. Thus, its vector equation is

$$\Pi_3 : \quad \mathbf{r} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \overrightarrow{OM} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{2}\left[\overrightarrow{OA} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \overrightarrow{OA'} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right] = \frac{1}{2}(2+0) = 1,$$

where we used the fact that A' lies on Π_2 and M lies on Π_3.

Note that $\overrightarrow{AM}$ is parallel to the normal vector $(1, -1, 1)^\mathsf{T}$, so

$$\overrightarrow{OM} = \overrightarrow{OA} + \overrightarrow{AM} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

for some $s \in \mathbb{R}$. Since M lies on Π_3, we must have

$$\left[\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right] \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = 2 + 3s = 1 \implies s = -\frac{1}{3}.$$

Thus,

$$\overrightarrow{OM} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 8 \\ 4 \\ -1 \end{pmatrix}.$$

* * * * *

Problem 7. The planes p_1, p_2 and p_3 have equations $x = 1$, $2x + y + az = 5$ and $x + 2y + z = b$, where a and b are real constants. Given that p_1 and p_2 intersect at the line l, show that the vector equation of l, in terms of a, is $\mathbf{r} = \mathbf{i} + (3 - \lambda a)\mathbf{j} + \lambda\mathbf{k}$, where λ is a real parameter.

(a) The acute angle between l and p_3 is $60°$. Without using a calculator, find the possible values of a.

(b) Given that the shortest distance from the origin to p_3 is $\sqrt{6}/3$ and without solving for the value of b, determine the possible position vectors of the foot of perpendicular from the origin to p_3.

(c) What can be said about a and b if p_1, p_2 and p_3 do not have any points in common?

Solution. Note that p_1, p_2 and p_3 have vector equations

$$p_1 : \quad \mathbf{r} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1, \qquad p_2 : \quad \mathbf{r} \cdot \begin{pmatrix} 2 \\ 1 \\ a \end{pmatrix} = 5, \qquad p_3 : \quad \mathbf{r} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = b.$$

Consider the intersection of p_1 and p_2. Substituting $x = 1$ into the equation for p_2, we get $y = 3 - az$. Thus,

$$l : \quad \mathbf{r} = \begin{pmatrix} 1 \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 3 - az \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ -a \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ -a \\ 1 \end{pmatrix},$$

where $\lambda = z$ is a real parameter.

Part (a). We have

$$\frac{\sqrt{3}}{2}=\frac{\left|(0,\,-a,\,1)^{\mathsf{T}}\cdot(1,\,2,\,1)^{\mathsf{T}}\right|}{\left|(0,\,-a,\,1)^{\mathsf{T}}\right|\left|(1,\,2,\,1)^{\mathsf{T}}\right|}=\frac{|1-2a|}{\sqrt{a^2+1}\sqrt{6}}.$$

This yields

$$|1-2a|=\frac{3}{\sqrt{2}}\sqrt{a^2+1}\implies(1-2a)^2=\frac{9}{2}\left(a^2+1\right),$$

which simplifies to

$$a^2+8a+7=(a+7)(a+1)=0.$$

Thus, the possible values of a are $a=-1$ or $a=-7$.

Part (b). Let N be the foot of perpendicular from the origin to p_3. Then $|ON|=s\,(1,\,2,\,1)^{\mathsf{T}}$ for some $s\in\mathbb{R}$. The given condition implies

$$\frac{\sqrt{6}}{3}=\left|\overrightarrow{ON}\right|=\left|s\begin{pmatrix}1\\2\\1\end{pmatrix}\right|=|s|\sqrt{6}\implies|s|=\frac{1}{3},$$

so $s=\pm1/3$, thus

$$\overrightarrow{ON}=\frac{1}{3}\begin{pmatrix}1\\2\\1\end{pmatrix}\quad\text{or}\quad\overrightarrow{ON}=-\frac{1}{3}\begin{pmatrix}1\\2\\1\end{pmatrix}.$$

Part (c). If the three planes do not have any points in common, it must be that l does not intersect p_3. Thus,

$$\left[\begin{pmatrix}1\\3\\0\end{pmatrix}+\lambda\begin{pmatrix}0\\-a\\1\end{pmatrix}\right]\cdot\begin{pmatrix}1\\2\\1\end{pmatrix}=7+\lambda\,(1-2a)\neq b$$

for all $\lambda\in\mathbb{R}$. This implies that $1-2a=0$ so $a=1/2$, and $b\neq7$.

* * * * *

Problem 8 (🌶). The points A and B have position vectors $\mathbf{a}$ and $\mathbf{b}$ respectively. The plane π, with vector equation $\mathbf{r}=\mathbf{b}+\lambda\mathbf{u}+\mu\mathbf{v}$, where λ and μ are real parameters, contains B but not A.

(a) Show that the perpendicular distance of A from π is p, where

$$p=\frac{|(\mathbf{u}\times\mathbf{v})\cdot(\mathbf{b}-\mathbf{a})|}{|\mathbf{u}\times\mathbf{v}|}.$$

(b) The perpendicular from A to π meets π at C, and D is the point on AB such that CD is perpendicular to AB. Show that $AD=p^2/AB$ and hence, or otherwise, show that the position vector of D is

$$\mathbf{a}+\left(\frac{p}{|\mathbf{b}-\mathbf{a}|}\right)^2(\mathbf{b}-\mathbf{a}).$$

In the case where $\mathbf{a}=-\mathbf{i}+7\mathbf{j}+8\mathbf{k}$, $\mathbf{b}=2\mathbf{i}+7\mathbf{j}+5\mathbf{k}$, $\mathbf{u}=\mathbf{i}-2\mathbf{j}+2\mathbf{k}$ and $\mathbf{v}=3\mathbf{i}+2\mathbf{j}+2\mathbf{k}$, find the value of p, and show that

$$\overrightarrow{CD}=\frac{8\sqrt{2}}{9}\mathbf{x}+\frac{4}{9}\mathbf{y},$$

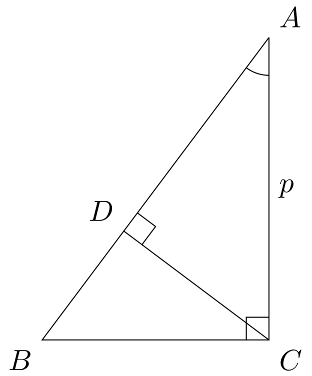
where $\mathbf{x}$ and $\mathbf{y}$ are the unit vectors of $\overrightarrow{CB}$ and $\overrightarrow{CA}$ respectively.

Solution.

Part (a). Note that the normal vector of π is $\mathbf{n} = \mathbf{u} \times \mathbf{v}$. Thus, the perpendicular distance of A from π is

$$p = \frac{\mathbf{n} \times \overrightarrow{AB}}{|\mathbf{n}|} = \frac{|(\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{b} - \mathbf{a})|}{|\mathbf{u} \times \mathbf{v}|}.$$

Part (b).



Consider the above diagram. Observe that $\triangle ACB$ is similar to $\triangle ADC$, so

$$\frac{AD}{AC} = \frac{AC}{AB} \implies AD = \frac{AC^2}{AB}.$$

But AC is the perpendicular distance from A to π, so $AC = p$ and $AD = p^2/AB$ as desired.

Note that

$$\frac{AD}{AB} = \frac{p^2}{AB^2},$$

thus

$$\frac{AD}{DB} = \frac{AD}{AB - AD} = \frac{1}{AB/AD - 1} = \frac{1}{\frac{AB^2}{p^2} - 1} = \frac{p^2}{AB^2 - p^2}.$$

Thus, by the Ratio Theorem,

$$\overrightarrow{OD} = \frac{p^2\mathbf{b} + \left(AB^2 - p^2\right)\mathbf{a}}{p^2 + (AB^2 - p^2)} = \frac{AB^2}{AB^2}\mathbf{a} + \left(\frac{p}{AB}\right)^2 (\mathbf{b} - \mathbf{a}) = \mathbf{a} + \left(\frac{p}{|\mathbf{b} - \mathbf{a}|}\right)^2 (\mathbf{b} - \mathbf{a}).$$

We have

$$\mathbf{u} \times \mathbf{v} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} \times \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix} = 4\begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} \quad \text{and} \quad \overrightarrow{AB} = \begin{pmatrix} 2 \\ 7 \\ 5 \end{pmatrix} - \begin{pmatrix} -1 \\ 7 \\ 8 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

thus

$$p = \frac{12\left|(-2,\, 1,\, 2)^\mathsf{T} \times (1,\, 0,\, -1)^\mathsf{T}\right|}{4\left|(-2,\, 1,\, 2)^\mathsf{T}\right|} = 4 \quad \text{and} \quad AB = \left|\overrightarrow{AB}\right| = 3\sqrt{2}.$$

This gives

$$\frac{AD}{DB} = \frac{p^2}{AB^2 - p^2} = 4^2\left(3\sqrt{2}\right)^2 - 16 = \frac{16}{2} = \frac{8}{1}.$$

By the Ratio Theorem,

$$\overrightarrow{CD} = \frac{8\overrightarrow{CB} + \overrightarrow{CA}}{9} = \frac{8}{9}\,(CB)\,\mathbf{x} + \frac{1}{9}\,(CA)\,\mathbf{y}.$$

Note that $CA = p = 4$. Meanwhile, using the Pythagorean theorem, we see that

$$AB^2 = BC^2 + CA^2 \implies CB^2 = AB^2 - CA^2 = \left(3\sqrt{2}\right)^2 - 4^2 = 2,$$

so

$$\overrightarrow{CD} = \frac{8\sqrt{2}}{9}\mathbf{x} + \frac{4}{9}\mathbf{y}.$$

Assignment A9

Problem 1. The equation of the plane Π_1 is $y + z = 0$ and the equation of the line l is $\frac{x-5}{2} = \frac{y-2}{-1} = \frac{z-2}{3}$. Find

(a) the position vector of the point of intersection of l and Π_1,

(b) the length of the perpendicular from the origin to l,

(c) the Cartesian equation for the plane Π_2 which contains l and the origin,

(d) the acute angle between the planes Π_1 and Π_2, giving your answer correct to the nearest $0.1°$.

Solution. Note that Π_1 has equation $\mathbf{r} \cdot (0,\, 1,\, 1)^\mathsf{T} = 0$ and l has equation $\mathbf{r} = (5,\, 2,\, 2)^\mathsf{T} + \lambda\,(2,\, -1,\, 3)^\mathsf{T}$, $\lambda \in \mathbb{R}$.

Part (a). Let P be the point of intersection of Π_1 and l. Then $\overrightarrow{OP} = (5,\, 2,\, 2)^\mathsf{T} + \lambda\,(2,\, -1,\, 3)^\mathsf{T}$ for some $\lambda \in \mathbb{R}$. Also, $\overrightarrow{OP} \cdot (0,\, 1,\, 1)^\mathsf{T} = 0$. Hence,

$$\left[\begin{pmatrix} 5 \\ 2 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}\right] \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 0 \implies 4 + 2\lambda = 0 \implies \lambda = -2.$$

Thus,

$$\overrightarrow{OP} = \begin{pmatrix} 5 \\ 2 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ -4 \end{pmatrix}.$$

Part (b). The perpendicular distance from the origin to l is

$$\frac{\left|(5,\, 2,\, 2)^\mathsf{T} \times (2,\, -1,\, 3)^\mathsf{T}\right|}{\left|(2,\, -1,\, 3)^\mathsf{T}\right|} = \frac{1}{\sqrt{14}} \left|\begin{pmatrix} 8 \\ -11 \\ -9 \end{pmatrix}\right| = \frac{\sqrt{266}}{\sqrt{14}} = \sqrt{19} \text{ units.}$$

Part (c). Observe that Π_2 is parallel to $(5,\, 2,\, 2)^\mathsf{T}$ and $(2,\, -1,\, 3)^\mathsf{T}$. Thus, $\mathbf{n} = (5,\, 2,\, 2)^\mathsf{T} \times (2,\, -1,\, 3)^\mathsf{T} = (8,\, -11,\, -9)^\mathsf{T}$. Since Π_2 contains the origin, $d = 0$. Hence, Π_2 has vector equation $\mathbf{r} \cdot (8,\, -11,\, -9)^\mathsf{T} = 0$, which translates to $8x - 11y - 9z = 0$.

Part (d). Let the acute angle be θ.

$$\cos\theta = \frac{\left|(0,\, 1,\, 1)^\mathsf{T} \cdot (8,\, -11,\, -9)^\mathsf{T}\right|}{\left|(0,\, 1,\, 1)^\mathsf{T}\right| \left|(8,\, -11,\, -9)^\mathsf{T}\right|} = \frac{20}{\sqrt{2}\sqrt{266}} \implies \theta = 29.9° \text{ (1 d.p.).}$$

* * * * *

Problem 2. The plane Π_1 has equation $\mathbf{r} \cdot (-\mathbf{i} + 2\mathbf{k}) = -4$ and the points A and P have position vectors $4\mathbf{i}$ and $\mathbf{i} + \alpha\mathbf{j} + \mathbf{k}$ respectively, where $\alpha \in \mathbb{R}$.

(a) Show that A lies on Π_1, but P does not.

(b) Find, in terms of α, the position vector of N, the foot of perpendicular of P on Π_1.

The plane Π_2 contains the points A, P and N.

(c) Show that the equation of Π_2 is $\mathbf{r} \cdot (2\alpha\mathbf{i} + 5\mathbf{j} + \alpha\mathbf{k}) = 8\alpha$ and write down the equation of l, the line of the intersection of Π_1 and Π_2.

The plane Π_3 has equation $\mathbf{r} \cdot (\mathbf{i} + \mathbf{j} + 2\mathbf{k}) = 4$.

(d) By considering l, or otherwise, find the value of α for which the three planes intersect in a line.

Solution. Note that $\Pi_1 : \mathbf{r} \cdot (-1,\, 0,\, 2)^{\mathsf{T}} = -4$, $\overrightarrow{OA} = (4,\, 0,\, 0)^{\mathsf{T}}$ and $\overrightarrow{OP} = (1,\, \alpha,\, 1)^{\mathsf{T}}$.

Part (a). Since $\overrightarrow{OA} \cdot (-1,\, 0,\, 2)^{\mathsf{T}} = (4,\, 0,\, 0)^{\mathsf{T}} \cdot (-1,\, 0,\, 2)^{\mathsf{T}} = -4$, A lies on Π_1. On the other hand, since $\overrightarrow{OP} \cdot (-1,\, 0,\, 2)^{\mathsf{T}} = (1,\, \alpha,\, 1)^{\mathsf{T}} \cdot (-1,\, 0,\, 2)^{\mathsf{T}} = 1 \neq -4$, P does not lie on Π_1.

Part (b). Note that $\overrightarrow{NP} = \lambda\,(-1,\, 0,\, 2)^{\mathsf{T}}$ for some $\lambda \in \mathbb{R}$, and $\overrightarrow{ON} \cdot (-1,\, 0,\, 2)^{\mathsf{T}} = -4$. Hence,

$$\overrightarrow{NP} = \overrightarrow{OP} - \overrightarrow{ON} = \begin{pmatrix}1\\ \alpha\\ 1\end{pmatrix} - \overrightarrow{ON} = \lambda \begin{pmatrix}-1\\ 0\\ 2\end{pmatrix}.$$

Thus,

$$\left[\begin{pmatrix}1\\ \alpha\\ 1\end{pmatrix} - \overrightarrow{ON}\right] \cdot \begin{pmatrix}-1\\ 0\\ 2\end{pmatrix} = \lambda \begin{pmatrix}-1\\ 0\\ 2\end{pmatrix} \cdot \begin{pmatrix}-1\\ 0\\ 2\end{pmatrix} \implies 1 - (-4) = 5\lambda \implies \lambda = 1.$$

Hence, $\overrightarrow{NP} = (-1,\, 0,\, 2)^{\mathsf{T}}$, whence $\overrightarrow{ON} = \overrightarrow{OP} - \overrightarrow{NP} = (2,\, \alpha,\, -1)^{\mathsf{T}}$.

Part (c). Note that Π_2 is parallel to $\overrightarrow{NP} = (-1,\, 0,\, 2)^{\mathsf{T}}$ and $\overrightarrow{AN} = \overrightarrow{ON} - \overrightarrow{OA} = (-2,\, \alpha,\, -1)^{\mathsf{T}}$. Since $(-1,\, 0,\, 2)^{\mathsf{T}} \times (-2,\, \alpha,\, -1)^{\mathsf{T}} = -\,(2\alpha,\, 5,\, \alpha)^{\mathsf{T}}$, we take $\mathbf{n} = (2\alpha,\, 5,\, \alpha)^{\mathsf{T}}$, whence $d = (4,\, 0,\, 0)^{\mathsf{T}} \cdot (2\alpha,\, 5,\, \alpha)^{\mathsf{T}} = 8\alpha$. Thus, Π_2 has vector equation $\mathbf{r} \cdot (2\alpha,\, 5,\, \alpha)^{\mathsf{T}} = 8\alpha$ which translates to $\mathbf{r} \cdot (2\alpha\mathbf{i} + 5\mathbf{j} + \alpha\mathbf{k}) = 8\alpha$.

Meanwhile, the line of intersection between Π_1 and Π_2 has equation

$$l : \begin{pmatrix}4\\ 0\\ 0\end{pmatrix} + \mu \begin{pmatrix}-2\\ \alpha\\ -1\end{pmatrix}, \quad \mu \in \mathbb{R}.$$

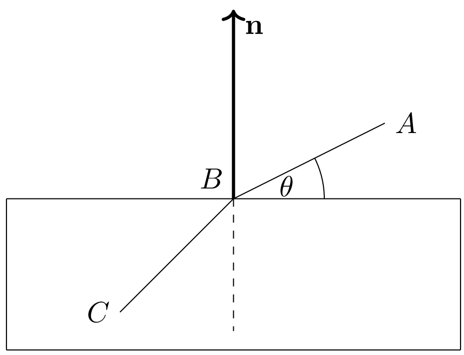
Part (d). If the three planes intersect in a line, they must intersect at l. Hence, l lies on Π_3.

$$\left[\begin{pmatrix}4\\ 0\\ 0\end{pmatrix} + \mu \begin{pmatrix}-2\\ \alpha\\ -1\end{pmatrix}\right] \cdot \begin{pmatrix}1\\ 1\\ 2\end{pmatrix} = 4 \implies 4 + (\alpha - 4)\mu = 4 \implies (\alpha - 4)\mu = 0.$$

Since $(\alpha - 4)\mu = 0$ must hold for all $\mu \in \mathbb{R}$, we must have $\alpha = 4$.

* * * * *

Problem 3. When a light ray passes from air to glass, it is deflected through an angle. The light ray ABC starts at point $A(1, 2, 2)$ and enters a glass object at point $B(0, 0, 2)$. The surface of the glass object is a plane with normal vector $\mathbf{n}$. The diagram shows a cross-section of the glass object in the plane of the light ray and $\mathbf{n}$.



(a) Find a vector equation of the line AB.

The surface of the glass object is a plane with equation $x + z = 2$. AB makes an acute angle θ with the plane.

(b) Calculate the value of θ, giving your answer in degrees.

The line BC makes an angle of 45° with the normal to the plane, and BC is parallel to the unit vector $(-2/3,\, p,\, q)^\mathsf{T}$.

(c) By considering a vector perpendicular to the plane containing the light ray and $\mathbf{n}$, or otherwise, find the values of p and q.

The light ray leaves the glass object through a plane with equation $3x + 3z = -4$.

(d) Find the exact thickness of the glass object, taking one unit as one cm.

(e) Find the exact coordinates of the point at which the light ray leaves the glass object.

Solution. Let Π_G be the plane representing the surface of the glass object.
Part (a). Note that $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = (0,\, 0,\, 2)^\mathsf{T} - (1,\, 2,\, 2)^\mathsf{T} = -\,(1,\, 2,\, 0)^\mathsf{T}$. Hence,

$$l_{AB} : \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Part (b). Observe that Π_G has equation $\mathbf{r} \cdot (1,\, 0,\, 1)^\mathsf{T} = 2$. Hence,

$$\sin\theta = \frac{\left|(1,\, 0,\, 1)^\mathsf{T} \cdot (1,\, 2,\, 0)^\mathsf{T}\right|}{\left|(1,\, 0,\, 1)^\mathsf{T}\right| \left|(1,\, 2,\, 0)^\mathsf{T}\right|} = \frac{1}{\sqrt{2}\sqrt{5}} \implies \theta = 71.6^\circ \text{ (1 d.p.)}.$$

Part (c). Since line BC makes an angle of 45° with $\mathbf{n}_G$,

$$\sin 45^\circ = \frac{\left|(1,\, 0,\, 1)^\mathsf{T} \cdot (-2/3,\, p,\, q)^\mathsf{T}\right|}{\left|(1,\, 0,\, 1)^\mathsf{T}\right| \left|(-2/3,\, p,\, q)^\mathsf{T}\right|} \implies \frac{1}{\sqrt{2}} = \frac{|q - 2/3|}{\sqrt{2} \cdot 1} \implies \left|q - \frac{2}{3}\right| = 1.$$

Hence, $q = -1/3$. Note that we reject $q = 5/3$ since $(-2/3,\, p,\, q)^\mathsf{T}$ is a unit vector, which implies that $|q| \leq 1$.

Let Π_L be the plane containing the light ray. Note that Π_L is parallel to $\overrightarrow{AB}$ and $\overrightarrow{BC}$. Hence, $\mathbf{n}_L = (1,\, 2,\, 0)^\mathsf{T} \times (-2/3,\, p,\, q)^\mathsf{T} = \frac{1}{3}(6q,\, -3q,\, 3p + 4)^\mathsf{T}$. Since Π_L contains $\mathbf{n}_G$, we have that $\mathbf{n}_L \perp \mathbf{n}_G$, whence $\mathbf{n}_L \cdot \mathbf{n}_G = 0$. This gives us

$$\begin{pmatrix} 6q \\ -3q \\ 3p + 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 0 \implies 6q + 3p + 4 = 0 \implies 6\left(-\frac{1}{3}\right) + 3p + 4 = 0 \implies p = -\frac{2}{3}.$$

Part (d). Let Π_G' be the plane with equation $3x + 3z = -4$. Observe that Π_G is parallel to Π_G'. Also note that $(-4/3, 0, 0)$ is a point on Π_G'. Hence, the distance between Π_G and Π_G' is given by

$$\frac{\left|2 - (-4/3,\, 0,\, 0)^\mathsf{T} \cdot (1,\, 0,\, 1)^\mathsf{T}\right|}{\left|(1,\, 0,\, 1)^\mathsf{T}\right|} = \frac{10}{3\sqrt{2}} \text{ cm}.$$

Part (e). Observe that $(-2/3,\, p,\, q)^{\mathsf{T}} = (-2/3,\, -2/3,\, -1/3)^{\mathsf{T}} = -\frac{1}{3}(2,\, 2,\, 1)^{\mathsf{T}}$, whence the line BC has equation $\mathbf{r} = (0,\, 0,\, 2)^{\mathsf{T}} + \mu\,(2,\, 2,\, 1)^{\mathsf{T}}$, $\mu \in \mathbb{R}$. Let P be the intersection between line BC and Π_G'. Also note that $\overrightarrow{OP} = (0,\, 0,\, 2)^{\mathsf{T}} + \mu\,(2,\, 2,\, 1)^{\mathsf{T}}$ for some $\mu \in \mathbb{R}$, and $\overrightarrow{OP} \cdot (3,\, 0,\, 3)^{\mathsf{T}} = -4$. Hence,

$$\left[\begin{pmatrix}0\\0\\2\end{pmatrix} + \mu\begin{pmatrix}2\\2\\1\end{pmatrix}\right] \cdot \begin{pmatrix}3\\0\\3\end{pmatrix} = -4 \implies 6 - 9\mu = -4 \implies \mu = -\frac{10}{9}.$$

Hence, $\overrightarrow{OP} = (0,\, 0,\, 2)^{\mathsf{T}} - \frac{10}{9}(2,\, 2,\, 1)^{\mathsf{T}} = (-20/9,\, -20/9,\, 8/9)^{\mathsf{T}}$. The coordinates of the point are hence $(-20/9, -20/9.8/9)$.

A10.1 Complex Numbers - Complex Numbers in Cartesian Form

Tutorial A10.1

Problem 1. Given that $z = 3 - 2\mathrm{i}$ and $w = 1 + 4\mathrm{i}$, express in the form $a + b\mathrm{i}$, where $a, b \in \mathbb{R}$:

(a) $z + 2w$

(b) zw

(c) z/w

(d) $(w - w^*)^3$

(e) z^4

Solution.
Part (a).

$$z + 2w = (3 - 2\mathrm{i}) + 2(1 + 4\mathrm{i}) = 3 - 2\mathrm{i} + 2 + 8\mathrm{i} = 5 + 6\mathrm{i}.$$

Part (b).

$$zw = (3 - 2\mathrm{i})(1 + 4\mathrm{i}) = 3 + 12\mathrm{i} - 2\mathrm{i} + 8 = 11 + 10\mathrm{i}.$$

Part (c).

$$\frac{z}{w} = \frac{3 - 2\mathrm{i}}{1 + 4\mathrm{i}} = \frac{(3 - 2\mathrm{i})(1 - 4\mathrm{i})}{(1 + 4\mathrm{i})(1 - 4\mathrm{i})} = \frac{3 - 12\mathrm{i} - 2\mathrm{i} - 8}{1^2 + 4^2} = \frac{-5 - 14\mathrm{i}}{17} = -\frac{5}{17} - \frac{14}{17}\mathrm{i}.$$

Part (d).

$$(w - w^*)^3 = [2\,\mathrm{Im}(w)\,\mathrm{i}]^3 = (8\mathrm{i})^3 = -512\mathrm{i}.$$

Part (e).

$$\begin{aligned} z^4 = (3 - 2\mathrm{i})^4 &= 3^4 + 4 \cdot 3^3(-2\mathrm{i}) + 6 \cdot 3^2(-2\mathrm{i})^2 + 4 \cdot 3(-2\mathrm{i})^3 + (-2\mathrm{i})^4 \\ &= 81 - 216\mathrm{i} - 216 + 96\mathrm{i} + 16 = -119 - 120\mathrm{i}. \end{aligned}$$

* * * * *

Problem 2. Is the following true or false in general?

(a) $\mathrm{Im}(zw) = \mathrm{Im}(z)\,\mathrm{Im}(w)$

(b) $\mathrm{Re}(zw) = \mathrm{Re}(z)\,\mathrm{Re}(w)$

Solution. Let $z = a+bi$ and $w = c+di$. Then $zw = (a+bi)(c+di) = (ac-bd)+(ad+bc)i$.
Part (a). Observe that

$$\mathrm{Im}(zw) = ad + bc \neq bd = \mathrm{Im}(z)\,\mathrm{Im}(w)\,.$$

Hence, the statement is false in general.
Part (b). Observe that

$$\mathrm{Re}(zw) = ac - bd \neq ac = \mathrm{Re}(z)\,\mathrm{Re}(w)\,.$$

Hence, the statement is false in general.

Problem 3.

(a) Find the complex number z such that $\frac{z-2}{z} = 1 + \mathrm{i}$.

(b) Given that $u = 2 + \mathrm{i}$ and $v = -2 + 4\mathrm{i}$, find in the form $a + b\mathrm{i}$, where $a, b \in \mathbb{R}$, the complex number z such that $\frac{1}{z} = \frac{1}{u} + \frac{1}{v}$.

Solution.

Part (a).

$$\frac{z-2}{z} = 1 + \mathrm{i} \implies z - 2 = z + \mathrm{i}z \implies \mathrm{i}z = -2 \implies z = -\frac{2}{\mathrm{i}} = 2\mathrm{i}.$$

Part (b).

$$\frac{1}{z} = \frac{1}{u} + \frac{1}{v} \implies z = \frac{1}{1/u + 1/v} = \frac{uv}{u+v} = \frac{(2+\mathrm{i})(-2+4\mathrm{i})}{(2+\mathrm{i}) + (-2+4\mathrm{i})} = \frac{-8+6\mathrm{i}}{5\mathrm{i}} = \frac{6}{5} + \frac{8}{5}\mathrm{i}.$$

* * * * *

Problem 4. The complex numbers z and w are $1 + a\mathrm{i}$ and $b - 2\mathrm{i}$ respectively, where a and b are real and a is negative. Given that $zw^* = 8i$, find the exact values of a and b.

Solution. Note that

$$zw^* = (1 + a\mathrm{i})(b + 2\mathrm{i}) = (b - 2a) + (2 + ab)\mathrm{i}.$$

Comparing real and imaginary parts, we have $b - 2a = 0 \implies b = 2a$ and $2 + ab = 8$. Hence, $2a^2 = 6$, giving $a = -\sqrt{3}$ and $b = -2\sqrt{3}$.

* * * * *

Problem 5. Find, in the form $x + \mathrm{i}y$, the two complex numbers z satisfying both of the equations

$$\frac{z}{z^*} = \frac{3}{5} + \frac{4}{5}\mathrm{i} \quad \text{and} \quad zz^* = 5.$$

Solution. Multiplying both equations together, we have $z^2 = 3 + 4\mathrm{i}$. Let $z = x + \mathrm{i}y$, with $x, y \in \mathbb{R}$. We thus have $z^2 = x^2 - y^2 + 2xy\mathrm{i} = 3 + 4\mathrm{i}$. Comparing real and imaginary parts, we obtain the following system:

$$x^2 - y^2 = 3, \quad 2xy = 4.$$

Squaring the second equation yields $x^2y^2 = 4$. From the first equation, we have $x^2 = 3+y^2$. Thus, $y^2(3 + y^2) = 4 \implies y^2 = 1 \implies y = \pm 1 \implies x = \pm 2$. Hence, $z = 2 + \mathrm{i}$ or $z = -2 - \mathrm{i}$.

Problem 6.

(a) Given that $\mathrm{i}w + 3z = 2 + 4\mathrm{i}$ and $w + (1 - \mathrm{i})z = 2 - \mathrm{i}$, find z and w in the form of $x + \mathrm{i}y$, where x and y are real numbers.

(b) Determine the value of k such that $z = \frac{1-k\mathrm{i}}{\sqrt{3}+\mathrm{i}}$ is purely imaginary, where $k \in \mathbb{R}$.

Solution.

Part (a). Let $w = a + b\mathrm{i}$ and $z = c + d\mathrm{i}$. Then

$$\mathrm{i}w + 3z = \mathrm{i}(a + b\mathrm{i}) + 3(c + d\mathrm{i}) = (-b + 3c) + (a + 3d)\mathrm{i} = 2 + 4\mathrm{i}$$

and

$$w + (1 - \mathrm{i})z = (a + b\mathrm{i}) + (1 - \mathrm{i})(c + d\mathrm{i}) = (a + c + d) + (b - c + d)\mathrm{i} = 2 - \mathrm{i}.$$

Comparing the real and imaginary parts of both equations yields the following system:

$$\begin{cases} -b + 3c = 2 \\ a + 3d = 4 \\ a + c + d = 2 \\ b - c + d = -1 \end{cases}$$

which has the unique solution $a = 1$, $b = -2$, $c = 0$ and $d = 1$. Hence, $w = 1 - 2\mathrm{i}$ and $z = \mathrm{i}$.

Part (b).

$$z = \frac{1 - k\mathrm{i}}{\sqrt{3} + \mathrm{i}} = \frac{(1 - k\mathrm{i})(\sqrt{3} - \mathrm{i})}{\sqrt{3}^2 + 1^2} = \frac{1}{4}(\sqrt{3} - \mathrm{i} - k\sqrt{3}\mathrm{i} - k) = \frac{1}{4}\left[(\sqrt{3} - k) - (1 + k\sqrt{3})\mathrm{i}\right].$$

Since z is purely imaginary, $\mathrm{Re}(z) = 0$. Hence, $\frac{1}{4}(\sqrt{3} - k) = 0 \implies k = \sqrt{3}$.

* * * * *

Problem 7.

(a) The complex number $x + \mathrm{i}y$ is such that $(x + iy)^2 = \mathrm{i}$. Find the possible values of the real numbers x and y, giving your answers in exact form.

(b) Hence, find the possible values of the complex number w such that $w^2 = -\mathrm{i}$.

Solution.

Part (a). Note that $(x + \mathrm{i}y)^2 = x^2 - y^2 + 2xy\mathrm{i} = \mathrm{i}$. Comparing real and imaginary parts, we have

$$x^2 - y^2 = 0, \quad 2xy = 1.$$

Note that the second equation implies that both x and y have the same sign. Hence, from the first equation, we have $x = y$. Thus, $x^2 = y^2 = 1/2 \implies x = y = \pm 1/\sqrt{2}$.

Part (b).

$$w^2 = -\mathrm{i} \implies (w^*)^2 = \mathrm{i} \implies w^* = \pm\frac{1}{\sqrt{2}} \pm \frac{1}{\sqrt{2}}\mathrm{i} \implies w = \pm\frac{1}{\sqrt{2}} \mp \frac{1}{\sqrt{2}}\mathrm{i}.$$

Problem 8.

(a) The roots of the equation $z^2 = -8\mathrm{i}$ are z_1 and z_2. Find z_1 and z_2 in Cartesian form $x + \mathrm{i}y$, showing your working.

(b) Hence, or otherwise, find in Cartesian form the roots w_1 and w_2 of the equation $w^2 + 4w + (4 + 2\mathrm{i}) = 0$.

Solution.

Part (a). Let $z = x + \mathrm{i}y$ where $x, y \in \mathbb{R}$. Then $(x + \mathrm{i}y)^2 = x^2 - y^2 + 2xy\mathrm{i} = -8\mathrm{i}$. Comparing real and imaginary parts, we have the following system:

$$x^2 - y^2 = 0, \quad 2xy = 8.$$

From the second equation, we know that x and y have opposite signs. Hence, from the first equation, we have that $x = -y$. Thus, $x^2 = 4 \implies x = \pm 2 \implies y = \mp 2$. Thus, $z = \pm 2(1 - \mathrm{i})$, whence $z_1 = 2 - 2\mathrm{i}$ and $z_2 = -2 + 2\mathrm{i}$.

Part (b).

$$\begin{aligned} w^2 + 4w + (4 + 2\mathrm{i}) = 0 &\implies (w + 2)^2 = -2\mathrm{i} \implies (2w + 4)^2 = -8\mathrm{i} \\ &\implies 2w + 4 = \pm 2(1 - \mathrm{i}) \implies w = 2 \pm (1 - \mathrm{i}). \end{aligned}$$

* * * * *

Problem 9. One of the roots of the equations $2x^3 - 9x^2 + 2x + 30 = 0$ is $3 + \mathrm{i}$. Find the other roots of the equation.

Solution. Let $P(x) = 2x^3 - 9x^2 + 2x + 30$. Since $P(x)$ is a polynomial with real coefficients, by the conjugate root theorem, we have that $(3 + \mathrm{i})^* = 3 - \mathrm{i}$ is also a root of $P(x)$. Let α be the third root of $P(x)$. Then

$$P(x) = 2x^3 - 9x^2 + 2x + 30 = 2(x - \alpha)\left[x - (3 + i)\right]\left[x - (3 - i)\right].$$

Comparing constants,

$$2(-\alpha)(-3 - \mathrm{i})(-3 + \mathrm{i}) = 30 \implies \alpha = -\frac{15}{(-3 - \mathrm{i})(-3 + \mathrm{i})} = -\frac{3}{2}.$$

Hence, the other roots of the equation are $3 - \mathrm{i}$ and $-3/2$.

* * * * *

Problem 10. Obtain a cubic equation having 2 and $\frac{5}{4} - \frac{\sqrt{7}}{4}\mathrm{i}$ as two of its roots, in the form $az^3 + bz^2 + cz + d = 0$, where a, b, c and d are real integral coefficients to be determined.

Solution. Let $P(z) = az^3 + bz^2 + cz + d$. Since $P(z)$ is a polynomial with real coefficients, by the conjugate root theorem, we have that $\left(\frac{5}{4} - \frac{\sqrt{7}}{4}\mathrm{i}\right)^* = \frac{5}{4} + \frac{\sqrt{7}}{4}\mathrm{i}$ is also a root of $P(z)$. We can thus write $P(z)$ as

$$\begin{aligned} P(z) &= k(z - 2)\left[z - \left(\frac{5}{4} - \frac{\sqrt{7}}{4}\mathrm{i}\right)\right]\left[z - \left(\frac{5}{4} + \frac{\sqrt{7}}{4}\mathrm{i}\right)\right] \\ &= k(z - 2)\left[\left(z - \frac{5}{4}\right)^2 + \left(\frac{\sqrt{7}}{4}\right)^2\right] = k(z - 2)\left(z^2 - \frac{5}{2}z + 2\right) \\ &= \frac{1}{2}k(2z^3 - 9z^2 + 14z - 8), \end{aligned}$$

where k is an arbitrary real number. Taking $k = 2$, we have $P(z) = 2z^3 - 9z^2 + 14z - 8$, whence $a = 2$, $b = -9$, $c = 14$ and $d = -8$.

* * * * *

Problem 11.

(a) Verify that $-1 + 5\mathrm{i}$ is a root of the equation $w^2 + (-1 - 8\mathrm{i})w + (-17 + 7\mathrm{i}) = 0$. Hence, or otherwise, find the second root of the equation in Cartesian form, $p + \mathrm{i}q$, showing your working.

(b) The equation $z^3 - 5z^2 + 16z + k = 0$, where k is a real constant, has a root $z = 1 + a\mathrm{i}$, where a is a positive real constant. Find the values of a and k, showing your working.

Solution.

Part (a). Let $P(w) = w^2 + (-1 - 8\mathrm{i})w + (-17 + 7\mathrm{i})$. Consider $P(-1 + 5\mathrm{i})$.

$$\begin{aligned} P(-1+5\mathrm{i}) &= (-1+5\mathrm{i})^2 + (-1-8\mathrm{i})(-1+5\mathrm{i}) + (-17+7\mathrm{i}) \\ &= (1 - 10\mathrm{i} - 25) + (1 - 5\mathrm{i} + 8\mathrm{i} + 40) + (-17 + 7\mathrm{i}) = 0. \end{aligned}$$

Hence, $-1 + 5\mathrm{i}$ is a root of $w^2 + (-1 - 8\mathrm{i})w + (-17 + 7\mathrm{i}) = 0$.

Let α be the other root of the equation. By Vieta's formula, we have

$$\alpha + (-1 + 5\mathrm{i}) = -\left(\frac{-1 - 8\mathrm{i}}{1}\right) = 1 + 8\mathrm{i} \implies \alpha = 2 + 3\mathrm{i}.$$

Part (b). Let $P(z) = z^3 - 5z^2 + 16z + k$. Then $P(1 + a\mathrm{i}) = 0$. Note that

$$\begin{aligned} P(1+a\mathrm{i}) &= (1+a\mathrm{i})^3 - 5(1+a\mathrm{i})^2 + 16(1+a\mathrm{i}) + k \\ &= \left[1 + 3a\mathrm{i} - 3a^2 - a^3\mathrm{i}\right] - 5(1 + 2a\mathrm{i} - a^2) + (16 + 16a\mathrm{i}) + k \\ &= (12 + k + 2a^2) + (9 - a^2)a\mathrm{i}. \end{aligned}$$

Comparing real and imaginary parts, we have $a(9 - a^2) = 0 \implies a = 3$ (since $a > 0$) and $12 + k + 2a^2 = 0 \implies k = -30$.

Self-Practice A10.1

Problem 1. By writing $z = x + \mathrm{i}y$, $x, y \in \mathbb{R}$, solve the simultaneous equations

$$z^2 + zw - 2 = 0 \quad \text{and} \quad z^* = \frac{w}{1+\mathrm{i}},$$

where z^* is the conjugate of z.

Solution. From the second equation, we see that $w = (1+\mathrm{i})z^*$. Substituting this into the first equation yields

$$z^2 + zz^*(1+\mathrm{i}) - 2 = 0.$$

Let $z = x + \mathrm{i}y$, where $x, y \in \mathbb{R}$. Then

$$(x+\mathrm{i}y)^2 + \left(x^2+y^2\right)(1+\mathrm{i}) - 2 = 0.$$

Simplifying, we get

$$2\left(x^2-1\right) + (x+y)^2\,\mathrm{i} = 0.$$

Comparing real and imaginary parts, we require $x^2 - 1 = 0$ and $x + y = 0$, so $x = \pm 1$ and $y = -x = \mp 1$, so $z = \pm 1 \mp \mathrm{i}$.

When $z = 1 - \mathrm{i}$, we have $w = (1+\mathrm{i})^2 = 2\mathrm{i}$. When $z = -1 + \mathrm{i}$, we have $w = (-1+\mathrm{i})(1+\mathrm{i}) = -2$.

* * * * *

Problem 2. Given that the complex numbers w and z satisfy the equations

$$w^* + 2z = \mathrm{i} \quad \text{and} \quad w + (1-2\mathrm{i})z = 3 + 3\mathrm{i},$$

find w and z in the form $a + b\mathrm{i}$, where a and b are real.

Solution. From the first equation, we obtain $w = -\mathrm{i} - 2z^*$. Substituting this into the second equation, we see that

$$(-\mathrm{i} - 2z^*)(1-2\mathrm{i})\,z = 3 + 3\mathrm{i}.$$

Let $z = a + b\mathrm{i}$, where $a, b \in \mathbb{R}$. Then

$$[-\mathrm{i} - 2(a - b\mathrm{i})] + (1-2\mathrm{i})(a + b\mathrm{i}) = 3 + 3\mathrm{i},$$

which upon simplification yields

$$(2b - a) + \mathrm{i}\,(3b - 2a) = 3 + 4\mathrm{i}.$$

Comparing real and imaginary parts, we require $2b - a = 3$ and $3b - 2a = 3$, which gives $a = 1$ and $b = 2$. Thus, $z = 1 + 2\mathrm{i}$ and $w = -\mathrm{i} - 2(1 - 2\mathrm{i}) = -2 + 3\mathrm{i}$.

* * * * *

Problem 3.

(a) Determine the complex numbers u and v for which

$$z^2 + (6-2\mathrm{i})z = (z-u)^2 - v, \quad \forall z \in \mathbb{C}.$$

(b) Write down the square roots of $7 - 24\mathrm{i}$. Hence, solve the quadratic equation $z^2 + (6-2\mathrm{i})z = -1 - 18\mathrm{i}$.

Solution.
Part (a). Completing the square, we see that

$$z^2 - (6-2\mathrm{i})\,z = (z + (3-\mathrm{i}))^2 - (3-\mathrm{i})^2\,,$$

so $u = -(3-\mathrm{i}) = -3+\mathrm{i}$ and $v = (3-\mathrm{i})^2 = 8-6\mathrm{i}$.
Part (b). Using G.C., $\pm\sqrt{7-24\mathrm{i}} = \pm\,(4-3\mathrm{i})$. From (a), we see that

$$(z + (3-\mathrm{i}))^2 - (8-6\mathrm{i}) = z^2 + (6-2\mathrm{i})\,z = -1-18\mathrm{i},$$

thus

$$(z + (3-\mathrm{i}))^2 = -1-18\mathrm{i}+8-6\mathrm{i} = 7-24\mathrm{i},$$

so

$$z + (3-\mathrm{i}) = \pm\,(4-3\mathrm{i})\,.$$

Finally, we obtain $z = 1-2\mathrm{i}$ or $z = -7+4\mathrm{i}$.

* * * * *

Problem 4. If $z = i$ is a root of the equation $z^3+(1-3\mathrm{i})z^2-(2+3\mathrm{i})z-2=0$, determine the other roots. Hence, find the roots of the equation $w^3+(1+3\mathrm{i})w^2+(3\mathrm{i}-2)w-2=0$.

Solution. By inspection,

$$(-1)^3 + (1-3\mathrm{i})(-1)^2 - (2+3\mathrm{i})(-1) - 2 = 0,$$

so $z = -1$ is a root. Let α be the other root. By Vieta's formula, $\mathrm{i} + (-1) + \alpha = -\,(1-3\mathrm{i}) \implies \alpha = 2\mathrm{i}$. Thus, the roots are $z = \mathrm{i}$, $z = 2\mathrm{i}$ and $z = -1$.

Conjugating the cubic in w, we see that

$$(w^*)^3 + (1-3\mathrm{i})\,(w^*)^2 + (-2-3\mathrm{i})w^* - 2 = 0,$$

so

$$w^* = \mathrm{i}, 2\mathrm{i}, -1 \implies w = -\mathrm{i}, -2\mathrm{i}, -1.$$

* * * * *

Problem 5. Show that the equation $z^4-2z^3+6z^2-8z+8=0$ has a root of the form $k\mathrm{i}$, where k is real. Hence, solve the equation $z^4-2z^3+6z^2-8z+8=0$.

Solution. Let $z = k\mathrm{i}$. Then

$$(k\mathrm{i})^4 - 2(k\mathrm{i})^3 + 6(k\mathrm{i})^2 - 8(k\mathrm{i}) + 8 = \left(k^4-6k^2+8\right) + \mathrm{i}\left(2k^3-8k\right) = 0.$$

We hence require

$$k^4-6k^2+8=0 \quad \text{and} \quad 2k^3-8k=0.$$

By inspection $k = 2$ satisfies both equation, so $z = 2\mathrm{i}$ is a root.

Since the coefficients of the quartic are all real, by the conjugate root theorem, $z = -2\mathrm{i}$ is also a root. Let $P(z)$ be a degree two polynomial such that the quartic factorizes as

$$z^4-2z^3+6z^2-8z+8 = (z-2\mathrm{i})\,(z+2\mathrm{i})\,P(z).$$

Then

$$P(z) = \frac{z^4-2z^3+6z^2-8z+8}{z^2+4} = z^2-2z+2.$$

Solving $P(z) = 0$, we get $z = 1 \pm \mathrm{i}$, so the roots to the quartic are $z = 2\mathrm{i}$, $-2\mathrm{i}$, $1+\mathrm{i}$, $1-\mathrm{i}$.

Problem 6. Verify that $-2+\mathrm{i}$ is a root of the equation $z^4+24z+55=0$. Hence, determine the other roots.

Solution. Substituting $z=-2+\mathrm{i}$, we see that

$$(-2+\mathrm{i})^4+24\,(-2+\mathrm{i})+55=0,$$

so it is a root. Since the coefficients of $z^4+24z+55$ are real, by the conjugate root theorem, $z=-2-\mathrm{i}$ is also a root. Let $P(z)$ be a degree two polynomial such that the quartic factorizes as

$$z^4+24z+55=(z-(-2+\mathrm{i}))\,(z-(-2-\mathrm{i}))\,P(z).$$

Then

$$P(z)=\frac{z^4+24z+55}{z^2+4z+5}=z^2+4z+11.$$

Solving $P(z)=0$, we get $z=2\pm\sqrt{7}\mathrm{i}$. Hence, the roots of the quartic are $z=-2+\mathrm{i}$, $-2-\mathrm{i}$, $2+\sqrt{7}\mathrm{i}$, $2-\sqrt{7}\mathrm{i}$.

Assignment A10.1

Problem 1. The complex number w is such that $ww^* + 2w = 3 + 4\mathrm{i}$, where w^* is the complex conjugate of w. Find w in the form $a + \mathrm{i}b$, where a and b are real.

Solution. Note $ww^* = (\operatorname{Re} w)^2 + (\operatorname{Im} w)^2 \in \mathbb{R}$.

Taking the imaginary part of the given equation,

$$\operatorname{Im}(ww^* + 2w) = \operatorname{Im}(3 + 4i) \implies 2 \operatorname{Im} w = 4 \implies \operatorname{Im} w = 2.$$

Taking the real part of the given equation,

$$\begin{aligned} \operatorname{Re}(ww^* + 2w) = \operatorname{Re}(3 + 4i) &\implies \left[(\operatorname{Re} w)^2 + (\operatorname{Im} w)^2\right] + 2 \operatorname{Re} w = 3 \\ \implies (\operatorname{Re} w)^2 + 2 \operatorname{Re}(w) + 1 = 0 &\implies (\operatorname{Re} w + 1)^2 = 0 \implies \operatorname{Re}(w) = -1. \end{aligned}$$

Hence, $w = -1 + 2\mathrm{i}$.

* * * * *

Problem 2. Express $(3 - \mathrm{i})^2$ in the form $a + \mathrm{i}b$.

Hence, or otherwise, find the roots of the equation $(z + \mathrm{i})^2 = -8 + 6\mathrm{i}$.

Solution. We have

$$(3 - \mathrm{i})^2 = 3^2 - 6\mathrm{i} + \mathrm{i}^2 = 8 - 6\mathrm{i}.$$

Consider $(z + \mathrm{i})^2 = -8 + 6\mathrm{i}$. Note that $-(z + \mathrm{i})^2 = (\mathrm{i}z - 1)^2$.

$$\begin{aligned} (z + \mathrm{i})^2 = -8 + 6\mathrm{i} &\implies (\mathrm{i}z - 1)^2 = 8 - 6\mathrm{i} \implies \mathrm{i}z - 1 = \pm(3 - \mathrm{i}) \\ \implies z = \frac{1}{\mathrm{i}}(1 \pm (3 - \mathrm{i})) &= -\mathrm{i}(1 \pm (3 - \mathrm{i})) = -1 - 4\mathrm{i} \text{ or } 1 + 2\mathrm{i}. \end{aligned}$$

* * * * *

Problem 3.

(a) It is given that $z_1 = 1 + \sqrt{3}\mathrm{i}$. Find the value of z_1^3, showing clearly how you obtain your answer.

(b) Given that $1 + \sqrt{3}\mathrm{i}$ is a root of the equation

$$2z^3 + az^2 + bz + 4 = 0$$

find the values of the real numbers a and b. Hence, solve the above equation.

Solution.

Part (a). We have

$$z_1^3 = \left(1 + \sqrt{3}\mathrm{i}\right)^3 = 1 + 3\left(\sqrt{3}\mathrm{i}\right) + 3\left(\sqrt{3}\mathrm{i}\right)^2 + \left(\sqrt{3}\mathrm{i}\right)^3 = 1 + 3\sqrt{3}\mathrm{i} - 9 - 3\sqrt{3}\mathrm{i} = -8.$$

Part (b). Since $1 + \sqrt{3}\mathrm{i}$ is a root of the given equation, we have

$$\begin{aligned} &2\left(1 + \sqrt{3}\mathrm{i}\right)^3 + a\left(1 + \sqrt{3}\mathrm{i}\right)^2 + b\left(1 + \sqrt{3}\mathrm{i}\right) + 4 = 0 \\ \implies &-16 + a\left(-2 + 2\sqrt{3}\mathrm{i}\right) + b\left(1 + \sqrt{3}\mathrm{i}\right) + 4 = 0 \implies (-2a + b) + \sqrt{3}(2a + b)\mathrm{i} = 12. \end{aligned}$$

Comparing real and imaginary parts, we obtain $-2a + b = 12$ and $2a + b = 0$, whence $a = -3$ and $b = 6$.

Since the coefficients of $2z^3 + az^2 + bz + 4$ are all real, the second root is $(1+\sqrt{3}\mathrm{i})^* = 1-\sqrt{3}\mathrm{i}$. Let the third root be α. By Vieta's formula,

$$(1+\sqrt{3}\mathrm{i})(1-\sqrt{3}\mathrm{i})\alpha = -\frac{4}{2} \implies 4\alpha = -2 \implies \alpha = -\frac{1}{2}.$$

The roots of the equation are hence $1+\sqrt{3}\mathrm{i}$, $1-\sqrt{3}\mathrm{i}$ and $-\frac{1}{2}$.

* * * * *

Problem 4. The complex number z is such that $az^2 + bz + a = 0$ where a and b are real constants. It is given that $z = z_0$ is a solution to this equation where $\operatorname{Im}(z_0) \neq 0$.

(a) Verify that $z = \frac{1}{z_0}$ is the other solution. Hence, show that $|z_0| = 1$.

Take $\operatorname{Im}(z_0) = 1/2$ for the rest of the question.

(b) Find the possible complex numbers for z_0.

(c) If $\operatorname{Re}(z_0) > 0$, find b in terms of a.

Solution.

Part (a).

$$a\left(\frac{1}{z_0}\right)^2 + b\left(\frac{1}{z_0}\right) + a = \left(\frac{1}{z_0}\right)^2\left(a + bz_0 + az_0^2\right) = 0$$

Hence, $z = 1/z_0$ is a root of the given equation.

Since $a, b \in \mathbb{R}$, by the conjugate root theorem, $z_0^* = 1/z_0$. Hence,

$$z_0 z_0^* = 1 \implies \operatorname{Re}(z_0)^2 + \operatorname{Im}(z_0)^2 = |z_0|^2 = 1 \implies |z_0| = 1.$$

Part (b). Let $z_0 = x + \frac{1}{2}\mathrm{i}$. Then

$$\left|x + \frac{1}{2}i\right| = 1 \implies x^2 + \left(\frac{1}{2}\right)^2 = 1^2 \implies x^2 = \frac{3}{4} \implies x = \pm\frac{\sqrt{3}}{2}.$$

Hence, $z_0 = \frac{\sqrt{3}}{2} + \frac{1}{2}\mathrm{i}$ or $z_0 = -\frac{\sqrt{3}}{2} + \frac{1}{2}\mathrm{i}$.

Part (c). Since $\operatorname{Re}(z_0) > 0$, we have $z_0 = \frac{\sqrt{3}}{2} + \frac{1}{2}\mathrm{i}$. By Vieta's formula,

$$-\frac{b}{a} = z_0 + \frac{1}{z_0} = z_0 + z_0^* = 2\operatorname{Re}(z_0) = \sqrt{3} \implies b = -\sqrt{3}a.$$

A10.2 Complex Numbers - Complex Numbers in Polar Form

Tutorial A10.2

Problem 1. Is the following true or false in general?

(a) $\left|w^2\right| = |w|^2$

(b) $|z + 2w| = |z| + |2w|$

Solution.
Part (a). Let $w = r\mathrm{e}^{\mathrm{i}\theta}$, where $r, \theta \in \mathbb{R}$. Note that $\left|\mathrm{e}^{\mathrm{i}\theta}\right| = \left|\mathrm{e}^{2\mathrm{i}\theta}\right| = 1$.

$$\left|w^2\right| = \left|r^2\mathrm{e}^{2\mathrm{i}\theta}\right| = r^2\left|\mathrm{e}^{2\mathrm{i}\theta}\right| = r^2 = r^2\left|\mathrm{e}^{\mathrm{i}\theta}\right|^2 = \left|r\mathrm{e}^{\mathrm{i}\theta}\right|^2 = |w|^2 .$$

The statement is hence true in general.
Part (b). Take $z = 1$ and $w = -1$.

$$|z + 2w| = |1 - 2| = 1 \neq 3 = |1| + |2(-1)| = |z| + |2w| .$$

The statement is hence false in general.

* * * * *

Problem 2. Express the following complex numbers z in polar form $r(\cos\theta + \mathrm{i}\sin\theta)$ with exact values.

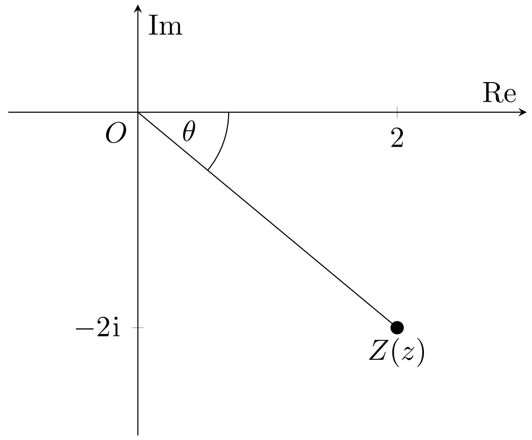
(a) $z = 2 - 2\mathrm{i}$

(b) $z = -1 + \mathrm{i}\sqrt{3}$

(c) $z = -5\mathrm{i}$

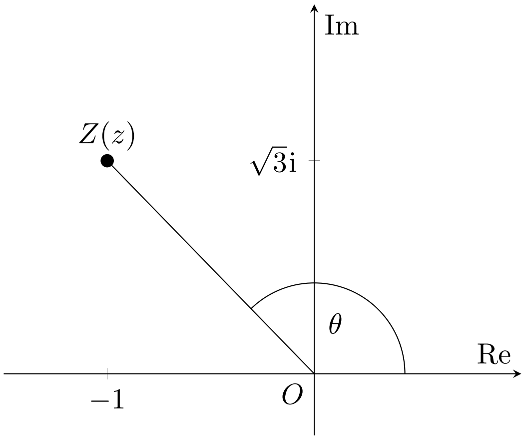
(d) $z = -2\sqrt{3} - 2\mathrm{i}$

Solution.
Part (a).



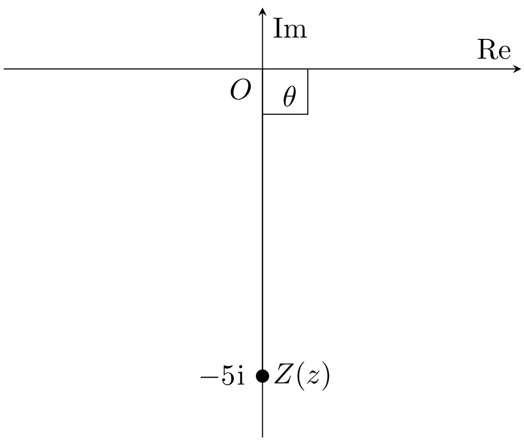
We have $r^2 = 2^2 + (-2)^2 \implies r = 2\sqrt{2}$ and $\tan\theta = -2/2 \implies \theta = -\pi/4$. Hence, $2 - 2\mathrm{i} = 2\sqrt{2}\left[\cos\left(-\frac{\pi}{4}\right) + \mathrm{i}\sin\left(-\frac{\pi}{4}\right)\right]$.

Part (b).



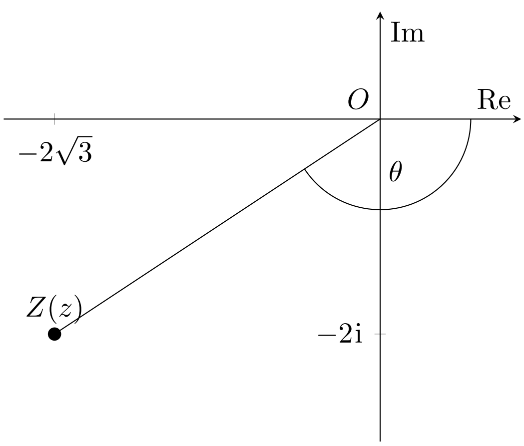
We have $r^2 = (-1)^2 + (\sqrt{3})^2 \implies r = 2$ and $\tan t = \sqrt{3}/(-1) \implies \theta = 2\pi/3$. Hence, $-1 + \sqrt{3}\mathrm{i} = 2\left[\cos\left(\frac{2\pi}{3}\right) + \mathrm{i}\sin\left(\frac{2\pi}{3}\right)\right]$.

Part (c).



We have $r = 5$ and $\theta = -\pi/2$. Hence, $-5\mathrm{i} = 5\left[\cos\left(-\frac{\pi}{2}\right) + \mathrm{i}\sin\left(-\frac{\pi}{2}\right)\right]$.

Part (d).



We have $r^2 = (-2\sqrt{3})^2 + (-2)^2 \implies r = 4$ and $\tan t = -2/(-2\sqrt{3}) \implies \theta = -5\pi/6$. Hence, $-2\sqrt{3} - 2\mathrm{i} = 4\left[\cos\left(-\frac{5\pi}{6}\right) + \mathrm{i}\sin\left(-\frac{5\pi}{6}\right)\right]$.

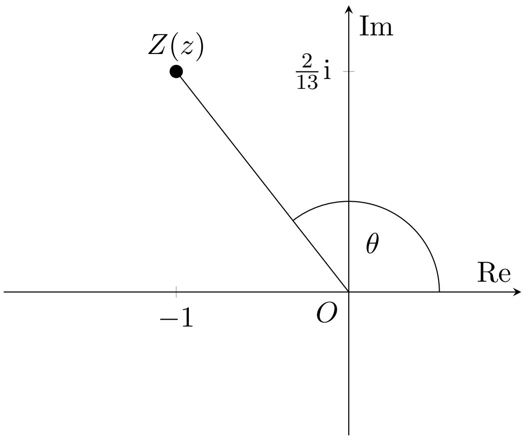
Problem 3. Express the following complex numbers z in exponential form $r\mathrm{e}^{\mathrm{i}\theta}$.

(a) $z = -1 + \frac{2}{13}\mathrm{i}$

(b) $z = \cos 50° - \mathrm{i}\sin 50°$

Solution.

Part (a).



We have $r^2 = (-1)^2 + \left(\frac{2}{13}\right)^2 \implies r = 1.01$ (3 s.f.) and $\tan t = \frac{2/13}{-1} \implies \theta = 2.99$ (3 s.f.). Hence, $-1 + \frac{2}{13}\mathrm{i} = 1.01\mathrm{e}^{2.99\mathrm{i}}$.

Part (b). We have $r = 1$ and $\theta = -50° = -\frac{5}{18}\pi$. Hence, $\cos 50° + i\sin 50° = \mathrm{e}^{-\mathrm{i}\frac{5}{18}\pi}$.

* * * * *

Problem 4. Express the following complex numbers z in Cartesian form.

(a) $z = 7\mathrm{e}^{1-5\mathrm{i}}$

(b) $z = 6\left(\cos\frac{\pi}{8} - \mathrm{i}\sin\frac{\pi}{8}\right)$

Solution.

Part (a). We have

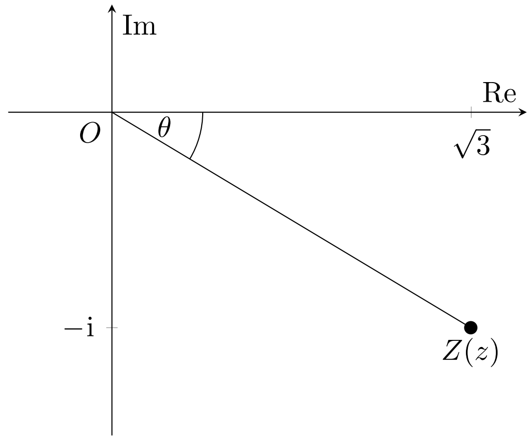
$$z = 7\mathrm{e}^{1-5\mathrm{i}} = 7\mathrm{e}\cdot\mathrm{e}^{-5\mathrm{i}} = 7\mathrm{e}\left[\cos(-5) + \mathrm{i}\sin(-5)\right] = 5.40 + 18.2\mathrm{i} \text{ (3 s.f.)}.$$

Part (b). We have

$$z = 6\left(\cos\frac{\pi}{8} - \mathrm{i}\sin\frac{\pi}{8}\right) = 5.54 - 2.30\mathrm{i} \text{ (3 s.f.)}.$$

Problem 5. Given that $z = \sqrt{3} - \mathrm{i}$, find the exact modulus and argument of z. Hence, find the exact modulus and argument of $1/z^2$ and z^{10}.

Solution.



We have $r^2 = (\sqrt{3})^2 + (-1)^2 \implies r = 2$ and $\tan\theta = -1/\sqrt{3} \implies \theta = -\pi/6$. Hence, $|z| = 2$ and $\arg z = -\pi/6$.

Note that $\left|1/z^2\right| = |z|^{-2} = 1/4$. Also, $\arg\left(1/z^2\right) = -2\arg z = \pi/3$.

Note that$\left|z^{10}\right| = |z|^1 0 = 1024$. Also, $\arg z^{10} = 10\arg z = -5\pi/3 \equiv \pi/3$.

* * * * *

Problem 6. If $\arg(z - 1/2) = \pi/5$, determine $\arg(2z - 1)$.

Solution.

$$\arg(2z-1) = \arg\left(\frac{1}{2}\left[z - \frac{1}{2}\right]\right) = \arg\left(z - \frac{1}{2}\right) = \frac{\pi}{5}.$$

* * * * *

Problem 7. In an Argand diagram, points P and Q represent the complex numbers $z = 1 + \mathrm{i}$ and $w = 1 + 2\mathrm{i}$ respectively, and O is the origin.

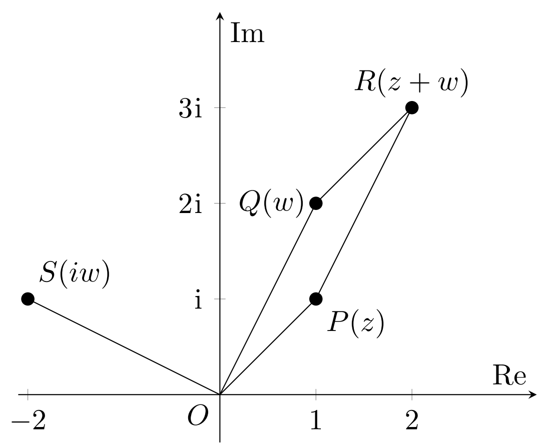
(a) Mark on the Argand diagram the points P and Q, and the points R and S which represent $z + w$ and iw respectively.

(b) What is the geometrical shape of $OPRQ$?

(c) State the angle SOP.

Solution.

Part (a).



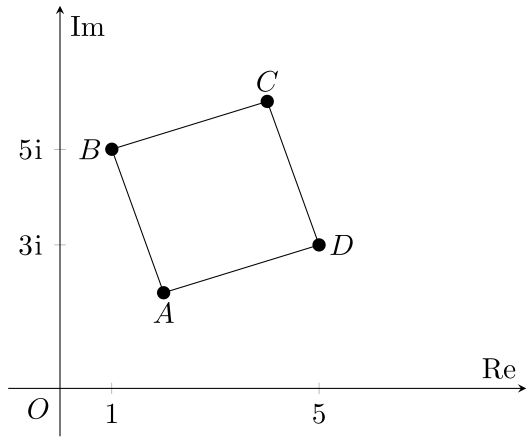
Part (b). $OPRQ$ is a parallelogram.

Part (c). $\angle SOP = \pi/2$.

* * * * *

Problem 8. B and D are points in the Argand diagram representing the complex numbers $1+5\mathrm{i}$ and $5+3\mathrm{i}$ respectively. Given that BD is a diagonal of the square $ABCD$, calculate the complex numbers represented by A and C.

Solution.



Let $A(x+\mathrm{i}y)$. Since $AB \perp AD$, we have $b-a = \mathrm{i}(d-a)$.

$$b - a = \mathrm{i}(d-a) \implies (1+5\mathrm{i}) - (x+\mathrm{i}y) = i\left[(5+3\mathrm{i}) - (x+\mathrm{i}y)\right]$$
$$\implies (1-x) + (5-y)\mathrm{i} = (-3+y) + (5-x)\mathrm{i} \implies (x+y) + (y-x)\mathrm{i} = 4.$$

Comparing real and imaginary parts, we obtain $x = y = 2$. Hence, $A(2+2\mathrm{i})$.

Let $C(u+\mathrm{i}v)$. Since $CB \perp CD$, we have $d-c = \mathrm{i}(b-c)$.

$$d - c = \mathrm{i}(b-c) \implies (5+3\mathrm{i}) - (u+\mathrm{i}v) = \mathrm{i}\left[(1+5\mathrm{i}) - (u+\mathrm{i}v)\right]$$
$$\implies (5-u) + (3-v)\mathrm{i} = (-5+v) + (1-u)\mathrm{i} \implies (u+v) + (v-u)\mathrm{i} = 10+2\mathrm{i}.$$

Comparing real and imaginary parts, we obtain $u = 4$ and $v = 6$. Hence, $C(4+6\mathrm{i})$.

* * * * *

Problem 9.

(a) Given that $u = 2\left(\cos\frac{\pi}{6} + \mathrm{i}\sin\frac{\pi}{6}\right)$ and $w = 4\left(\cos\frac{\pi}{3} - \mathrm{i}\sin\frac{\pi}{3}\right)$, find the modulus and argument of u^*/w^3 in exact form.

(b) Let z be the complex number $-1+\mathrm{i}\sqrt{3}$. Find the value of the real number a such that $\arg\left(z^2+az\right) = -\pi/2$.

Solution.

Part (a). Note that $|u| = 2$, $\arg u = \pi/6$, $|w| = 4$ and $\arg w = -\pi/3$. Hence,

$$\left|\frac{u^*}{w^3}\right| = \frac{|u^*|}{|w^3|} = \frac{|u|}{|w|^3} = \frac{2}{4^3} = \frac{1}{32}$$

and

$$\arg\frac{u^*}{w^3} = \arg u^* - \arg w^3 = -\arg u - 3\arg w = -\frac{\pi}{6} - 3\left(-\frac{\pi}{3}\right) = \frac{5}{6}\pi.$$

Part (b). Since $\arg\left(z^2+az\right)=-\pi/2$, we have that z^2+az is purely imaginary, with a negative imaginary part. Since

$$z^2+az=\left(-1+\mathrm{i}\sqrt{3}\right)^2+a\left(-1+\mathrm{i}\sqrt{3}\right)=\left(-2-2\sqrt{3}\mathrm{i}\right)+a\left(-1+\mathrm{i}\sqrt{3}\right).$$

Hence,

$$\operatorname{Re}\left(z^2+az\right)=0 \implies -2-a=0 \implies a=-2.$$

* * * * *

Problem 10. The complex number w has modulus r and argument θ, where $0<\theta<\pi/2$, and w^* denotes the conjugate of w. State the modulus and argument of p, where $p=w/w^*$. Given that p^5 is real and positive, find the possible values of θ.

Solution. Clearly, $|p|=1$ and $\arg p=2\theta$.

Since p^5 is real and positive, we have $\arg p^5=2\pi n$, where $n\in\mathbb{Z}$. Thus, $\arg p=2\pi n/5=2\theta \implies \theta=\pi n/5$. Since $0<\theta<\pi/2$, the possible values of θ are $\pi/5$ and $2\pi/5$.

* * * * *

Problem 11. The complex number w has modulus $\sqrt{2}$ and argument $-3\pi/4$, and the complex number z has modulus 2 and argument $-\pi/3$. Find the modulus and argument of wz, giving each answer exactly.

By first expressing w and z in the form $x+\mathrm{i}y$, find the exact real and imaginary parts of wz.

Hence, show that $\sin\frac{\pi}{12}=\frac{\sqrt{3}-1}{2\sqrt{2}}$.

Solution. Note that

$$|wz|=|w|\,|z|=2\sqrt{2}$$

and

$$\arg(wz)=\arg w+\arg z=-\frac{3}{4}\pi-\frac{1}{3}\pi=-\frac{13}{12}\pi\equiv\frac{11}{12}\pi.$$

Also,

$$w=\sqrt{2}\left[\cos\left(-\frac{3}{4}\pi\right)+\mathrm{i}\sin\left(-\frac{3}{4}\pi\right)\right]=\sqrt{2}\left(-\frac{1}{\sqrt{2}}-\frac{1}{\sqrt{2}}\mathrm{i}\right)=-1-\mathrm{i}$$

and

$$z=2\left[\cos\left(-\frac{\pi}{3}\right)+\mathrm{i}\sin\left(-\frac{\pi}{3}\right)\right]=2\left(\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i}\right)=1-\sqrt{3}\mathrm{i}.$$

Hence,

$$wz=(-1-\mathrm{i})(1-\sqrt{3}\mathrm{i})=(-1+\sqrt{3}-\mathrm{i}-\sqrt{3})=(-1-\sqrt{3})+(\sqrt{3}-1)\mathrm{i},$$

whence $\operatorname{Re}(wz)=-1-\sqrt{3}$ and $\operatorname{Im}(wz)=\sqrt{3}-1$.

From the first part, we have that $wz=2\sqrt{2}\left[\cos\left(\frac{11}{12}\pi\right)+\mathrm{i}\sin\left(\frac{11}{12}\pi\right)\right]$. Thus, $\operatorname{Im}(wz)=2\sqrt{2}\sin\left(\frac{11}{12}\pi\right)=2\sqrt{2}\sin\frac{\pi}{12}$. Equating the result for $\operatorname{Im}(wz)$ found in the second part, we have

$$2\sqrt{2}\sin\frac{\pi}{12}=\sqrt{3}-1 \implies \sin\frac{\pi}{12}=\frac{\sqrt{3}-1}{2\sqrt{2}}.$$

Problem 12. Given that $\frac{5+z}{5-z} = \mathrm{e}^{\mathrm{i}\theta}$, show that z can be written as $5\mathrm{i}\tan\frac{\theta}{2}$.

Solution. Note that

$$\frac{5+z}{5-z} = \mathrm{e}^{\mathrm{i}\theta} \implies 5+z = \mathrm{e}^{\mathrm{i}\theta}(5-z) \implies z + \mathrm{e}^{\mathrm{i}\theta}z = 5\mathrm{e}^{\mathrm{i}\theta} - 5 \implies z = 5\left(\frac{\mathrm{e}^{\mathrm{i}\theta}-1}{\mathrm{e}^{\mathrm{i}\theta}+1}\right).$$

Hence,

$$z = 5\left(\frac{\mathrm{e}^{\mathrm{i}\theta}-1}{\mathrm{e}^{\mathrm{i}\theta}+1}\right) = 5\left(\frac{\mathrm{e}^{\mathrm{i}\theta/2}-\mathrm{e}^{-\mathrm{i}\theta/2}}{\mathrm{e}^{\mathrm{i}\theta/2}+\mathrm{e}^{-\mathrm{i}\theta/2}}\right) = 5\left(\frac{2\mathrm{i}\sin(\theta/2)}{2\cos(\theta/2)}\right) = 5\mathrm{i}\tan\frac{\theta}{2}.$$

* * * * *

Problem 13. The polynomial $P(z)$ has real coefficients. The equation $P(z) = 0$ has a root $r\mathrm{e}^{\mathrm{i}\theta}$, where $r > 0$ and $0 < \theta < \pi$.

(a) Write down a second root in terms of r and θ, and hence show that a quadratic factor of $P(z)$ is $z^2 - 2rz\cos\theta + r^2$.

(b) Given that 3 roots of the equation $z^6 = -64$ are $2\mathrm{e}^{\mathrm{i}\frac{\pi}{6}}$, $2\mathrm{e}^{\mathrm{i}\frac{\pi}{2}}$ and $2\mathrm{e}^{-\mathrm{i}\frac{5\pi}{6}}$, express $z^6 + 64$ as a product of three quadratic factors with real coefficients, giving each factor in non-trigonometric form.

(c) Represent all roots of $z^6 = -64$ on an Argand diagram and interpret the geometrical shape formed by joining the roots.

Solution.

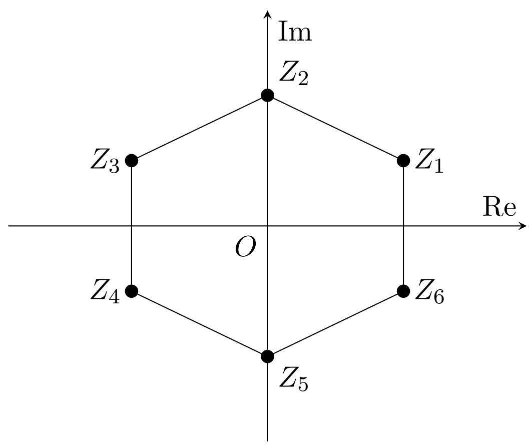
Part (a). Since $P(z)$ has real coefficients, by the conjugate root theorem, $\left(r\mathrm{e}^{\mathrm{i}\theta}\right)^* = r\mathrm{e}^{-\mathrm{i}\theta}$ is also a root of $P(z)$. By the factor theorem, a quadratic factor of $P(z)$ is

$$(z - r\mathrm{e}^{\mathrm{i}\theta})(z - r\mathrm{e}^{-\mathrm{i}\theta}) = z^2 - rz(\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}) + r^2\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\theta} = z^2 - 2rz\cos\theta + r^2.$$

Part (b). Let $r_1 = r_2 = r_3 = 2$ and $\theta_1 = \pi/6$, $\theta_2 = \pi/2$ and $\theta_3 = -5\pi/6$.

$$\begin{aligned} z^6 + 64 &= \left(z^2 - 2r_1z\cos\theta_1 + r_1^2\right)\left(z^2 - 2r_2z\cos\theta_2 + r_2^2\right)\left(z^2 - 2r_3z\cos\theta_3 + r_3^2\right) \\ &= \left(z^2 - 4z\cos\left(\frac{\pi}{6}\right) + 4\right)\left(z^2 - 4z\cos\left(\frac{\pi}{2}\right) + 4\right)\left(z^2 - 4z\cos\left(-\frac{5}{6}\pi\right) + 4\right) \\ &= \left(z^2 - 2\sqrt{3}z + 4\right)\left(z^2 + 4\right)\left(z^2 + 2\sqrt{3}z + 4\right) \end{aligned}$$

Part (c).



The geometrical shape formed is a regular hexagon.

Self-Practice A10.2

Problem 1. The complex numbers $2e^{i\pi/12}$ and $2e^{i(5\pi/12)}$ are represented by the points A and B respectively in an Argand diagram with origin O. Show that the triangle OAB is equilateral.

Solution. Note that $OA = OB = 2$ and

$$\angle BOA = \arg\left(2e^{i(5\pi/12)}\right) - \arg\left(2e^{i(\pi/12)}\right) = \frac{\pi}{3}.$$

It follows that $\triangle OAB$ is equilateral.

* * * * *

Problem 2. The complex numbers z and w are such that

$$|z| = 2, \quad \arg(z) = -\frac{2\pi}{3}, \qquad \text{and} \qquad |w| = 5, \quad \arg(w) = \frac{3\pi}{4}.$$

(a) Find the exact values of the modulus and argument of w/z^2. Hence, represent z, w and w/z^2 clearly in an Argand diagram.

(b) Express w/z^2 in the exponential form. Hence, or otherwise, find the smallest positive integer n such that $(w/z^2)^n$ is a real number.

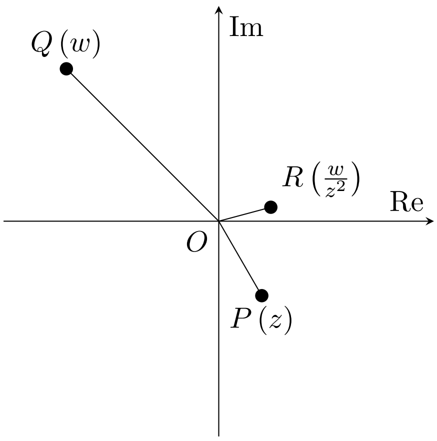
Solution.

Part (a). We have

$$\left|\frac{w}{z^2}\right| = \frac{|w|}{|z|^2} = \frac{5}{2^2} = \frac{5}{4}$$

and

$$\arg\left(\frac{w}{z^2}\right) = \arg(w) - 2\arg(z) = \frac{3\pi}{4} - 2\left(-\frac{2\pi}{3}\right) = \frac{\pi}{12}.$$



Part (b). For $(w/z^2)^n$ to be real, its argument must be an integer multiple of π, i.e.

$$\arg\left(\frac{w}{z^2}\right)^n = n\arg\left(\frac{w}{z^2}\right) = \frac{n\pi}{12} = k\pi \implies n = 12k$$

for some $k \in \mathbb{Z}$. It is clear that the smallest value n can be is 12 (occurring when $k = 1$).

* * * * *

Problem 3. Express $\frac{\cot\theta + i}{\cot\theta - i}$ in the exponential form.

Solution. We have
$$\frac{\cot\theta + \mathrm{i}}{\cot\theta - \mathrm{i}} = \frac{\cos\theta + \mathrm{i}\sin\theta}{\cos\theta - \mathrm{i}\sin\theta} = \frac{\mathrm{e}^{\mathrm{i}\theta}}{\mathrm{e}^{-\mathrm{i}\theta}} = \mathrm{e}^{2\mathrm{i}\theta}.$$

* * * * *

Problem 4. Do not use a calculator in answering this question.
Two complex numbers are $z_1 = 2\left(\cos\frac{\pi}{18} - \mathrm{i}\sin\frac{\pi}{18}\right)$ and $z_2 = 2\mathrm{i}$.

(a) Show that
$$\frac{z_1^2}{z_1^*} + z_2 = \sqrt{3} + \mathrm{i}.$$

(b) A third complex number, z_3, is such that
$$\left(\frac{z_1^2}{z_1^*} + z_2\right) z_3 \in \mathbb{R} \quad \text{and} \quad \left|\left(\frac{z_1^2}{z_1^*} + z_2\right) z_3\right| = \frac{2}{3}.$$
Find the possible values of z_3 in the form of $r(\cos\theta + \mathrm{i}\sin\theta)$, where $r > 0$ and $-\pi < \theta \leq \pi$.

Solution.
Part (a). Note that
$$z_1 = 2\left(\cos\left(-\frac{\pi}{18}\right) + \mathrm{i}\sin\left(-\frac{\pi}{18}\right)\right) = 2\mathrm{e}^{-\mathrm{i}\pi/18}.$$
Thus,
$$\frac{z_1^2}{z_1^*} + z_2 = \frac{z_1^3}{|z_1|^2} + z_2 = \frac{2^3\mathrm{e}^{-\mathrm{i}\pi/6}}{2^2} + 2\mathrm{i} = 2\left[\cos\left(-\frac{\pi}{6}\right) + \mathrm{i}\sin\left(-\frac{\pi}{6}\right)\right] + 2\mathrm{i} = \sqrt{3} + \mathrm{i}.$$
Part (b). Let $w = z_1^2/z_1^* + z_2$. Note that
$$|w| = \sqrt{\sqrt{3}^2 + 1^2} = 1 \quad \text{and} \quad \arg(w) = \arctan\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6},$$
so $w = 2\mathrm{e}^{\mathrm{i}\pi/6}$. Let $z_3 = r\mathrm{e}^{\mathrm{i}\theta}$. Since wz_3 is real, its argument must be an integer multiple of π, i.e.
$$\arg(wz_3) = \arg(w) + \arg(z_3) = \frac{\pi}{6} + \theta = k\pi \implies \theta = \frac{\pi(6k-1)}{6}$$
for some $k \in \mathbb{Z}$. The only solutions for θ within the specified range $(-\pi, \pi)$ are $\theta = -\pi/6$ and $\theta = 5\pi/6$. Further, we have
$$\frac{2}{3} = |wz_3| = |w|\,|z_3| = 2r \implies r = \frac{1}{3}.$$
Thus,
$$z_3 = \frac{1}{3}\left(\cos\left(-\frac{\pi}{6}\right) + \mathrm{i}\sin\left(-\frac{\pi}{6}\right)\right) \quad \text{or} \quad z_3 = \frac{1}{3}\left(\cos\frac{5\pi}{6} + \mathrm{i}\sin\frac{5\pi}{6}\right).$$

* * * * *

Problem 5. Do not use a calculator in answering this question.
The complex numbers z and w satisfy the following equations:
$$w - z = 1 - \sqrt{3}, \qquad \mathrm{i}z + w = \left(\sqrt{3} + 1\right)\mathrm{i}.$$
Find w in the form $r\mathrm{e}^{\mathrm{i}\theta}$, where $r > 0$ and $-\pi < \theta \leq \pi$. Give r and θ in exact form.

Hence, find the three smallest positive whole number values of n for which $(iw)^n$ is an imaginary number.

Solution. Multiplying the second equation by i yields

$$\mathrm{i}w - z = -\left(1+\sqrt{3}\right).$$

Along with the first equation, this gives

$$w - \mathrm{i}w = \left(1-\sqrt{3}\right) + \left(1+\sqrt{3}\right) = 2 \implies w = \frac{2}{1-\mathrm{i}} = \frac{2\left(1+\mathrm{i}\right)}{2} = 1+\mathrm{i} = \sqrt{2}\mathrm{e}^{\mathrm{i}\pi/4}.$$

For $(\mathrm{i}w)^n$ to be purely imaginary, its argument must be a half-integer multiple of π, i.e.

$$\arg((\mathrm{i}w)^n) = n\left(\arg(\mathrm{i}) + \arg(w)\right) = n\left(\frac{\pi}{2} + \frac{\pi}{4}\right) = \left(k + \frac{1}{2}\right)\pi \implies n = \frac{4k+2}{3}$$

for some $k \in \mathbb{Z}$. The first three smallest positive values of n are hence $n = 2, 6, 10$ (occurring when $k = 1, 4, 7$ respectively).

* * * * *

Problem 6 (🌶). It is given that $z = \cos\theta + \mathrm{i}\sin\theta$, where $0 < \theta < \pi/2$.

(a) Show that $\mathrm{e}^{\mathrm{i}(\theta-\pi/2)} = \sin\theta - \mathrm{i}\cos\theta$.

(b) Hence, or otherwise, show that $\arg\left(1-z^2\right) = \theta - \pi/2$ and find the modulus of $1-z^2$.

(c) Hence, represent the complex number $1-z^2$ on an Argand diagram.

(d) Given that $\frac{z^*}{z^3(1-z^2)}$ is real, find the possible values of θ.

Solution.

Part (a). By trigonometric identities, we readily have

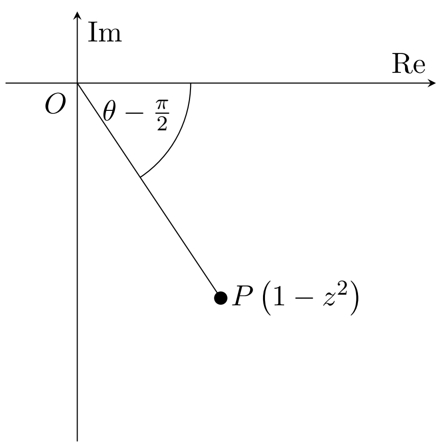
$$\mathrm{e}^{\mathrm{i}(t-\pi/2)} = \cos\left(\theta - \frac{\pi}{2}\right) + \mathrm{i}\sin\left(\theta - \frac{\pi}{2}\right) = \sin\theta - \mathrm{i}\cos\theta.$$

Part (b). Note that $z = r\mathrm{e}^{\mathrm{i}\theta}$. Thus,

$$\begin{aligned} 1 - z^2 &= -\left(\mathrm{e}^{2\mathrm{i}\theta} - 1\right) = -\mathrm{e}^{\mathrm{i}\theta}\left(\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{-\mathrm{i}\theta}\right) = -\mathrm{e}^{\mathrm{i}\theta}\left(2\mathrm{i}\sin\theta\right) \\ &= \left(2\sin\theta\right)\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\pi/2} = \left(2\sin\theta\right)\mathrm{e}^{\mathrm{i}(\theta-\pi/2)}. \end{aligned}$$

Thus, $\arg\left(1-z^2\right) = \theta - \pi/2$ and $\left|1-z^2\right| = 2\sin\theta$.

Part (c).



Part (d). Note that

$$\arg\left(\frac{z^*}{z^3\left(1-z^2\right)}\right) = \arg(z^*) - 3\arg(z) - \arg\left(1-z^2\right) = (-\theta) - 3\theta - \left(\theta - \frac{\pi}{2}\right) = -5\theta + \frac{\pi}{2}.$$

Since $\frac{z^*}{z^3(1-z^2)}$ is real, its argument is an integer multiple of π, i.e.

$$-5\theta + \frac{\pi}{2} = k\pi \implies \theta = \frac{\pi\left(1-2k\right)}{10}$$

for some $k \in \mathbb{Z}$. Since $\theta \in (0, \pi/2)$, the only possible values of θ are $\theta = \pi/10$ and $\theta = 3\pi/10$ (corresponding to $k = 0$ and $k = -1$ respectively).

Assignment A10.2

Problem 1. On an Argand diagram, mark and label clearly the points P and Q representing the complex numbers p and q respectively, where

$$p = \cos\frac{\pi}{4} + \mathrm{i}\sin\frac{\pi}{4}, \qquad q = 2\cos\frac{\pi}{4} + 2\mathrm{i}\sin\frac{\pi}{4}.$$

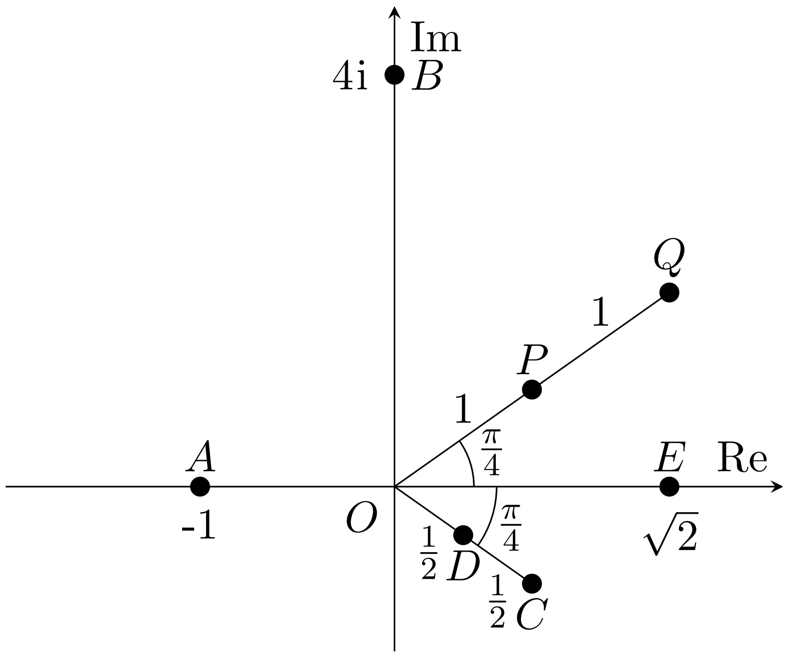
Find the moduli and arguments of the complex numbers a, b, c, d and e, where $a = p^4$, $b = q^2$, $c = -ip$, $d = 1/q$, $e = p + p^*$.

On your Argand diagram, mark and label the points A, B, C, D and E representing these complex numbers.

Find the area of triangle COQ.

Find the modulus and argument of $p^{13/3}q^{45/2}$.

Solution.



Note that $p = \mathrm{e}^{\mathrm{i}\pi/4}$ and $q = 2\mathrm{e}^{\mathrm{i}\pi/4}$.

$$a = p^4 = \left(\mathrm{e}^{\mathrm{i}\pi/4}\right)^4 = \mathrm{e}^{\mathrm{i}\pi}, \quad b = q^2 = \left(2\mathrm{e}^{\mathrm{i}\pi/4}\right)^2 = 4\mathrm{e}^{\mathrm{i}\pi/2}$$

$$c = -\mathrm{i}p = \mathrm{e}^{-\mathrm{i}\pi/2}\mathrm{e}^{\mathrm{i}\pi/4} = \mathrm{e}^{-\mathrm{i}\pi/4}, \quad d = \frac{1}{q} = \frac{1}{2}\mathrm{e}^{-\mathrm{i}\pi/4}$$

$$e = p + p^* = 2\,\mathrm{Re}\,p = 2\cos\left(\frac{\pi}{4}\right) = \sqrt{2}$$

z	$\lvert z\rvert$	$\arg z$
a	1	π
b	4	$\pi/2$
c	1	$-\pi/4$
d	$1/2$	$-\pi/4$
e	$\sqrt{2}$	0

Since $\angle COQ = \pi/2$, we have $[\triangle COQ] = \frac{1}{2}(2)(1) = 1$ units^2.

We have

$$p^{\frac{13}{3}}q^{\frac{45}{2}} = \left(\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}\right)^{\frac{13}{3}}\left(2\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}\right)^{\frac{45}{2}} = 2^{\frac{45}{2}}\mathrm{e}^{\mathrm{i}\frac{161\pi}{24}} = 2^{\frac{45}{2}}\mathrm{e}^{\mathrm{i}\frac{17\pi}{24}}.$$

Hence, $\left|p^{13/3}q^{45/2}\right| = \mathrm{e}^{45/2}$ and $\arg\left(p^{13/3}q^{45/2}\right) = \frac{17}{24}\pi$.

* * * * *

Problem 2. The complex number q is given by $q = \frac{\mathrm{e}^{\mathrm{i}2\theta}}{1-\mathrm{e}^{\mathrm{i}2\theta}}$, where $0 < \theta < 2\pi$. In either order,

(a) find the real part of q,

(b) show that the imaginary part of q is $\frac{1}{2}\cot\theta$.

Solution. We have

$$q = \frac{\mathrm{e}^{\mathrm{i}2\theta}}{1-\mathrm{e}^{\mathrm{i}2\theta}} = \frac{\mathrm{e}^{\mathrm{i}\theta}}{\mathrm{e}^{-\mathrm{i}\theta}-\mathrm{e}^{\mathrm{i}\theta}} = \frac{\cos\theta + \mathrm{i}\sin\theta}{-2\mathrm{i}\sin\theta} = -\frac{1}{2} - \frac{1}{2\mathrm{i}}\cot\theta = -\frac{1}{2} + \frac{\mathrm{i}}{2}\cot\theta.$$

Hence, $\operatorname{Re} q = -\frac{1}{2}$ and $\operatorname{Im} q = \frac{1}{2}\cot\theta$.

* * * * *

Problem 3. The complex numbers z and w are such that $z = 4\left(\cos\frac{3}{4}\pi + \mathrm{i}\sin\frac{3}{4}\pi\right)$ and $w = 1 - \mathrm{i}\sqrt{3}$. z^* denotes the conjugate of z.

(a) Find the modulus r and the argument θ of w^2/z^*, where $r > 0$ and $-\pi < \theta < \pi$.

(b) Given that $\left(w^2/z^*\right)^n$ is purely imaginary, find the set of values that n can take.

Solution.

Part (a). Note that $z = 4\mathrm{e}^{\mathrm{i}3\pi/4}$ and $w = 2\left(\frac{1}{2} - \mathrm{i}\frac{\sqrt{3}}{2}\right) = 2\mathrm{e}^{-\mathrm{i}\pi/3}$. Hence,

$$\frac{w^2}{z^*} = \frac{\left(2\mathrm{e}^{-\mathrm{i}\frac{\pi}{3}}\right)^2}{4\mathrm{e}^{-\mathrm{i}\frac{3\pi}{4}}} = \frac{4\mathrm{e}^{-\mathrm{i}\frac{2\pi}{3}}}{4\mathrm{e}^{-\mathrm{i}\frac{3\pi}{4}}} = \mathrm{e}^{\mathrm{i}\frac{\pi}{12}}.$$

Thus, $r = 1$ and $\theta = \pi/12$.

Part (b). Note that $\left(w^2/z^*\right)^n = \left(\mathrm{e}^{\mathrm{i}\pi/12}\right)^n = \mathrm{e}^{\mathrm{i}n\pi/12}$. Since $\left(w^2/z^*\right)^n$ is purely imaginary, we have $\arg\left(w^2/z^*\right)^n = \pi/2 + \pi k$, where $k \in \mathbb{Z}$. Thus, $n\pi/12 = \pi/2 + \pi k$, whence $n = 6 + 12k$. Hence, $\{n \in \mathbb{Z} : n = 6 + 12k,\ k \in Z\}$.

* * * * *

Problem 4. The complex number w has modulus $\sqrt{2}$ and argument $\pi/4$ and the complex number z has modulus $\sqrt{2}$ and argument $5\pi/6$.

(a) By first expressing w and z in the form $x + \mathrm{i}y$, find the exact real and imaginary parts of $w + z$.

(b) On the same Argand diagram, sketch the points P, Q, R representing the complex numbers z, w, and $z+w$ respectively. State the geometrical shape of the quadrilateral $OPRQ$.

(c) Referring the Argand diagram in part (b), find $\arg(w + z)$ and show that $\tan\frac{11}{24}\pi = \frac{a+\sqrt{2}}{\sqrt{6}+b}$, where a and b are constants to be determined.

Solution.

Part (a). Note that

$$w = \sqrt{2}\mathrm{e}^{\mathrm{i}\pi/4} = \sqrt{2}\left(\cos\frac{\pi}{4} + \mathrm{i}\sin\frac{\pi}{4}\right) = \sqrt{2}\left(\frac{1}{\sqrt{2}} + \mathrm{i}\frac{1}{\sqrt{2}}\right) = 1 + \mathrm{i}$$

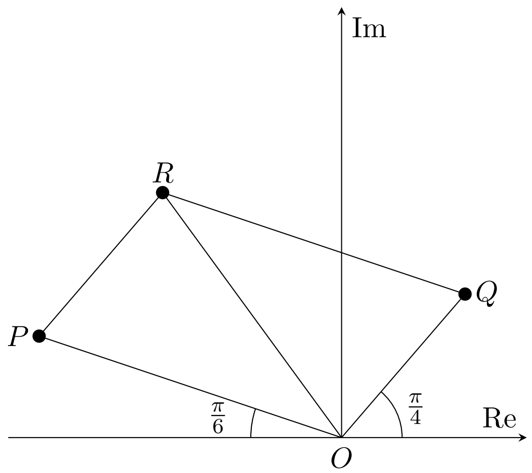
and

$$z = \sqrt{2}\mathrm{e}^{\mathrm{i}5\pi/6} = \sqrt{2}\left(\cos\frac{5}{6}\pi + \mathrm{i}\sin\frac{5}{6}\pi\right) = \sqrt{2}\left(-\frac{\sqrt{3}}{2} + \mathrm{i}\frac{1}{2}\right) = -\frac{\sqrt{3}}{\sqrt{2}} + \mathrm{i}\frac{1}{\sqrt{2}}.$$

Hence,

$$w + z = (1+\mathrm{i}) + \left(-\frac{\sqrt{3}}{\sqrt{2}} + \mathrm{i}\frac{1}{\sqrt{2}}\right) = \left(1 - \frac{\sqrt{3}}{\sqrt{2}}\right) + \mathrm{i}\left(1 + \frac{1}{\sqrt{2}}\right).$$

Part (b).



$OPRQ$ is a rhombus.

Part (c). Note that $\angle POQ = \pi - \frac{\pi}{6} - \frac{\pi}{4} = \frac{7}{12}\pi$. Since $|z| = |w|$, we have $OP = OQ$, whence $\angle ROQ = \frac{1}{2} \cdot \frac{7}{12}\pi = \frac{7}{24}\pi$. Hence, $\arg(w+z) = \frac{\pi}{4} + \frac{7}{24}\pi = \frac{13}{24}\pi$. Thus,

$$\tan\left(\frac{13}{24}\pi\right) = \frac{1 + 1/\sqrt{2}}{1 - \sqrt{3}/\sqrt{2}} = \frac{\sqrt{2}+1}{\sqrt{2}-\sqrt{3}} = \frac{2+\sqrt{2}}{2-\sqrt{6}}$$

However, $\tan\left(\frac{13}{24}\pi\right) = -\tan\left(\pi - \frac{13}{24}\right) = -\tan\left(\frac{11}{24}\pi\right)$. Hence,

$$\tan\left(\frac{11}{24}\pi\right) = -\frac{2+\sqrt{2}}{2-\sqrt{6}} = \frac{2+\sqrt{2}}{\sqrt{6}-2},$$

whence $a = 2$ and $b = -2$.

* * * * *

Problem 5. The complex number z is given by $z = 2(\cos\beta + \mathrm{i}\sin\beta)$ where $0 < \beta < \frac{\pi}{2}$.

(a) Show that $\frac{z}{4-z^2} = (k\csc\beta)\mathrm{i}$, where k is positive real constant to be determined.

(b) State the argument of $\frac{z}{4-z^2}$, giving your reasons clearly.

(c) Given the complex number $w = -\sqrt{3} + \mathrm{i}$, find the three smallest positive integer values of n such that $\left(\frac{z}{4-z^2}\right)(w^*)^n$ is a real number.

Solution.

Part (a). Observe that $z = 2(\cos\beta + \mathrm{i}\sin\beta) = 2\mathrm{e}^{\mathrm{i}\beta}$. Hence,

$$\frac{z}{4-z^2} = \frac{2\mathrm{e}^{\mathrm{i}\beta}}{4-4\mathrm{e}^{\mathrm{i}2\beta}} = \frac{1}{2}\left(\frac{1}{\mathrm{e}^{-\mathrm{i}\beta}-\mathrm{e}^{\mathrm{i}\beta}}\right) = \frac{1}{2}\left(\frac{1}{-2\mathrm{i}\sin\beta}\right) = \left(\frac{1}{4}\csc\beta\right)\mathrm{i},$$

thus $k = 1/4$.

Part (b). Since $0 < \beta < \pi/2$, we know that $\csc\beta > 0$. Hence, $\mathrm{Im}\left(\frac{z}{4-z^2}\right) > 0$. Furthermore, $\mathrm{Re}\left(\frac{z}{4-z^2}\right) = 0$. Thus, $\arg\left(\frac{z}{4-z^2}\right) = \pi/2$.

Part (c). Note that $w = -\sqrt{3} + \mathrm{i} = 2\left(-\frac{\sqrt{3}}{2} + \frac{1}{2}\mathrm{i}\right) = 2\mathrm{e}^{-\mathrm{i}5\pi/6}$. Hence,

$$\arg\left(\left(\frac{z}{4-z^2}\right)(w^*)^n\right) = \frac{\pi}{2} + n\left(-\frac{5\pi}{6}\right) = \pi\left(\frac{1}{2} - \frac{5n}{6}\right).$$

For $\left(\frac{z}{4-z^2}\right)(w^*)^n$ to be a real number, we require $\arg\left(\left(\frac{z}{4-z^2}\right)(w^*)^n\right) = \pi k$, where $k \in \mathbb{Z}$. Hence,

$$\pi\left(\frac{1}{2} - \frac{5}{6}n\right) = \pi k \implies \frac{1}{2} - \frac{5}{6}n = k \implies 3 - 5n = 6k \implies n \equiv 3 \pmod 6.$$

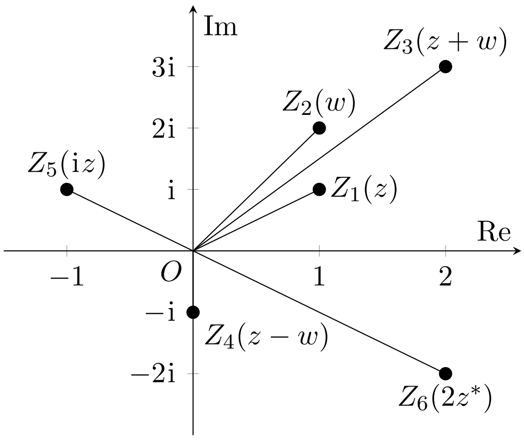
Hence, the three smallest possible values of n are 3, 9 and 15.

A10.3 Complex Numbers - Geometrical Effects and De Moivre's Theorem

Tutorial A10.3

Problem 1. Given that $z = 1 + \mathrm{i}$ and $w = 1 + 2\mathrm{i}$, mark on an Argand diagram, the positions representing: z, w, $z + w$, $z - w$, $\mathrm{i}z$ and $2z^*$.

Solution.



* * * * *

Problem 2.

(a) Write down the exact values of the modulus and the argument of the complex number $\frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i}$.

(b) The complex numbers z and w satisfy the equation

$$z^2 - zw + w^2 = 0.$$

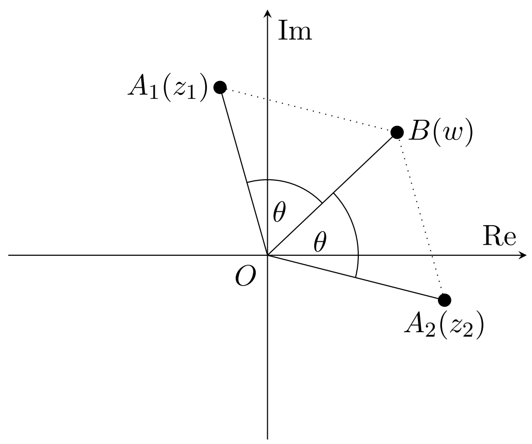
Find z in terms of w. In an Argand diagram, the points O, A and B represent the complex numbers 0, z and w respectively. Show that $\triangle OAB$ is an equilateral triangle.

Solution.

Part (a). We have $r^2 = \left(\frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2 \implies r = 1$ and $\tan\theta = \frac{\sqrt{3}/2}{1/2} \implies \theta = \frac{\pi}{3}$. Hence, $\left|\frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i}\right| = 1$ and $\arg\left(\frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i}\right) = \frac{\pi}{3}$.

Part (b). From the quadratic formula, we have

$$z = \frac{w \pm \sqrt{w^2 - 4w^2}}{2} = w\left(\frac{1}{2} \pm \frac{\sqrt{3}}{2}\mathrm{i}\right).$$



Since $\left|\frac{1}{2} \pm \frac{\sqrt{3}}{2}\mathrm{i}\right| = 1$, we have that $OB = OA_1 = OA_2$. Further, since $\arg\left(\frac{1}{2} \pm \frac{\sqrt{3}}{2}\mathrm{i}\right) = \pm\pi/3$, we know $\angle A_1OB = \angle A_2OB = \pi/3$, whence $\triangle A_1OB$ and $\triangle A_2OB$ are both equilateral.

* * * * *

Problem 3. Find the exact roots of the equations

(a) $z^3 = 1$

(b) $(z-1)^4 = -16$

in the form $x + \mathrm{i}y$.

Solution.

Part (a). Note that

$$z^3 = 1 = \mathrm{e}^{\mathrm{i}2\pi n} \implies z = \mathrm{e}^{\mathrm{i}2\pi n/3} = \cos\frac{2\pi n}{3} + \mathrm{i}\sin\frac{2\pi n}{3},$$

for $n \in \mathbb{Z}$. Evaluating z in the $n = 0, 1, 2$ cases, we see that the roots of $z^3 = 1$ are

$$z = 1,\ -\frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i},\ -\frac{1}{2} - \frac{\sqrt{3}}{2}\mathrm{i}.$$

Part (b). Note that $(z-1)^4 = -16 = 16\mathrm{e}^{\mathrm{i}\pi(2n+1)}$. Hence,

$$z = 1 + 2\mathrm{e}^{\mathrm{i}\pi(2n+1)/4} = 1 + 2\left[\cos\left(\frac{2n+1}{4}\pi\right) + \mathrm{i}\sin\left(\frac{2n+1}{4}\pi\right)\right],$$

where $n \in \mathbb{Z}$. Evaluating z in the $n = 0, 1, 2, 3$ cases, we see that the roots of $(z-1)^4 = -16$ are

$$z = (1+\sqrt{2}) + \mathrm{i}\sqrt{2},\ (1-\sqrt{2}) + \mathrm{i}\sqrt{2},\ (1-\sqrt{2}) - \mathrm{i}\sqrt{2},\ (1+\sqrt{2}) - \mathrm{i}\sqrt{2}.$$

* * * * *

Problem 4.

(a) Write down the 5 roots of the equation $z^5 - 1 = 0$ in the form $r\mathrm{e}^{\mathrm{i}\theta}$, where $r > 0$ and $-\pi < \theta \le \pi$.

(b) Show that the roots of the equation $(5+z)^5 - (5-z)^5 = 0$ can be written in the form $5\mathrm{i}\tan\frac{k\pi}{5}$, where $k = 0, \pm1, \pm2$.

Solution.

Part (a). Note that

$$z^5 = 1 = \mathrm{e}^{\mathrm{i}2\pi n} \implies z = \mathrm{e}^{\mathrm{i}2\pi n/5}.$$

Since $-\pi < \theta \leq \pi$, we have

$$z = \mathrm{e}^{-\mathrm{i}4\pi/5},\ \mathrm{e}^{-\mathrm{i}2\pi/5},\ 1,\ \mathrm{e}^{\mathrm{i}2\pi/5},\ \mathrm{e}^{\mathrm{i}4\pi/5}.$$

Part (b). Note that

$$(5+z)^5 - (5-z)^5 = 0 \implies \left(\frac{5+z}{5-z}\right)^5 - 1 = 0 \implies \frac{5+z}{5-z} = \mathrm{e}^{\mathrm{i}2k\pi/5}.$$

Solving for z, we get

$$z = 5\left(\frac{\mathrm{e}^{\mathrm{i}2k\pi/5} - 1}{\mathrm{e}^{\mathrm{i}2k\pi/5} + 1}\right) = 5\left(\frac{\mathrm{e}^{\mathrm{i}k\pi/5} - \mathrm{e}^{-\mathrm{i}k\pi/5}}{\mathrm{e}^{\mathrm{i}k\pi/5} + \mathrm{e}^{-\mathrm{i}k\pi/5}}\right) = 5\left[\frac{2\mathrm{i}\sin(k\pi/5)}{2\cos(k\pi/5)}\right] = 5\mathrm{i}\tan\frac{k\pi}{5}.$$

* * * * *

Problem 5. De Moivre's theorem for a positive integral exponent states that

$$(\cos\theta + \mathrm{i}\sin\theta)^n = \cos n\theta + \mathrm{i}\sin n\theta.$$

Use de Moivre's theorem to show that

$$\cos 7\theta = 64\cos^7\theta - 112\cos^5\theta + 56\cos^3\theta - 7\cos\theta.$$

Hence, obtain the roots of the equation

$$128x^7 - 224x^5 + 112x^3 - 14x + 1 = 0$$

in the form $\cos q\pi$, where q is a rational number.

Solution. Taking $n = 7$, we have $\cos 7\theta + \mathrm{i}\sin 7\theta = (\cos\theta + \mathrm{i}\sin\theta)^7$, whence $\cos 7\theta = \operatorname{Re}(\cos\theta + \mathrm{i}\sin\theta)^7$. Let $c = \cos\theta$ and $s = \sin\theta$. By the binomial theorem,

$$\cos 7\theta = \operatorname{Re}(c + \mathrm{i}s)^7 = \operatorname{Re}\sum_{k=0}^{7}\binom{7}{k}\mathrm{i}^k s^k c^{7-k}.$$

Note that $\operatorname{Re} i^k$ is given by

$$\operatorname{Re} i^k = \begin{cases} 0, & k = 1, 3 \pmod 4 \\ 1, & k = 0 \pmod 4 \\ -1, & k = 2 \pmod 4 \end{cases}$$

We hence have

$$\begin{aligned} \cos 7\theta &= c^7 - 21c^5s^2 + 35c^3s^4 - 7cs^6 = c^7 - 21c^5\left(1-c^2\right) + 35c^3\left(1-c^2\right)^2 - 7c\left(1-c^2\right)^3 \\ &= 64c^7 - 112c^5 + 56c^3 - 7c = 64\cos^7\theta - 112\cos^5 + 56\cos^3\theta - 7\cos\theta. \end{aligned}$$

Observe that we can manipulate the given equation into

$$128x^7 - 224x^5 + 112x^3 - 14x + 1 = 0 \implies 64x^7 - 112x^5 + 56x^3 - 7x = -\frac{1}{2}.$$

Under the substitution $x = \cos\theta$, we see that

$$\cos 7\theta = -\frac{1}{2} \implies 7\theta = \frac{2}{3}\pi + 2\pi n \implies \theta = \frac{2\pi}{21}(3n+1),$$

where $n \in \mathbb{Z}$. Taking $0 \le n < 7$,

$$\begin{aligned} x &= \cos\frac{2\pi}{21}, \cos\frac{8\pi}{21}, \cos\frac{14\pi}{21}, \cos\frac{20\pi}{21}, \cos\frac{26\pi}{21}, \cos\frac{32\pi}{21}, \cos\frac{38\pi}{21} \\ &= \cos\frac{2\pi}{21}, \cos\frac{4\pi}{21}, \cos\frac{8\pi}{21}, \cos\frac{10\pi}{21}, \cos\frac{14\pi}{21}, \cos\frac{16\pi}{21}, \cos\frac{20\pi}{21}. \end{aligned}$$

* * * * *

Problem 6. By considering $\sum_{n=1}^{N} z^{2n-1}$, where $z = \mathrm{e}^{\mathrm{i}\theta}$, or by any method, show that

$$\sum_{n=1}^{N} \sin(2n-1)\theta = \frac{\sin^2 N\theta}{\sin\theta},$$

provided $\sin\theta \neq 0$.

Solution. Observe that

$$\sum_{n=1}^{N} \sin(2n-1)\theta = \operatorname{Im}\sum_{n=1}^{N} [\cos(2n-1)\theta + \mathrm{i}\sin(2n-1)\theta] = \operatorname{Im}\sum_{n=1}^{N} z^{2n-1}.$$

Since

$$\begin{aligned} \sum_{n=1}^{N} z^{2n-1} &= \frac{1}{z}\sum_{n=1}^{N}\left(z^2\right)^n = \frac{1}{z}\left(\frac{z^2\left[\left(z^2\right)^N - 1\right]}{z^2-1}\right) = \frac{z^{2N}-1}{z-z^{-1}} \\ &= z^N\left(\frac{z^N - z^{-N}}{z-z^{-1}}\right) = z^N\left(\frac{2\mathrm{i}\sin N\theta}{2\mathrm{i}\sin\theta}\right) = z^N\left(\frac{\sin N\theta}{\sin\theta}\right), \end{aligned}$$

we have

$$\sum_{n=1}^{N} \sin(2n-1)\theta = \left(\frac{\sin N\theta}{\sin\theta}\right)\operatorname{Im}\left(z^N\right) = \left(\frac{\sin N\theta}{\sin\theta}\right)\sin N\theta = \frac{\sin^2 N\theta}{\sin\theta}.$$

* * * * *

Problem 7. By considering the series $\sum_{n=0}^{N}\left(\mathrm{e}^{2\mathrm{i}\theta}\right)^n$, show that, provided $\sin\theta \neq 0$,

$$\sum_{n=0}^{N} \cos 2n\theta = \frac{\sin(N+1)\theta\cos N\theta}{\sin\theta}$$

and deduce that

$$\sum_{n=0}^{N} \sin^2 n\theta = \frac{N}{2} + \frac{1}{2} - \frac{\sin(N+1)\theta\cos N\theta}{2\sin\theta}.$$

Solution. Let $z = \mathrm{e}^{\mathrm{i}\theta}$. Then

$$\sum_{n=0}^{N} \cos 2n\theta = \operatorname{Re}\sum_{n=0}^{N}(\cos 2n\theta + \mathrm{i}\sin 2n\theta) = \operatorname{Re}\sum_{n=0}^{N}\mathrm{e}^{\mathrm{i}2n\theta} = \operatorname{Re}\sum_{n=0}^{N}\left(z^2\right)^n.$$

Observe that

$$\sum_{n=0}^{N}\left(z^2\right)^n=\frac{\left(z^2\right)^{N+1}-1}{z^2-1}=\frac{z^{N+1}}{z}\left(\frac{z^{N+1}-z^{-(N+1)}}{z-z^{-1}}\right)=z^N\left(\frac{\sin(N+1)\theta}{\sin\theta}\right).$$

Hence,

$$\sum_{n=0}^{N}\cos 2n\theta=\left(\frac{\sin(N+1)\theta}{\sin\theta}\right)\operatorname{Re}\left(z^N\right)=\frac{\sin(N+1)\theta\cos N\theta}{\sin\theta}.$$

Recall that $\cos 2n\theta=1-2\sin^2 n\theta \implies \sin^2 n\theta=\frac{1}{2}(1-2\cos 2n\theta)$. Thus,

$$\sum_{n=0}^{N}\sin^2 n\theta=\frac{1}{2}\sum_{n=0}^{N}(1-\cos 2n\theta)=\frac{N+1}{2}-\frac{\sin(N+1)\theta\cos N\theta}{2\sin\theta}.$$

* * * * *

Problem 8. Given that $z=\mathrm{e}^{\mathrm{i}\theta}$, show that $z^k+1/z^k=2\cos k\theta$, $k\in\mathbb{Z}$.

Hence, show that $\cos^8\theta=\frac{1}{128}\left(\cos 8\theta+8\cos 6\theta+28\cos 4\theta+56\cos 2\theta+35\right)$.

Find, correct to three decimal places, the values of θ such that $0<\theta<\frac{1}{2}\pi$ and $\cos 8\theta+8\cos 6\theta+28\cos 4\theta+56\cos 2\theta+1=0$.

Solution. Note that

$$\begin{aligned}z^k+\frac{1}{z^k}&=z^k+z^{-k}=\left(\mathrm{e}^{\mathrm{i}\theta}\right)^k+\left(\mathrm{e}^{\mathrm{i}\theta}\right)^{-k}=\mathrm{e}^{\mathrm{i}k\theta}+\mathrm{e}^{-\mathrm{i}k\theta}\\&=[\cos(k\theta)+\mathrm{i}\sin(k\theta)]+[\cos(-k\theta)+\mathrm{i}\sin(-k\theta)]=2\cos(k\theta)\,.\end{aligned}$$

Observe that

$$\begin{aligned}\cos^8\theta&=\frac{1}{256}(2\cos\theta)^8=\frac{1}{256}(z+z^{-1})^8=\frac{1}{256}z^{-8}\left(z^2+1\right)^8\\&=\frac{1}{256}\left(z^{-8}+8z^{-6}+28z^{-4}+56z^{-2}+70+56z^2+28z^4+8z^6+z^8\right)\\&=\frac{1}{128}\left[\left(\frac{z^8+z^{-8}}{2}\right)+8\left(\frac{z^6+z^{-6}}{2}\right)+28\left(\frac{z^4+z^{-4}}{2}\right)+56\left(\frac{z^2+z^{-2}}{2}\right)+\frac{70}{2}\right]\\&=\frac{1}{128}\left(\cos 8\theta+8\cos 6\theta+28\cos 4\theta+56\cos 2\theta+35\right).\end{aligned}$$

Note that we rewrite the equation as

$$\cos 8\theta+8\cos 6\theta+28\cos 4\theta+56\cos 2\theta+35=128\cos^8\theta=34.$$

Thus,

$$\cos\theta=\sqrt[8]{\frac{34}{128}}\implies\theta=0.560\ (3\text{ s.f.}).$$

Self-Practice A10.3

Problem 1. Express $\frac{\cot\theta+\mathrm{i}}{\cot\theta-\mathrm{i}}$ in the exponential form. Hence, show that one of the roots of the equation

$$z^4 = \frac{\sqrt{3}+\mathrm{i}}{\sqrt{3}-\mathrm{i}}$$

is $\mathrm{e}^{\mathrm{i}\pi/12}$, and find three more roots in the exponential form.

Solution. Note that

$$\frac{\cot\theta+\mathrm{i}}{\cot\theta-\mathrm{i}} = \frac{\cos\theta+\mathrm{i}\sin\theta}{\cos\theta-\mathrm{i}\sin\theta} = \frac{\mathrm{e}^{\mathrm{i}\theta}}{\mathrm{e}^{-\mathrm{i}\theta}} = \mathrm{e}^{2\mathrm{i}\theta}.$$

Note that $\cot\theta = \sqrt{3} \implies \theta = \pi/6$, so

$$z^4 = \frac{\sqrt{3}+\mathrm{i}}{\sqrt{3}-\mathrm{i}} = \mathrm{e}^{2\mathrm{i}\pi/6} = \mathrm{e}^{\mathrm{i}\pi(1/3+2k)},$$

where $k \in \mathbb{Z}$. Taking fourth roots,

$$z = \mathrm{e}^{\mathrm{i}\pi(1/12+k/2)}.$$

Taking $k = 0, 1, 2, 3$, we see that the four roots are

$$z = \mathrm{e}^{\mathrm{i}\pi/12},\ \mathrm{e}^{7\mathrm{i}\pi/12},\ \mathrm{e}^{13\mathrm{i}\pi/12},\ \mathrm{e}^{19\mathrm{i}\pi/12}.$$

* * * * *

Problem 2. Find the cube roots of the complex number $1+\mathrm{i}\sqrt{3}$. Give your answers exactly, in the form $r\mathrm{e}^{\mathrm{i}\theta}$. Hence, solve the equation $z^6 - 2z^3 + 4 = 0$. Give your answers exactly, in the form $r\mathrm{e}^{\mathrm{i}\theta}$.

Solution. Consider

$$z^3 = 1+\mathrm{i}\sqrt{3} = 2\mathrm{e}^{\mathrm{i}\pi/3} = 2\mathrm{e}^{\mathrm{i}\pi(1/3+2k)},$$

where $k \in \mathbb{Z}$. Taking roots,

$$z = 2^{1/3}\mathrm{e}^{\mathrm{i}\pi(1/9+2k/3)}.$$

Taking $k = -1, 0, 1$, the cube roots of $1+\mathrm{i}\sqrt{3}$ are

$$z = 2^{1/3}\mathrm{e}^{-8\mathrm{i}\pi/9}, 2^{1/3}\mathrm{e}^{\mathrm{i}\pi/9}, 2^{1/3}\mathrm{e}^{7\mathrm{i}\pi/9}.$$

Consider $z^6 - 2z^3 + 4 = 0$. Then

$$z^3 = 1 \pm \sqrt{3}.$$

From the positive branch, we get the aforementioned roots. Since the coefficients of the sextic are all real, by the conjugate root theorem, the six roots are

$$z = 2^{1/3}\mathrm{e}^{-8\mathrm{i}\pi/9},\ 2^{1/3}\mathrm{e}^{-7\mathrm{i}\pi/9},\ 2^{1/3}\mathrm{e}^{-\mathrm{i}\pi/9},\ 2^{1/3}\mathrm{e}^{\mathrm{i}\pi/9},\ 2^{1/3}\mathrm{e}^{7\mathrm{i}\pi/9},\ 2^{1/3}\mathrm{e}^{8\mathrm{i}\pi/9}.$$

* * * * *

Problem 3. Express $8\left(\sqrt{3}-\mathrm{i}\right)$ in the form $r(\cos\theta+\mathrm{i}\sin\theta)$, where $r>0$ and $-\pi \le \theta \le \pi$, giving θ in terms of π. Hence, obtain the roots of the equation $z^4 = 8\left(\sqrt{3}-\mathrm{i}\right)$ in the same form.

Solution. Note that

$$8\left(\sqrt{3}-\mathrm{i}\right)=16\left(\frac{\sqrt{3}}{2}-\frac{1}{2}\mathrm{i}\right)=16\left[\cos\left(-\frac{\pi}{6}\right)+\mathrm{i}\sin\left(-\frac{\pi}{6}\right)\right]=16\mathrm{e}^{-\mathrm{i}\pi/6}.$$

We are given

$$z^4=8\left(\sqrt{3}-\mathrm{i}\right)=16\mathrm{e}^{-\mathrm{i}\pi/6}=16\mathrm{e}^{\mathrm{i}\pi(-1/6+2k)},$$

for $k\in\mathbb{Z}$. Taking roots,

$$z=2\mathrm{e}^{\mathrm{i}\pi(-1/24+k/2)}.$$

Taking $k=-1,0,1,2$, the roots are

$$z=2\mathrm{e}^{-13\mathrm{i}\pi/24},\ 2\mathrm{e}^{-\mathrm{i}\pi/24},\ 2\mathrm{e}^{11\mathrm{i}\pi/24},\ 2\mathrm{e}^{23\mathrm{i}\pi/24}.$$

* * * * *

Problem 4. Write down, in any form, the five complex numbers which satisfy the equation $z^5-1=0$. Hence, show that the five complex numbers which satisfy the equation

$$\left(\frac{2w+1}{w}\right)^5=1$$

are

$$\frac{-2+\cos\left(\frac{2}{5}\pi k\right)-\mathrm{i}\sin\left(\frac{2}{5}\pi k\right)}{5-4\cos\left(\frac{2}{5}\pi k\right)},$$

where $k=0,1,2,3,4$.

Solution. The fifth roots of unity are given by

$$z=\mathrm{e}^{2k\mathrm{i}\pi/5},$$

where $k=0,1,2,3,4$.

We have

$$\frac{2w+1}{w}=\mathrm{e}^{2k\mathrm{i}\pi/5}\implies w=\frac{1}{\mathrm{e}^{2k\mathrm{i}\pi/5}-2}=\frac{\mathrm{e}^{-2k\mathrm{i}\pi/5}-2}{\left|\mathrm{e}^{2k\mathrm{i}\pi/5}-2\right|^2}.$$

Note that

$$\left|\mathrm{e}^{2k\mathrm{i}\pi/5}-2\right|^2=\left(\mathrm{e}^{2k\mathrm{i}\pi/5}-2\right)\left(\mathrm{e}^{-2k\mathrm{i}\pi/5}-2\right)=1-2\left(2\cos\frac{2k\pi}{5}\right)+4=5-4\cos\frac{2k\pi}{5}.$$

Thus,

$$w=\frac{-2+\cos\left(\frac{2}{5}\pi k\right)-\mathrm{i}\sin\left(\frac{2}{5}\pi k\right)}{5-4\cos\left(\frac{2}{5}\pi k\right)}$$

for $k=0,1,2,3,4$.

* * * * *

Problem 5.

(a) Show that, for all complex numbers z and all real numbers α,

$$\left(z-\mathrm{e}^{\mathrm{i}\alpha}\right)\left(z-\mathrm{e}^{-\mathrm{i}\alpha}\right)=z^2-2z\cos\alpha+1.$$

(b) Write down, in any form, the seven complex numbers which satisfy the equation $z^7-1=0$.

(c) Hence, show that, for all complex numbers z,

$$z^7 - 1 = (z-1)\left[z^2 - 2z\cos\frac{2\pi}{7} + 1\right]\left[z^2 - 2z\cos\frac{4\pi}{7} + 1\right]\left[z^2 - 2z\cos\frac{6\pi}{7} + 1\right].$$

Solution.

Part (a). We have

$$\left(z - \mathrm{e}^{\mathrm{i}\alpha}\right)\left(z - \mathrm{e}^{-\mathrm{i}\alpha}\right) = z^2 - \left(\mathrm{e}^{\mathrm{i}\alpha} + \mathrm{e}^{-\mathrm{i}\alpha}\right)z + 1 = z^2 - 2z\cos\alpha + 1.$$

Part (b). The seventh roots of unity are

$$z = \mathrm{e}^{2k\mathrm{i}\pi/7}$$

where $k = -3, -2, -1, 0, 1, 2, 3$.

Part (c). Let $P_k = z - \mathrm{e}^{2k\mathrm{i}\pi/7}$. Observe that

$$P_k P_{-k} = \left(z - \mathrm{e}^{2k\mathrm{i}\pi/7}\right)\left(z - \mathrm{e}^{2k\mathrm{i}\pi/7}\right) = z^2 - 2z\cos\frac{2k\pi}{7} + 1.$$

Hence,

$$\begin{aligned} z^7 - 1 &= P_0\left(P_1P_{-1}\right)\left(P_2P_{-2}\right)\left(P_3P_{-3}\right) \\ &= (z-1)\left[z^2 - 2z\cos\frac{2\pi}{7} + 1\right]\left[z^2 - 2z\cos\frac{4\pi}{7} + 1\right]\left[z^2 - 2z\cos\frac{6\pi}{7} + 1\right]. \end{aligned}$$

* * * * *

Problem 6. Use De Moivre's theorem to show that

$$\cos 6\theta = 32\cos^6\theta - 48\cos^4\theta + 18\cos^2\theta - 1.$$

Deduce that, for all θ,

$$0 \le \cos^6\theta - \frac{3}{2}\cos^4\theta + \frac{9}{16}\cos^2\theta \le \frac{1}{16}.$$

Solution. Let $c = \cos\theta$ and $s = \sin\theta$. Then

$$\begin{aligned} \cos 6\theta &= \operatorname{Re}\mathrm{e}^{6\mathrm{i}\theta} = \operatorname{Re}\left(c + \mathrm{i}s\right)^6 \\ &= \binom{6}{0}c^6 - \binom{6}{2}c^4s^2 + \binom{6}{4}c^2s^4 - \binom{6}{6}s^6 \\ &= c^6 - 15c^4\left(1 - c^2\right) + 15c^2\left(1 - c^2\right)^2 - \left(1 - c^2\right)^3 \\ &= c^6 - 15c^4\left(1 - c^2\right) + 15c^2\left(1 - 2c^2 + c^4\right) - \left(1 - 3c^2 + 3c^4 - c^6\right) \\ &= 32c^6 - 48c^4 + 18c^2 - 1 \\ &= 32\cos^6\theta - 48\cos^4\theta + 18\cos^2\theta - 1. \end{aligned}$$

Observe that

$$-\frac{1}{32} \le \frac{\cos 6\theta}{32} = \cos^6\theta - \frac{3}{2}\cos^4\theta + \frac{9}{16}\cos^2\theta - \frac{1}{32} \le \frac{1}{32},$$

so

$$0 \le \cos^6\theta - \frac{3}{2}\cos^4\theta + \frac{9}{16}\cos^2\theta \le \frac{1}{16}.$$

Problem 7. Show that for $z \neq -1$,

$$z - z^2 + z^3 - \cdots + z^7 = \frac{z + z^8}{1 + z}.$$

Hence, by substituting $z = \mathrm{e}^{\mathrm{i}\theta}$, show that

$$\sum_{k=1}^{7} (-1)^{k-1} \sin k\theta = \frac{\sin 4\theta \cos \frac{7}{2}\theta}{\cos \frac{1}{2}\theta},$$

where θ is not an odd multiple of π.

Solution. Observe that $z - z^2 + z^3 - \cdots + z^7$ is a geometric series with common ratio $-z$, so it evaluates to

$$z - z^2 + z^3 - \cdots + z^7 = z\left(\frac{1 - (-z)^7}{1 - (-z)}\right) = \frac{z - z^8}{1 + z},$$

with the condition $z \neq -1$.

We have

$$\begin{aligned}\sum_{k=1}^{7} (-1)^{k-1} \sin k\theta &= \sum_{k=1}^{7} (-1)^{k-1} \operatorname{Im} z^k = \operatorname{Im} \sum_{k=1}^{7} (-1)^{k-1} z^k = \operatorname{Im} \frac{z + z^8}{1 + z} \\ &= \operatorname{Im} \frac{z^{9/2}\left(z^{7/2} + z^{-7/2}\right)}{z^{1/2}\left(z^{1/2} + z^{-1/2}\right)} = \frac{2\cos\left(\frac{7}{2}\theta\right)}{2\cos\left(\frac{1}{2}\theta\right)} \operatorname{Im} z^4 = \frac{\cos\left(\frac{7}{2}\theta\right)\sin 4\theta}{\cos \frac{1}{2}\theta}.\end{aligned}$$

Note that $z = \mathrm{e}^{\mathrm{i}\theta} \neq -1 = \mathrm{e}^{\mathrm{i}\pi(2k+1))}$ for $k \in \mathbb{Z}$, so $\theta \neq (2k+1)\pi$, i.e. θ cannot be an odd multiple of π.

Assignment A10.3

Problem 1.

(a) Solve $z^4 = -4 - 4\sqrt{3}\mathrm{i}$, expressing your answers in the form $r\mathrm{e}^{\mathrm{i}\theta}$, where $r > 0$ and $-\pi < \theta \leq \pi$.

(b) Sketch the roots on an Argand diagram.

(c) Hence, solve $w^4 = -1 + \sqrt{3}\mathrm{i}$, expressing your answers in a similar form.

Solution.

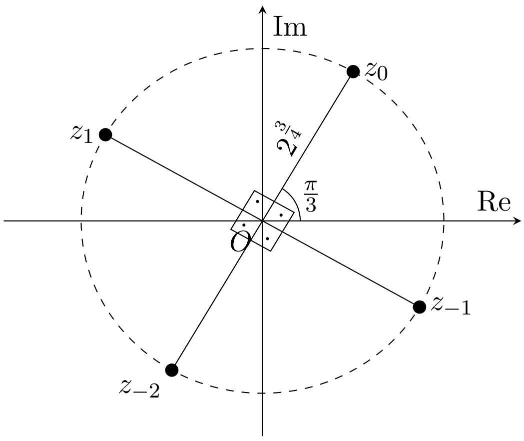
Part (a). Observe that $-4 - 4\sqrt{3}\mathrm{i} = 8\left(-\frac{1}{2} - \frac{\sqrt{3}}{2}\mathrm{i}\right) = 8\mathrm{e}^{\mathrm{i}\frac{4}{3}\pi + 2k\pi\mathrm{i}}$ for all $k \in \mathbb{Z}$. Hence,

$$z^4 = 8\mathrm{e}^{\mathrm{i}\frac{4}{3}\pi + 2k\pi\mathrm{i}} \implies z = 8^{\frac{1}{4}}\mathrm{e}^{\mathrm{i}\frac{1}{3}\pi + \frac{1}{2}k\pi\mathrm{i}} = 2^{\frac{3}{4}}\mathrm{e}^{\mathrm{i}\frac{2+3k}{6}\pi}.$$

Taking $k = -2, -1, 0, 1$, we see that the roots are

$$z_{-2} = 2^{\frac{3}{4}}\mathrm{e}^{-\mathrm{i}\frac{2}{3}\pi}, \quad z_{-1} = 2^{\frac{3}{4}}\mathrm{e}^{-\mathrm{i}\frac{1}{6}\pi}, \quad z_0 = 2^{\frac{3}{4}}\mathrm{e}^{\mathrm{i}\frac{1}{3}\pi}, \quad z_1 = 2^{\frac{3}{4}}\mathrm{e}^{\mathrm{i}\frac{5}{6}\pi}.$$

Part (b).



Part (c). Observe that $w^4 = -1 + \sqrt{3}\mathrm{i} = \frac{1}{4}(-4 + 4\sqrt{3}\mathrm{i}) = 2^{-2}(z^*)^4$. Hence, $w = 2^{-1/2}z^*$. Thus, the roots are

$$w_{-2} = 2^{\frac{1}{4}}\mathrm{e}^{\mathrm{i}\frac{2}{3}\pi}, \quad w_{-1} = 2^{\frac{1}{4}}\mathrm{e}^{\mathrm{i}\frac{1}{6}\pi}, \quad w_0 = 2^{\frac{1}{4}}\mathrm{e}^{-\mathrm{i}\frac{1}{3}\pi}, \quad w_1 = 2^{\frac{1}{4}}\mathrm{e}^{-\mathrm{i}\frac{5}{6}\pi}.$$

* * * * *

Problem 2. Let

$$C = 1 - \binom{2n}{1}\cos\theta + \binom{2n}{2}\cos 2\theta - \binom{2n}{3}\cos 3\theta + \ldots + \cos 2n\theta$$
$$S = -\binom{2n}{1}\sin\theta + \binom{2n}{2}\sin 2\theta - \binom{2n}{3}\sin 3\theta + \ldots + \sin 2n\theta$$

where n is a positive integer.

Show that $C = (-4)^n\cos(n\theta)\sin^{2n}(\theta/2)$, and find the corresponding expression for S.

Solution. Clearly,

$$C = \sum_{k=0}^{2n} \binom{2n}{k} (-1)^k \cos k\theta, \quad S = \sum_{k=0}^{2n} \binom{2n}{k} (-1)^k \sin k\theta.$$

Hence,

$$\begin{aligned} C + \mathrm{i}S &= \sum_{k=0}^{2n} \binom{2n}{k} (-1)^k (\cos k\theta + \mathrm{i} \sin k\theta) = \sum_{k=0}^{2n} \binom{2n}{k} (-\mathrm{e}^{\mathrm{i}\theta})^k = \left(1 - \mathrm{e}^{\mathrm{i}\theta}\right)^{2n} \\ &= \left(\mathrm{e}^{\mathrm{i}\theta/2}\right)^{2n} \left(\mathrm{e}^{-\mathrm{i}\theta/2} - \mathrm{e}^{\mathrm{i}\theta/2}\right)^{2n} = \mathrm{e}^{\mathrm{i}n\theta} \left(2\mathrm{i} \sin \frac{\theta}{2}\right)^{2n} = \mathrm{e}^{\mathrm{i}n\theta} (-4)^n \sin^{2n} \frac{\theta}{2} \\ &= (\cos n\theta + \mathrm{i} \sin n\theta)(-4)^n \sin^{2n} \frac{\theta}{2}. \end{aligned}$$

Comparing real and imaginary parts, we have

$$C = (-4)^n \cos(n\theta) \sin^{2n} \frac{\theta}{2}, \quad S = (-4)^n \sin(n\theta) \sin^{2n} \frac{\theta}{2}.$$

* * * * *

Problem 3. Given that $z = \cos\theta + \mathrm{i}\sin\theta$, show that

(a) $z - 1/z = 2\mathrm{i}\sin\theta$,

(b) $z^n + z^{-n} = 2\cos n\theta$.

Hence, show that

$$\sin^6\theta = \frac{1}{32}(10 - 15\cos 2\theta + 6\cos 4\theta - \cos 6\theta)$$

Find a similar expression for $\cos^6\theta$, and hence express $\cos^6\theta - \sin^6\theta$ in the form $a\cos 2\theta + b\cos 6\theta$.

Solution.

Part (a). Note that

$$z - \frac{1}{z} = z - z^{-1} = \mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{-\mathrm{i}\theta} = [\cos\theta + \mathrm{i}\sin\theta] - [\cos(-\theta) + \mathrm{i}\sin(-\theta)] = 2\mathrm{i}\sin\theta.$$

Part (b). Note that

$$z^n + z^{-n} = \mathrm{e}^{\mathrm{i}n\theta} + e^{-\mathrm{i}n\theta} = [\cos n\theta + \mathrm{i}\sin n\theta] + [\cos(-n\theta) + \mathrm{i}\sin(n\theta)] = 2\cos n\theta.$$

Observe that

$$\begin{aligned} \sin^6\theta &= \frac{1}{(2\mathrm{i})^6}(2\mathrm{i}\sin\theta)^6 = -\frac{1}{64}(z - z^{-1})^6 \\ &= -\frac{1}{64}\left(z^6 - 6z^4 + 15z^2 - 20 + 15z^{-2} - 6z^{-4} + z^{-6}\right) \\ &= -\frac{1}{32}\left[-\frac{20}{2} + 15\left(\frac{z^2 + z^{-2}}{2}\right) - 6\left(\frac{z^4 + z^{-4}}{2}\right) + \left(\frac{z^6 + z^{-6}}{2}\right)\right] \\ &= \frac{1}{32}\left(10 - 15\cos 2\theta + 6\cos 4\theta - \cos 6\theta\right). \end{aligned}$$

Similarly,

$$\begin{aligned}
\cos^6\theta &= \frac{1}{2^6}(2\cos\theta)^6 = \frac{1}{64}(z+z^{-1})^6 \\
&= \frac{1}{64}\left[z^6+6z^4+15z^2+20+15z^{-2}+6z^{-4}+z^{-6}\right] \\
&= \frac{1}{32}\left[\frac{20}{2}+15\left(\frac{z^2+z^{-2}}{2}\right)+6\left(\frac{z^4+z^{-4}}{2}\right)+\left(\frac{z^6+z^{-6}}{2}\right)\right] \\
&= \frac{1}{32}\left(10+15\cos 2\theta+6\cos 4\theta+\cos 6\theta\right).
\end{aligned}$$

Hence,

$$\cos^6\theta-\sin^6\theta = \frac{1}{32}(30\cos 2\theta+2\cos 6\theta) = \frac{15}{16}\cos 2\theta+\frac{1}{16}\cos 6\theta,$$

whence $a = 15/16$ and $b = 1/16$.

A10.4 Complex Numbers - Loci in Argand Diagram

Tutorial A10.4

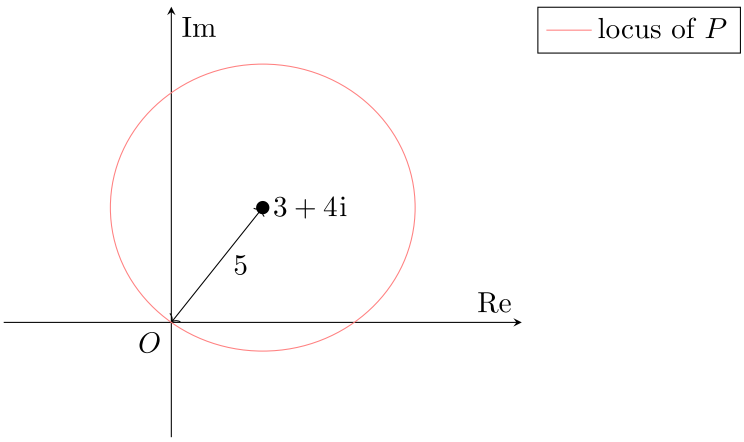
Problem 1. A complex number z is represented in an Argand diagram by the point P. Sketch, on separate Argand diagrams, the locus of P. Describe geometrically the locus of P and determine its Cartesian equation.

(a) $|2z - 6 - 8\mathrm{i}| = 10$

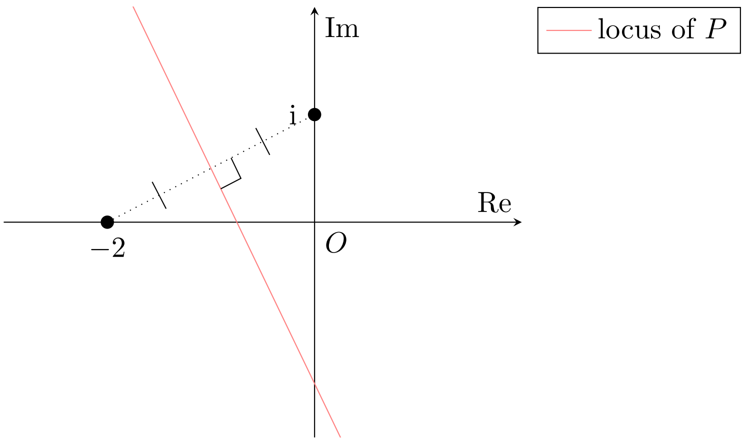
(b) $|z + 2| = |z - \mathrm{i}|$

(c) $\arg(z + 2 - \mathrm{i}) = -\pi/4$

Solution.
Part (a). Note that $|2z - 6 - 8\mathrm{i}| = 10 \implies |z - (3 + 4\mathrm{i})| = 5$.

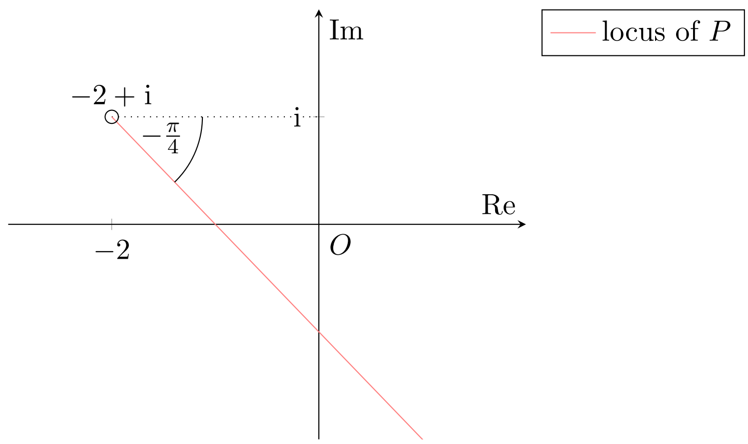


The locus of P is a circle with centre $(3, 4)$ and radius 5. Its Cartesian equation is $(x - 3)^2 + (y - 4)^2 = 5^2$.
Part (b). Note that $|z + 2| = |z - \mathrm{i}| \implies |z - (-2)| = |z - \mathrm{i}|$.



The locus of P is the perpendicular bisector of the line segment joining $(-2,0)$ and $(0,1)$. Its Cartesian equation is $y = -2x - 1.5$.

Part (c). Note that $\arg(z+2-\mathrm{i}) = -\pi/4 \implies \arg(z-(-2+\mathrm{i})) = -\pi/4$.



The locus of P is the half-line starting from $(-2,1)$ and inclined at an angle $-\pi/4$ to the positive real axis. Its Cartesian equation is $y = -x - 1$

* * * * *

Problem 2. Sketch the following loci on separate Argand diagrams.

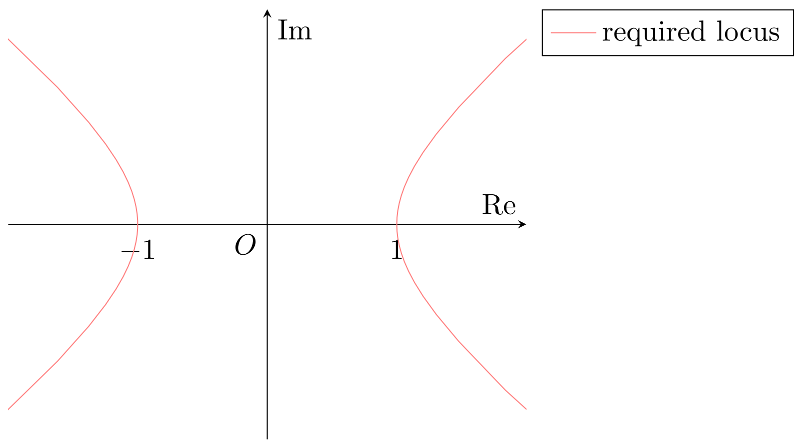
(a) $\mathrm{Re}(z^2) = 1$

(b) $|6 - \mathrm{i}z| = 2$,

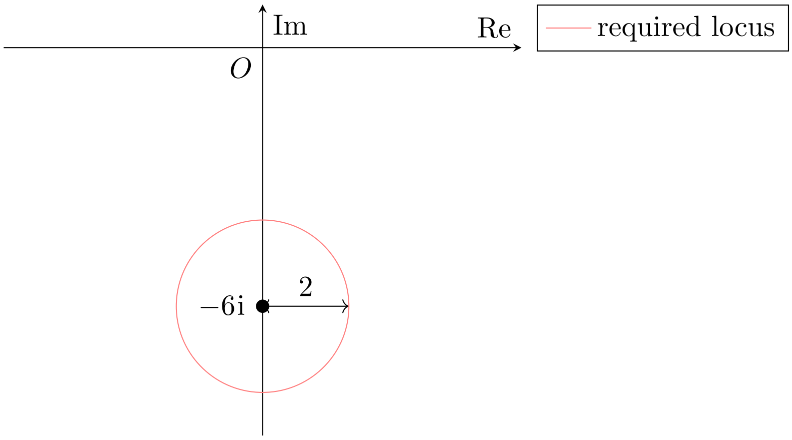
(c) $\arg\left(\frac{\mathrm{i}z}{1-\sqrt{3}\mathrm{i}}\right) = \pi$

Solution.

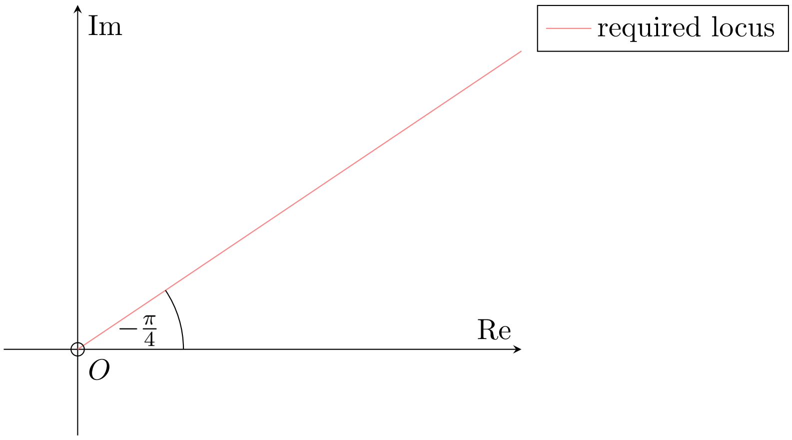
Part (a). Let $z = r(\cos\theta + \mathrm{i}\sin\theta)$. Then $\mathrm{Re}(z^2) = 1 \implies r^2\cos 2\theta = 1 \implies r^2 = \sec 2\theta$.



Part (b). Note $|6 - \mathrm{i}z| = 2 \implies |-\mathrm{i}(z+6i)| = 2 \implies |z+6\mathrm{i}| = 2 \implies |z-(6\mathrm{i})| = 2$.



Part (c). Note $\arg\left(\frac{\mathrm{i}z}{1-\sqrt{3}\mathrm{i}}\right) = \pi \implies \frac{\pi}{2} + \arg(z) - \left(-\frac{\pi}{3}\right) \implies \arg(z) = \frac{\pi}{6}$.



* * * * *

Problem 3. Sketch, on separate Argand diagrams, the set of points satisfying the following inequalities.

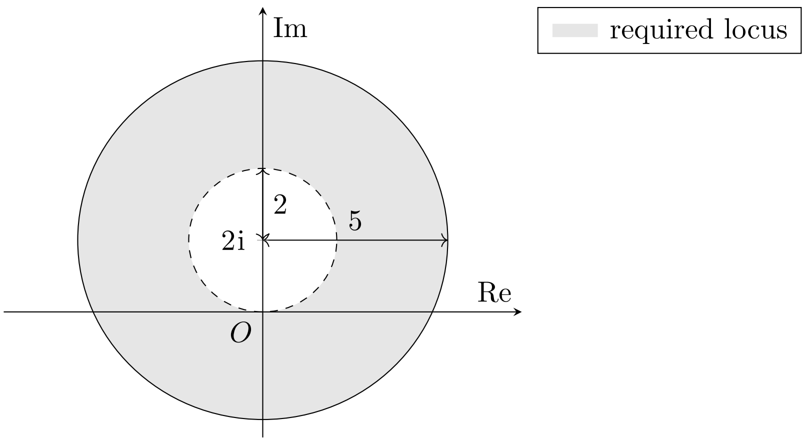
(a) $2 < |z - 2\mathrm{i}| \leq |3 - 4\mathrm{i}|$

(b) $|z + \mathrm{i}| > |z + 1 - \mathrm{i}|$

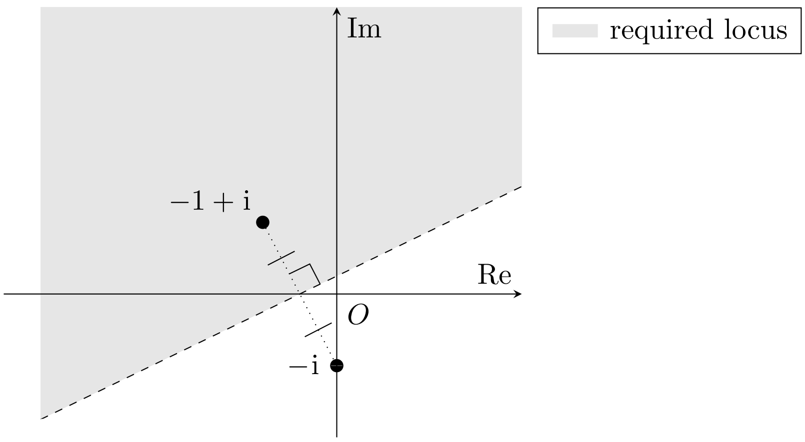
(c) $\frac{\pi}{4} < \arg\left(\frac{1}{z}\right) \leq \frac{\pi}{2}$

Solution.

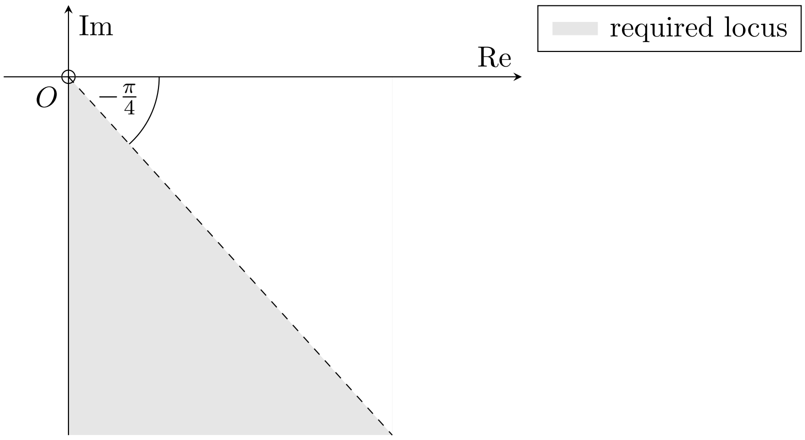
Part (a). Note $2 < |z - 2\mathrm{i}| \leq |3 - 4\mathrm{i}| \implies 2 < |z - 2\mathrm{i}| \leq 5$.



Part (b). Note $|z+\mathrm{i}| > |z+1-\mathrm{i}| \implies |z-(-\mathrm{i})| > |z-(-1+\mathrm{i})|$.



Part (c). Note $\frac{\pi}{4} < \arg\left(\frac{1}{z}\right) \leq \frac{\pi}{2} \implies \frac{\pi}{4} < -\arg(z) \leq \frac{\pi}{2} \implies -\frac{\pi}{2} \geq \arg(z) > -\frac{\pi}{4}$.



* * * * *

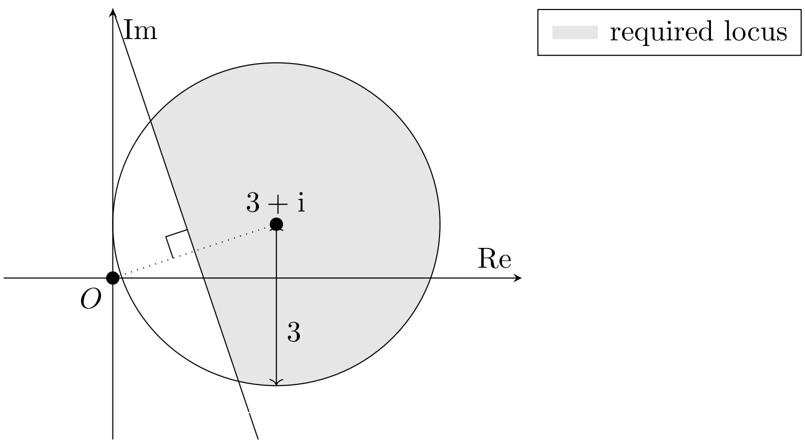
Problem 4. Sketch on separate Argand diagrams for (a) and (b) the set of points representing all complex numbers z satisfying both of the following inequalities.

(a) $|z-3-\mathrm{i}| \leq 3$ and $|z| \geq |z-3-\mathrm{i}|$

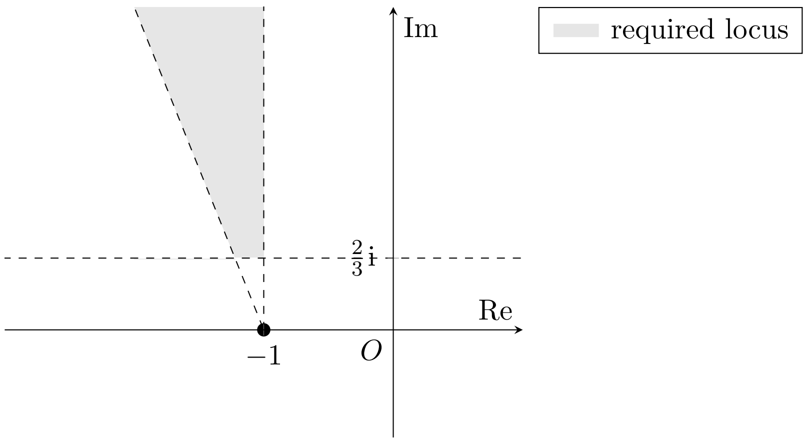
(b) $\frac{\pi}{2} < \arg(z+1) \leq \frac{2}{3}\pi$ and $3\,\mathrm{Im}(z) > 2$

Solution.

Part (a). Note $|z-3-\mathrm{i}| \leq 3 \implies |z-(3+\mathrm{i})| \leq 3$ and $|z| \geq |z-3-\mathrm{i}| \implies |z| \geq |z-(3+\mathrm{i})|$.



Part (b). Note $\frac{\pi}{2} < \arg(z+1) < \frac{2}{3}\pi \implies \frac{\pi}{2} < \arg(z-(-1)) < \frac{2}{3}\pi$ and $3\operatorname{Im}(z) > 2 \implies \operatorname{Im}(z) > \frac{2}{3}$.



* * * * *

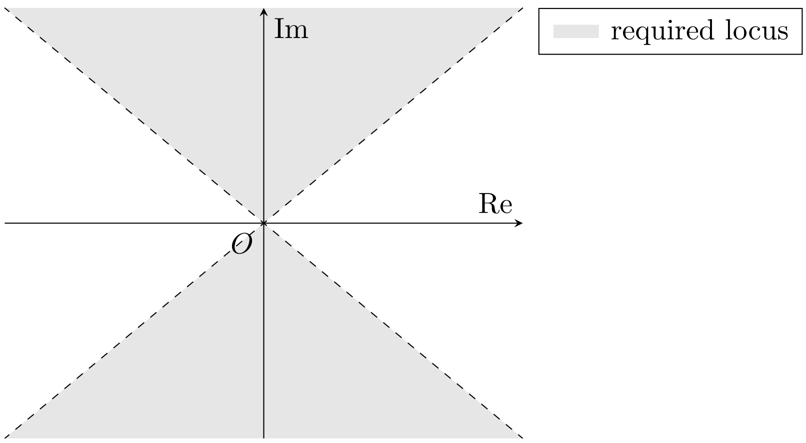
Problem 5. Illustrate, in separate Argand diagrams, the set of points z for which

(a) $\operatorname{Re}(z^2) < 0$

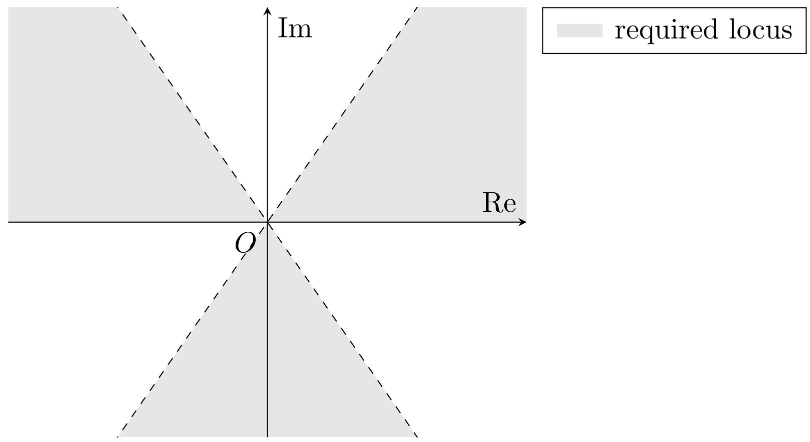
(b) $\operatorname{Im}(z^3) > 0$

Solution.

Part (a). Let $z = r(\cos\theta + \mathrm{i}\sin\theta)$, $0 \le \theta < 2\pi$. Then $\operatorname{Re}(z^2) < 0 \implies r^2\cos 2\theta < 0 \implies \cos 2\theta < 0 \implies 2\theta \in \left(\frac{1}{2}\pi, \frac{3}{2}\pi\right) \cup \left(\frac{5}{2}\pi, \frac{7}{2}\pi\right) \implies \theta \in \left(\frac{1}{4}\pi, \frac{3}{4}\pi\right) \cup \left(\frac{5}{4}\pi, \frac{7}{4}\pi\right)$.



Part (b). Let $z = r(\cos\theta + \mathrm{i}\sin\theta)$, $0 \le \theta < 2\pi$. Then $\operatorname{Im}(z^3) > 0 \implies r^3 \sin 3\theta > 0 \implies \sin 3\theta > 0 \implies 3\theta \in (0, \pi) \cup (2\pi, 3\pi) \cup (4\pi, 5\pi) \implies \theta \in \left(0, \frac{1}{3}\pi\right) \cup \left(\frac{2}{3}\pi, \pi\right) \cup \left(\frac{4}{3}\pi, \frac{5}{3}\pi\right)$.



* * * * *

Problem 6. The complex number z satisfies $|z + 4 - 4\mathrm{i}| = 3$.

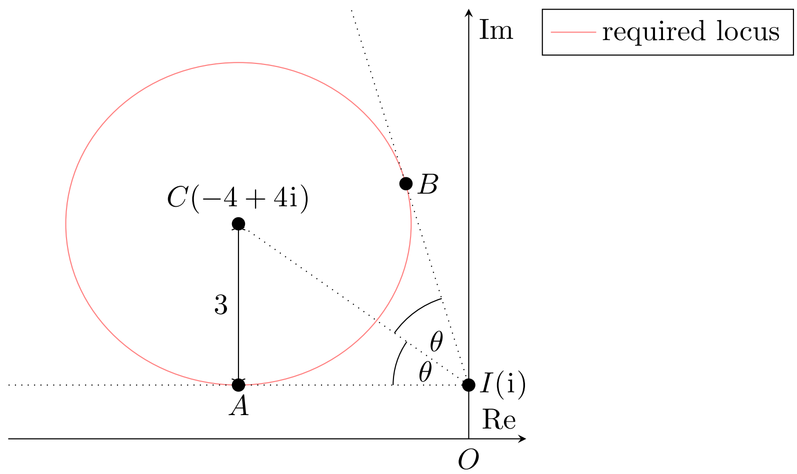
(a) Describe, with the aid of a sketch, the locus of the point which represents z in an Argand diagram.

(b) Find the least possible value of $|z - \mathrm{i}|$.

(c) Find the range of values of $\arg(z - \mathrm{i})$.

Solution.

Part (a). Note $|z + 4 - 4\mathrm{i}| = 3 \implies |z - (-4 + 4\mathrm{i})| = 3$.



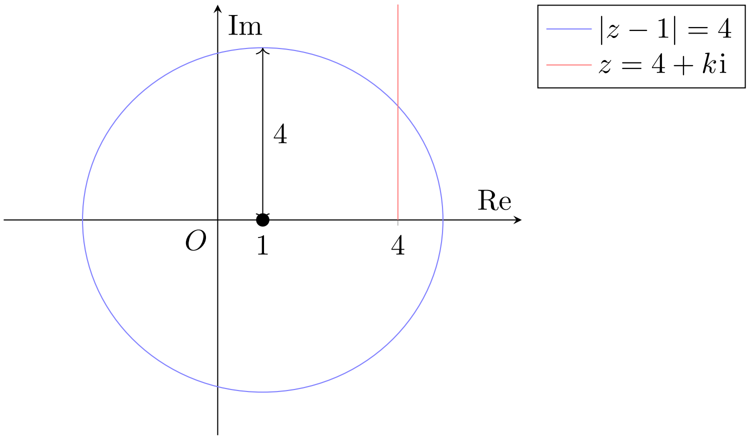
Part (b). Observe that the distance CI is equal to the sum of the radius of the circle and $\min|z - \mathrm{i}|$. Hence,

$$\min|z - \mathrm{i}| = \sqrt{(-4 - 0)^2 + (4 - 1)^2} - 3 = 2.$$

Part (c). Let A and B be points on the circle such that AI and BI are tangent to the circle. Let $\angle CIA = \theta$. Then $\tan\theta = \frac{3}{4} \implies \theta = \arctan\frac{3}{4}$. By symmetry, we also have $\angle CIB = \theta$, whence $\angle AIB = 2\theta = 2\arctan\frac{3}{4}$. Hence, $\min\arg(z - \mathrm{i}) = \pi - 2\arctan\frac{3}{4}$ (at B) and $\max\arg(z - \mathrm{i}) = \pi$ (at A). Thus, $\pi - 2\arctan\frac{3}{4} \le \arg(z - \mathrm{i}) \le \pi$.

Problem 7. Sketch, on the same Argand diagram, the two loci representing the complex number z for which $z = 4 + k\mathrm{i}$, where k is a positive real variable, and $|z-1| = 4$. Write down, in the form $x + \mathrm{i}y$, the complex number satisfying both conditions.

Solution.



Note that z is of the form $4 + k\mathrm{i}$, $k \in \mathbb{R}^+$. Since $|z-1| = 4$, we have $|3 + k\mathrm{i}| = 4 \implies 3^2 + k^2 = 4 \implies k = \sqrt{7}$. Note that we reject $k = -\sqrt{7}$ since $k > 0$. Thus, $z = 4 + \sqrt{7}\mathrm{i}$.

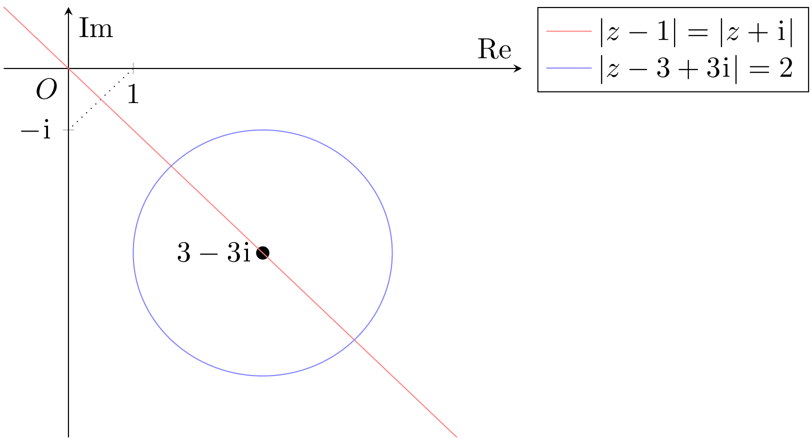
* * * * *

Problem 8. Describe, in geometrical terms, the loci given by $|z-1| = |z+\mathrm{i}|$ and $|z-3+3\mathrm{i}| = 2$ and sketch both loci on the same diagram.

Obtain, in the form $a + \mathrm{i}b$, the complex numbers representing the points of intersection of the loci, giving the exact values of a and b.

Solution. Note that $|z-1| = |z+\mathrm{i}| \implies |z-1| = |z-(-\mathrm{i})|$ and $|z-3+3\mathrm{i}| = 2 \implies |z-(3-3\mathrm{i})| = 2$.

The locus given by $|z-1| = |z+\mathrm{i}|$ is the perpendicular bisector of the line segment joining 1 and $-i$. The locus given by $|z-3+3\mathrm{i}| = 2$ is a circle with centre $3-3\mathrm{i}$ and radius 2.



Observe that the locus of $|z-1| = |z+\mathrm{i}|$ has Cartesian equation $y = -x$ and the locus of $|z-3+3\mathrm{i}| = 2$ has Cartesian equation $(x-3)^2 + (y+3)^2 = 2^2$. Solving both equations simultaneously, we have

$$(x-3)^2 + (y+3)^2 = (x-3)^2 + (3-x)^2 = 2^2 \implies x^2 - 6x + 7 = 0$$
$$\implies x = 3 \pm \sqrt{2} \implies y = -3 \mp \sqrt{2}.$$

Hence, the complex numbers representing the points of intersections of the loci are $(3+\sqrt{2})+(-3-\sqrt{2})\mathrm{i}$ and $(3-\sqrt{2})+(-3+\sqrt{2})\mathrm{i}$.

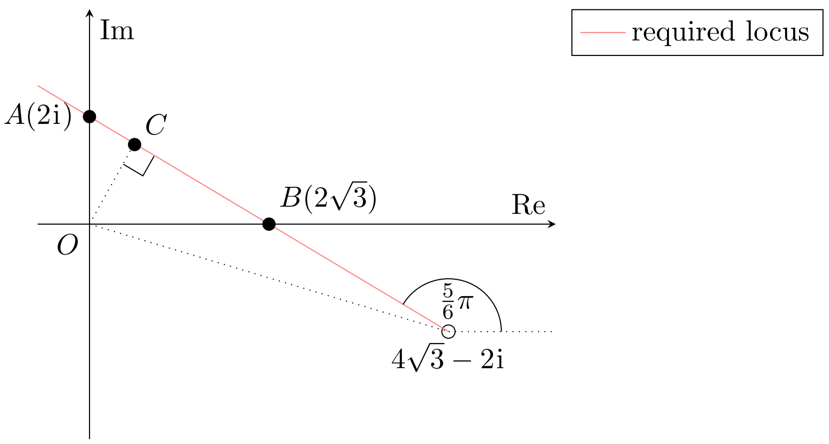
* * * * *

Problem 9. Sketch the locus for $\arg\big(z-(4\sqrt{3}-2\mathrm{i})\big)=\frac{5}{6}\pi$ in an Argand diagram.

(a) Verify that the points $2\mathrm{i}$ and $2\sqrt{3}$ lie on it.

(b) Find the minimum value of $|z|$ and the range of values of $\arg(z)$.

Solution.



Part (a). Note that

$$\arg\Big(2\mathrm{i}-(4\sqrt{3}-2\mathrm{i})\Big)=\arg\Big(-\sqrt{3}+\mathrm{i}\Big)=\arctan\frac{1}{-\sqrt{3}}=\frac{5}{6}\pi$$

and

$$\arg\Big(2\sqrt{3}-(4\sqrt{3}-2\mathrm{i})\Big)=\arg\Big(-\sqrt{3}+\mathrm{i}\Big)=\arctan\frac{1}{-\sqrt{3}}=\frac{5}{6}\pi.$$

Hence, the points $2\mathrm{i}$ and $2\sqrt{3}$ satisfy the equation $\arg\big(z-(4\sqrt{3}-2\mathrm{i})\big)=\frac{5}{6}\pi$ and thus lie on its locus.

Part (b). Let $A(2\mathrm{i})$ and $B(2\sqrt{3})$. Let C be the point on the required locus such that $OC\perp AB$. Observe that $\triangle OAB$, $\triangle COB$ and $\triangle CAO$ are all similar to one another. Hence,

$$\frac{OC}{CB}=\frac{AO}{BO}=\frac{1}{\sqrt{3}}\implies AC=\frac{1}{\sqrt{3}}OC,\quad \frac{OC}{CA}=\frac{BO}{OA}=\frac{\sqrt{3}}{1}\implies BC=\sqrt{3}OC.$$

Hence, $AB=AC+CB=\Big(\sqrt{3}+\frac{1}{\sqrt{3}}\Big)OC$, whence

$$\min|z|=OC=\frac{AB}{\sqrt{3}+1/\sqrt{3}}=\frac{\sqrt{2^2+(2\sqrt{3})^2}}{\sqrt{3}+1\sqrt{3}}=\frac{4\sqrt{3}}{4}=\sqrt{3}.$$

Observe that $\max\arg(z)=\frac{5}{6}\pi$ and $\min\arg(z)=\min\arg\big(4\sqrt{3}-2\mathrm{i}\big)=\arctan\frac{-2}{4\sqrt{3}}=-\arctan\frac{1}{2\sqrt{3}}$. Thus, $-\arctan\frac{1}{2\sqrt{3}}<\arg(z)\leq\frac{5}{6}\pi$.

Problem 10. The complex number z satisfies $|z-3-3\mathrm{i}| \geq |z-1-\mathrm{i}|$ and $\frac{\pi}{6} < \arg(z) \leq \frac{\pi}{3}$.

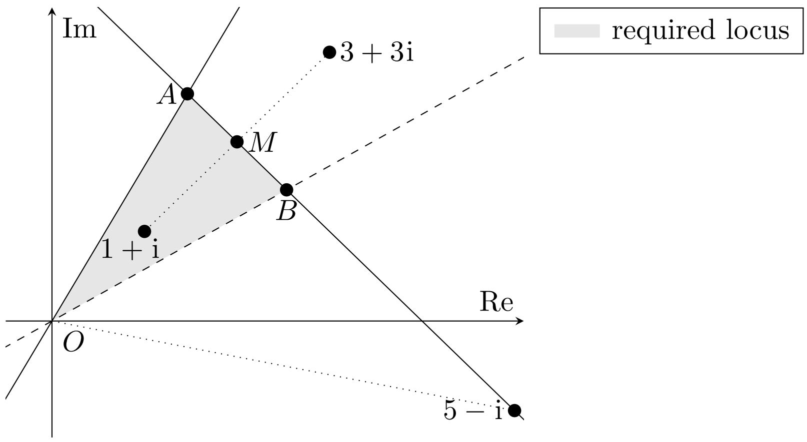
(a) On an Argand diagram, sketch the region in which the point representing z can lie.

(b) Find the area of the region in part (a).

(c) Find the range of values of $\arg(z-5+\mathrm{i})$.

Solution.

Part (a). Note that $|z-3-3\mathrm{i}| \leq |z-1-\mathrm{i}| \implies |z-(3+3\mathrm{i})| \leq |z-(1+\mathrm{i})|$.



Part (b). Note that the locus of $|z-3-3\mathrm{i}| = |z-1-\mathrm{i}|$ has Cartesian equation $y = -x+4$, while the loci of $\frac{\pi}{6} = \arg(z)$ and $\arg(z) = \frac{\pi}{3}$ have Cartesian equations $y = \frac{1}{\sqrt{3}}x$ and $y = \sqrt{3}x$ respectively. Let A and B be the intersections between $y=-x+4$ with $y=\sqrt{3}x$ and $y = \frac{1}{\sqrt{3}}x$ respectively.

At A, we have $y = \sqrt{3}x = -x+4$, whence $A\left(\frac{4}{1+\sqrt{3}}, \frac{4\sqrt{3}}{1+\sqrt{3}}\right)$. Thus,

$$OA = \sqrt{\left(\frac{4}{1+\sqrt{3}}\right)^2 + \left(\frac{4\sqrt{3}}{1+\sqrt{3}}\right)^2} = \frac{8}{1+\sqrt{3}}.$$

By symmetry, we also have $OA = OB$. Finally, since $\angle AOB = \frac{\pi}{3} - \frac{\pi}{6} = \frac{\pi}{6}$,

$$[\triangle AOB] = \frac{1}{2}(OA)(OB)\sin\angle AOB = \frac{1}{2}\left(\frac{8}{1+\sqrt{3}}\right)^2 \frac{1}{2} = \frac{16}{\left(1+\sqrt{3}\right)^2} = 4\left(1-\sqrt{3}\right)^2.$$

Part (c). Observe that $\min\arg(z-(5-\mathrm{i})) = \frac{3}{4}\pi$ and $\max\arg(z-(5-\mathrm{i})) = \arctan\frac{-1}{5} + \pi = \pi - \arctan\frac{1}{5}$. Hence, $\frac{3}{4}\pi \leq \arg(z-5+\mathrm{i}) < \pi - \arctan\frac{1}{5}$.

Problem 11. Sketch on an Argand diagram the set of points representing all complex numbers z satisfying both inequalities

$$|\mathrm{i}z - 2\mathrm{i} - 2| \leq 2 \qquad \text{and} \qquad \mathrm{Re}(z) > \left|1 + \sqrt{3}\mathrm{i}\right|$$

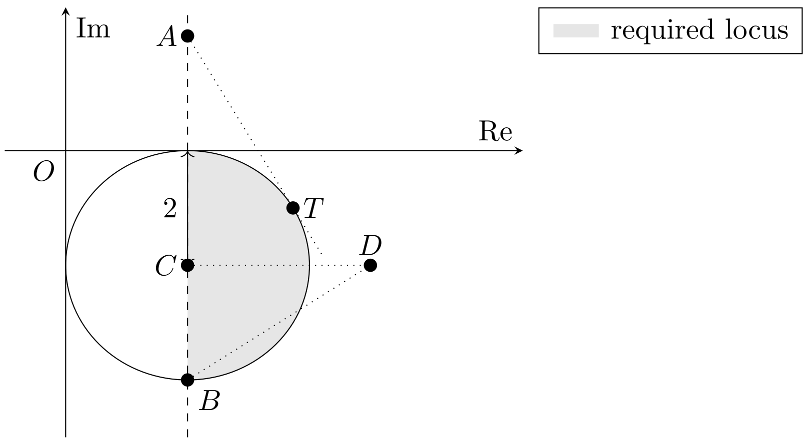
Find

(a) the range of $\arg(z - 2 - 2\mathrm{i})$,

(b) the complex number z where $\arg(z - 2 - 2\mathrm{i})$ is a maximum.

The locus of the complex number w is defined by $|w - 5 + 2\mathrm{i}| = k$, where k is a real and positive constant. Find the range of values of k such that the loci of w and z will intersect.

Solution. Note $|\mathrm{i}z - 2i - 2| \leq 2 \implies |\mathrm{i}(z - 2 + 2i)| \leq 2 \implies |z - (2 - 2\mathrm{i})| \leq 2$ and $\mathrm{Re}(z) > \left|1 + \sqrt{3}\mathrm{i}\right| = 2$.



Part (a). Note $|z - 2 - 2\mathrm{i}| = \arg(z - (2 + 2\mathrm{i}))$. Let $A(2+2\mathrm{i})$ and $C(2-2\mathrm{i})$. Let T be the point at which AT is tangent to the circle. Then $\angle ATC = \frac{\pi}{2}$, $AC = 4$ and $TC = 2$. Hence, $\angle CAT = \arcsin \frac{2}{4} = \frac{\pi}{6}$. Thus, $\min \arg(z - 2 - 2\mathrm{i}) = -\frac{\pi}{2}$ and $\max \arg(z - 2 - 2\mathrm{i}) = \min \arg(z - 2 - 2\mathrm{i}) + \angle CAT = -\frac{\pi}{2} + \frac{\pi}{6} = -\frac{\pi}{3}$. Hence, $-\frac{\pi}{2} < \arg(z - 2 - 2i) \leq -\frac{\pi}{3}$.

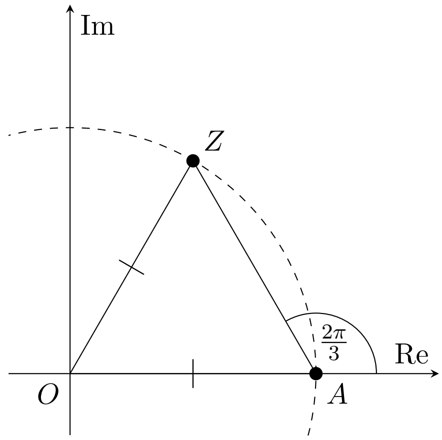
Part (b). Relative to C, T is given by $2\left(\cos \frac{\pi}{6} + \mathrm{i} \sin \frac{\pi}{6}\right) = \sqrt{3} + \mathrm{i}$. Thus, $T = (\sqrt{3} + \mathrm{i}) + (2 - 2\mathrm{i}) = 2 + \sqrt{3} - \mathrm{i}$.

Note $|w - 5 + 2\mathrm{i}| = k \implies |w - (5 - 2\mathrm{i})| = k$. Let $D(5 - 2\mathrm{i})$. Observe that CD is given by the sum of the radius of the circle and $\min k$. Hence, $\min k = 3 - 2 = 1$. Let $B(2 - 4\mathrm{i})$. Then $\max k$ is given by the distance between B and D. By the Pythagorean Theorem, we have $\max k = \sqrt{(5-2)^2 + (-2-(-4))^2} = \sqrt{13}$. Thus, $1 \leq k \leq \sqrt{13}$.

Self-Practice A10.4

Problem 1. If $\arg(z-2) = 2\pi/3$ and $|z| = 2$, determine $\arg(z)$.

Solution. Let $A(2+0\mathrm{i})$ and $Z(z)$.



Observe that $\angle OAZ = \pi - 2\pi/3 = \pi/3$. Since $OA = OZ = 2$ it follows that $\triangle OAZ$ is equilateral, so $\arg(z) = \angle AOZ = \pi/3$.

* * * * *

Problem 2. z is a complex number such that $\arg(z-1) = \pi/3$ and $\arg(z-\mathrm{i}) = \pi/6$. By finding the Cartesian equations of the two half-lines, or otherwise, find the value of $\arg(z)$.

Solution. Let $z = x + \mathrm{i}y$, where $x, y \in \mathbb{R}$. Then

$$\arg(z-1) = \arctan\frac{y}{x-1} = \frac{\pi}{3} \implies \frac{y}{x-1} = \sqrt{3} \implies y = \sqrt{3}x - \sqrt{3}$$

and

$$\arg(z-\mathrm{i}) = \arctan\frac{y-1}{x} = \frac{\pi}{6} \implies \frac{y-1}{x} = \frac{1}{\sqrt{3}} \implies y = 1 + \frac{1}{\sqrt{3}}x.$$

Equating the two, we have

$$\sqrt{3}x - \sqrt{3} = 1 + \frac{1}{\sqrt{3}}x \implies x = \frac{1+\sqrt{3}}{\sqrt{3} - 1/\sqrt{3}} = \frac{\sqrt{3}+3}{2}.$$

Thus,

$$y = \sqrt{3}\,(x-1) = \sqrt{3}\left(\frac{\sqrt{3}+3}{2} - 1\right) = \frac{3+\sqrt{3}}{2},$$

so $x = y$ and

$$\arg z = \arctan\frac{y}{x} = \arctan 1 = \frac{\pi}{4}.$$

* * * * *

Problem 3. The complex number z is given by $z = r\mathrm{e}^{\mathrm{i}\theta}$, where $r > 0$ and $0 \leq \theta \leq \pi/2$.

(a) Given that $w = \left(1 - \mathrm{i}\sqrt{3}\right)z$, find $|w|$ in terms of r and $\arg w$ in terms of θ.

(b) Given that r has a fixed value, draw an Argand diagram to show the locus of z as θ varies. On the same Argand diagram, show the corresponding locus of w. You should identify the modulus and argument of the end-point of each locus.

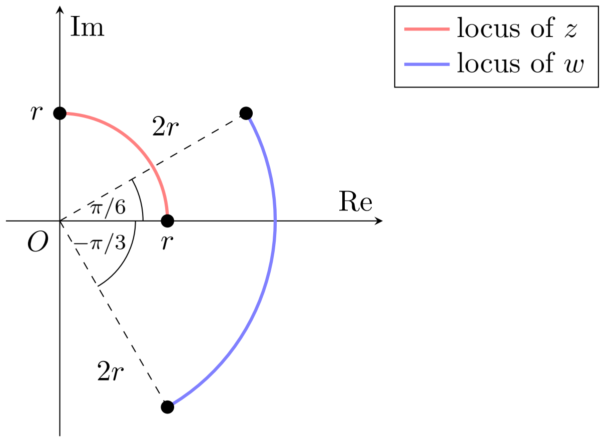
Solution.

Part (a). Note that $1 - \mathrm{i}\sqrt{3} = 2\mathrm{e}^{-\pi/3}$. Thus,

$$w = \left(1 - \sqrt{3}\mathrm{i}\right) z = \left(2\mathrm{e}^{-\pi/3}\right)\left(r\mathrm{e}^{\mathrm{i}\theta}\right) = 2r\mathrm{e}^{\mathrm{i}(\theta - \pi/3)}.$$

Hence, $|w| = 2r$ and $\arg(w) = \theta - \pi/3$.

Part (b).

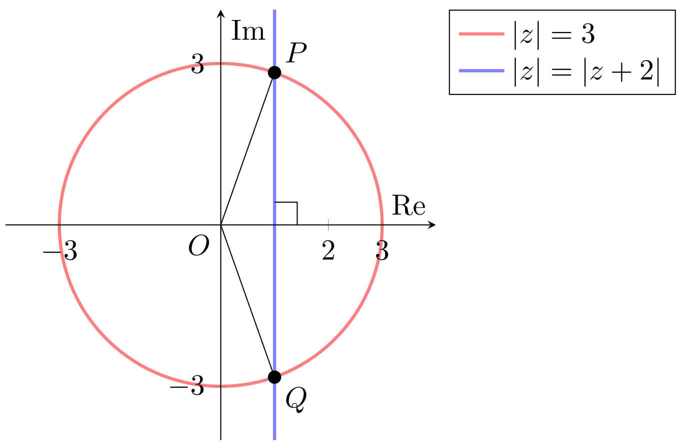


* * * * *

Problem 4. The complex number z satisfies the equation $|z| = |z+2|$. Show that the real part of z is -1. The complex number z also satisfies the equation $|z| = 3$. The two possible values of z are represented by the points P and Q in an Argand diagram. Draw a sketch showing the positions of P and Q, and calculate the two possible values of $\arg z$, giving your answers in radians correct to 3 significant figures.

It is given that P and Q lie on the locus $|z-a| = b$, where a and b are real, and $b > 0$. Give a geometrical description of this locus, and hence find the least possible value of b and the corresponding value of a.

Solution. Observe that the locus of $|z| = |z+2|$ is the perpendicular bisector of $(0,0)$ and $(2,0)$, which has Cartesian equation $x = 1$, $y \in \mathbb{R}$. Thus, the real part of z (i.e. x) is always 1.



From the diagram,

$$\arg(z) = \pm \arccos \frac{1}{3} = \pm 1.23 \text{ (3 s.f.)}.$$

The locus of $|z-a| = b$ is a circle of radius b centred at the point representing a.

For b to be at a minimum, PQ must be the diameter of the circle. By the Pythagorean Theorem,

$$3^2 = \left(\frac{PQ}{2}\right)^2 + 1^2 \implies PQ = 32.$$

Thus,

$$\min b = \frac{PQ}{2} = \frac{\sqrt{32}}{2} = 2\sqrt{2}.$$

The point representing a is then the midpoint of P and Q, i.e. $a = 1$.

* * * * *

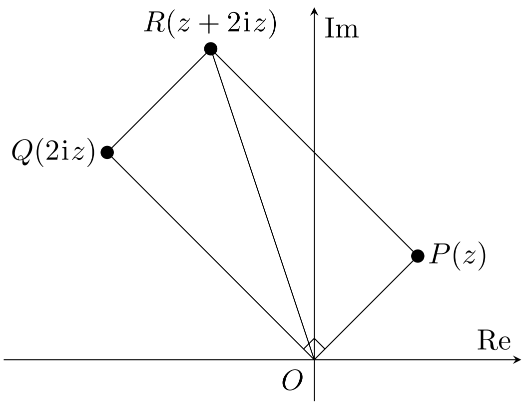
Problem 5. The complex number z is given by $z = x + \mathrm{i}y$, where $x > 0$ and $y > 0$. Sketch an Argand diagram, with origin O, showing points P, Q and R representing z, $2\mathrm{i}z$ and $(z + 2\mathrm{i}z)$ respectively. State the size of angle POQ, and describe briefly the geometrical relationship between O, P, Q and R.

(a) Given that $x = 2y$, show that R lies on the imaginary axis.

(b) Given that $y = 2x$, show that the point representing z^2 is collinear with the origin and the point R.

(c) Given that $|z| \leq 2$ and $\arctan \frac{1}{2} \leq \arg z \leq \arctan 2$, calculate the area of the region in which the point P can lie.

Solution.



$\angle POQ = \pi/2$, and $OPQR$ forms a rectangle.

Part (a). Given $x = 2y$, we have

$$z + 2\mathrm{i}z = z\,(1 + 2\mathrm{i}) = (2y + \mathrm{i}y)\,(1 + 2\mathrm{i}) = 5y\mathrm{i},$$

which is purely imaginary. Hence, R lies on the imaginary axis.

Part (b). Given $y = 2x$, we have

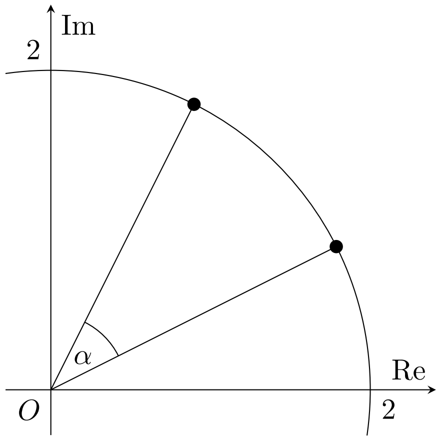
$$z = x + 2\mathrm{i}x = x\,(1 + 2\mathrm{i}) \implies \arg\!\left(z^2\right) = 2\arg(1 + 2\mathrm{i})\,.$$

Meanwhile,

$$\arg(z + 2\mathrm{i}z) = \arg(z) + \arg(1 + 2\mathrm{i}) = \arg(1 + 2\mathrm{i}) + \arg(1 + 2\mathrm{i}) = 2\arg(1 + 2\mathrm{i})\,.$$

Since z^2 and $z + 2\mathrm{i}z$ have identical arguments, the points representing them must be collinear with the origin.

Part (c).



From the above figure, we see that $\alpha = \arctan 2 - \arctan(1/2)$. The area of the region in which P can lie in is thus

$$\text{Area} = \pi(2)^2 \times \frac{\arctan 2 - \arctan(1/2)}{2\pi} = 1.29 \text{ units}^2.$$

* * * * *

Problem 6. A complex number z satisfies $|z - a| = a$, $a \in \mathbb{R}^+$.

(a) The point P represents the complex number w, where $w = 1/z$, in an Argand diagram. Show that the locus of P is a straight line.

(b) Sketch both loci on the same diagram and show that the two loci do not intersect if $0 < a < 1/2$.

(c) For $a = 1/2$, find the range of values of $\arg(z - 1/a)$, giving your answer correct to $0.1°$. State the limit of $\arg(z - 1/a)$ when a approaches 0.

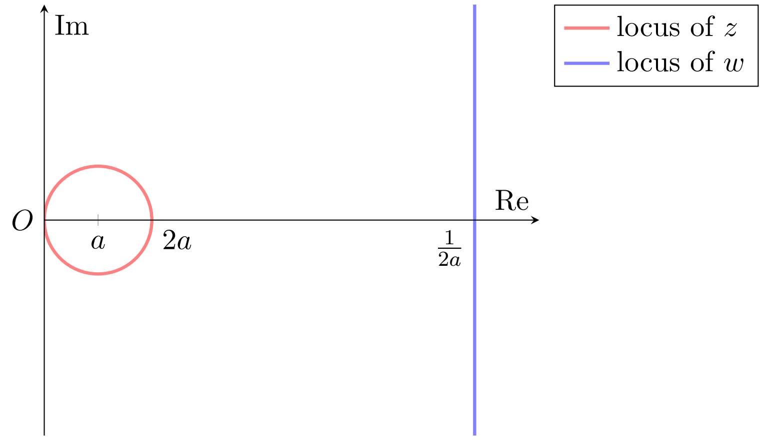
Solution.

Part (a). We have

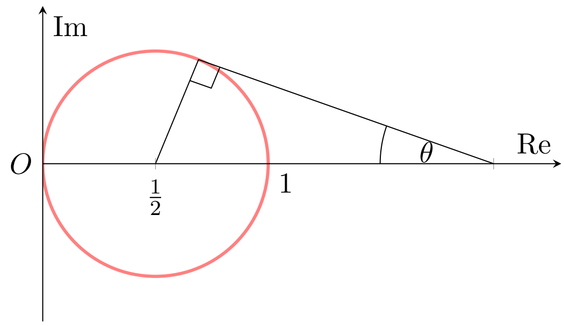
$$|z - a| = \left|\frac{1}{w} - a\right| = \left|\frac{1 - aw}{w}\right| = a \implies |1 - aw| = |aw| \implies |w| = \left|\frac{1 - aw}{a}\right| = \left|\frac{1}{a} - w\right|.$$

Hence, the locus of P is the perpendicular bisector of the origin and $(1/a, 0)$. Equivalently, it is the vertical line passing through $(1/2a, 0)$.

Part (b). If $0 < a < 1/2$, then $1/a > 2$, so the real part of any point on the locus of w is $1/2a > 1$. The largest real part of any point on the locus of z is $a + a = 2a < 1$. Thus, both loci will not intersect.



Part (c).



From the diagram,

$$\sin\theta = \frac{1/2}{2 - 1/2} \implies \theta = \arcsin\frac{1}{3}.$$

Thus,

$$160.5^\circ = \pi - \theta \leq \arg(z - 2) \leq \pi + \theta = 199.5^\circ.$$

As $a \to 0$, $\arg(z - 1/a) \to \pi$.

* * * * *

Problem 7. Sketch, on an Argand diagram, the locus representing the complex number z for which

$$|z - 4 - 3\mathrm{i}| = 2.$$

(a) Given that a is the least possible value of $|z|$, find a.

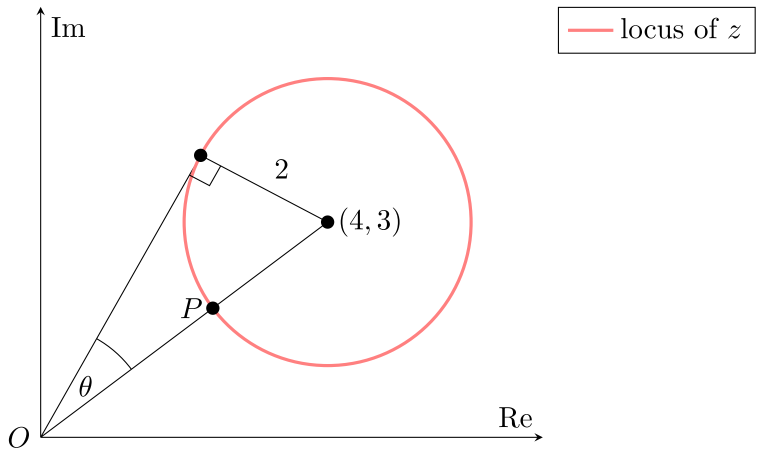
(b) The complex number p is such that

$$|p - 4 - 3\mathrm{i}| = 2 \quad \text{and} \quad |p| = a.$$

State the exact value of $\arg p$.

(c) Deduce the greatest value of $\arg(z/p)$, giving your answer correct to 2 decimal places.

Solution.



Part (a). Clearly, $a = \sqrt{4^2 + 3^2} - 2 = 3$.

Part (b). Clearly, $\arg p = \arctan(3/4)$.

Part (c). Observe that

$$\max \arg \frac{z}{p} = \theta = \arcsin \frac{2}{5} = 23.58°.$$

* * * * *

Problem 8 (🌶). On an Argand diagram, the point U represents the complex number z, and the points V and W represent the complex numbers z^2 and $z^2 + 1$ respectively.

(a) (i) Given that $\arg(z) = \alpha$, where $\pi/4 < \alpha < \pi/2$, so that U lies on the half-line L_1 with equation $y = x \tan \alpha$ for $x > 0$, show that V lies on the half-line L_2 with equation $y = x \tan 2\alpha$ for $x < 0$. Find the equation of the locus L_3 of W.

(ii) The points E and F represent the values of z for which W coincides with U. Find the value of α for which the common point of L_1 and L_3 is either E or F.

(b) Given instead that $|z| = k$, where $k > 0$, so that U lies on a circle C, show that W lies on a circle C', and find its centre and radius. Find the value of k for which the common points of C and C' are E and F.

Solution.

Part (a).

Part (a)(i). Note that $\arg z^2 = 2 \arg z = 2\alpha$. Let $z^2 = x + iy$, where $x, y \in \mathbb{R}$. Then $\arg(z^2) = \arctan(y/x)$. Equating the two yields

$$\arctan \frac{y}{x} = 2\alpha \implies y = x \tan(2\alpha).$$

Note that $\arg(z^2) = 2\alpha \in (\pi/2, \pi)$, so $x = \operatorname{Re}(z^2) < 0$.

L_3 is precisely L_2 shifted one unit in the positive real axis. Hence, the equation of L_3 is $y = (x - 1)\tan(2\alpha)$.

Part (a)(ii). Since W coincides with U, we have $z = z^2 + 1$. Solving, we get

$$z = \frac{1 + \sqrt{3}i}{2}.$$

Note that we reject the negative branch since $\arg z \in (\pi/4, \pi/2)$ implies $\operatorname{Im} z > 0$. Thus,

$$\alpha = \arg z = \arctan \frac{\sqrt{3}}{1} = \frac{\pi}{3}.$$

Part (b). Observe that

$$\left|\left(z^2 + 1\right) - 1\right| = \left|z^2\right| = k^2,$$

so W lies on a circle with radius k^2 and centre $(1, 0)$. For E and F to lie on C, we require

$$k = \left|\frac{1 \pm \sqrt{3}\mathrm{i}}{2}\right| = 1.$$

Assignment A10.4

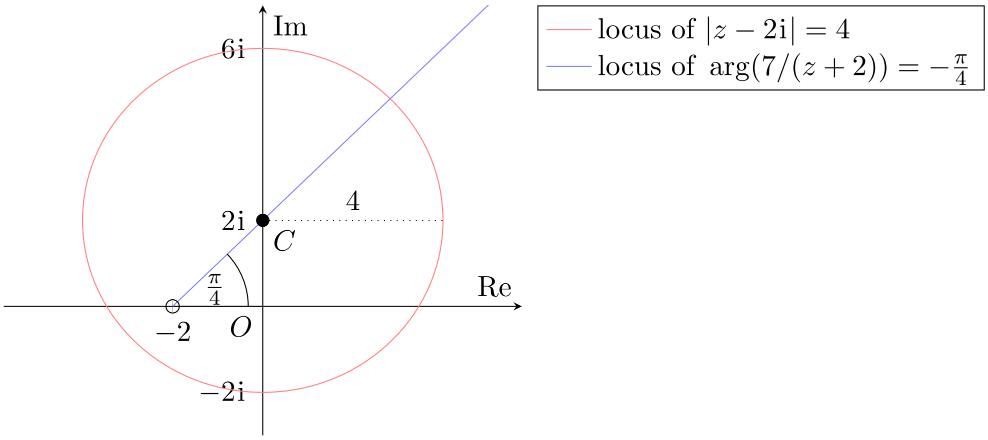
Problem 1. On a single Argand diagram, sketch the following loci.

(a) $|z - 2\mathrm{i}| = 4$.

(b) $\arg\left(\frac{7}{z+2}\right) = -\frac{\pi}{4}$.

Hence, or otherwise, find the exact value of z satisfying both equations in part (a) and (b).

Solution. Note that $\arg\left(\frac{7}{z+2}\right) = -\frac{\pi}{4} \implies \arg(z - (-2)) = \frac{\pi}{4}$.



Solving both equations simultaneously,

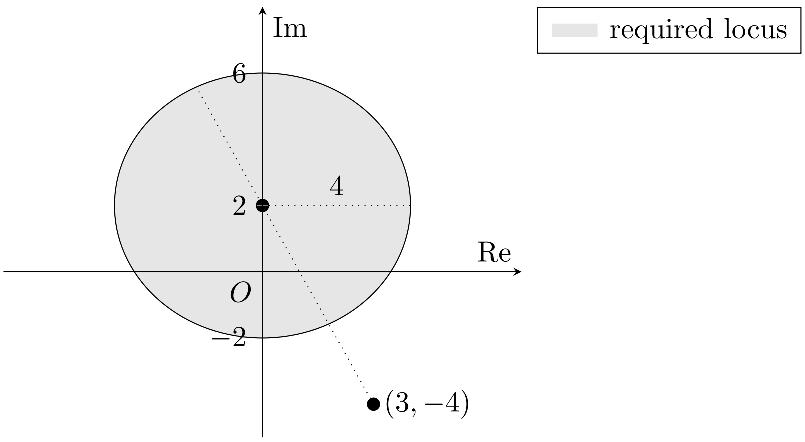
$$z = 2\mathrm{i} + \left(\cos\frac{\pi}{4} + \mathrm{i}\sin\frac{\pi}{4}\right) = 2\mathrm{i} + \frac{\sqrt{2}}{2} + \mathrm{i}\frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{2} + \left(2 + \frac{\sqrt{2}}{2}\right)\mathrm{i}.$$

* * * * *

Problem 2. Given that $|z - 2\mathrm{i}| \leq 4$, illustrate the locus of the point representing the complex number z in an Argand diagram.

Hence, find the greatest and least possible value of $|z - 3 + 4\mathrm{i}|$, given that $|z - 2\mathrm{i}| \leq 4$.

Solution.



Note that $|z - 3 + 4\mathrm{i}| = |z - (3 - 4\mathrm{i})|$ represents the distance between z and the point $(3, -4)$. By Pythagoras' Theorem, the distance between the centre of the circle $(0, 2)$

and $(3,-4)$ is $\sqrt{(0-3)^2+(2+4)^2} = 3\sqrt{5}$. Hence, $\max|z-3+4\mathrm{i}| = 3\sqrt{5}+4$, while $\min|z-3+4\mathrm{i}| = 3\sqrt{5}-4$. Thus, $\max|z-3+4\mathrm{i}| = 3\sqrt{5}+4$, $\min|z-3+4\mathrm{i}| = 3\sqrt{5}-4$.

* * * * *

Problem 3. The point A on an Argand diagram represents the fixed complex number a, where $0 < \arg a < \frac{\pi}{2}$. The complex numbers z and w are such that $|z-2\mathrm{i}a| = |a|$ and $|w| = |w+\mathrm{i}a|$.

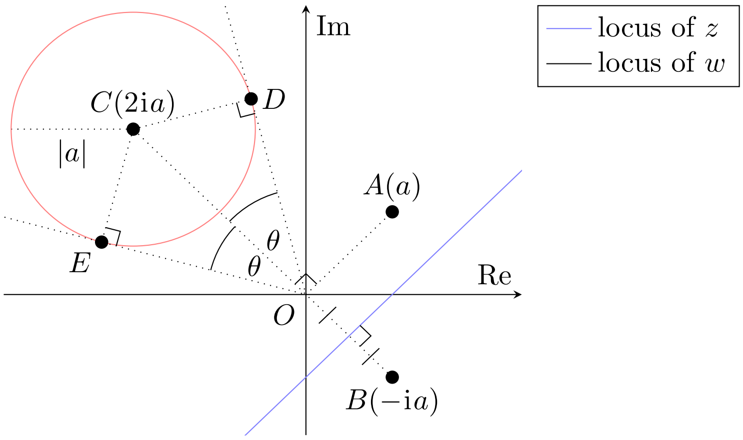
Sketch, on a single diagram, the loci of the point representing z and w.

Find

(a) the minimum value of $|z-w|$ in terms of $|a|$,

(b) the range of values of $\arg\frac{1}{z}$ in terms of $\arg a$.

Solution. Note that $|w| = |w+ia| \implies |w-0| = |w-(-ia)|$.



Part (a). Let $B(-\mathrm{i}a)$ and $C(2\mathrm{i}a)$. Note that $W\left(-\frac{1}{2}\mathrm{i}a\right)$ lies on the locus of w as well as the line passing through OC. Since CW is perpendicular to the locus of w, it follows that the minimum value of $|z-w|$ is given by

$$CW - |a| = \left|2\mathrm{i}a + \frac{1}{2}\mathrm{i}a\right| - |a| = \frac{5}{2}|a|\,|\mathrm{i}| - |a| = \frac{3}{2}|a|.$$

Part (b). Let D and E be such that OD and OE are tangent to the circle given by the locus of z. Let $\angle COD = \theta$. Observe that $\sin\theta = \frac{CD}{CO} = \frac{|a|}{|2\mathrm{i}a|} = \frac{1}{2}$, whence $\theta = \frac{\pi}{6}$. Since $\angle COA = \arg\mathrm{i} = \frac{\pi}{2}$, it follows that $\angle DOA = \frac{\pi}{2} - \theta = \frac{\pi}{2} - \arcsin\frac{1}{2} = \frac{\pi}{3}$. Thus, $\min\arg z = \arg a + \angle DOA = \arg a + \frac{\pi}{3}$. Meanwhile, $\angle COE = \angle COD = \theta$, whence $\max\arg z = \arg a + \frac{\pi}{2} + \theta = \arg a + \frac{2}{3}\pi$. Since $\arg\frac{1}{z} = -\arg z$, we thus have $\arg\frac{1}{z} \in \left[-\left(\arg a + \frac{2}{3}\pi\right), -\left(\arg a + \frac{\pi}{3}\right)\right]$.

* * * * *

Problem 4.

(a) Solve the equation

$$z^7 - (1+\mathrm{i}) = 0,$$

giving the roots in the form $r\mathrm{e}^{\mathrm{i}\alpha}$, where $r > 0$ and $-\pi < \alpha \le \pi$.

(b) Show the roots on an Argand diagram.

(c) The roots represented by z_1 and z_2 are such that $0 < \arg z_1 < \arg z_2 < \frac{\pi}{2}$. Explain why the locus of all points z such that $|z-z_1| = |z-z_2|$ passes through the origin. Draw this locus on your Argand diagram and find its Cartesian equation.

(d) Describe the transformation that will map the points representing the roots of the equation $z^7 - (1 + \mathrm{i}) = 0$ to the points representing the roots of the equation $(z - 2)^7 - (1 + \mathrm{i}) = 0$ on the Argand diagram.

Solution.

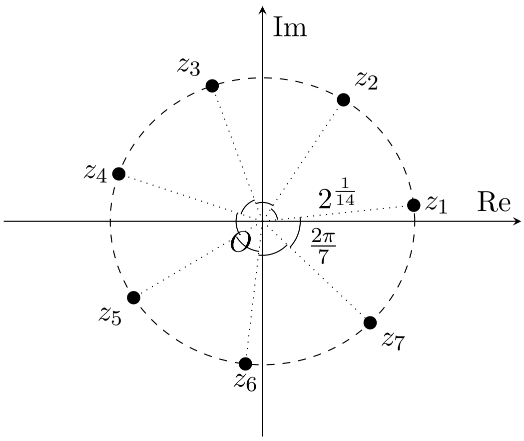
Part (a). Note that $1 + \mathrm{i} = 2^{\frac{1}{2}} \mathrm{e}^{\mathrm{i}\pi(\frac{1}{4}+2k)}$, where $k \in \mathbb{Z}$. Hence,

$$z^7 = 1 + \mathrm{i} = 2^{\frac{1}{2}} \mathrm{e}^{\mathrm{i}\pi(\frac{1}{4}+2k)} \implies z = 2^{\frac{1}{14}} \mathrm{e}^{\mathrm{i}\pi(\frac{1}{4}+2k)/7} = 2^{\frac{1}{14}} \mathrm{e}^{\mathrm{i}\pi(1+8k)/28}.$$

Taking $k \in \{-3, -2, \ldots, 2, 3\}$, we have

$$z = 2^{\frac{1}{14}} \mathrm{e}^{-\mathrm{i}\pi\frac{23}{28}},\ 2^{\frac{1}{14}} \mathrm{e}^{-\mathrm{i}\pi\frac{15}{28}},\ 2^{\frac{1}{14}} \mathrm{e}^{-\mathrm{i}\pi\frac{7}{28}},\ 2^{\frac{1}{14}} \mathrm{e}^{\mathrm{i}\pi\frac{1}{28}},\ 2^{\frac{1}{14}} \mathrm{e}^{\mathrm{i}\pi\frac{9}{28}},\ 2^{\frac{1}{14}} \mathrm{e}^{\mathrm{i}\pi\frac{17}{28}},\ 2^{\frac{1}{14}} \mathrm{e}^{\mathrm{i}\pi\frac{25}{28}}.$$

Part (b).



Part (c). Since $|z_1| = |z_2| = 2^{\frac{1}{14}}$, the distance between z_1 and the origin and the distance between z_2 and the origin are equal. Since the locus of $|z - z_1| = |z - z_2|$ represents all points equidistant from z_1 and z_2, it passes through the origin.

Observe that the midpoint of z_1 and z_2 will have argument $\frac{1}{2}\left(\frac{1}{28}\pi + \frac{9}{28}\pi\right) = \frac{5}{28}\pi$. Thus, the Cartesian equation of the locus of z is given by $y = \tan(5\pi/28)\, x$.

Part (d). Translate the points 2 units in the positive real direction.

A11 Permutations and Combinations

Tutorial A11

Problem 1. In a particular country, the alphabet contains 25 letters. A car registration number consists of two different letters of the alphabet followed by an integer n such that $100 \leq n \leq 999$. Find the number of possible car registration numbers.

Solution. Note that the number of possible n is $999 - 100 + 1 = 900$. Hence, the number of possible car registration numbers is given by ${}^{25}C_2 \cdot 900 = 540000$.

* * * * *

Problem 2. A girl wishes to phone a friend but cannot remember the exact number. She knows that it is a five-digit number that is even, and that it consists of the digits 2, 3, 4, 5, and 6 in some order. Using this information, find the greatest number of wrong telephone numbers she could try.

Solution. Since the number is odd, there are only 3 possibilities for the last digit. Hence, the maximum wrong numbers she could try is $3 \cdot 4! - 1 = 71$.

* * * * *

Problem 3. How many ways are there to select a committee of

(a) 3 students

(b) 5 students

out of a group of 8 students?

Solution.
Part (a). There are ${}^{8}C_3 = 56$ ways.
Part (b). There are ${}^{8}C_5 = 56$ ways.

* * * * *

Problem 4. How many ways are there for 2 men, 2 women and 2 children to sit a round table?

Solution. Since the men, women and children are all distinct, there are $(2+2+2-1)! = 120$ ways.

* * * * *

Problem 5. Find the number of different arrangements of the eight letters of the word NONSENSE if

(a) there is no restriction on the arrangement,

(b) the two letters E are together,

(c) the two letters E are not together,

(d) the letters N are all separated,

(e) only two of the letters N are together.

Solution.

Part (a). Note that N, S and E are repeated 3, 2, and 2 times respectively. Thus, the total number of arrangements is given by $\frac{8!}{3!\,2!\,2!} = 1680$.

Part (b). Consider the two E's as one unit. Altogether, there are 7 units. Hence, the required number of arrangements is given by $\frac{7!}{3!\,2!} = 420$.

Part (c). From part (a) and part (b), the required number of arrangements is given by $1680 - 420 = 1260$.

Part (d). There are $\frac{5!}{2!\,2!}$ ways to arrange the non-N letters, and ${}^6C_3$ ways to slot in the 3 N's into the 6 gaps in between the non-N letters. Thus, the required number of arrangements is given by $\frac{5!}{2!\,2!} \cdot {}^6C_3 = 600$.

Part (e). Consider the three N's as one unit. Altogether there are 6 units. Hence, the number of arrangements where all 3 N's are together is given by $\frac{6!}{2!\,2!} = 180$. Thus, from parts (a) and (d), the required number of arrangements is given by $1680-600-180 = 900$.

* * * * *

Problem 6. Find the number of teams of 11 that can be select from a group of 15 players

(a) if there is no restriction on choice,

(b) if the youngest two players and at most one of the oldest two players are to be included.

Solution.

Part (a). The number of teams is given by ${}^{15}C_{11} = 1365$.

Part (b). Given that the youngest two players are always included, we are effectively finding the number of teams of 9 from a group of 13 players with the restriction that at most one of the oldest two players are to be included.

Disregarding the restriction, the total number of teams is given by ${}^{13}C_9 = 715$.

Consider now that number of teams where both of the 2 oldest players are included. This is given by ${}^{11}C_7 = 330$.

Thus, the required number of teams is $715 - 330 = 385$.

* * * * *

Problem 7. A ten-digit number is formed by writing down the digits 0, 1, ..., 9 in some order. No number is allowed to start with 0. Find how many such numbers are

(a) odd,

(b) less than 2 500 000 000.

Solution.

Part (a). Since the number is odd, there are 5 possibilities for the last digit. Furthermore, since no number is allowed to start with 0, there are $10 - 2 = 8$ possibilities for the first digit. The remaining 8 digits are free. Hence, the required number of numbers is $5 \cdot 8 \cdot 8! = 1612800$.

Part (b). *Case 1*: *Number starts with 1.* Since there are no further restrictions, the number of valid numbers in this case is 9!.

Case 2: *Number starts with 2.* Given the restriction that the number be less than 2 500 000 000, the second digit must be strictly less than 5, thus giving 4 possibilities for the second digit. The remaining 8 digits are free, for a total number of valid numbers of $4 \cdot 8!$.

Thus, the required number of numbers is $9! + 4 \cdot 8! = 524160$.

* * * * *

Problem 8. Eleven cards each bear a single letter, and together, they can be made to spell the word "EXAMINATION".

(a) Three cards are selected from the eleven cards, and the order of selection is not relevant. Find how many possible selections can be made

(i) if the three cards all bear different letters,

(ii) if two of the three cards bear the same letter.

(b) Two cards bearing the letter N have been taken away. Find the number of different arrangements for the remaining cards that can be made with no two adjacent letters the same.

Solution.

Part (a).

Part (a)(i). Observe that there are 8 distinct letters in "EXAMINATION". Hence, the number of possible selections is ${}^8C_3 = 56$.

Part (a)(ii). Note that there are 3 letters that appear twice in "EXAMINATION". Hence, the number of possible selections is given by ${}^3C_1 \cdot {}^7C_1 = 21$.

Part (b). Note that there are now 2 letters that appear twice, namely A and I. Hence, the total number of possible arrangements is $\frac{9!}{2!\,2!}$.

Consider "AA" and "II" as one unit each. Altogether, there are 7 units. The number of arrangements with two pairs of adjacent letters that are the same is hence given by 7!.

Consider "AA" as one unit, and suppose the two I's are not adjacent to each other. Observe that the non-I letters comprise 6 units, hence giving 6! ways of arranging them. Also observe that there are ${}^7C_2$ ways to slot in the two I's (which guarantee that they are not adjacent to each other). There are hence $6! \cdot {}^7C_2$ possible arrangements in this case. A similar argument follows for the case where the two I's are adjacent but the A's are not.

From the above discussion, it follows that the required number of arrangements is given by $\frac{9!}{2!\,2!} - 7! - 2 \cdot 6! \cdot {}^7C_2 = 55440$.

* * * * *

Problem 9. Find how many three-letter code words can be formed from the letters of the word:

(a) PEAR.

(b) APPLE.

(c) BANANA.

Solution.

Part (a). Since all 4 letters are distinct, the number of code-words is given by ${}^4P_3 = 24$.

Part (b). Tally of letters: 2 'P', 1 'A', 1 'L', 1 'E' (5 letters, 4 distinct).

Case 1: *All letters distinct.* Since there are 4 distinct letters, the number of code-words in this case is ${}^4P_3 = 24$.

Case 2: *2 letters the same, 1 different.* Note that 'P' is the only letter repeated more than once. Reserving two spaces for 'P' leaves one space left for three remaining letters. Hence, there are ${}^1C_1 \cdot {}^3C_1 = 3$ different combinations that can be formed, with $\frac{3!}{2!} = 3$ ways to arrange each combination. Hence, the number of code-words in this case is $3 \cdot 3 = 9$.

Thus, the total number of code-words is $24 + 9 = 33$.

Part (c). Tally of letters: 3 'A', 2 'N', 1 'B' (6 letters, 3 distinct).

Case 1: *All letters distinct.* Since there are only 3 distinct letters, the number of code-words in this case is ${}^3P_3 = 6$.

Case 2: *2 letters the same, 1 different.* Observe that both 'A' and 'N' are repeated more than once. Reserving 2 spaces for either letter leaves one space left for the two remaining letters. Hence, there are ${}^2C_1 \cdot {}^2C_1 = 4$ different combinations that can be formed, with $\frac{3!}{2!} = 3$ ways to arrange each combination. Hence, the number of code-words in this case is $4 \cdot 3 = 12$.

Case 3: *All letters the same.* Observe that 'A' is the only letter repeated thrice. Hence, the number of code-words in this case is 1.

Altogether, the total number of code-words is $6 + 12 + 1 = 19$.

* * * * *

Problem 10. A group of diplomats is to be chosen to represent three islands, K, L and M. The group is to consist of 8 diplomats and is chosen from a set of 12 diplomats consisting of 3 from K, 4 from L and 5 from M. Find the number of ways in which the group can be chosen if it includes

(a) 2 diplomats from K, 3 from L and 3 from M,

(b) diplomats from L and M only,

(c) at least 4 diplomats from M,

(d) at least 1 diplomat from each island.

Solution.

Part (a). Note that there are ${}^3C_2$ ways to select 2 diplomats from K, ${}^4C_3$ ways to select 3 diplomats from L, and ${}^5C_3$ ways to select 3 diplomats from M. Thus, the number of possible groups is given by ${}^3C_2 \cdot {}^4C_3 \cdot {}^5C_3 = 120$.

Part (b). There are a total of 9 diplomats from L and M. Hence, the number of possible groups is ${}^9C_8 = 9$.

Part (c). *Case 1*: *4 diplomats from M.* Note that there are ${}^5C_4$ combinations for the 4 diplomats from M. Furthermore, since M contributes 4 diplomats, K and L must contribute the other 4 diplomats. Since K and L have a total of 7 diplomats, this gives a total of ${}^5C_4 \cdot {}^7C_4$ possibilities.

Case 2: *5 diplomats from M.* Since M has 5 diplomats, there is only one way for M to send 5 diplomats (all of them have to be chosen). Meanwhile, K and L must contribute the other 3 diplomats from a pool of 7. This gives a total of ${}^7C_3$ possibilities.

Altogether, there are ${}^5C_4 \cdot {}^7C_4 + {}^7C_3 = 210$ total possibilities.

Part (d). Observe that K and M have a total of 8 diplomats. Hence, there is only one possibility where the group only consists of diplomats from K and M.

Since K and L have a total of 7 diplomats, it is impossible for the group to only come from K and L.

From part (b), we know that there are 9 ways where the group consists only of diplomats from L and M.

Note that there are a total of ${}^{12}C_8$ possible ways to choose the group.

Altogether, the required number of possibilities is given by ${}^{12}C_8 - 9 - 1 = 485$.

Problem 11. Alisa and Bruce won a hamper at a competition. The hamper comprises 9 different items.

(a) How many ways can the 9 items be divided among Alisa and Bruce if each of them gets at least one item each?

(b) How many ways can a set of 3 or more items be selected from the 9 items?

Solution.

Part (a). Note that the total number of ways to distribute the items is given by $2^9 = 512$. Also note that the only way either of them does not receive an item is when the other party gets all the items. This can only occur twice (once when Alisa receives nothing, and once when Bruce receives nothing). Thus, the number of ways where both of them gets at least one item each is $512 - 2 = 510$.

Part (b). Observe that the number of ways to choose a set of n items from the original 9 is given by ${}^9C_n$. Hence, the required number of ways is given by $512 - ({}^9C_0 + {}^9C_1 + {}^9C_2) = 466$.

* * * * *

Problem 12. In how many ways can 12 different books be distributed among students A, B, C and D

(a) if A gets 5, B gets 4, C gets 2 and D gets 1?

(b) if each student gets 3 books each?

Solution.

Part (a). At the start, A gets to pick 5 books from the 12 available books. There are ${}^{12}C_5$ ways to do so. Next, B gets to pick 4 books from the $12 - 5 = 7$ remaining books. There are ${}^7C_4$ ways to do so. Similarly, there are ${}^3C_2$ ways for C to pick his book, and ${}^1C_1$ ways for D to pick his. Hence, there are a total of ${}^{12}C_5 \cdot {}^7C_4 \cdot {}^3C_2 \cdot {}^1C_1 = 83160$ ways for the 12 books to be distributed.

Part (b). Following a similar argument as in part (a), the number of ways the 12 books can be distributed is given by ${}^{12}C_3 \cdot {}^9C_3 \cdot {}^6C_3 \cdot {}^3C_3 = 369600$.

* * * * *

Problem 13. 3 men, 2 women and 2 children are arranged to sit around a round table with 7 non-distinguishable seats. Find the number of ways if

(a) (i) the 3 men are to be together,

(ii) the 3 men are to be together, and the seats are numbered,

(b) no 2 men are to be adjacent to each other,

(c) only 2 men are adjacent to each other.

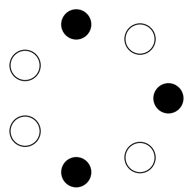
Solution.

Part (a).

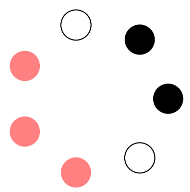
Part (a)(i). Consider the 3 men as one unit. Altogether, there are a total of 5 units, which gives a total of $(5 - 1)! = 4!$ ways for the 5 units to be arranged around the table. Since there are $3!$ ways to arrange the men, there are a total of $4! \cdot 3! = 144$ arrangements.

Part (a)(ii). Since there are a total of 7 distinguishable seats, the total number of arrangements is 7 times that of the number of arrangements with non-distinguishable seats. From part (a), this gives $144 \cdot 7 = 1008$ total arrangements.

Part (b). Observe that there is only one possible layout for no 2 men to be adjacent to each other (as shown in the diagram below). Since there are 4! ways to arrange the non-men, and 3! ways to arrange the men, there are a total of $4! \cdot 3! = 144$ arrangements.



Part (c). Observe that there are 3 possible layouts for only 2 men to be adjacent to each other (as shown in the diagram below). Since there are 4! ways to arrange the non-men, and 3! ways to arrange the men, there are a total of $3 \cdot 4! \cdot 3! = 432$ arrangements.



* * * * *

Problem 14. Find the number of ways for 4 men and 4 boys to be seated alternately if they sit

(a) in a row,

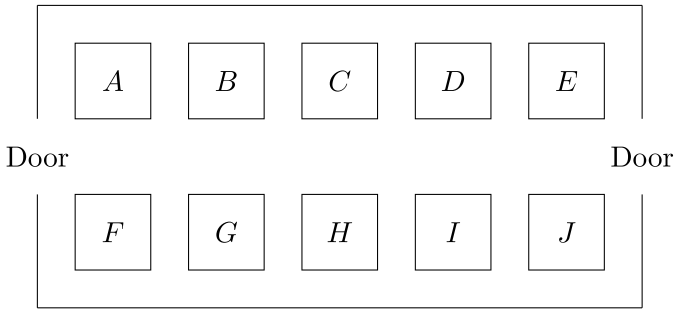
(b) at a round table.

Solution.

Part (a). Note that there are 2 possible layouts: one where a man sits at the start of the row, and one where a boy sits at the start of the row. Since there are 4! ways to arrange both the men and boys, there are a total of $2 \cdot 4! \cdot 4! = 1152$ arrangements.

Part (b). Given the rotational symmetry of the circle, there is now only one possible layout. Fixing one man, there are 3! ways to arrange the other men and 4! ways to arrange the boys, giving a total of $3! \cdot 4! = 144$ arrangements.

Problem 15. A rectangular shed, with a door at each end, contains ten fixed concrete bases marked A, B, C, $\ldots$, J, five on each side (see diagram). Ten canisters, each containing a different chemical, are placed with one canister on each base. In how many ways can the canisters be placed on the bases?



Find the number of ways in which the canisters can be placed

(a) if 2 particular canisters must not be placed on any of the 4 bases A, E, F and J next to a door,

(b) if 2 particular canisters must not be placed next to each other on the same side.

Solution. There are $10! = 3628800$ ways to place the canisters on the bases.

Part (a). Observe that there are ${}^6P_2$ possible placements for the two particular canisters. Since the other 8 canisters have no restrictions, the total number of ways to place the canisters is given by ${}^6P_2 \cdot 8! = 1209600$.

Part (b). Consider the number of ways the two particular canisters can be placed adjacently. There are $2 \cdot (5 - 1) = 8$ possible arrangements per side, giving a total of $2 \cdot 8 = 16$ possible arrangements. Since the other 8 canisters have no restrictions, the total number of ways to place the canisters is given by $16 \cdot 8! = 645120$. The required number of ways is thus given by $3628800 - 645120 = 2983680$.

Self-Practice A11

Problem 1. Find the number of three-letter codewords that can be made using the letters of the word "THREE" if at least one of the letters is E.

Solution.

Case 1: *Exactly 1 'E'.* There are ${}^3C_2 \times 3! = 18$ ways to form the codeword.

Case 2: *2 'E's.* There are ${}^3C_1 \times 3!/2! = 9$ ways to form the codeword.

Thus, there are a total of $18 + 9 = 27$ ways to form a codeword containing at least one 'E'.

* * * * *

Problem 2. Eight people go to the theatre and sit in a particular group of eight adjacent reserved seats in the front row. Three of the eight belong to one family and sit together.

(a) If the other five people do not mind where they sit, find the number of possible seating arrangements for all eight people.

(b) If the other five people do not mind where they sit, except that two of them refuse to sit together, find the number of possible seating arrangements for all eight people.

Solution.

Part (a). Treat the family as one unit. Altogether, there are 6 units. There are 6! ways to arrange the 6 units, and there are 3! ways to arrange the family within their unit. Hence, there are a total of $6! \times 3! = 4320$ possible arrangements for all eight people.

Part (b). We arrange the family unit and the three non-conflicting people first. There are 4! ways to do so. Next, we slot in the two conflicting people. There are ${}^5P_2$ ways to do so. Lastly, we arrange the family members, of which there are 3! ways to do so. Altogether, there are $4! \times {}^5P_2 \times 3! = 2880$ possible arrangements.

* * * * *

Problem 3. A panel of judges in an essay competition has to select, and place in order of merit, 4 winning entries from a total entry of 20. Find the number of ways in which this can be done.

As a first step in the selection, 5 finalists are selected, without being placed in order. Find the number of ways in which this can be done.

All 20 essays are subsequently assessed by three panels of judges for content, accuracy and style, respectively, and three special prizes are awarded, one by each panel. Find the number of ways in which this can be done, assuming that an essay may win more than one prize.

Solution. There are ${}^{20}P_4 = 116280$ ways to select and place the four winning entries.

There are ${}^{20}C_5 = 15504$ ways to select the five finalists.

There are $\left({}^{20}C_1\right)^3 = 8000$ ways to give out the three prizes.

* * * * *

Problem 4.

(a) A bookcase has four shelves with ten books on each shelf. Find the number of different selections that can be made by taking two books from each shelf (i.e. 8 books in all). Find also the number of different selections that can be made by taking eight books from each shelf (i.e. 32 books in all.)

(b) Eight cards each have a single digit written on them. The digits are 2, 2, 4, 5, 7, 7, 7, 7 respectively. Find the number of different 7-digit numbers that can be formed by placing seven of the cards side by side.

Solution.

Part (a). The number of difference selections in both scenarios is given by

$$\left({}^{10}C_2\right)^9 = \left({}^{10}C_8\right)^9 = 4100625.$$

Part (b).

Case 1: *A '4' or '5' is not selected.* Of the seven digits available, there are two '2's and four '7's. The number of arrangements is hence

$$\frac{7!}{2!4!} \times 2 = 210.$$

Case 2: *A '7' is excluded.* Of the seven digits available, there are two '2's and three '7's. The number of arrangements is hence

$$\frac{7!}{2!3!} = 420.$$

Case 3: *A '2' is excluded.* Of the seven digits available, there are four '7's. The number of arrangements is hence

$$\frac{7!}{4!} = 210.$$

Altogether, there are $210 + 420 + 210 = 840$ different 7-digit numbers that can be formed.

* * * * *

Problem 5. A team in a particular sport consists of 1 goalkeeper, 4 defenders, 2 midfielders and 4 attackers. A certain club has 3 goalkeepers, 8 defenders, 5 midfielders and 6 attackers.

(a) How many different teams can be formed by the club?

One of the midfielders in the club is the brother of one of the attackers in the club.

(b) How many different teams can be formed which include exactly one of the two brothers?

The two brothers leave the club. The club manager decides that one of the remaining midfielders can play either as a midfielder or a defender.

(c) How many different teams can now be formed by the club?

Solution.

Part (a).

Position	Goalkeeper	Defender	Midfielder	Attacker
No. Available	3	8	5	6
No. to Select	1	4	2	4

The number of teams that can be formed is

$${}^3C_1 \times {}^8C_4 \times {}^5C_2 \times {}^6C_4 = 31500.$$

Part (b). *Case 1*. Suppose the midfielder brother is included.

Position	Goalkeeper	Defender	Midfielder	Attacker
No. Available	3	8	4	6
No. to Select	1	4	1	4

The number of teams that can be formed in this case is

$$^3C_1 \times {}^8C_4 \times {}^4C_1 \times {}^6C_4 = 4200.$$

Case 2. Suppose the attacker brother is included.

Position	Goalkeeper	Defender	Midfielder	Attacker
No. Available	3	8	5	5
No. to Select	1	4	2	3

The number of teams that can be formed in this case is

$$^3C_1 \times {}^8C_4 \times {}^5C_2 \times {}^5C_3 = 12600.$$

Altogether, there are $4200 + 12600 = 16800$ ways to form a team where exactly one brother plays.

Part (c). *Case 1*. The midfielder appears as a midfielder.

Position	Goalkeeper	Defender	Midfielder	Attacker
No. Available	3	8	3	5
No. to Select	1	4	1	4

The number of teams that can be formed in this case is

$$^3C_1 \times {}^8C_4 \times {}^3C_1 \times {}^5C_4 = 3150.$$

Case 2. The midfielder appears as a defender.

Position	Goalkeeper	Defender	Midfielder	Attacker
No. Available	3	8	3	5
No. to Select	1	3	2	4

The number of teams that can be formed in this case is

$$^3C_1 \times {}^9C_3 \times {}^3C_2 \times {}^5C_4 = 2520.$$

Case 3. The midfielder does not play.

Position	Goalkeeper	Defender	Midfielder	Attacker
No. Available	3	8	3	5
No. to Select	1	4	2	4

The number of teams that can be formed in this case is

$$^3C_1 \times {}^8C_4 \times {}^3C_2 \times {}^5C_4 = 3150.$$

Altogether, there are $3150 + 2520 + 3150 = 8820$ ways to form a team.

Problem 6. A group of 12 people consists of 6 married couples.

(a) The group stands in a line.

(i) Find the number of different possible orders.

(ii) Find the number of different possible orders in which each man stands next to his wife.

(b) The group stands in a circle.

(i) Find the number of different possible arrangements.

(ii) Find the number of different possible arrangements if men and women alternate.

(iii) Find the number of different possible arrangements if each man stands next to his wife and men and women alternate.

Solution.

Part (a).

Part (a)(i). There are $12! = 479001600$ different possible orders.

Part (a)(ii). Group each couple as one unit, for a total of 6 units. There are $6!$ ways to arrange the 6 units, and 2 ways to arrange each couple within their unit. Thus, there are a total of

$$6! \times 2^6 = 46080$$

different possible orders.

Part (b).

Part (b)(i). There are $11! = 39916800$ different possible orders.

Part (b)(ii). Fix one man. There are then

$$6 \times 5 \times 5 \times 4 \times 4 \times 3 \times 3 \times 2 \times 2 \times 1 \times 1 = 86400$$

ways to arrange all other 11 people.

Part (b)(iii). Group each couple as one unit, for a total of 6 units. There are $(6-1)!$ ways to arrangement the 6 units. Since men and women alternate, we either have 'man-woman' or 'woman-man' within each unit. Thus, there are a total of

$$(6-1)! \times 2 = 240$$

different possible orders.

* * * * *

Problem 7 (). A delegation of four students is to be selected from five badminton players, m floorball players, where $m > 3$, and six swimmers to attend the opening ceremony of the 2017 National Games. A pair of twins is among the floorball players. The delegation is to consist of at least one player from each sport.

(a) Show that the number of ways to select the delegation in which neither of the twins is selected is $k(m-2)(m+6)$, where k is an integer to be determined.

(b) Given that the number of ways to select a delegation in which neither of the twins is selected is more than twice the number of ways to select a delegation which includes exactly one of the twins, find the least value of m.

The pair of twins, one badminton player, one swimmer and two teachers, have been selected to attend a welcome lunch at the opening ceremony. Find the number of ways in which the group can be seated at a round table with distinguishable seats if the pair of twins is to be seated together and the teachers are separated.

Solution.

Part (a).

Case 1: *2 badminton players.* There are

$$^{5}C_2 \times (m-2) \times 6 = 60(m-2)$$

ways to form a delegation without the twins in this case.

Case 2: *2 floorball players.* There are

$$5 \times {}^{m-2}C_2 \times 6 = 30 \times \frac{(m-2)(m-3)}{2} = 15(m-3)(m-2)$$

ways to form a delegation without the twins in this case.

Case 3: *2 swimmers.* There are

$$5 \times (m-2) \times {}^{6}C_2 = 75(m-2)$$

ways to form a delegation without the twins in this case.

Altogether, there are a total of

$$60(m-2) + 15(m-3)(m-2) + 75(m-2) = 15(m+6)(m-2)$$

ways to form a delegation without the twins, so $k = 15$.

Part (b).

Case 1: *2 badminton players.* There are

$$^{5}C_2 \times 2 \times 6 = 120$$

ways to form a delegation with exactly one twin in this case.

Case 2: *2 floorball players.* There are

$$5 \times 2\,(m-2) \times 6 = 60(m-2)$$

ways to form a delegation with exactly one twin in this case.

Case 3: *2 swimmers.* There are

$$5 \times 2 \times {}^{6}C_2 = 150$$

ways to form a delegation with exactly one twin in this case.

Altogether, there are a total of

$$120 + 60(m-2) + 150 = 60m + 150$$

ways to form a delegation with exactly one twin. From the given condition,

$$2\,(60m + 150) < 15(m-2)(m+6),$$

hence the least m is 9.

First, consider the case where there are no restrictions on the teachers. Group the twins together as one unit for a total of 5 units. Since the seats are distinguishable, there are 5! ways to arrange the 5 units, and 2 ways to arrange the twins within their unit. In total, there are $5! \times 2 = 240$ arrangements without restrictions.

Now, consider the case where the teachers are together. Group the twins together, and group the teachers together for a total of 4 units. Since the seats are distinguishable, there are 4! ways to arrangement the 4 units. There are 2 ways each to arrange the twins and teachers within their unit. In total, there are $4! \times 2^2 = 96$ arrangements where the teachers are together.

Thus, there are $240 - 96 = 144$ arrangements where the teachers are separated.

Assignment A11

Problem 1. Find the number of different arrangements of seven letters in the word ADVANCE. Find the number of these arrangements which begin and end with "A" and in which "C" and "D" are always together.

Find the number of 4-letter code words that can be made from the letters of the word ADVANCE, using

(a) neither of the "A"s,

(b) both of the "A"s.

Solution. Tally of letters: 2 "A"s, 1 "D", 1 "V", 1 "N", 1 "C", 1 "E" (7 total, 6 distinct)

$$\text{Number of different arrangements} = \frac{7!}{2!} = 2520.$$

Since both "A"s are at the extreme ends, we are effectively finding the number of arrangements of the word "DVNCE" such that "C" and "D" are always together.

Let "C" and "D" be one unit. Altogether, there are 4 units. Hence,

$$\text{Required number of arrangements} = 4! \cdot 2 = 48.$$

Part (a). Without both "A"s, there are only 5 available letters to form the code words. This gives ${}^5C_4$ ways to select the 4 letters of the code word. Since each of the 5 remaining letters are distinct, there are 4! possible ways to arrange each word. This gives ${}^5C_4 {\cdot} 4! = 120$ such code words.

Part (b). With both "A"s included, we need another 2 letters from the 5 non-"A" letters. This gives ${}^5C_2$ ways to select the 4 letters of the code word. Since the 2 non-"A" letters are distinct, but the "A"s are repeated, there are $\frac{4!}{2!}$ possible ways to arrange each code word. This gives ${}^5C_2 \cdot \frac{4!}{2!} = 120$ such code words.

* * * * *

Problem 2. A box contains 8 balls, of which 3 are identical (and so are indistinguishable from one another) and the other 5 are different from each other. 3 balls are to be picked out of the box; the order in which they are picked out does not matter. Find the number of different possible selections of 3 balls.

Solution. Note that there are 6 distinct balls in the box.

Case 1: *No identical balls chosen.* No. of selections = ${}^6C_3$

Case 2: *2 identical balls chosen.* No. of selections = ${}^5C_1$

Case 3: *3 identical balls chosen.* No. of selections = ${}^3C_3$

Hence, the total number of selections is given by ${}^6C_3 + {}^5C_1 + {}^3C_3 = 26$.

* * * * *

Problem 3. The management board of a company consists of 6 men and 4 women. A chairperson, a secretary and a treasurer are chosen from the 10 members of the board. Find the number of ways the chairperson, the secretary and the treasurer can be chosen so that

(a) they are all women,

(b) at least one is a woman and at least one is a man.

The 10 members of the board sit at random around a round table. Find the number of ways that

(c) the chairperson, the secretary and the treasurer sit in three adjacent places.

(d) the chairperson, the secretary and the treasurer are all separated from each other by at least one other person.

(Extension) What if the seats around the table are numbered? Try parts (c) and (d) again.

Solution.

Part (a). Since there are 4 women and 3 distinct roles, the required number of ways is given by ${}^4P_3 = 24$.

Part (b). Note that the number of ways that all three positions are men is given by ${}^6P_3$, while the number of ways to choose without restriction is given by ${}^{10}P_3$. Hence, the required number of ways is given by ${}^{10}P_3 - {}^6P_3 - 24 = 576$.

Part (c). Consider the three positions as one unit. This gives 8 units altogether. There are hence $(8-1)! \cdot 3! = 30240$ ways.

Part (d). Seat the seven other people first. There are $(7-1)!$ ways to do so. Then, slot in the three positions in the 7 slots. There are ${}^7C_3 \cdot 3!$ ways to do so. Hence, the required number of ways is given by $(7-1)! \cdot {}^7C_3 \cdot 3! = 151200$.

Extension. Since the seats are numbered, the number of ways scales up by the number of seats, i.e. 10. Hence, the number of ways becomes 302400 and 1512000.

A12 Probability

Tutorial A12

Problem 1. A and B are two independent events such that $\mathbb{P}[A] = 0.2$ and $\mathbb{P}[B] = 0.15$. Evaluate the following probabilities.

(a) $\mathbb{P}[A \mid B]$,

(b) $\mathbb{P}[A \cap B]$,

(c) $\mathbb{P}[A \cup B]$.

Solution.
Part (a). Since A and B are independent, $\mathbb{P}[A \mid B] = \mathbb{P}[A] = 0.2$.
Part (b). Since A and B are independent, $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B] = 0.2 \cdot 0.15 = 0.03$.
Part (c). $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B] = 0.2 + 0.15 - 0.03 = 0.32$.

* * * * *

Problem 2. Two events A and B are such that $\mathbb{P}[A] = \frac{8}{15}$, $\mathbb{P}[B] = \frac{1}{3}$ and $\mathbb{P}[A \mid B] = \frac{1}{5}$. Calculate the probabilities that

(a) both events occur,

(b) only one of the two events occurs,

(c) neither event occurs.

Determine if event A and B are mutually exclusive or independent.

Solution.
Part (a).

$$\mathbb{P}[A \cap B] = \mathbb{P}[B]\,\mathbb{P}[A \mid B] = \frac{1}{3} \cdot \frac{1}{5} = \frac{1}{15}.$$

Part (b).

$$\begin{aligned}\mathbb{P}[\text{only one occurs}] &= \mathbb{P}[A \cup B] - \mathbb{P}[A \cap B] = \mathbb{P}[A] + \mathbb{P}[B] - 2\,\mathbb{P}[A \cap B] \\ &= \frac{8}{15} + \frac{1}{3} - 2\left(\frac{1}{15}\right) = \frac{11}{15}.\end{aligned}$$

Part (c).

$$\mathbb{P}[\text{neither occurs}] = 1 - \mathbb{P}[\text{at least one occurs}] = 1 - \left(\frac{1}{15} - \frac{11}{15}\right) = \frac{1}{5}.$$

Since $\mathbb{P}[A] = \frac{8}{15} \neq \frac{1}{5} = \mathbb{P}[A \mid B]$, it follows that A and B are not independent. Also, since $\mathbb{P}[A \cap B] = \frac{1}{15} \neq 0$, the two events are also not mutually exclusive.

Problem 3. Two events A and B are such that $\mathbb{P}[A] = \mathbb{P}[B] = p$ and $\mathbb{P}[A \cup B] = \frac{5}{9}$.

(a) Given that A and B are independent, find a quadratic equation satisfied by p.

(b) Hence, find the value of p and the value of $\mathbb{P}[A \cap B]$.

Solution.

Part (a). Since A and B are independent, we have $\mathbb{P}[A \mid B] = \mathbb{P}[A] = p$. Hence,

$$\begin{aligned} p = \mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} &= \frac{\mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cup B]}{\mathbb{P}[B]} = \frac{p + p - 5/9}{p} = 2 - \frac{5}{9p} \\ &\implies 9p^2 = 18p - 5 \implies 9p^2 - 18p + 5 = 0. \end{aligned}$$

Part (b). Observe that $9p^2 - 18p + 5 = (3p-1)(3p-5)$. Thus, $p = \frac{1}{3}$. Note that $p \neq \frac{5}{3}$ since $0 < p \leq 1$.

Since A and B are independent, $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B] = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$.

* * * * *

Problem 4. Two players A and B regularly play each other at chess. When A has the first move in a game, the probability of A winning that game is 0.4 and the probability of B winning that game is 0.2. When B has the first move in a game, the probability of B winning that game is 0.3 and the probability of A winning that game is 0.2. Any game of chess that is not won by either player ends in a draw.

(a) Given that A and B toss a fair coin to decide who has the first move in a game, find the probability of the game ending in a draw.

(b) To make their games more enjoyable, A and B agree to change the procedure for deciding who has the first move in a game. As a result of their new procedure, the probability of A having the first move in any game is p. Find the value of p which gives A and B equal chances of winning each game.

Solution.

Part (a).

$$\begin{aligned} \mathbb{P}[\text{draw}] &= \mathbb{P}[A \text{ first}]\,\mathbb{P}[\text{draw} \mid A \text{ first}] + \mathbb{P}[B \text{ first}]\,\mathbb{P}[\text{draw} \mid B \text{ first}] \\ &= 0.5 \cdot (1 - 0.4 - 0.2) + 0.5 \cdot (1 - 0.3 - 0.2) = 0.45. \end{aligned}$$

Part (b). Observe that

$$\begin{aligned} \mathbb{P}[A \text{ wins}] &= \mathbb{P}[A \text{ first}]\,\mathbb{P}[A \text{ wins} \mid A \text{ first}] + \mathbb{P}[B \text{ first}]\,\mathbb{P}[A \text{ wins} \mid B \text{ first}] \\ &= p \cdot 0.4 + (1-p) \cdot 0.2 = 0.2p + 0.2 \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}[B \text{ wins}] &= \mathbb{P}[A \text{ first}]\,\mathbb{P}[B \text{ wins} \mid A \text{ first}] + \mathbb{P}[B \text{ first}]\,\mathbb{P}[B \text{ wins} \mid B \text{ first}] \\ &= p \cdot 0.2 + (1-p) \cdot 0.3 = -0.1p + 0.3 \end{aligned}$$

Consider $\mathbb{P}[A \text{ wins}] = \mathbb{P}[B \text{ wins}]$. Then $0.2p + 0.2 = -0.1p + 0.3 \implies p = \frac{1}{3}$.

Problem 5. Two fair dices are thrown, and events A, B and C are defined as follows:

- A: the sum of the two scores is odd,
- B: at least one of the two scores is greater than 4,
- C: the two scores are equal.

Find, showing your reasons clearly in each case, which two of these three events are

(a) mutually exclusive,

(b) independent.

Find also $\mathbb{P}[C \mid B]$, making your method clear.

Solution.

Part (a). Let the scores of the first and second die be p and q respectively. Suppose A occurs. Then p and q are of different parities (e.g. p even $\implies$ q odd). Thus, p and q cannot be equal. Hence, C cannot occur, whence A and C are mutually exclusive.

Part (b). Let the scores of the first and second die be p and q respectively. Observe that p is independent of q, and vice versa. Hence, the parity of q is not affected by the parity of p. Thus, $\mathbb{P}[A] = \mathbb{P}[p \text{ even}]\,\mathbb{P}[q \text{ odd}] + \mathbb{P}[p \text{ odd}]\,\mathbb{P}[q \text{ even}] = \frac{3}{6} \cdot \frac{3}{6} + \frac{3}{6} \cdot \frac{3}{6} = \frac{1}{2}$.

We also have $\mathbb{P}[B] = 1 - \mathbb{P}[\text{neither } p \text{ nor } q \text{ is greater than } 4] = 1 - \left(\frac{4}{6}\right)^2 = \frac{20}{36}$.

$p \backslash q$	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

We now consider $\mathbb{P}(A \cap B)$. From the table of outcomes above, it is clear that $\mathbb{P}(A \cap B) = \frac{10}{36} = \mathbb{P}[A]\,\mathbb{P}[B]$. Hence, A and B are independent.

* * * * *

Problem 6. For events A and B, it is given that $\mathbb{P}[A] = 0.7$, $\mathbb{P}[B] = 0.6$ and $\mathbb{P}[A \mid B'] = 0.8$. Find

(a) $\mathbb{P}[A \cap B']$,

(b) $\mathbb{P}[A \cup B]$,

(c) $\mathbb{P}[B' \mid A]$.

For a third event C, it is given that $\mathbb{P}[C] = 0.5$ and that A and C are independent.

(d) Find $\mathbb{P}[A' \cap C]$.

(e) Hence, find an inequality satisfied by $\mathbb{P}[A' \cap B \cap C]$ in the form

$$p \le \mathbb{P}\big[A' \cap B \cap C\big] \le q,$$

where p and q are constants to be determined.

Solution.

Part (a).

$$\mathbb{P}\big[A \cap B'\big] = \mathbb{P}\big[B'\big]\,\mathbb{P}\big[A \mid B'\big] = (1-0.6)\cdot 0.8 = 0.32.$$

Part (b).

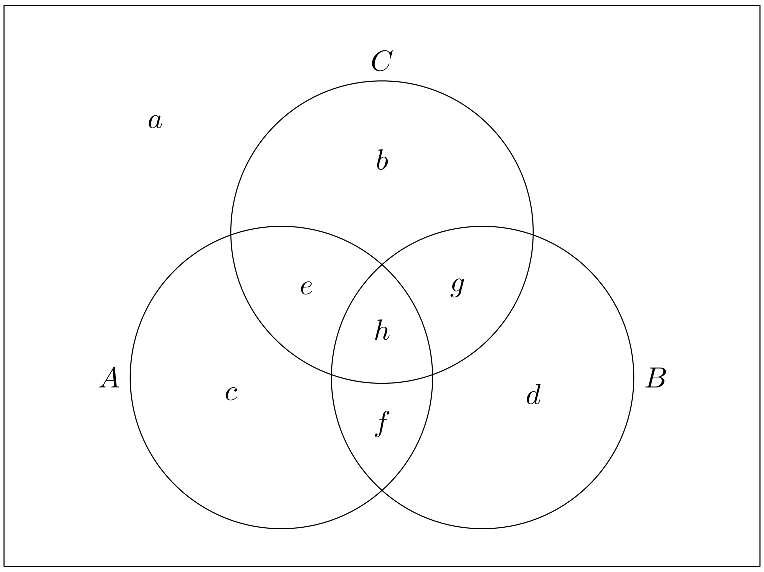
$$\begin{aligned}\mathbb{P}[A \cup B] &= \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B] = \mathbb{P}[A] + \mathbb{P}[B] - \big[\mathbb{P}[A] - \mathbb{P}\big[A \cap B'\big]\big]\\ &= 0.7 + 0.6 - (0.7 - 0.32) = 0.92.\end{aligned}$$

Part (c).

$$\mathbb{P}\big[B' \mid A\big] = \frac{\mathbb{P}[B' \cap A]}{\mathbb{P}[A]} = \frac{0.32}{0.7} = \frac{16}{35}.$$

Part (d). Since A and C are independent, $\mathbb{P}[A \cap C] = \mathbb{P}[A]\,\mathbb{P}[C]$. Hence, $\mathbb{P}[A' \cap C] = \mathbb{P}[C] - \mathbb{P}[A \cap C] = 0.5 - 0.7 \cdot 0.5 = 0.15$.

Part (e). Consider the following Venn diagram.



Note that $\mathbb{P}[A' \cap B \cap C] = g$. Firstly, from part (d), we have $b + g = \mathbb{P}[A' \cap C] = 0.15$. Hence, $g \leq 0.15$. Secondly, from part (b), we have $a + b = 1 - \mathbb{P}[A \cup B] = 1 - 0.92 = 0.08$. Hence, $b \leq 0.08 \implies g \geq 0.07$. Lastly, we know that $\mathbb{P}[A' \cap B] = \mathbb{P}[A \cup B] - \mathbb{P}[A] = 0.92 - 0.7 = 0.22$. Hence, $d + g = 0.22 \implies g \leq 0.22$.

Thus, $0.07 \leq g \leq 0.15$, whence $0.07 \leq \mathbb{P}[A' \cap B \cap C] \leq 0.15$.

* * * * *

Problem 7. Camera lenses are made by two companies, A and B. 60% of all lenses are made by A and the remaining 40% by B. 5% of the lenses made by A are faulty. 7% of the lenses made by B are faulty.

(a) One lens is selected at random. Find the probability that

  (i) it is faulty,

  (ii) it was made by A, given that it is faulty.

(b) Two lenses are selected at random. Find the probability that both were made by A, given that exactly one is faulty.

(c) Ten lenses are selected at random. Find the probability that exactly two of them are faulty.

Solution.

Part (a).

Part (a)(i).

$$\mathbb{P}[\text{faulty}] = \mathbb{P}[A \cup \text{faulty}] + \mathbb{P}[B \cup \text{faulty}] = 0.6 \cdot 0.05 + 0.4 \cdot 0.07 = 0.058.$$

Part (a)(ii).

$$\mathbb{P}[A \mid \text{faulty}] = \frac{\mathbb{P}[A \cap \text{faulty}]}{\mathbb{P}[\text{faulty}]} = \frac{0.6 \cdot 0.05}{0.058} = \frac{15}{19}.$$

Part (b).

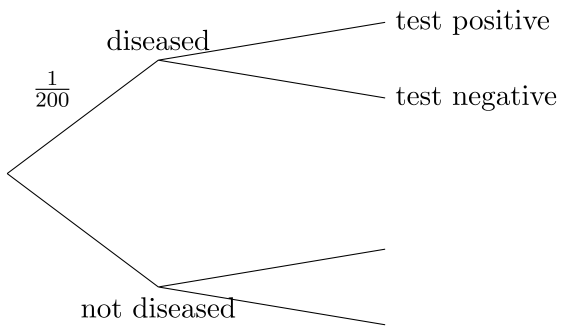
$$\mathbb{P}[\text{both } A \mid \text{one faulty}] = \frac{\mathbb{P}[\text{both } A \cup \text{one faulty}]}{\mathbb{P}[\text{one faulty}]} = \frac{[0.6 \cdot 0.05] \cdot [0.6 \cdot (1 - 0.05)]}{0.058 \cdot (1 - 0.058)} = \frac{1425}{4553}.$$

Part (c).

$$\mathbb{P}[\text{two faulty}] = 0.058^2(1 - 0.058)^8 \cdot \frac{10!}{2!8!} = 0.0939 \text{ (3 s.f.)}$$

* * * * *

Problem 8. A certain disease is present in 1 in 200 of the population. In a mass screening programme a quick test of the disease is used, but the test is not totally reliable. For someone who does have the disease there is a probability of 0.9 that the test will prove positive, whereas for someone who does not have the disease there is a probability of 0.02 that the test will prove positive.



(a) One person is selected at random and test.

(i) Copy and complete the tree diagram, which illustrates one application of the test.

(ii) Find the probability that the person has the disease and the test is positive.

(iii) Find the probability that the test is negative.

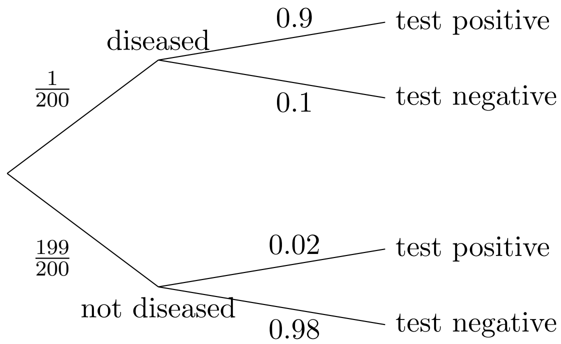
(iv) Given that the test is positive, find the probability that the person has the disease.

(b) People for whom the test proves positive are recalled and re-tested. Find the probability that a person has the disease if the second test also proves positive.

Solution.

Part (a).

Part (a)(i).



Part (a)(ii).

$$\mathbb{P}[\text{diseased} \cap \text{positive}] = \frac{1}{200} \cdot 0.9 = 0.0045.$$

Part (a)(iii).

$$\mathbb{P}[\text{negative}] = \frac{1}{200} \cdot 0.1 + \frac{199}{200} \cdot 0.98 = 0.9756.$$

Part (a)(iv).

$$\mathbb{P}[\text{diseased} \mid \text{positive}] = \frac{\mathbb{P}[\text{diseased} \cap \text{positive}]}{\mathbb{P}[\text{positive}]} = \frac{0.0045}{1 - 0.9756} = 0.184.$$

Part (b).

$$\begin{aligned}\text{Required probability} &= \frac{\mathbb{P}[\text{diseased} \cap \text{both positive}]}{\mathbb{P}[\text{both positive}]} \\ &= \frac{\mathbb{P}[\text{diseased} \cap \text{both positive}]}{\mathbb{P}[\text{diseased} \cap \text{both positive}] + \mathbb{P}[\text{not diseased} \cap \text{both positive}]} \\ &= \frac{1/200 \cdot 0.9^2}{1/200 \cdot 0.9^2 + 199/200 \cdot 0.02^2} = \frac{2025}{2224}.\end{aligned}$$

* * * * *

Problem 9. In a probability experiment, three containers have the following contents.

- A jar contains 2 white dice and 3 black dice.
- A white box contains 5 red balls and 3 green balls.
- A black box contains 4 red balls and 3 green balls.

One die is taken at random from the jar. If the die is white, two balls are taken from the white box, at random and without replacement. If the die is black, two balls are taken from the black box, at random and without replacement. Events W and M are defined as follows:

- W: A white die is taken from the jar.
- M: One red ball and one green ball are obtained.

Show that $\mathbb{P}[M \mid W] = \frac{15}{28}$.

Find, giving each of your answers as an exact fraction in its lowest terms,

(a) $\mathbb{P}[M \cap W]$,

(b) $\mathbb{P}[W \mid M]$,

(c) $\mathbb{P}[W \cup M]$.

All the dice and balls are now placed in a single container, and four objects are taken at random, each object being replaced before the next one is taken. Find the probability that one object of each colour is obtained.

Solution. Since W has occurred, both red and green balls must come from the white box. Note that there are two ways for M to occur: first a red then a green, or first a green then a red. Hence, $\mathbb{P}[M \mid W] = \frac{5}{8} \cdot \frac{3}{7} + \frac{3}{8} \cdot \frac{5}{7} = \frac{15}{28}$ as desired.

Part (a).

$$\mathbb{P}[M \cap W] = \mathbb{P}[W]\,\mathbb{P}[M \mid W] = \frac{2}{5} \cdot \frac{15}{28} = \frac{3}{14}.$$

Part (b). Let B represent the event that a black die is taken from the jar. Then

$$\begin{aligned}\mathbb{P}[M] &= \mathbb{P}[M \cap W] + \mathbb{P}[M \cap B] = \mathbb{P}[M \cap W] + \mathbb{P}[B]\,\mathbb{P}[M \mid B] \\ &= \frac{3}{14} + \frac{3}{5}\left(\frac{4}{7} \cdot \frac{3}{6} + \frac{3}{7} \cdot \frac{4}{6}\right) = \frac{39}{70}.\end{aligned}$$

Hence, $\mathbb{P}[W \mid M] = \frac{\mathbb{P}[W \cap M]}{\mathbb{P}[M]} = \frac{3/14}{39/70} = \frac{5}{13}$.

Part (c).

$$\mathbb{P}[W \cup M] = \mathbb{P}[W] + \mathbb{P}[M] - \mathbb{P}[W \cap M] = \frac{2}{5} + \frac{39}{70} - \frac{3}{14} = \frac{26}{35}.$$

Note that the container has 2 white objects, 3 black objects, 9 red objects and 6 green objects, for a total of 20 objects. The probability that one object of each colour is taken is thus given by

$$\frac{2}{20} \cdot \frac{3}{20} \cdot \frac{9}{20} \cdot \frac{6}{20} \cdot 4! = \frac{243}{5000}.$$

* * * * *

Problem 10. A man writes 5 letters, one each to A, B, C, D and E. Each letter is placed in a separate envelope and sealed. He then addresses the envelopes, at random, one each to A, B, C, D and E.

(a) Find the probability that the letter to A is in the correct envelope and the letter to B is in an incorrect envelope.

(b) Find the probability that the letter to A is in the correct envelope, given that the letter to B is in an incorrect envelope.

(c) Find the probability that both the letters to A and B are in incorrect envelopes.

Solution.

Part (a).

$$\mathbb{P}[A \text{ correct} \cap B \text{ incorrect}] = \frac{1}{5} \times \frac{3}{4} = \frac{3}{20}.$$

Part (b).

$$\mathbb{P}[A \text{ correct} \mid B \text{ incorrect}] = \frac{\mathbb{P}[A \text{ correct} \cap B \text{ incorrect}]}{\mathbb{P}[B \text{ incorrect}]} = \frac{3/20}{4/5} = \frac{3}{16}.$$

Part (c).

$$\begin{aligned}\mathbb{P}[A \text{ incorrect} \cap B \text{ incorrect}] &= \mathbb{P}[B \text{ incorrect}]\, \mathbb{P}[A \text{ incorrect} \mid B \text{ incorrect}] \\ &= \frac{4}{5}\left(1 - \frac{3}{16}\right) = \frac{13}{20}.\end{aligned}$$

* * * * *

Problem 11. A bag contains 4 red counters and 6 green counters. Four counters are drawn at random from the bag, without replacement. Calculate the probability that

(a) all the counters drawn are green,

(b) at least one counter of each colour is drawn,

(c) at least two green counters are drawn,

(d) at least two green counters are drawn, given that at least one counter of each colour is drawn.

State with a reason whether the events "at least two green counters are drawn" and "at least one counter of each colour is drawn" are independent.

Solution.

Part (a).

$$\mathbb{P}[\text{all green}] = \frac{{}^6C_4}{10!/(4!\,6!)} = \frac{1}{14}.$$

Part (b).

$$\mathbb{P}[\text{one of each colour}] = 1 - \mathbb{P}[\text{all green}] - \mathbb{P}[\text{all red}] = 1 - \frac{1}{14} - \frac{{}^4C_4}{10!/(4!\,6!)} = \frac{97}{105}.$$

Part (c).

$$\mathbb{P}[\text{at least 2 green}] = 1 - \mathbb{P}[\text{no green}] - \mathbb{P}[\text{one green}] = 1 - \frac{1}{210} - \frac{{}^6C_1 \cdot {}^4C_3}{10!/(4!\,6!)} = \frac{37}{42}.$$

Part (d).

$$\mathbb{P}[\text{at least 2 green} \mid \text{one of each colour}] = \frac{{}^6C_3 \cdot {}^4C_1 + {}^6C_2 \cdot {}^4C_2}{10!/(4!\,6!) - {}^6C_4 - {}^4C_4} = \frac{85}{97}.$$

Since $\mathbb{P}[\text{at least 2 green}] = \frac{37}{42} \neq \frac{85}{97} = \mathbb{P}[\text{at least 2 green} \mid \text{one of each colour}]$, the two events are not independent.

Problem 12. A group of fifteen people consists of one pair of sisters, one set of three brothers and ten other people. The fifteen people are arranged randomly in a line.

(a) Find the probability that the sisters are next to each other.

(b) Find the probability that the brother are not all next to one another.

(c) Find the probability that either the sisters are next to each other or the brothers are all next to one another or both.

(d) Find the probability that the sisters are next to each other given that the brothers are not all next to one another.

Solution.

Part (a). Let the two sisters be one unit. There are hence 14 units altogether, giving $14! \cdot 2!$ arrangements with the restriction. Since there are a total of 15! arrangements without the restriction, the required probability is $\frac{14! \cdot 2!}{15!} = \frac{2}{15}$.

Part (b). Consider the case where all brothers are next to one another. Counting the brothers as one unit gives 13 units altogether. There are hence $13! \cdot 3!$ arrangements with this restriction. Since there are a total of 15! arrangements without the restriction, the probability that all three brothers are not together is given by $\frac{13! \cdot 3!}{15!} = \frac{34}{35}$.

Part (c). Consider the case where both the sisters are adjacent, and all three brothers are next to one another. Counting the sisters as one unit, and counting the brothers as one unit gives 12 units altogether. There are hence $12! \cdot 2! \cdot 3!$ arrangements with this restriction. Since there are a total of 15! arrangements without the restriction, we have

$$\mathbb{P}[\text{sisters together} \cap \text{brothers together}] = \frac{12! \cdot 2! \cdot 3!}{15!} = \frac{2}{455}.$$

Hence,

$$\begin{aligned} &\mathbb{P}[\text{sisters together} \cup \text{brothers together}] \\ &= \mathbb{P}[\text{sisters together}] + \mathbb{P}[\text{brothers together}] - \mathbb{P}[\text{sisters together} \cap \text{brothers together}] \\ &= \frac{2}{15} + \left(1 - \frac{1}{35}\right) - \frac{2}{455} = \frac{43}{273}. \end{aligned}$$

Part (d). Note that

$$\begin{aligned} &\mathbb{P}[\text{sisters together} \cap \text{brothers not together}] \\ &= \mathbb{P}[\text{sisters together}] - \mathbb{P}[\text{sisters together} \cap \text{brothers together}] \\ &= \frac{2}{15} - \frac{2}{455} = \frac{176}{1365}. \end{aligned}$$

Hence, the required probability can be calculated as

$$\begin{aligned} \mathbb{P}[\text{sisters together} \mid \text{brothers not together}] &= \frac{\mathbb{P}[\text{sisters together} \cap \text{brothers not together}]}{\mathbb{P}[\text{brothers not together}]} \\ &= \frac{176/1365}{34/35} = \frac{88}{663}. \end{aligned}$$

Self-Practice A12

Problem 1. Two events A and B are such that $\mathbb{P}[A] = 0.6$, $\mathbb{P}[B] = 0.3$, $\mathbb{P}[A \mid B] = 0.2$. Calculate the probabilities that

(a) both events occur,

(b) at least one of the two events occurs,

(c) exactly one of the events occur.

Solution.

Part (a). We have

$$\mathbb{P}[A \cap B] = \mathbb{P}[B]\,\mathbb{P}[A \mid B] = (0.2)(0.3) = 0.06.$$

Part (b). We have

$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B] = 0.6 + 0.3 - 0.06 = 0.84.$$

Part (c). The required probability is given by

$$\mathbb{P}[A \cup B] - \mathbb{P}[A \cap B] = 0.84 - 0.06 = 0.78.$$

* * * * *

Problem 2. For events A and B, it is given that $\mathbb{P}[A] = 0.7$, $\mathbb{P}[B \mid A'] = 0.8$, $\mathbb{P}[A \mid B'] = 0.88$. Find

(a) $\mathbb{P}[B \cap A']$,

(b) $\mathbb{P}[A' \cap B']$,

(c) $\mathbb{P}[A \cap B]$.

Solution.

Part (a). We have

$$\mathbb{P}\big[B \cap A'\big] = \mathbb{P}\big[A'\big]\,\mathbb{P}\big[B \mid A'\big] = (1 - 0.7)\,(0.8) = 0.24.$$

Part (b). Note that $\mathbb{P}[B' \mid A'] = 1 - \mathbb{P}[B \mid A']$. Hence,

$$\mathbb{P}\big[B' \cap A'\big] = \mathbb{P}\big[A'\big]\,\mathbb{P}\big[B' \mid A'\big] = (1 - 0.7)\,(1 - 0.8) = 0.06.$$

Part (c). Let $x = \mathbb{P}[A \cap B]$. Then

$$\mathbb{P}\big[A \cap B'\big] = \mathbb{P}[A] - x = 0.7 - x,$$

so

$$0.88 = \mathbb{P}\big[A \mid B'\big] = \frac{\mathbb{P}[A \cap B']}{\mathbb{P}[B']} = \frac{\mathbb{P}[A \cap B']}{\mathbb{P}[A \cap B'] + \mathbb{P}[A' \cap B']} = \frac{0.7 - x}{(0.7 - x) + 0.06},$$

which yields $x = \mathbb{P}[A \cap B] = 0.26$ upon simplification.

* * * * *

Problem 3. A group of student representatives is to be chosen from three schools, R, S and T. The group is to consist of 10 students and is chosen from a set of 15 students consisting of 3 from R, 4 from S and 8 from T. Find the probability that the group consists of

(a) students from S and T only,

(b) at least one student from each school.

Solution.

Part (a). There are ${}^{15}C_{10}$ ways to form a group without restriction, and there are ${}^{4+8}C_{10}$ ways to form a group consisting of students from only S and T. Thus, the desired probability is ${}^{4+8}C_{10}/\,{}^{15}C_{10} = 2/91$.

Part (b). Consider the complement, i.e. the event that at least one school has no representative. The only way this can happen is if the group consists of students from R and T only, or from S and T only. Thus, the required probability is

$$1 - \frac{{}^{3+8}C_{10} + {}^{4+8}C_{10}}{{}^{15}C_{10}} = \frac{38}{39}.$$

* * * * *

Problem 4. A box contains 25 apples, of which 20 are red and 5 are green. Of the red apples, 3 contain maggots and of the green apples, 1 contains maggots. Two apples are chosen at random from the box. Find, in any order,

(a) the probability that both apples contain maggots.

(b) the probability that both apples are red and at least one contains maggots.

(c) the probability that at least one apple contains maggots, given that both apples are red.

(d) the probability that both apples are red given that at least one apple is red.

Solution.

Part (a). The required probability is

$$\frac{4}{25} \times \frac{3}{24} = \frac{1}{50}.$$

Part (b). The required probability is

$$\mathbb{P}[\text{both red}] - \mathbb{P}[\text{both red and no maggot}] = \frac{20}{25} \times \frac{19}{24} - \frac{17}{25} \times \frac{16}{24} = \frac{9}{50}.$$

Part (c). The required probability is

$$\frac{\mathbb{P}[\text{both red and at least one maggot}]}{\mathbb{P}[\text{both red}]} = \frac{9/50}{(20/25) \times (19/24)} = \frac{27}{95}.$$

Part (d). The required probability is

$$\frac{\mathbb{P}[\text{both red}]}{\mathbb{P}[\text{at least one red}]} = \frac{\mathbb{P}[\text{both red}]}{1 - \mathbb{P}[\text{both green}]} = \frac{(20/25) \times (19/24)}{1 - (5/25)(4/24)} = \frac{19}{29}.$$

* * * * *

Problem 5. A bag contains 15 tokens that are indistinguishable apart from their colours. 2 of the tokens are blue and the rest are either red or green. Participants are required to draw the tokens randomly, one at a time, from the bag without replacement.

(a) Given that the probability that a participant draws 2 red tokens on the first 2 draws is 1/35, show that there are 3 red tokens in the bag.

(b) Find the probability that a participant draws a red or green token on the second draw.

Events A and B are defined as follows.

- A: A participant draws his/her second red token on the third draw.
- B: A participant draws a blue token on the second draw.

(c) Find $\mathbb{P}[A \cup B]$.

(d) Determine if A and B are independent events.

Solution.

Part (a). Let r be the number of red tokens. Then

$$\mathbb{P}[2 \text{ red tokens on second draw}] = \frac{r}{15} \times \frac{r-1}{14} = \frac{1}{35} \implies r^2 - r - 6 = 0.$$

Solving, we get $r = 3$. Note that we reject $r = -2$ since $r \geq 0$.

Part (b). We have

$$\mathbb{P}[\text{red/green on 2nd draw}] = \frac{13}{15} \times \frac{12}{14} + \frac{2}{15} \times \frac{13}{14} = \frac{13}{15}.$$

Part (c). We have

$$\mathbb{P}[A] = \frac{3}{15} \times \frac{12}{14} \times \frac{2}{13} + \frac{12}{15} \times \frac{3}{14} \times \frac{2}{13} = \frac{24}{455}$$

and

$$\mathbb{P}[B] = 1 - \mathbb{P}[\text{red/green on 2nd draw}] = 1 - \frac{13}{15} = \frac{2}{15}.$$

Note that the event $A \cap B$ can only occur if the first and third draws are red, and the second draw is blue. Thus,

$$\mathbb{P}[A \cap B] = \frac{3}{15} \times \frac{2}{14} \times \frac{2}{13} = \frac{2}{455}.$$

Thus,

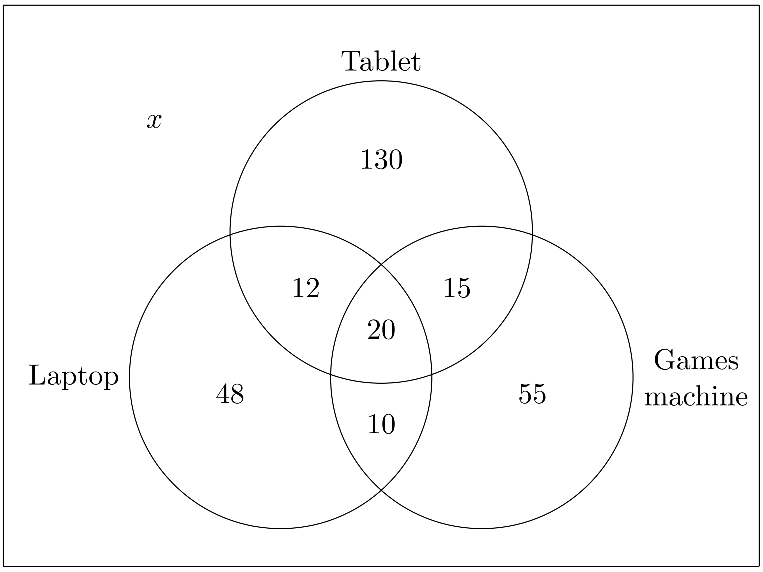
$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B] = \frac{24}{455} + \frac{2}{15} - \frac{2}{455} = \frac{248}{1365}.$$

Part (d). Observe that

$$\mathbb{P}[A]\,\mathbb{P}[B] = \frac{24}{455} \times \frac{2}{15} = \frac{16}{2275} \neq \frac{248}{1365} = \mathbb{P}[A \cap B]\,.$$

Thus, A and B are not independent.

Problem 6.



A group of students is asked whether they own any of a laptop, a tablet and a games machine. The numbers owning different combinations are shown in the Venn diagram. The number of students owning none of these is x. One of the students is chosen at random.

- L is the event that the student owns a laptop.
- T is the event that the student owns a tablet.
- G is the event that the student owns a game machine.

(a) Write down expressions for $\mathbb{P}[L]$ and $\mathbb{P}[G]$ in terms of x. Given that L and G are independent, show that $x = 10$.

Using this value of x, find

(b) $\mathbb{P}[L \cup T]$,

(c) $\mathbb{P}[T \cap G']$,

(d) $\mathbb{P}[L \mid G]$.

Two students from the whole group are chosen at random.

(e) Find the probability that both of these students each owns exactly two out of the three items (laptop, tablet, games machine).

Solution.

Part (a). We have

$$\mathbb{P}[L] = \frac{90}{290+x} \quad \text{and} \quad \mathbb{P}[G] = \frac{100}{290+x} \quad \text{and} \quad \mathbb{P}[L \cap G] = \frac{30}{290+x}.$$

Since L and G are independent,

$$\mathbb{P}[L]\,\mathbb{P}[G] = \mathbb{P}[L \cap G] \implies \frac{90}{290+x} \times \frac{100}{290+x} = \frac{30}{290+x}.$$

Clearing denominators and simplifying, we get

$$(290+x)^2 - 300\,(290+x) = 0,$$

which implies $290 + x = 300$, whence $x = 10$.

Part (b). We have

$$\mathbb{P}[L \cup T] = \frac{235}{300} = \frac{47}{60}.$$

Part (c). We have

$$\mathbb{P}\left[T \cap G'\right] = \frac{143}{300} = \frac{71}{150}.$$

Part (d). We have

$$\mathbb{P}[L \mid G] = \frac{\mathbb{P}[L \cap G]}{\mathbb{P}[G]} = \frac{30/300}{100/300} = \frac{3}{10}.$$

Part (e). Observe that the number of students that own exactly two items is $12+15+10 = 37$. The required probability is hence

$$\frac{37}{300} \times \frac{36}{299} = \frac{111}{7475}.$$

* * * * *

Problem 7. A group of students takes an examination in Science. A student who fails the examination at the first attempt is allowed one further attempt. For a randomly chosen student, the probability of passing the examination at the first attempt is p. If the student fails the examination at the first attempt, the probability of passing at the second attempt is 0.3 more than the probability of passing the examination at the first attempt.

(a) Show that the probability that a randomly chosen student passes the examination is $0.3 + 1.7p - p^2$.

Find the value of p such that the probability that a randomly chosen student passes the examination on the first attempt given that the student passes is 0.6.

Two students are randomly chosen.

(b) (i) Find the probability that one passes the examination on the first attempt and the other passes the examination on the second attempt, leaving your answer in terms of p.

(ii) Find the value of p such that the value of the probability in part (i) is maximum.

Solution.

Part (a). We have

$$\mathbb{P}[\text{pass}] = p + (1-p)(p+0.3) = -p^2 + 1.7p + 0.3.$$

Part (b). We have

$$\mathbb{P}[\text{pass on 1st attempt} \mid \text{pass}] = \frac{\mathbb{P}[\text{pass on 1st attempt}]}{\mathbb{P}[\text{pass}]} = \frac{p}{-p^2 + 1.7p + 0.3}.$$

Equating this to 0.6, we get $p = 0.565$ or $p = -0.531$, which we reject since $p \in [0, 1]$. Thus, the desired probability is $p = 0.565$.

Part (c). The required probability is

$$2\left[p \times (1-p)(p+0.3)\right] = -2p^3 + 1.4p^2 + 0.6p.$$

Part (d). Let $f(p) = -2p^3 + 1.4p^2 + 0.6p$. For stationary points,

$$\frac{\mathrm{d}f}{\mathrm{d}p} = -6p^2 + 2.8p + 0.6 = 0,$$

which occurs when $p = 0.626$. Note that we reject $p = -0.160$ since $p \in [0,1]$. At $p = 0.626$, the second derivative is

$$\left.\frac{\mathrm{d}^2 f}{\mathrm{d}p^2}\right|_{p=0.626} = \left.(-12p + 2.8)\right|_{p=0.626} = -4.712 < 0,$$

thus the probability is maximum at $p = 0.626$.

* * * * *

Problem 8. In Haha College, 70% of the students watch the show *Jogging Man* and 60% of the students watch the show *Voice of Me*. 40% of those who do not watch the show *Voice of Me* watch the show *Jogging Man*. Find the probability that a student chosen at random from the college

(a) watches both shows,

(b) watches exactly one show,

(c) watches the show *Voice of Me* given that the student does not watch the show *Jogging Man*.

State, with a reason, whether the events 'watches *Jogging Man*' and 'watches *Voice of Me*' are independent.

Solution. Let J be the event that a student watches *Jogging Man*, and V be the event that a student watches *Voice of Me*. We have $\mathbb{P}[J] = 0.7$, $\mathbb{P}[V] = 0.6$ and $\mathbb{P}[J \mid V'] = 0.4$.

Part (a). We have

$$\mathbb{P}\big[J \cap V'\big] = \mathbb{P}\big[V'\big]\,(J)\,V' = (1 - 0.6)\,(0.4) = 0.16.$$

Thus, the probability that the student watches both shows is

$$\mathbb{P}[J \cap V] = \mathbb{P}[J] - \mathbb{P}\big[J \cap V'\big] = 0.7 - 0.16 = 0.54.$$

Part (b). The probability that the student watches exactly one show is

$$\mathbb{P}[J \cup V] - \mathbb{P}[J \cap V] = (\mathbb{P}[J] + \mathbb{P}[V] - \mathbb{P}[J \cap V]) - \mathbb{P}[J \cap V] = (0.7 + 0.6 - 0.54) - 0.54 = 0.22.$$

Part (c). The desired probability is

$$\mathbb{P}\big[V \mid J'\big] = \frac{\mathbb{P}[V \cap J']}{\mathbb{P}[J']} = \frac{\mathbb{P}[V] - \mathbb{P}[V \cap J]}{1 - \mathbb{P}[J]} = \frac{0.6 - 0.54}{1 - 0.7} = 0.2.$$

Note that

$$\mathbb{P}[J \cap V] = 0.54 \neq (0.7)(0.6) = \mathbb{P}[J]\,\mathbb{P}[V]\,,$$

thus J and V are not independent.

* * * * *

Problem 9. For events A and B, it is given that $\mathbb{P}[A] = 2/3$ and $\mathbb{P}[B] = 1/2$.

(a) State an inequality satisfied by $\mathbb{P}[A \cap B]$.

It is given further that A and B are independent. Find

(b) $\mathbb{P}[A \cap B]$,

(c) $\mathbb{P}[A' \cup B]$.

Solution.

Part (a). Note that

$$\mathbb{P}[A \cap B] \leq \min\{\mathbb{P}[A], \mathbb{P}[B]\} = \frac{1}{2}.$$

Further,

$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B] = \frac{7}{6} - \mathbb{P}[A \cap B],$$

so $\mathbb{P}[A \cap B] \geq 1/6$. Putting both inequalities together, we have

$$\frac{1}{6} \leq \mathbb{P}[A \cap B] \leq \frac{1}{2}.$$

Part (b). Since A and B are independent,

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B] = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}.$$

Part (c). We have

$$\mathbb{P}\left[A' \cup B\right] = 1 - \mathbb{P}\left[A \cap B'\right] = 1 - (\mathbb{P}[A] - \mathbb{P}[A \cap B]) = 1 - \left(\frac{2}{3} - \frac{1}{3}\right) = \frac{2}{3}.$$

* * * * *

Problem 10 (🌶). A fast food restaurant gives away a free action figure for every child's meal bought. There are five different action figures and each figure is equally likely to be given away with a child's meal. A customer intends to collect all five different figures by buying child's meals.

(a) Find the probability that the first 4 child's meals bought by the customer all had different action figures.

(b) Two of the five action figures are X and Y. Find the probability that the first 4 action figures obtained result in the customer having at least one X or one Y or both.

(c) Find the probability that the first 4 child's meals bought by the customer had exactly two different action figures.

(d) At a certain stage, the customer collected 4 of the five action figures. Given that the probability of the customer completing the set by at most n meals is larger than 0.95, find the least value of n.

Solution.

Part (a). Without restriction, there are 5^4 ways to get action figures from four orders. If all four action figures are distinct, there are only ${}^5P_4$ ways to do so. The required probability is thus ${}^5P_4/5^4 = 24/125$.

Part (b). Consider the complement, i.e. the event that the first 4 action figures do not contain any X or Y. Since there are now only three possible action figures available, there are 3^4 ways for this to happen. The required probability is then

$$1 - \frac{3^4}{5^4} = \frac{544}{625}.$$

Part (c). Let X and Y be the pair of action figures that the customer obtained. There are ${}^5C_2$ possible pairs. Since there are only two possible action figures available, there are

2^4 ways to obtain at most two different action figures. Since there are 2 ways to get only one action figure (either all four orders are X's or Y's), the desired probability is

$$\frac{{}^5C_2\left(2^4-2\right)}{5^4}=\frac{28}{125}.$$

Part (d). Consider the complement, i.e. the event that the customer does not complete his set by at most n meals. This probability is given by $(4/5)^n$, so the desired inequality is

$$1-\left(\frac{4}{5}\right)^n \geq 0.95.$$

Using G.C., the least n is 14.

Assignment A12

Problem 1.

(a) Events A and B are such that $\mathbb{P}[A] = 0.4$, $\mathbb{P}[B] = 0.3$ and $\mathbb{P}[A \cup B] = 0.5$.

(i) Determine whether A and B are mutually exclusive.

(ii) Determine whether A and B are independent.

(b) In a competition, 2 teams (A and B) will play each other in the best of 3 games. That is, the first team to win 2 games will be the winner and the competition will end. In the first game, both teams have equal chances of winning. In subsequent games, the probability of team A winning team B given that team A won in the previous game is p and the probability of team A winning team B given that team A lost in the previous game is $\frac{1}{3}$.

(i) Illustrate the information with an appropriate tree diagram.

(ii) Find the value of p such that team A has equal chances of winning and losing the competition.

Solution.

Part (a).

Part (a)(i). Note that

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cup B] = 0.4 + 0.3 - 0.5 = 0.2.$$

Since $\mathbb{P}[A \cap B] = 0.2 \neq 0$, A and B are not mutually exclusive.

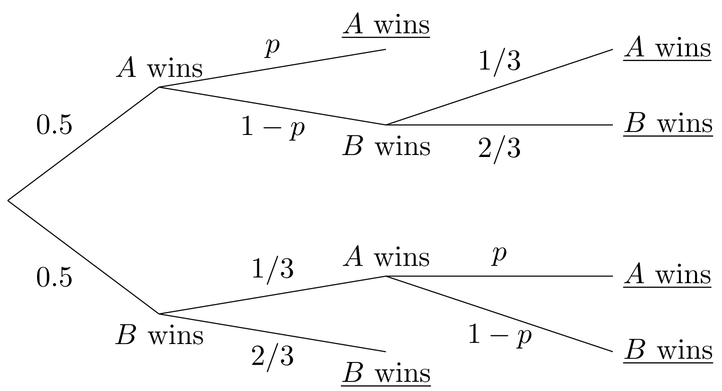
Part (a)(ii). Note that

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{0.2}{0.3} = \frac{2}{3}.$$

Since $\mathbb{P}[A] = 0.4 \neq \frac{2}{3} = \mathbb{P}[A \mid B]$, A and B are not independent.

Part (b).

Part (b)(i).



Part (b)(ii). Consider

$$\mathbb{P}[A \text{ wins competition}] = \left[\frac{1}{2} \cdot p\right] + \left[\frac{1}{2} \cdot (1-p) \cdot \frac{1}{3}\right] + \left[\frac{1}{2} \cdot \frac{1}{3} \cdot p\right] = \frac{p}{2} + \frac{1}{6} = \frac{1}{2}.$$

We hence need $p = \frac{2}{3}$ for A to have equal chances of winning and losing.

Problem 2. A Personal Identification Number (PIN) consists of 4 digits in order, where each digit ranges from 0 to 9. Susie has difficulty remembering her PIN. She tries to remember her PIN and writes down what she thinks it is. The probability that the first digit is correct is 0.8 and the probability that the second digit is correct is 0.86. The probability that the first two digits are both correct is 0.72. Find

(a) the probability that the second digit is correct given that the first digit is correct,

(b) the probability that the first digit is correct, and the second digit is incorrect,

(c) the probability that the second digit is incorrect given that the first digit is incorrect.

Solution. Let $1D$ be the event that the first digit is correct, and $2D$ be the event that the second digit is correct. We have $\mathbb{P}[1D] = 0.8$, $\mathbb{P}[2D] = 0.86$, and $\mathbb{P}[1D \cap 2D] = 0.72$.

Part (a).

$$\mathbb{P}[2D \mid 1D] = \frac{\mathbb{P}[2D \cap 1D]}{\mathbb{P}[1D]} = \frac{0.72}{0.8} = 0.9.$$

Part (b).

$$\mathbb{P}\big[1D \cap 2D'\big] = \mathbb{P}[1D] - \mathbb{P}[1D \cap 2D] = 0.8 - 0.72 = 0.08.$$

Part (c).

$$\mathbb{P}\big[2D' \mid 1D'\big] = \frac{\mathbb{P}[2D' \cap 1D']}{\mathbb{P}[1D']} = \frac{1 - \mathbb{P}[1D \cup 2D]}{1 - \mathbb{P}[1D]}$$
$$= \frac{1 - [\mathbb{P}[1D] + \mathbb{P}[2D] - \mathbb{P}[1D \cap 2D]]}{1 - \mathbb{P}[1D]} = \frac{1 - (0.8 + 0.86 - 0.72)}{1 - 0.8} = 0.3.$$

* * * * *

Problem 3. An international tour group consists of the following seventeen people: a pair of twin sisters and their boyfriends, all from Canada; three policewomen from China; a married couple and their two daughters from Singapore, and a large family from Indonesia, consisting of a man, his wife, his parents and his two sons.

Four people from the group are randomly chosen to play a game. Find the probability that

(a) the four people are all of different nationalities,

(b) the four people are all the same gender,

(c) the four people are all of different nationalities, given that they are all the same gender.

Solution.

TALLY	Male	Female	SUBTOTAL
Canada	2	2	4
China	0	3	3
Singapore	1	3	4
Indonesia	4	2	6
SUBTOTAL	7	10	17

Part (a).

$$\mathbb{P}[\text{all different nationalities}] = \frac{4}{17} \cdot \frac{3}{16} \cdot \frac{4}{15} \cdot \frac{6}{14} \cdot 4! = \frac{72}{595}.$$

Part (b).

$$\mathbb{P}[\text{all same gender}] = \frac{{}^{7}C_4 + {}^{10}C_4}{{}^{17}C_4} = \frac{7}{68}.$$

Part (c).

$$\mathbb{P}[\text{all different nationalities} \mid \text{all female}] = \frac{2}{17} \cdot \frac{3}{16} \cdot \frac{3}{15} \cdot \frac{2}{14} \cdot 4! = \frac{9}{595}$$

Note that $\mathbb{P}[\text{all different nationalities} \mid \text{all male}]$ since there are no males from China, whence

$$\begin{aligned}&\mathbb{P}[\text{all different nationalities} \mid \text{all same gender}]\\ &= \frac{\mathbb{P}[\text{all different nationalities} \cap \text{all same gender}]}{\mathbb{P}[\text{all same gender}]} = \frac{9/595 + 0}{7/68} = \frac{36}{245}.\end{aligned}$$

A14A Discrete Random Variables

Tutorial A14A

Problem 1. An unbiased die is in the form of a regular tetrahedron and has its faces numbered 1, 2, 3, 4. When the die is thrown on to a horizontal table, the number on the fact in contact with the table is noted. Two such dice are thrown and the score X is found by multiplying these numbers together. Obtain the probability distribution of X. Find the values of

(a) $\mathbb{P}[X > 8]$,

(b) $\mathbb{E}[X]$,

(c) $\text{Var}[X]$.

Solution. The following table displays all possible outcomes.

	1	2	3	4
1	1	2	3	4
2	2	4	6	8
3	3	6	9	12
4	4	8	12	16

Hence, the probability distribution is

x	1	2	3	4	6	8	9	12	16
$\mathbb{P}[X = x]$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

Part (a).

$$\mathbb{P}[X > 8] = \mathbb{P}[X = 9] + \mathbb{P}[X = 12] + \mathbb{P}[X = 16] = \frac{1}{16} + \frac{2}{16} + \frac{1}{16} = \frac{1}{4}.$$

Part (b). Using G.C., $\mathbb{E}[X] = 6.25$.

Part (c). Using G.C., $\text{Var}[X] = (4.14578)^2 = 17.2$.

* * * * *

Problem 2. A computer can give independent observations of a random variable X with probability distribution given by $\mathbb{P}[X = 0] = \frac{3}{4}$ and $\mathbb{P}[X = 2] = \frac{1}{4}$. It is programmed to output a value for the random variable Y defined by $Y = X_1 + X_2$, where X_1 and X_2 are two observations of X.

Tabulate the probability distribution of Y and show that $\mathbb{E}[Y] = 1$.

The random variable T is defined by $T = Y^2$. Find $\mathbb{E}[T]$ and show that $\text{Var}[T] = \frac{63}{4}$.

Solution. Quite clearly, we have

$$\mathbb{P}[Y = 0] = \left(\frac{3}{4}\right)^2 = \frac{9}{16}, \quad \mathbb{P}[Y = 2] = 2\left(\frac{3}{4}\right)\left(\frac{1}{4}\right) = \frac{3}{8}, \quad \mathbb{P}[Y = 4] = \left(\frac{1}{4}\right)^2 = \frac{1}{16}.$$

Thus, the probability distribution of Y is given by

y	0	2	4
$\mathbb{P}[Y = y]$	$\frac{9}{16}$	$\frac{3}{8}$	$\frac{1}{16}$

Thus,

$$\mathbb{E}[Y] = 0\left(\frac{9}{16}\right) + 2\left(\frac{3}{8}\right) + 4\left(\frac{1}{16}\right) = 1.$$

Note that $\mathbb{E}[T] = \mathbb{E}\left[Y^2\right]$ and $\mathbb{E}\left[T^2\right] = \mathbb{E}\left[Y^4\right]$. Hence,

$$\mathbb{E}[T] = 0^2\left(\frac{9}{16}\right) + 2^2\left(\frac{3}{8}\right) + 4^2\left(\frac{1}{16}\right) = \frac{5}{2}$$

and

$$\mathbb{E}\left[T^2\right] = 0^4\left(\frac{9}{16}\right) + 2^4\left(\frac{3}{8}\right) + 4^4\left(\frac{1}{16}\right) = 22.$$

Thus,

$$\mathrm{Var}[T] = \mathbb{E}\left[T^2\right] - \mathbb{E}[T]^2 = 22 - \left(\frac{5}{2}\right)^2 = \frac{63}{4}.$$

* * * * *

Problem 3. The discrete random variable X takes values -1, 0, 1 with probabilities $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$ respectively. The variable $\bar{X}$ is the mean of a random sample of 3 values of X (i.e. X_1, X_2 and X_3 are independent random variables).

Tabulate the probability distribution of $\bar{X}$, and use your values to calculate $\mathrm{Var}\left[\bar{X}\right]$. Hence, verify that $\mathrm{Var}\left[\bar{X}\right] = \frac{1}{3}\mathrm{Var}[X]$ in this case.

Solution. By symmetry, we have $\mathbb{P}\left[\bar{X} = -n\right] = \mathbb{P}\left[\bar{X} = n\right]$. Now, notice that the only way to get a total score of 3 is to have $X_1 = X_2 = X_3 = 1$. Thus,

$$\mathbb{P}\left[\bar{X} = 1\right] = \mathbb{P}\left[\bar{X} = -1\right] = \left(\frac{1}{4}\right)^3 = \frac{1}{64}.$$

Similarly, the only way to get a total score of 2 is to have two 1's and one 0. Thus,

$$\mathbb{P}\left[\bar{X} = \frac{2}{3}\right] = \mathbb{P}\left[\bar{X} = -\frac{2}{3}\right] = \binom{3}{1}\left(\frac{1}{4}\right)^2\left(\frac{1}{2}\right) = \frac{3}{32}.$$

Now note that there are two ways to achieve a total score of 1: have two 1's and one -1, or have two 0's and one 1. This gives

$$\mathbb{P}\left[\bar{X} = \frac{1}{3}\right] = \mathbb{P}\left[\bar{X} = -\frac{1}{3}\right] = \binom{3}{1}\left(\frac{1}{4}\right)^2\left(\frac{1}{4}\right) + \binom{3}{1}\left(\frac{1}{2}\right)^2\left(\frac{1}{4}\right) = \frac{15}{64}.$$

Lastly, by the complement principle, we have

$$\mathbb{P}\left[\bar{X} = 0\right] = 1 - 2\left(\frac{1}{64} + \frac{3}{32} + \frac{15}{64}\right) = \frac{5}{16}.$$

Hence, the probability distribution of X is given by

$\bar{x}$	-1	$-\frac{2}{3}$	$-\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$	1
$\mathbb{P}\left[\bar{X} = \bar{x}\right]$	$\frac{1}{64}$	$\frac{3}{32}$	$\frac{15}{64}$	$\frac{5}{16}$	$\frac{15}{64}$	$\frac{3}{32}$	$\frac{1}{64}$

We now calculate $\mathrm{Var}[\bar{X}]$. Observe that the means of X and $\bar{X}$ are 0 by symmetry. Hence,

$$\mathrm{Var}[\bar{X}] = \mathbb{E}[\bar{X}^2] = 2\left[1^2\left(\frac{1}{64}\right) + \left(\frac{2}{3}\right)^2\left(\frac{3}{32}\right) + \left(\frac{1}{3}\right)^2\left(\frac{15}{64}\right)\right] = \frac{1}{6}.$$

Now, note that

$$\mathrm{Var}[X] = \mathbb{E}[X^2] = 2\left[1^2\left(\frac{1}{4}\right)\right] = \frac{1}{2}.$$

Thus,

$$\mathrm{Var}[\bar{X}] = \frac{1}{3}\,\mathrm{Var}[X].$$

* * * * *

Problem 4. The probability of obtaining a head when a particular type of coin is tossed is p. The random variable X is the number of heads obtained when three such coins are tossed.

(a) Draw up a table showing the probability distribution of X.

(b) Prove that $\mathbb{E}\left[\frac{1}{3}X\right] = p$.

(c) Given that $p = \frac{1}{3}$, and denoting by E the event that $X > 1$, find the probability that in 100 throws of the three coins, E will not occur more than 30 times.

Solution.

Part (a). Observe that

$$\mathbb{P}[X = n] = \binom{3}{n} p^n (1-p)^n.$$

Hence, the probability distribution of X is given by

x	0	1	2	3
$\mathbb{P}[X = x]$	$(1-p)^3$	$3p(1-p)^2$	$3p^2(1-p)$	p^3

Part (b). Note that

$$\mathbb{E}[X] = \sum_{n=0}^{3} n\binom{3}{n} p^n (1-p)^{3-n}.$$

Differentiating with respect to p, we get

$$0 = \sum_{n=0}^{3}\binom{3}{n}\left[np^{n-1}(1-p)^{3-n} - (3-n)p^n(1-p)^{3-n-1}\right].$$

Rearranging, we have

$$\left(\frac{1}{p} + \frac{1}{1-p}\right)\underbrace{\sum_{n=0}\binom{3}{n} np^n (1-p)^{3-n}}_{\mathbb{E}[X]} = \frac{3}{1-p}\underbrace{\sum_{n=0}\binom{3}{n} p^n (1-p)^{3-n}}_{1}.$$

Thus,

$$\mathbb{E}[X] = \frac{\frac{3}{1-p}}{\frac{1}{p} + \frac{1}{1-p}} = 3p \implies \mathbb{E}\left[\frac{1}{3}X\right] = \frac{1}{3}\,\mathbb{E}[X] = \frac{1}{3}(3p) = p.$$

Part (c). Note that

$$\mathbb{P}[E] = \mathbb{P}[X > 1] = \mathbb{P}[X = 2] + \mathbb{P}[X = 3] = 3\left(\frac{1}{3}\right)^2\left(1 - \frac{1}{3}\right) + \left(\frac{1}{3}\right)^3 = \frac{7}{27}.$$

Now, observe that

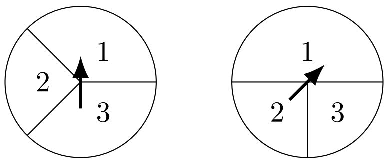
$$\mathbb{P}[\#E = n] = \binom{100}{n}\left(\frac{7}{27}\right)^n\left(1 - \frac{7}{27}\right)^{100-n}.$$

Thus,

$$\mathbb{P}[\#E \leq 30] = \sum_{n=0}^{30}\binom{100}{n}\left(\frac{7}{27}\right)^n\left(1 - \frac{7}{27}\right)^{100-n} = 0.851.$$

* * * * *

Problem 5.



A circular card is divided into 3 sectors 1, 2, 3 and having angles 135°, 90° and 135° respectively. On a second circular card, sectors scoring 1, 2, 3 have angles 180°, 90° and 90° respectively (see diagram). Each card has a pointer pivoted at its centre. After being set in motion, the pointers come to rest independently in random positions. Find the probability that

(a) the score on each card is 1,

(b) the score on at least one of the cards is 3.

The random variable X is the larger of the two scores if they are different, and their common value if they are the same. Show that $\mathbb{P}[X = 2] = \frac{9}{32}$.

Show that $\mathbb{E}[X] = \frac{75}{32}$ and find $\operatorname{Var}[X]$.

Solution.

Part (a). Clearly,

$$\mathbb{P}[\text{both scores are 1}] = \frac{135}{360} \cdot \frac{180}{360} = \frac{3}{16}.$$

Part (b). Likewise,

$$\mathbb{P}[\text{one score is 3}] = \frac{135}{360} + \frac{90}{360} - \frac{135}{360} \cdot \frac{90}{360} = \frac{17}{32}.$$

Observe that the event $X = 1$ is equivalent to both scores being 1, whence we have $\mathbb{P}[X = 1] = \frac{3}{16}$ from part (a). From part (b), we also have $\mathbb{P}[X = 3] = \frac{17}{32}$. Thus,

$$\mathbb{P}[X = 2] = 1 - \mathbb{P}[X = 1] - \mathbb{P}[X = 3] = 1 - \frac{3}{16} - \frac{17}{32} = \frac{9}{32}.$$

Note that

$$\mathbb{E}[X] = 1\left(\frac{3}{16}\right) + 2\left(\frac{9}{32}\right) + 3\left(\frac{17}{32}\right) = \frac{75}{32}$$

and

$$\mathbb{E}\left[X^2\right] = 1^2\left(\frac{3}{16}\right) + 2^2\left(\frac{9}{32}\right) + 3^2\left(\frac{17}{32}\right) = \frac{195}{32}.$$

Thus,

$$\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{195}{32} - \left(\frac{75}{32}\right)^2 = \frac{615}{1024}.$$

* * * * *

Problem 6. Alfred and Bertie play a game, each starting with cash amounting to \$100. Two dice are thrown. If the total score if 5 or more, then Alfred pays $\$x$, where $0 < x \leq 8$, to Bertie. If the total score if 4 or less, then Bertie pays $\$(x+8)$ to Alfred.

(a) Show that the expectation of Alfred's cash after the first game is $\$\frac{1}{3}(304 - 2x)$.

(b) Find the expectation of Alfred's cash after six games.

(c) Find the value of x for the game to be fair.

(d) Given that $x = 3$, find the variance of Alfred's cash after the first game.

Solution.

Part (a). Note that

$$\mathbb{P}[\text{score} < 5] = \frac{3+2+1}{6^2} = \frac{1}{6} \implies \mathbb{P}[\text{score} \geq 5] = 1 - \frac{1}{6} = \frac{5}{6}.$$

Let $\$a_n$ be the expectation of Alfred's cash after n games. Suppose Alfred and Bertie play one more game (i.e. $n+1$ total games). Then

$$a_{n+1} = \frac{5}{6}(a_n - x) + \frac{1}{6}(a_n + x + 8) = a_n + \frac{2}{3}(2 - x).$$

a_n is in AP with common difference $\frac{2}{3}(2-x)$ and is thus given by

$$a_n = a_0 + n\left[\frac{2}{3}(2-x)\right] = 100 + \frac{2n}{3}(2-x).$$

Hence, the expectation of Alfred's cash after the first game is

$$a_1 = 100 + \frac{2(1)}{3}(2-x) = \frac{1}{3}(304 - 2x).$$

Part (b). The expectation of Alfred's cash after six games is

$$a_6 = 100 + \frac{2(6)}{3}(2-x) = 108 - 4x.$$

Part (c). For the game to be fair, $a_0 = a_1 = a_2 = \cdots$, i.e. the common difference is 0. Hence, $x = 2$.

Part (d). Let the random variable X be Alfred's cash after one game. Since the payouts are unaffected by a_0, we take $a_0 = 0$. When $x = 3$, $\mathbb{E}(X) = -\frac{2}{3}$. Hence,

$$\operatorname{Var}[X] = \frac{5}{6}\left(3 - \frac{2}{3}\right)^2 + \frac{1}{6}\left(3 + 8 + \frac{2}{3}\right)^2 = \frac{245}{9}.$$

Problem 7. A random variable X has the probability distribution given in the following table.

x	2	3	4	5
$\mathbb{P}[X = x]$	p	$\frac{2}{10}$	$\frac{3}{10}$	q

(a) Given that $\mathbb{E}[X] = 4$, find p and q.

(b) Show that $\mathrm{Var}[X] = 1$.

(c) Find $\mathbb{E}[|X - 4|]$.

(d) Ten independent observations of X are taken. Find the probability that the value 3 is obtained at most three times.

Solution.

Part (a). We have

$$\mathbb{E}[X] = 2p + 3\left(\frac{2}{10}\right) + 4\left(\frac{3}{10}\right) + 5q = 4 \implies 2p + 5q = 2.2.$$

Additionally, we know that the probabilities must sum to 1:

$$p + \frac{2}{10} + \frac{3}{10} + q = 1 \implies p + q = 0.5.$$

We hence get a system of two linear equations. Solving, we have $p = 1/10$ and $q = 2/5$.

Part (b). Note that

$$\mathbb{E}\left[X^2\right] = 2^2\left(\frac{1}{10}\right) + 3^2\left(\frac{2}{10}\right) + 4^2\left(\frac{3}{10}\right) + 5^2\left(\frac{2}{5}\right) = 17.$$

Thus,

$$\mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = 17 - 4^2 = 1.$$

Part (c).

x	0	1	2
$\mathbb{P}[\lvert X - 4\rvert = x]$	$\frac{3}{10}$	$\frac{2}{5} + \frac{2}{10}$	$\frac{1}{10}$

Hence,

$$\mathbb{E}[|X - 4|] = 0\left(\frac{3}{10}\right) + 1\left(\frac{2}{5} + \frac{2}{10}\right) + 2\left(\frac{2}{10}\right) = 0.8.$$

Part (d). Observe that the probability that we get exactly n 3's is given by

$$\mathbb{P}[n \text{ 3's}] = \binom{10}{n}\left(\frac{2}{10}\right)^n\left(1 - \frac{2}{10}\right)^{10-n}.$$

Hence, the required probability is

$$\text{Required probability} = \sum_{n=0}^{3}\binom{10}{n}\left(\frac{2}{10}\right)^n\left(1 - \frac{2}{10}\right)^{10-n} = 0.879.$$

Self-Practice A14A

Problem 1. An unbiased disc has a single dot marked on one side and two dots marked on the other side. The disc and an unbiased die are thrown, and the random variable X is the sum of the number of dots showing on the disc and on the top of the die.

(a) Tabulate the probability distribution of X.

(b) Show that $\mathbb{P}[X \geq 4 \mid X \leq 7] = 8/11$.

(c) Write down $\mathbb{E}[X]$ and show that $\mathrm{Var}[X] = 19/6$.

* * * * *

Problem 2. The discrete random variable X denotes the number of "sixes" showing when two ordinary fair dice are thrown. Tabulate the probability distribution of X.

Two dice are thrown repeatedly. Find the probability that, in 5 throws, the result $X = 2$ occurs at least once.

The dice are thrown n times. Find the least value of n such that

$$\mathbb{P}[X = 2 \text{ occurs at least once in the } n \text{ throws}] > 0.9.$$

* * * * *

Problem 3. A writer who writes articles for a magazine finds that his proposed articles sometimes need to be revised before they are accepted for publication. The writer finds that the number of days, X, spent in revising a randomly chosen article can be modelled by the following discrete probability distribution.

x	0	1	2	4
$\mathbb{P}[X = x]$	0.8	0.1	0.05	0.05

Calculate $\mathbb{E}[X]$ and $\mathrm{Var}[X]$.

The writer prepares a series of 15 articles for the magazine. Find the expected value of the total time required for revisions to these articles.

The writer regards articles that need no revisions (i.e. $X = 0$) or which need only minor revisions (i.e. $X = 1$) as 'successful' articles, and those requiring major revisions (i.e. $X = 2$) or complete replacement (i.e. $X = 4$) as 'failures'. Assuming independence, find the probability that there will be fewer than 3 'failures' in the 15 articles in the series.

* * * * *

Problem 4. In a game, 2 red balls and 8 blue balls are placed in a bottle. The bottle is shaken and Mary draws 3 balls at random without replacement. The number of red balls that the draws is denoted by R. Find the probability distribution of R, and show that $\mathbb{P}[R \geq 1] = 8/15$.

Show that the expectation of R is $3/5$ and find the variance of R.

Mary scores 4 points for each red ball that she draws. The balls are now replaced in the bottle and the bottle is shaken again. John draws 3 balls at random and without replacement. He scores 1 point for each blue ball that he draws. Mary's score is denoted by M and John's score is denoted by J. Find the expectation and variance of $M - J$.

* * * * *

Problem 5 (🌶). A fair cubical die has three faces marked with a '1', two faces marked with a '2' and one face marked with a '3'.

(a) Calculate the expectation and variance of the score obtained when this die is thrown once.

(b) Deduce the expectation and variance of the score obtained in one throw of a second cubical die, which has one face marked '1', two faces marked '2' and three faces marked '3'.

(c) Two of the first type of die and one of the second type are thrown together, and X denotes the total score obtained. Denoting the expectation and variance of X by μ and σ^2 respectively, show that $\sigma^2 = 5/3$ and $\mathbb{P}[|X - \mu| > 2\sigma] = 1/18$.

Assignment A14A

Problem 1. On a long train journey, a statistician is invited by a gambler to play a die game. The game uses two ordinary dice which the statistician is to throw.

If the total score is 12, the statistician is paid \$6 by the gambler. If the total score is 8, the statistician is paid \$3 by the gambler. However, if both or either dice show a 1, the statistician pays the gambler \$2. The game is considered a draw if none of the 3 scenarios occur.

Let $\$X$ be the amount paid to the statistician by the gambler after the dice are thrown once.

(a) Determine the probability that

(i) $X = 6$,

(ii) $X = 3$,

(iii) $X = -2$.

(b) Find the expected value of X and show that, if the statistician played the game 100 times, his expected loss would be \$2.78, to the nearest cent.

(c) Find the amount $\$a$ that the \$6 would have to be changed to in order to make the game unbiased.

Solution.

Part (a).

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

From the table of outcomes, we clearly have

Part (a)(i).

$$\mathbb{P}[X = 6] = \frac{1}{36}.$$

Part (a)(ii).

$$\mathbb{P}[X = 3] = \frac{5}{36}.$$

Part (a)(iii).

$$\mathbb{P}[X = -2] = \frac{11}{36}.$$

Part (b). We have

$$\mathbb{E}[X] = (6)\left(\frac{1}{36}\right) + (3)\left(\frac{5}{36}\right) + (-2)\left(\frac{11}{36}\right) = -\frac{1}{36}.$$

Thus, the expected value of X after 100 games is

$$\mathbb{E}[X_1 + X_2 + \cdots + X_{100}] = -\frac{1}{36} \cdot 100 = -2.78.$$

Part (c). Replacing \$6 with \$$a$, the expected value of X becomes

$$\mathbb{E}[X] = (a)\left(\frac{1}{36}\right) + (3)\left(\frac{5}{36}\right) + (-2)\left(\frac{11}{36}\right) = \frac{1}{36}(a-7).$$

For the game to be unbiased, $\mathbb{E}[X] = 0$. Hence, $a = 7$.

* * * * *

Problem 2. Four rods of length 1, 2, 3, and 4 units are placed in a bag from which one rod is selected at random. The probability of selecting a rod of length l is kl.

(a) Find the value of k.

(b) Show that the expected value of X, the length of the selected rod, is 3 units and find the variance of X.

After a rod has been selected it is not replaced. The probabilities of selection for each of the three rods that remain are in the same ratio as they were before the first selection. A second rod is now selected from the bag. Let Y be the length of this rod.

(c) Show that $16\,\mathbb{P}[Y = 1 \mid X = 2] = 9\,\mathbb{P}[Y = 2 \mid X = 1]$.

(d) Show that $\mathbb{P}[X + Y = 3] = 17/370$.

Solution.

Part (a). The sum of probabilities must be 1. Hence,

$$1k + 2k + 3k + 4k = 1 \implies k = \frac{1}{10}.$$

Part (b). We have

$$\mathbb{E}[X] = 1\left(\frac{1}{10}\right) + 2\left(\frac{2}{10}\right) + 3\left(\frac{3}{10}\right) + 4\left(\frac{4}{10}\right) = 3.$$

Also,

$$\mathbb{E}\left[X^2\right] = 1^2\left(\frac{1}{10}\right) + 2^2\left(\frac{2}{10}\right) + 3^2\left(\frac{3}{10}\right) + 4^2\left(\frac{4}{10}\right) = 10.$$

Thus,

$$\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 = 10 - 3^2 = 1.$$

Part (c). Consider $\mathbb{P}[Y = 1 \mid X = 2]$. Since the rod of length 2 has already been chosen, we are left with the rods of length 1, 3, and 4. Thus,

$$\mathbb{P}[Y = 1 \mid X = 2] = \frac{1}{1+3+4} = \frac{1}{8}.$$

Consider $\mathbb{P}[Y = 2 \mid X = 1]$. Since the rod of length 1 has already been chosen, we are left with the rods of length 2, 3, and 4. Thus,

$$\mathbb{P}[Y = 2 \mid X = 1] = \frac{2}{2+3+4} = \frac{2}{9}.$$

Thus,

$$16\,\mathbb{P}[Y = 1 \mid X = 2] = 16\left(\frac{1}{8}\right) = 2 = 9\left(\frac{2}{9}\right) = 9\,\mathbb{P}[Y = 2 \mid X = 1].$$

Part (d). For $X+Y=3$, either $X=1$, $Y=2$ or $X=2$, $Y=1$. Thus,

$$\mathbb{P}[X+Y=3]=\mathbb{P}[X=1,Y=2]+\mathbb{P}[X=2,Y=1]=\left(\frac{1}{10}\right)\left(\frac{2}{9}\right)+\left(\frac{2}{10}\right)\left(\frac{1}{8}\right)=\frac{17}{360}.$$

* * * * *

Problem 3. The random variable has the following probability distribution:

x	1	2	3
$\mathbb{P}[X=x]$	θ	2θ	$1-3\theta$

(a) It is given that $0<\theta<1/3$. Show that $\mathbb{E}[X]=3-4\theta$, and find $\mathrm{Var}[X]$ in terms of θ.

The random variable S is the sum of n independent values of X.

(b) Write down $\mathbb{E}[S]$ and $\mathrm{Var}[S]$ in terms of θ and n.

The random variable T is defined by $T=a+bS$. The values of a and b are such that $\mathbb{E}[T]=\theta$ for all θ in the interval $0<\theta<1/3$. Show that

(c) $a=3/4$ and $b=-1/4n$,

(d) $\mathrm{Var}[T]=\theta(3-8\theta)/8n$.

Solution.

Part (a). We have

$$\mathbb{E}[X]=1(\theta)+2(2\theta)+3(1-3\theta)=3-4\theta.$$

Also,

$$\mathbb{E}\left[X^2\right]=1^2(\theta)+2^2(2\theta)+3^2(1-3\theta)=9-18\theta.$$

Hence,

$$\mathrm{Var}[\theta]=\mathbb{E}\left[X^2\right]-(\mathbb{E}[X])^2=(9-18\theta)-(3-4\theta)^2=6\theta-16\theta^2.$$

Part (b). We have

$$\mathbb{E}[S]=\mathbb{E}[X_1+X_2+\cdots+X_n]=n\,\mathbb{E}[X]=n\,(3-4\theta)\,.$$

Also,

$$\mathrm{Var}[S]=\mathrm{Var}[X_1+X_2+\cdots+X_n]=n\,\mathrm{Var}[X]=n\left(6\theta-16\theta^2\right).$$

Part (c). We have

$$\mathbb{E}[T]=\mathbb{E}[a+bS]=a+b\,\mathbb{E}[S]=a+bn\,(3-4\theta)=(a+3bn)-(4bn)\,\theta.$$

Since $\mathbb{E}[T]=\theta$, we have the system

$$a=3bn=0,\qquad 4bn=1,$$

whence $b=-1/4n$. Substituting this into the first equation yields

$$a+3\left(-\frac{1}{4n}\right)n=0\implies a=\frac{3}{4}.$$

Part (d). We have

$$\mathrm{Var}[T]=\mathrm{Var}[a+bS]=\mathrm{Var}[bS]=b^2\,\mathrm{Var}[S]=\left(-\frac{1}{4n}\right)^2\left[n\left(6\theta-16\theta^2\right)\right]=\frac{\theta(3-8\theta)}{8n}.$$

A14B Special Discrete Random Variables

Tutorial A14B

Problem 1. For each of the following situations, determine whether it can be modelled by a binomial distribution, geometric distribution, a Poisson distribution, or neither of the mentioned.

(a) The number of heads obtained when a biased coin is tossed three times.

(b) The number of phone calls received in a randomly chosen hour.

(c) The number of accidents occurring in a factory in a randomly chosen week.

(d) The number of accidents until the first fatal accident at a traffic junction.

(e) The number of red balls obtained when 3 balls are chosen from a bag containing 4 red, 3 green and 3 white balls

 (i) with replacement;

 (ii) without replacement.

(f) The number of typing errors on a randomly chosen page in a draft of a novel.

(g) The number of seeds in a chosen packet of 12 seeds that fail to germinate.

(h) The number of throws of a die until a six is obtained.

Solution.

Part	**Distribution**
(a)	Binomial
(b)	Poisson
(c)	Poisson
(d)	Geometric
(e)(i)	Binomial
(e)(ii)	-
(f)	Poisson
(g)	Binomial
(h)	Geometric

* * * * *

Problem 2. Explain why each of the following situations is not a good model for the proposed distribution.

(a) Using the Poisson distribution to model the number of cars sold at a particular car dealer in a randomly chosen year.

(b) Using the Binomial distribution to model the number of family members that will vote for Party A in the coming election.

(c) Using the Poisson distribution to model the number of people using a particular ATM during a randomly chosen day.

(d) Using the Geometric distribution to model the number of train trips for a particular train before the first breakdown.

Solution.

Part (a). Over the course of a year, the mean rate will likely not be uniform. For instance, the car dealer may only be open on weekdays, so the mean rate on weekdays is different from that on weekends.

Part (b). The probability that one will vote for Party A is not uniform.

Part (c). Over the course of a day, the mean rate will likely not be uniform. For instance, the mean rate at night will be less than the mean rate in the afternoon.

Part (d). The trials are not independent. For instance, wear and tear from previous trials will affect the probability that the next train will break down.

* * * * *

Problem 3. Calculate the probability that in a group of ten people,

(a) none has his or her birthday on a Saturday,

(b) at least two have their birthdays on Saturday,

(c) more than two but at most five have their birthdays on Saturday,

(d) less than four have their birthdays on other days except Saturday.

Find also the mean number of people whose birthday falls on Saturday.

Solution. Let X be the number of people with a birthday on Saturday. Note that $X \sim \mathrm{B}(10, 1/7)$.

Part (a). $\mathbb{P}[X = 0] = 0.214$.

Part (b). $\mathbb{P}[X \geq 2] = 1 - \mathbb{P}[X \leq 1] = 0.429$.

Part (c). $\mathbb{P}[2 < X <\leq 5] = \mathbb{P}[X = 3] + \mathbb{P}[X = 4] + \mathbb{P}[X = 5] = 0.161$.

Part (d). $\mathbb{P}[X > 6] = 1 - \mathbb{P}[X \leq 6] = 9.77 \times 10^{-5}$.

Since $n = 10$ and $p = 1/7$, the expected value of X is

$$\mathbb{E}[X] = np = \frac{10}{7}.$$

* * * * *

Problem 4. In a binomial experiment, the mean number of successful trials is 24 and the variance is 20. Find the number of trials conducted and the probability of success for each trial.

Solution. Let n be the number of trials and p be the probability of success of each trial. We have

$$\mu = np = 24 \quad \text{and} \quad \sigma^2 = np(1-p) = 20.$$

Thus,

$$p = 1 - \frac{20}{np} = \frac{1}{6} \implies n = \frac{24}{p} = 144.$$

* * * * *

Problem 5. If $Y \sim \mathrm{Po}(2.5)$, state the expected value, μ, and the standard deviation, σ, of Y. Use your GC to evaluate the following correct to 3 significant figures.

(a) $\mathbb{P}[Y = 3]$

(b) $\mathbb{P}[Y > 4.5]$

(c) $\mathbb{P}[Y \leq 5]$

(d) $\mathbb{P}[6 < Y < 10]$

(e) $\mathbb{P}[Y \text{ is } 0 \text{ or } 1]$

(f) $\mathbb{P}[|Y - \mu| < \sigma]$

Solution. Since Y follows a Poisson distribution, $\mu = \sigma^2 = 2.5$, whence $\sigma = \sqrt{2.5}$.
Part (a). $\mathbb{P}[Y = 3] = 0.214$.
Part (b). $\mathbb{P}[Y > 4.5] = 1 - \mathbb{P}[Y \leq 4] = 0.109$.
Part (c). $\mathbb{P}[Y \leq 5] = 0.958$.
Part (d). $\mathbb{P}[6 < Y < 10] = \mathbb{P}[Y \leq 9] - \mathbb{P}[Y \leq 6] = 0.0139$.
Part (e). $\mathbb{P}[Y \text{ is } 0 \text{ or } 1] = \mathbb{P}[Y = 0] + \mathbb{P}[Y = 1] = 0.287$.
Part (f). $\mathbb{P}[|Y - \mu| < \sigma] = \mathbb{P}\left[2.5 - \sqrt{2.5} < Y < 2.5 + \sqrt{2.5}\right] = \mathbb{P}[1 \leq Y \leq 4] = 0.809$.

* * * * *

Problem 6. Epple Company manufactures many E-phones. It is known that 1% of the E-phones manufactured are defective.

(a) A random sample of n phones was selected. Using an algebraic method, find the smallest value of n such that the probability that there is at least one defective phone in the sample is more than 0.95.

(b) A carton, which consists of 24 E-phones, will be rejected if there are at least two defective phones. Show that the probability that a randomly chosen carton is being rejected is 0.0239.

Solution.
Part (a). Let X be the number of defective phones in the sample. Then $X \sim \mathrm{B}(n, 0.01)$. Consider $\mathbb{P}[X \geq 1] \geq 0.95$:

$$\mathbb{P}[X \geq 1] \geq 0.95 \implies \mathbb{P}[X = 0] = 0.99^n \leq 0.05 \implies n \geq 298.1.$$

Since $n \in \mathbb{N}$, the least n is 299.
Part (b). Take $n = 24$. Then

$$\mathbb{P}[X \geq 2] = 1 - \mathbb{P}[X \leq 1] = 0.0239.$$

* * * * *

Problem 7. In an opinion poll before an election, a sample of 30 voters is obtained.

(a) The number of voters in the sample who support the Alliance Party is denoted by A. State, in context, what must be assumed for A to be well modelled by a binomial distribution.

Assume now that A has the distribution $\mathrm{B}(30, p)$.

(b) Given that $p = 0.15$, find $\mathbb{P}[A = 3 \text{ or } 4]$.

(c) For an unknown value of p, it is given that $\mathbb{P}[A = 15] = 0.06864$ correct to 5 decimal places. Show that p satisfies an equation of the form $p(1-p) = k$, where k is a constant to be determined. Hence, find the value of p to a suitable degree of accuracy, given that $p < 0.5$.

Solution.

Part (a). Votes must be made independently, and the probability of voting A is the same for all voters.

Part (b). $\mathbb{P}[A = 3 \text{ or } 4] = \mathbb{P}[A = 3] + \mathbb{P}[A = 4] = 0.373$.

Part (c). Note that $k = p(1-p) = \mathrm{Var}[A]$ is a constant. Consider $\mathbb{P}[A = 15]$:

$$\mathbb{P}[A = 15] = \binom{30}{15} p^{15}(1-p)^{15} = 0.06864 \implies k = p(1-p) = \sqrt[15]{\frac{0.06864}{\binom{30}{15}}} = 0.23790.$$

Expanding $k = p(1-p)$ into a quadratic in p, we have

$$p^2 - p + 0.23790 = 0 \implies p = 0.390 \text{ or } 0.610.$$

Since $p < 0.5$, we take $p = 0.390$.

* * * * *

Problem 8. In a large population, the proportion having blood group A is 35%. Specimens of blood from the first five people attending a clinic are to be tested. It can be assumed that these five people are a random sample from the population. The random variable X denotes the number of people in the sample we are found to have blood group A.

(a) Find $\mathbb{P}[X \leq 2]$, correct to 3 decimal places.

(b) Three such samples of five people are taken. Find

 (i) the probability that each of these three samples has more than two people with blood group A,

 (ii) the probability that one of these three samples has exactly one person with blood group A, another has exactly two people with blood group A, and the remaining sample has more than two people with blood group A.

(c) Ten such samples of five people were taken. Find the probability that seven samples have exactly one person with blood group A.

Solution. Note that $X \sim \mathrm{B}(5, 0.35)$.

Part (a). $\mathbb{P}[X \leq 2] = 0.76483$ (5 s.f.) $= 0.765$ (3 d.p.).

Part (b).

Part (b)(i). The required probability is given by

$$[\mathbb{P}[X > 2]]^3 = [1 - \mathbb{P}[X \leq 2]]^3 = 0.0130.$$

Part (b)(ii). The required probability is given by

$$3!\,[\mathbb{P}[X = 1]\,\mathbb{P}[X = 2]\,\mathbb{P}[X > 2]] = 0.148.$$

Part (c). Note that $\mathbb{P}[X = 1] = 0.31239$. Let Y be the number of samples with exactly one person with blood group A. Then $Y \sim \mathrm{B}(10, 0.31239)$. Hence, $\mathbb{P}[Y = 7] = 0.0113$.

* * * * *

Problem 9. Every student in Sunny Junior College owns a graphic calculator (GC). The probability of a student carrying a GC to school is 0.98. Assume that the number of students who carries a GC to school follows a binomial distribution.

(a) (i) Given that the probability that more than m students, in a random sample of 30 students, carry a GC to school is at most 0.99, find the least value of m.

(ii) Give a reason why in real life, the number of students who carries a GC to school may not follow a binomial distribution.

The latest operating system (OS) of the GC is required for the installation of a new application. On average, 3 out of 4 students have the latest OS in their GC. A class of 26 students is to report to their mathematics tutor, Mr Ng, to install the new application.

(b) Find the probability that the 15th student who reports to Mr Ng is the 9th student whose GC has the latest OS while the last student is the 10th student without the latest OS.

Solution.

Part (a).

Part (a)(i). Let X be the number of students who bring a GC to school. Then $X \sim \mathrm{B}(30, 0.98)$. Consider $\mathbb{P}[X > m] \leq 0.99$. Using GC, the least m is 27.

Part (a)(ii). The probability that a student brings their GC to school is not the same for all students, as different students may have different timetables.

Part (b). Number the students in the order in which they report to Mr Ng.

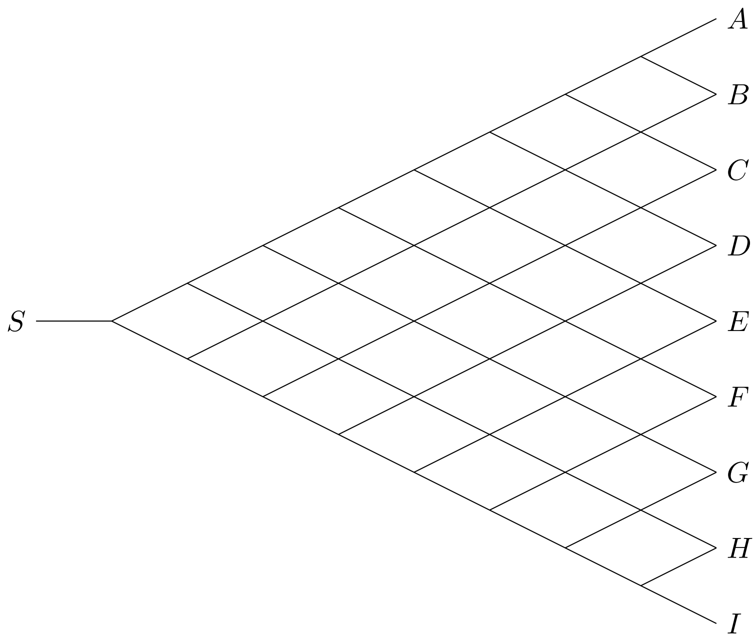
- Students 1–14: 8 students have the latest OS installed, remaining 6 students do not.
- Student 15: Has the latest OS installed.
- Students 16–25: 7 students have the latest OS installed, remaining 3 students do not.
- Student 26: Does not have the latest OS installed.

The probability of this happening is given by

$$\left[\binom{14}{8}\left(\frac{3}{4}\right)^8\left(\frac{1}{4}\right)^6\right]\left[\frac{3}{4}\right]\left[\binom{10}{3}\left(\frac{3}{4}\right)^7\left(\frac{1}{4}\right)^3\right]\left[\frac{1}{4}\right] = 0.00344.$$

* * * * *

Problem 10. In a computer game, a bug moves from left to right through a network of connected paths. The bug starts at S and, at each junction, randomly takes the left fork with probability p or the right fork with probability q, where $q = 1 - p$. The forks taken at each junction are independent. The bug finishes its journey at one of the 9 endpoints labelled A–I (see diagram below).



(a) Show that the probability that the bug finishes its journey at D is $56p^5q^3$.

(b) Given that the probability that the bug finishes its journey at D is greater than the probability at any one of the other endpoints, find exactly the possible range of values of p.

In another version of the game, the probability that, at each junction, the bug takes the left fork is $0.9p$, the probability that the bug takes the right fork is $0.9q$ and the probability that the bug is swallowed up by a 'black hole' is 0.1.

(c) Find the probability that, in this version of the game, the bug reaches one of the endpoints A–I, without being swallowed up by a black hole.

Solution.

Part (a). Relabel each endpoint from A–I to 0–8. Let the random variable X be the end-point that the bug ends up at. Clearly, to reach endpoint i, the bug must take i right forks and $8-i$ left forks. Hence, $X \sim \mathrm{B}(8, q)$ and the probability that the bug reaches endpoint 3 (i.e. endpoint D) is

$$\mathbb{P}[X = 3] = \binom{8}{3} q^3 (1-q)^{8-3} = 56p^5q^3.$$

Part (b). Since X follows a binomial distribution, it suffices to find the range of values of p that satisfy $\mathbb{P}[X = 2] < \mathbb{P}[X = 3] > \mathbb{P}[X = 4]$.

Case 1: $\mathbb{P}[X = 2] < \mathbb{P}[X = 3]$. Note that $\mathbb{P}(X = 2) = \binom{8}{2} q^2 (1-q)^{8-2} = 28p^6q^2$.

$$\mathbb{P}(X = 2) < \mathbb{P}(X = 3) \implies 28p^6q^2 < 56p^5q^3 \implies 28p < 56(1-p) \implies p < \frac{2}{3}.$$

Case 2: $\mathbb{P}[X = 3] > \mathbb{P}[X = 4]$. Note that $\mathbb{P}[X = 4] = \binom{8}{4} q^4 (1-q)^{8-4} = 70p^4q^4$.

$$\mathbb{P}(X = 3) > \mathbb{P}(X = 4) \implies 56p^5q^3 > 70p^4q^4 \implies 56p > 70(1-p) \implies p > \frac{5}{9}.$$

Hence, $5/9 < p < 2/3$.

Part (c). Note that the bug most take a total of 8 forks. Since the probability of not getting swallowed by a black hole at each fork is 0.9, the desired probability is clearly $0.9^8 = 0.430$ (3 s.f.).

* * * * *

Problem 11. The number of injuries, X, sustained by workers in a factory per week follows a Poisson distribution with standard deviation σ. Given that $3\,\mathbb{P}[X = 2] = 16\,\mathbb{P}[X = 4]$ determine the value of σ and hence state the mean of X.

(a) Find the probability that, in a randomly chosen week, there is at least one injury.

(b) Assuming that a month consists of four weeks, find the probability that, in a randomly chosen month, there are less than 4 injuries.

(c) Calculate the probability that there will be more than 1 but less than 4 injuries in each of two consecutive weeks.

Solution. We have

$$3\mathrm{e}^{-\mu}\frac{\mu^2}{2!} = 3\,\mathbb{P}[X = 2] = 16\,\mathbb{P}[X = 4] = 16\mathrm{e}^{-\mu}\frac{\mu^4}{4!} \implies \mu^2 = \sigma = 2.25 \implies \mu = 1.5.$$

Thus, $X \sim \mathrm{Po}(1.5)$.

Part (a). $\mathbb{P}[X \geq 1] = 1 - \mathbb{P}[X = 0] = 0.777$.

Part (b). Let Y be the number of injuries in a month. Then $Y \sim \mathrm{Po}(6)$. Hence, $\mathbb{P}[Y < 4] = \mathbb{P}[Y \leq 3] = 0.151$.

Part (c). The required probability is given by

$$[\mathbb{P}[1 < X < 4]]^2 = [\mathbb{P}[X = 2] + \mathbb{P}[X = 3]]^2 = 0.142.$$

* * * * *

Problem 12. During a weekday, heavy lorries pass a census point P on a village high street independently and at random times. The mean rate for westward travelling lorries is 2 in any 30-minute period and for eastward travelling lorries is 3 in any 30-min period. Find the probability

(a) that there will be no lorries passing P in a given 10-min period,

(b) that at least one lorry from each direction will pass P in a given 10-minute period,

(c) more than 2 westward travelling lorries will pass P between the time 1410 and 1445,

(d) that there will be exactly 4 lorries passing P in a given 20-minutes period

(e) at least 2 eastward travelling lorries passing P in a period of 20 minutes given that there are exactly 4 lorries passing P at that time.

Solution. Let W_k and E_k be the number of westward and eastward travelling lorries passing P in a k-minute period. Then

$$W_k \sim \mathrm{Po}\left(\frac{k}{15}\right) \quad \text{and} \quad E_k \sim \mathrm{Po}\left(\frac{k}{10}\right).$$

Part (a). $\mathbb{P}[W_{10} = 0]\,\mathbb{P}[E_{10} = 0] = 0.189$.

Part (b). $\mathbb{P}[W_{10} \geq 1]\,\mathbb{P}[E_{10} \geq 1] = [1 - \mathbb{P}[W_{10} = 0]]\,[1 - \mathbb{P}[E_{10} = 0]] = 0.308$.
Part (c). $\mathbb{P}[W_{35} \geq 2] = 1 - \mathbb{P}[W_{35} \leq 2] = 0.413$.
Part (d). Note that $W_{20} + E_{20} \sim \mathbb{P}\left[\frac{20}{15} + \frac{20}{10}\right] = \mathbb{P}\left[\frac{10}{3}\right]$. Thus, $\mathbb{P}[W_{20} + E_{20} = 4] = 0.184$.
Part (e). The desired probability is given by

$$1 - \frac{\mathbb{P}[W_{20} = 4]\,\mathbb{P}[E_{20} = 0] + \mathbb{P}[W_{20} = 3]\,\mathbb{P}[E_{20} = 1]}{\mathbb{P}[W_{20} + E_{20} = 4]} = 0.821.$$

* * * * *

Problem 13. A car rental company has n cars which may be hired on a daily basis. The demand for cars in a day follows a Poisson distribution with variance 1.5.

(a) For $n = 5$, for any one day,
  (i) find the probability that less than 3 cars are hired.
  (ii) find the probability that all the cars are hired.

(b) The probability that the demand for cars being met on any day is at least 0.95. Find the least value of n.

(c) The probability that no car is rented out on k consecutive days is less than 0.01. Find the least value of k.

(d) The probability that there are less than two cars rented out on h consecutive days is less than 0.005. Find the least value of h.

Solution. Let X be the number of cars hired in a given day. Then $X \sim \text{Po}(1.5)$.
Part (a).
Part (a)(i). $\mathbb{P}[X < 3] = \mathbb{P}[X \leq 2] = 0.809$ (3 s.f.).
Part (a)(ii). $\mathbb{P}[X \geq 5] = 1 - \mathbb{P}[X \leq 4] = 0.0186$.
Part (b). Consider $\mathbb{P}[X \leq n] \geq 0.95$. Using G.C., the least n is 4.
Part (c). Consider $\mathbb{P}[X = 0]^k \leq 0.01$. Using G.C., the least k is 4.
Part (d). Consider $\mathbb{P}[X < 2]^h = \mathbb{P}[X \leq 1]^h \leq 0.005$. Using G.C., the least h is 10.

* * * * *

Problem 14. A randomly chosen doctor in general practices sees, on average, one case of a broken nose per year and each case is independent of other similar cases.

(a) Regarding a month as a twelfth part of a year,
  (i) show that the probability that, between them, three such doctors see no cases of a broken nose in a period of one month is 0.779, correct to three significant figures,
  (ii) find the variance of the number of cases seen by three such doctors in a period of six months.

(b) Find the probability that, between them, three such doctors see at least three cases in one year.

(c) Find the probability that, of three such doctors, one sees three cases and the other two see no cases in one year.

Solution. Let $X_{t,n}$ be the number of cases of a broken nose seen by n doctors in t months. Then $X_{t,n} \sim \text{Po}(tn/12)$.

Part (a).

Part (a)(i). The required probability is

$$\mathbb{P}[X_{1,3} = 0] = 0.779 \text{ (3 s.f.)}.$$

Part (a)(ii). Since $X_{t,n}$ follows a Poisson distribution, $\mu = \sigma^2$. Hence,

$$\text{Var}[X_{6,3}] = \frac{(6)(3)}{12} = \frac{3}{2}.$$

Part (b). The required probability is

$$\mathbb{P}[X_{12,3} \geq 3] = 1 - \mathbb{P}[X_{12,3} \leq 2] = 0.577 \text{ (3 s.f.)}.$$

Part (c). The required probability is

$${}^3C_1\, \mathbb{P}[X_{12,1} = 3]\, \mathbb{P}[X_{12,2} = 0] = 0.0249 \text{ (3 s.f.)}.$$

* * * * *

Problem 15. During the winter in New York, the probability that snow will fall on any given day is 0.1. Taking November 1st as the first day of winter and assuming independence from day to day, find the probability that the first snow of winter will fall in New York on the last day of November (30th).

Given that no snow has fallen at New York during the whole of November, a teacher decides not to wait any longer to book a skiing holiday. The teacher decides to book for the earliest date for which the probability that snow will have fallen on, or before, that date is at least 0.9. Find the date of the booking.

Solution. The probability that the first snow of winter will fall on 30 November is given by

$$(0.9)^{29}(0.1) = 0.00471 \text{ (3 s.f.)}.$$

Let n be the number of days after 30 November. Consider $\mathbb{P}[X \leq n] \geq 0.9$, where $X \sim \text{Geo}(0.1)$. Using G.C., the least n is 22. Hence, the date of the booking is 22 December.

* * * * *

Problem 16. A salesman sells goods by telephone. The probability that any particular call achieves a sale is 1/12. The salesman continues to make calls until one call achieves a sale.

(a) State one assumption need for this to be modelled by a Geometric distribution.

(b) Given that a Geometric distribution is used to model this, find the probability that the call achieves a sale

 (i) is the fifth call made,

 (ii) does not occur in the first five calls.

(c) The salesman uses 5 minutes for each call, find the expected amount of time he has to spend to reach his first sale.

Solution.

Part (a). The calls must be independent.

Part (b). Let X be the number of calls made under the salesman achieves a sale. Then $X \sim \text{Geo}(1/12)$.

Part (b)(i). $\mathbb{P}[X = 5] = 0.0588$ (3 s.f.).

Part (b)(ii). $\mathbb{P}[X > 5] = (1 - 1/12)^5 = 0.647$.

Part (c). Note that $\mathbb{E}[X] = 1/p = 12$. Hence, the expected amount of time he has to spend to reach his first sale is $5 \times 12 = 60$ minutes.

* * * * *

Problem 17. If $X \sim \text{B}(n, p)$ and $Y \sim \text{Geo}(p)$, explain why $\mathbb{P}[Y = n] \leq \mathbb{P}[X = 1]$.

Solution. Both $X = 1$ and $Y = n$ represent the event that there is exactly one success in n trials. However, the event $Y = n$ has the added restriction that the success must come on the last trial, whereas the event $X = 1$ has no such restriction; the success can occur in any of the n trials. Hence, the event $Y = n$ is a subset of the event $X = 1$, thus $\mathbb{P}[Y = n] \leq \mathbb{P}[X = 1]$, with equality only when $n = 1$.

* * * * *

Problem 18. Serious delays on a certain railway line occurs at random, at an average rate of one per week. Show that the probability of at least 4 serious delays occurring during a particular 4-week period is 0.567, correct to 3 decimal places.

Taking a year to consist of thirteen 4-week periods, find the probability that, in a particular year, there are at least ten of these 4-week periods during which at least 4 serious delays occur.

Given that the probability of at least n serious delays occurring in a period of 6 weeks is greater than 0.795, find the largest possible integer value of n.

Solution. Let the number of serious delays in k weeks be $X_k \sim \text{Po}(k)$. We have

$$\mathbb{P}[X_4 \geq 4] = 1 - \mathbb{P}[X_4 \leq 3] = 0.56653 = 0.567 \text{ (3 d.p.)}.$$

Let Y be the number of 4-week periods during which at least 4 serious delays occur. Note that $Y \sim \text{B}(13, 0.56653)$. Hence,

$$\mathbb{P}[Y \geq 10] = 1 - \mathbb{P}[Y \leq 9] = 0.115 \text{ (3 s.f.)}.$$

Consider $\mathbb{P}[X_6 \geq n] \geq 0.795$. Using G.C., the greatest value of n is 4.

* * * * *

Problem 19. The demand for XO pies in a confectionary shop may be taken to follow a Poisson distribution with a mean of 0.4 pies per hour. The shop opens for 5 days in a week and does business for 8 hours per day.

(a) Find the probability that the demand for XO pies is at least 3 in a day.

(b) Find the probability that there is one day with demand for XO pies of at least 3, and another two days with demand 0.

(c) Find the probability that there is at most one day with zero demand for XO pies in a week.

(d) Given that the demand for XO pies is exactly 3 on a particular day, what is the probability that this occurred within the first hour of business.

Solution. Let the number of XO pies demanded in k hours be $X_k \sim \text{Po}(0.4k)$.

Part (a). $\mathbb{P}[X_8 \geq 3] = 1 - \mathbb{P}[X_8 \leq 2] = 0.620$ (3 s.f.).

Part (b). Note that $\mathbb{P}[X_8 = 0] = 0.40762$. The required probability is

$$^3C_1\,\mathbb{P}[X_8 \geq 3]\,\mathbb{P}[X_8 = 0] = 0.00309.$$

Part (c). Let the number of days in a week where there is 0 demand for XO pies be $Y \sim \text{B}(5, 0.40762)$. Then $\mathbb{P}[Y \leq 1] = 0.985$ (3 s.f.).

Part (d). If all three pies for the day were sold within the first hour of business, then no pies were sold in the remaining seven hours. Hence, the required probability is

$$\frac{\mathbb{P}[X_1 = 3]\,\mathbb{P}[X_1 = 0]^7}{\mathbb{P}[X_8 = 3]} = 0.00195.$$

* * * * *

Problem 20. Given the climate of the country and duration of transportation, the probability of a strawberry from a particular orchard turning rotten is believed to be 0.15. In a fruits wholesale centre where strawberries from that orchard are sold, they are packed and sold in trays of 20.

(a) Show that the probability that there are at most 5 rotten strawberries in a tray is 0.933.

(b) Find, to 3 decimal places, the probability that there are exactly 3 rotten strawberries in 2 randomly selected trays.

A cold desserts hawker bought 60 trays of strawberries from the wholesaler centre. Using a suitable approximation, find the probability that there are at least 4 trays with more than 5 rotten strawberries in each tray.

Solution. Let X_k be the number of rotten strawberries in k trays. We have $X_k \sim \text{B}(20k, 0.15)$.

Part (a). $\mathbb{P}[X_1 \leq 5] = 0.933$ (3 s.f.).

Part (b). $\mathbb{P}[X_2 = 3] = 0.0816$ (3 s.f.).

Note that $\mathbb{P}[X_1 > 5] = 1 - \mathbb{P}[X_1 \leq 5] = 0.067308$. Let Y be the number of trays with more than 5 rotten strawberries. We can approximate Y using a Poisson distribution since $n = 60$ is large and $p = 0.067308$ is small. We hence have $Y \sim \text{Po}(np) = \text{Po}(4.0385)$. Then $\mathbb{P}[Y \geq 4] = 1 - \mathbb{P}[Y \leq 3] = 0.574$ (3 s.f.).

Self-Practice A14B

Problem 1. A crossword puzzle is published in The Times each day of the week, except Sunday. A man is able to complete, on average, 8 out of 10 of the crossword puzzles.

(a) Find the expected value and the standard deviation of the number of completed crossword puzzles in a given week.

(b) Show that the probability that he will complete at least 5 in a given week is 0.655.

(c) Given that he completes the puzzle on Monday, find the probability that he will complete at least 4 in the rest of the week.

(d) Find the probability that, in a period of 4 weeks, he completes 4 or less in only one of the 4 weeks.

* * * * *

Problem 2. There is a lift on the ground floor of an old 50-storey building. This lift serves only the first 25 floors with an average number of breakdowns of 2 per week. On the 25th floor, there is another lift serving the 26th to the 50th floor with an average number of breakdowns of 0.5 per week, independent of the other lift. Find, correct to 3 decimal places, the probability that

(a) both lifts do not break down in a particular week,

(b) there are not more than 2 breakdowns altogether in a particular day,

(c) in a period of 7 days, there are 6 days on which there are not more than 2 breakdowns altogether.

* * * * *

Problem 3. The centre pages of the '8 days' magazine consist of 1 page of film and theatre reviews and 1 page of classified advertisements. The number of misprints in the reviews has a Poisson distribution with mean 2.3 and the number of misprints in the classified section has a Poisson distribution with variance 1.7.

Find the probabilities that, on the centre pages, there will be

(a) no misprints,

(b) 5 misprints.

Given that there are 5 misprints in the centre pages, find the probability that 2 of them occur in the classified section.

* * * * *

Problem 4. The student resource centre has 4 rooms which can be booked for students' activities in a day at a time. Requests of the booking of a room take place independently with a mean of 4 requests per day.

(a) Find the probability that not all requests for the booking of a room can be met on any particular day.

(b) On any particular day during which more than 2 requests are received for the booking of a room, find the probability that all requests for the booking of a room can be met on that day.

(c) Find the probability that there are at most 3 days in a particular week (taking 1 week to be 5 schooling days) during which not all requests for the booking of a room for that day can be met.

(d) Find the least number of rooms that the student resource centre should have so that, on any particular day, the probability that a request for the booking of a room for that day has to be refused is less than 0.05.

* * * * *

Problem 5. A jeweller sells rubies and diamonds. The average number of rubies and diamonds sold per week is 1.8 and 2.7 respectively.

(a) Find the probability that exactly two rubies are sold in a given week.

(b) Find the probability that exactly 4 diamonds are sold in a given two-week period.

(c) Find the probability that the total number of jewels sold in a given week is at least 4.

(d) Given that less than 3 jewels are sold in a given week, find the probability that the number of rubies sold is more than the diamonds sold.

* * * * *

Problem 6. The number of printing errors, Y, in a page of a book follows a Poisson distribution with standard deviation σ. Given that $3\,\mathbb{P}[Y = 2] = 16\,\mathbb{P}[Y = 4]$, determine the value of σ and hence state the mean of Y.

(a) The probability that there are no errors in k consecutive pages is less than 0.01. Determine the least k value.

(b) For a book of 100 pages, find the probability that at least one page has at least four errors.

* * * * *

Problem 7. In order to be offered a scholarship, a candidate has to pass two rounds of interview (it is assumed that all interviewers' decisions are independent).

In the first round, there will be a panel of 10 interviewers and the probability of each interviewer passing a candidate is 0.9. The candidate fails to qualify for the second round if more than one interviewer decides not to pass him or her.

(a) Find the probability that a candidate passes the first round of interview.

In the second round, there will be a panel of 5 interviewers and the probability of each interviewer passing a candidate is 0.8. The candidate is offered a scholarship only if all interviews pass him or her in the second round.

(b) Show that the probability that the candidate is offered the scholarship is 0.241, correct to three significant figures.

(c) There are n candidates going for the interviews. Find the smallest n such that there is at least a 98% chance of 2 or more candidates being offered the scholarship.

Problem 8. The two most common types of disciplinary offences in a particular boy school is keeping long hair and failure to wear the school badge. The mean number of disciplinary offences recorded per day involving long hair is 1.12. Assuming that each school week consists of five school days, the mean number of disciplinary offences recorded per school week involving failure to wear the school badge is 4.2. The number of cases for each disciplinary offence is assumed to have an independent Poisson distribution.

(a) Find the probability that at most 9 cases of disciplinary offence are recorded in a given school week.

(b) In a school week in which there are more than 7 cases of disciplinary offence involving long hair, find the probability that at most 9 cases of disciplinary offence are recorded.

(c) Calculate the probability that on a Thursday in a particular school week, it is the third day in the school week in which the discipline master caught at least 4 students having long hair in a day (you may assume that Monday is the first day of school for a school week).

(d) Explain why the Poisson distribution may not be a good model for the number of disciplinary cases involving long hair, in a school year.

* * * * *

Problem 9. In a small company, the employees send an average of 1.2 print jobs to the colour print and α print jobs to the laser printer per day. It is assumed that the print jobs are independent.

(a) Given that on 1 in 100 working days there are no print jobs for both printers, show that $\alpha = 3.41$ correct to 3 significant figures.

(b) Let E be the event that more than 3 print jobs were sent in on a working day. Find $\mathbb{P}[E]$. Hence, find the probability that the first occurrence of event E happens before the 5th working day.

(c) A typical working day consists of 8 hours of work. Find the probability that more than half of the total print jobs sent during a typical working day occurs within the first hour of work, given that there was a total of 3 print jobs for the day.

* * * * *

Problem 10. A firm investigated the number of employees suffering injuries whilst at work. The results recorded below were obtained for a 52-week period:

Number of employees injured in a week	0	1	2	3	4 or more
Number of weeks	31	17	3	1	0

Give reasons why one might expect this distribution to approximate to a Poisson distribution. Evaluate the mean and variance of the data and explain why this gives further evidence in favour of a Poisson distribution. Using the calculated value of the mean, find the theoretical frequencies of a Poisson distribution for the number of weeks in which 0, 1, 2, 3, 4 or more employees were injured.

Problem 11 (🌶). The Entrepreneur Club is in charge of selling the school's tee shirt. Based on the sales record of the club, it was found that the monthly demand for tee shirt size XS has a Poisson distribution with mean 2 and the monthly demand for tee shirt size XXL has a Poisson distribution with mean 3. The club kept a monthly stock of 3 and 4 for tee shirt sizes XS and XXL respectively.

(a) Calculate the probability that there is more than one XS size tee shirt being sold in a day, assuming there are 30 days in a month.

(b) Calculate the probability that the club will not meet the demand for either XS or XXL tee shirts in a month.

(c) Find the most probable number of XXL tee shirts sold in a month.

(d) Determine the least number of stock needed each month for the XS tee shirts in order to meet the demand with a probability of at least 0.95.

Assignment A14B

Problem 1. In a school with many students, an average of 1 out of 5 students is sick in a month.

(a) State, in this context, two conditions that must be met for the number of students who are sick in a month to be well modelled by a binomial distribution. Explain why each of these conditions may not be met.

For the remainder of this question, assume that these conditions are met.

(b) 25 students are randomly selected, one after another. Find the probability that the last student selected is the fifth student who is sick in a month.

(c) Find the least value of n such that in a sample of n randomly selected students, the probability that at least 6 are sick in a month is more than 0.95.

Solution.

Part (a). The students fall sick independently, e.g. there is no herd immunity to sickness, nor does sickness spread among the students.

Part (b). Since the last (25th) student is the fifth student who is sick, there must only be 4 sick students in the 24 students sampled before. Hence, the required probability is

$$^{24}C_4 \left(\frac{1}{5}\right)^4 \left(1 - \frac{1}{5}\right)^{24-4} \cdot \frac{1}{5} = 0.0392 \text{ (3 s.f.)}.$$

Part (c). Let X be the number of students that are sick. Then $X \sim \mathrm{B}(n, 1/5)$. Note that $\mathbb{P}[X \geq 6] \geq 0.95$ is equivalent to $\mathbb{P}[X \leq 5] \leq 0.05$. Using G.C., the least n that satisfies this inequality is $n = 50$.

* * * * *

Problem 2. A factory produced plastic cups. It is known that 10% of the cups have cracks. For quality control purposes, the factory adopts two kinds of checks on the day's produce.

A: Multiple checks are performed. In each check, individual cups are randomly selected and checked for cracks.

B: A random sample of 15 cups is first tested. The day's produce is rejected if more than four cups have cracks, and accepted if three or fewer have cracks. If exactly four cups have cracks, another random sample of 10 cups will be tested. The day's produce is accepted if none of the cups in the second sample have cracks and rejected otherwise.

(a) Find the probability that the first cracked cup appears before the 7th random check in a day under Scheme **A**.

(b) Find the probability that the batch of cups is rejected under Scheme **B**.

Solution.

Part (a). Let X be the number of cups checked before the first cracked cup appears under Scheme **A**. Then $X \sim \mathrm{Geo}(1/10)$. Hence, the required probability is

$$\mathbb{P}[X < 7] = \mathbb{P}[X \leq 6] = 0.469 \text{ (3 s.f.)}.$$

Part (b). Let Y and Z be the number of cracked cups in the first and second rounds of Scheme **B**, respectively. Then $Y \sim \mathrm{B}(15, 1/10)$ and $Z \sim \mathrm{B}(10, 1/10)$.

There are two ways for the batch of cups to be rejected:

1. Directly rejected in round 1 ($Y > 4$).
2. Proceeded to round 2 ($Y = 4$), and rejected in round 2 ($Z \geq 1$).

The probability that the batch of cups is rejected is hence given by

$$\mathbb{P}[Y > 4] + \mathbb{P}[Y = 4]\,\mathbb{P}[Z \geq 1] = 0.0406 \text{ (3 s.f.)}.$$

* * * * *

Problem 3. Floral Junior College faces two major disciplinary problems regarding their students, namely, littering and unauthorized flower-plucking. On average, there are two cases of littering per day and three cases of unauthorized flower-plucking per school week. You may assume that a school week consists of 5 days. The two types of offences are independent of each other and occur randomly.

(a) Find the probability that in a period of three consecutive days, there are at least 10 cases of littering or unauthorized flower-plucking.

(b) The school principal, Mrs Green, decides to activate the whole school to do a mass clean-up once the total number of new cases of littering exceeds m. Find the least value of m if the probability that the school is activated within the next three days is less than 0.7.

Solution. Let L_k and P_k be the number of cases of littering and flower-plucking in a period of k consecutive school days, respectively. Then $L_k \sim \text{Po}(2k)$ and $P_k \sim \text{Po}(3k/5)$.
Part (a). Since $L_3 + P_3 \sim \text{Po}(2(3) + 3(3)/5) = \text{Po}(7.8)$, the required probability is

$$\mathbb{P}[L_3 + P_3 \geq 10] = 1 - \mathbb{P}[L_3 + P_3 \leq 9] = 0.259 \text{ (3 s.f.)}.$$

Part (b). The probability that the school is activated within the next three days is given by $\mathbb{P}[L_3 > m]$. We thus consider the inequality $\mathbb{P}[L_3 > m] < 0.7$. Using G.C., the least m is 5.

* * * * *

Problem 4. In a sales campaign, a company gives each customer who purchases more than one hundred dollars' worth of goods a card with a picture of a film star on it. There are 10 different pictures, one each of 10 different film stars. On any occasion, the card received by any customer is equally likely to carry any one of the 10 pictures. Any customer who collects a complete set of all the 10 pictures gets a reward.

Suppose a customer has already collected r different pictures where $r = 1, 2, 3, \dots, 9$. Let X_r be the random variable denoting the additional number of cards that needs to be collected by the customer until he gets a card that carries a different picture from his r pictures.

(a) Find $\mathbb{P}[X_r = X]$, where $x \in \mathbb{Z}^+$ and state $\mathbb{E}[X_r]$.

(b) Prove that $\mathbb{P}[X_{r+1} + X_r \leq a \mid X_r = b] = \mathbb{P}[X_{r+1} \leq a - b]$ where $a > b$. Deduce that $\mathbb{P}[X_{r+1} + X_r \leq 4 \mid X_r = 2] = 1 - \left(\frac{r+1}{10}\right)^2$.

(c) Let Y be the random variable denoting the total number of cards that needs to be collected by a new customer until he obtains a complete set of 10 different picture cards. By expressing Y in terms of X_r, find the value of $\mathbb{E}[Y]$.

Solution.

Part (a). Since the customer already has r different pictures, the probability of the next card having a new picture is $1 - r/10$. Hence, $X_r \sim \text{Geo}(1 - r/10)$, whence

$$\mathbb{P}[X_r = X] = \left[1 - \left(1 - \frac{r}{10}\right)\right]^{X-1} \left(1 - \frac{r}{10}\right) = \left(\frac{r}{10}\right)^{X-1} - \left(\frac{r}{10}\right)^{X}$$

and

$$\mathbb{E}[X_r] = \frac{1}{1 - r/10} = \frac{10}{10 - r}.$$

Part (b). Note that

$$\begin{aligned}\mathbb{P}[X_{r+1} + X_r \le a \mid X_r = b] &= \frac{\mathbb{P}[X_{r+1} + X_r \le a \text{ and } X_r = b]}{\mathbb{P}[X_r = b]} \\ &= \frac{\mathbb{P}[X_{r+1} \le a - b \text{ and } X_r = b]}{\mathbb{P}[X_r = b]}.\end{aligned}$$

Since the events $X_{r+1} \le a - b$ and $X_r = b$ are independent, we get

$$\mathbb{P}[X_{r+1} + X_r \le a \mid X_r = b] = \frac{\mathbb{P}[X_{r+1} \le a - b]\,\mathbb{P}[X_r = b]}{\mathbb{P}[X_r = b]} = \mathbb{P}[X_{r+1} \le a - b]\,.$$

Taking $a = 4$ and $b = 2$,

$$\mathbb{P}[X_{r+1} + X_r \le 4 \mid X_r = 2] = \mathbb{P}[X_{r+1} \le 2] = 1 - \left[1 - \left(1 - \frac{r+1}{10}\right)\right]^2 = 1 - \left(\frac{r+1}{10}\right)^2.$$

Part (c). The customer has 0 pictures at first. To get 1 picture, he simply collects one card. Then, to get 2 different pictures, he collects another X_1 cards. To get 3 different pictures, he collects another X_2 cards. This continues until he collects all 10 pictures. Hence,

$$Y = 1 + X_1 + X_2 + \cdots + X_9.$$

The expectation of Y is hence

$$\mathbb{E}[Y] = \mathbb{E}[1 + X_1 + \cdots + X_9] = 1 + \sum_{r=1}^{9} \mathbb{E}[X_r] = 1 + \sum_{r=1}^{9} \frac{10}{10 - r} = 29.3 \text{ (3 s.f.)}.$$

A15A Continuous Random Variables

Tutorial A15A

Problem 1. The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{k}{1+x^2}, & -1 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Find the value of k and determine

(a) the mode of X;

(b) the expectation (mean) and variance of X;

(c) F, the cumulative density function of X and the median of X.

Solution. Since the sum of probabilities is 1, we have

$$1 = \int_{-\infty}^{\infty} f(x) \, \mathrm{d}x = \int_{-1}^{1} \frac{k}{1+x^2} \, \mathrm{d}x = k \left[\arctan x\right]_{-1}^{1} = \frac{\pi}{2} k,$$

whence $k = 2/\pi$.

Part (a). From the graph of $y = f(x)$, it is clear that $f(x)$ attains a maximum at $x = 0$, whence the mode of X is 0.

Part (b). Since $f(x)$ is symmetric about $x = 0$, we clearly have $\mathbb{E}[X] = 0$. Hence,

$$\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f(x) \, \mathrm{d}x = \frac{2}{\pi} \int_{-1}^{1} \frac{x^2}{1+x^2} \, \mathrm{d}x = 0.273 \text{ (3 s.f.)}.$$

Part (c). Note that

$$\int_{-1}^{x} f(t) \, \mathrm{d}t = \frac{2}{\pi} \int_{-1}^{x} \frac{\mathrm{d}t}{1+t^2} = \frac{2}{\pi} \left[\arctan t\right]_{-1}^{x} = \frac{2}{\pi} \arctan x + \frac{1}{2}.$$

Hence, $F(x)$, the cdf of X, is given by

$$F(x) = \begin{cases} 0, & x < -1, \\ \frac{2}{\pi} \arctan x + \frac{1}{2}, & -1 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

Since $f(x)$ is symmetric about $x = 0$, the median of X is 0.

Problem 2. X is a continuous random variable, taking values in the interval $0 < x \leq 1$, whose probability density function is given by $f(x) = 2(1-x)$. Calculate the expectations of X, $2X + 1$, and of X^3.

Solution. We have

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_0^1 2x(1-x)\,\mathrm{d}x = 0.333 \text{ (3 s.f.)}.\\
\mathbb{E}[2X+1] &= \int_{-\infty}^{\infty} (2x+1) f(x)\,\mathrm{d}x = \int_0^1 2(2x+1)(1-x)\,\mathrm{d}x = 1.67 \text{ (3 s.f.)}.\\
\mathbb{E}\big[X^3\big] &= \int_{-\infty}^{\infty} x^3 f(x)\,\mathrm{d}x = \int_0^1 2x^3(1-x)\,\mathrm{d}x = 0.1.
\end{aligned}$$

* * * * *

Problem 3. The continuous random variable X has cumulative distribution function F given by

$$F(x) = \begin{cases} 0, & x \leq 0, \\ \sqrt{x}, & 0 < x < 1, \\ 1, & x \geq 1. \end{cases}$$

Find the median of X. The probability density function of X is f. Write down an expression for $f(x)$ for $0 < x < 1$. Hence,

(a) show that $\mathbb{E}[X] = \frac{1}{3}$;

(b) find $\mathrm{Var}[X]$.

Show that the mean of $\sqrt{X}$ and the median of $\sqrt{X}$ are equal.

Solution. Let m be the median of X. Then

$$F(m) = \sqrt{m} = \frac{1}{2} \implies m = \frac{1}{4}.$$

Hence, $1/4$ is the median of X.

Differentiating the cdf, we see that the pdf of X is given by

$$f(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & 0 < x < 1, \\ 0, & \text{otherwise}. \end{cases}$$

Part (a). We have

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_0^1 \frac{x}{2\sqrt{x}}\,\mathrm{d}x = \frac{1}{2}\left[\frac{2}{3}x^{3/2}\right]_0^1 = \frac{1}{3}.$$

Part (b). Note that

$$\mathbb{E}\big[X^2\big] = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \int_0^1 \frac{x^2}{2\sqrt{x}}\,\mathrm{d}x = \frac{1}{2}\left[\frac{2}{5}x^{2/5}\right]_0^1 = \frac{1}{5}.$$

Hence,

$$\mathrm{Var}[X] = \mathbb{E}\big[X^2\big] - \mathbb{E}[X]^2 = \frac{1}{5} - \left(\frac{1}{3}\right)^2 = \frac{4}{45}.$$

Note that

$$\mathbb{E}\Big[\sqrt{X}\Big] = \int_{-\infty}^{\infty} \sqrt{x} f(x)\,\mathrm{d}x = \int_0^1 \frac{\sqrt{x}}{2\sqrt{x}}\,\mathrm{d}x = \frac{1}{2}.$$

Let s be the median of $\sqrt{X}$. Then

$$\frac{1}{2} = \mathbb{P}\left[\sqrt{X} < s\right] = \mathbb{P}\left[X < s^2\right] = \sqrt{s^2} = s.$$

Thus, both the mean and median of $\sqrt{X}$ are equal to 1/2.

* * * * *

Problem 4. The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} kx, & 0 \le x \le 1, \\ kx^2, & 1 < x \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Show that $k = 6/17$.

(b) Find the cumulative distribution function of X.

(c) Find, correct to two decimal places, the median, m, of X.

(d) Find, correct to two decimal places, $\mathbb{P}[|X - m| < 0.75]$.

Solution.

Part (a). Since the probabilities sum to 1, we have

$$1 = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = k\int_0^1 x\,\mathrm{d}x + k\int_1^2 x^2\,\mathrm{d}x = k\left[\frac{x^2}{2}\right]_0^1 + k\left[\frac{x^3}{3}\right]_1^2 = \frac{17}{6}k,$$

whence $k = 6/17$.

Part (b). Let F be the cdf of X. For $0 \le x \le 1$,

$$F(x) = \int_0^x \frac{6}{17}t\,\mathrm{d}t = \frac{3}{17}x^2.$$

For $1 < x \le 2$,

$$F(x) = F(1) + \int_1^x \frac{6}{17}t^2\,\mathrm{d}t = \frac{3}{17} + \frac{2}{17}\left(x^3 - 1\right) = \frac{1}{17}\left(2x^3 + 1\right).$$

Thus, F is given by

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{3}{17}x^2, & 0 \le x \le 1, \\ \frac{1}{17}\left(2x^3 + 1\right), & 1 < x \le 2, \\ 1, & x > 2. \end{cases}$$

Part (c). By inspection, $1 < m \le 2$. Hence,

$$\frac{1}{2} = F(m) = \frac{1}{17}\left(2m^3 + 1\right) \implies m = 1.55 \text{ (2 d.p.)}.$$

Part (d). Note that $|X - m| < 0.75$ is equivalent to $0.8 < X < 2.3$. Hence,

$$\mathbb{P}[|X - m| < 0.75] = \mathbb{P}[0.8 < X < 2.3] = F(2.3) - F(0.8) = 0.89 \text{ (2 d.p.)}.$$

Problem 5. The continuous random variable X has probability density function defined by

$$f(x) = \begin{cases} 0, & x < 0, \\ kx, & 0 \leq x \leq 2, \\ \frac{16k}{x^3}, & \text{otherwise.} \end{cases}$$

Calculate the value of k. Find the median value and the expectation of X.

Prove that the standard deviation of X is infinite.

Find the value of a such that $\mathbb{P}[X > a] = 0.005$.

Find the cumulative distribution function of X and sketch its graph.

Solution. Since the sum of probabilities is 1, we have

$$1 = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \int_0^2 kx\,\mathrm{d}x + \int_2^{\infty} \frac{16k}{x^3}\,\mathrm{d}x = \left[\frac{kx^2}{2}\right]_0^2 + \left[\frac{-8k}{x^2}\right]_2^{\infty} = 4k.$$

Hence, $k = 1/4$.

By inspection, the median of X is 2. The expectation of X is given by

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x = \int_0^2 \frac{x^2}{4}\,\mathrm{d}x + \int_2^{\infty} \frac{4}{x^2}\,\mathrm{d}x = 2.67 \text{ (3 s.f.)}.$$

Observe that $\mathbb{E}\left[X^2\right]$ diverges to ∞:

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \int_0^2 \frac{x^3}{4}\,\mathrm{d}x + \int_2^{\infty} \frac{4}{x}\,\mathrm{d}x > \int_2^{\infty} \frac{1}{x}\,\mathrm{d}x \to \infty.$$

Hence, the standard deviation σ of X also diverges to ∞:

$$\sigma = \sqrt{\mathrm{Var}[X]} = \sqrt{\mathbb{E}[X^2] - \mathbb{E}[X]^2} > \sqrt{\mathbb{E}[X^2]} \to \infty.$$

Let F be the cdf of X. We have $F(x) = 0$ for $x < 0$. For $0 \leq x \leq 2$,

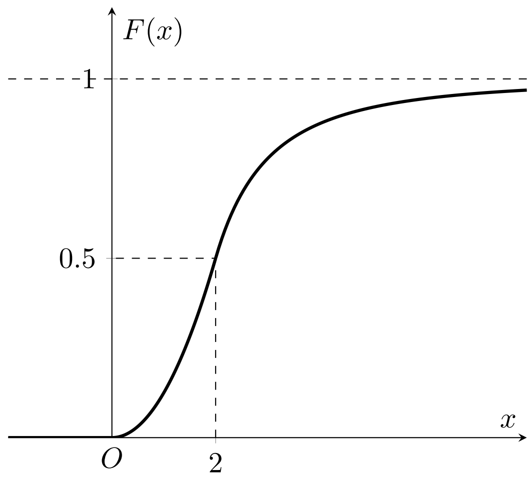
$$F(x) = F(0) + \int_0^x \frac{t}{4}\,\mathrm{d}t = 0 + \left[\frac{t^2}{8}\right]_0^x = \frac{x^2}{8}.$$

For $x > 2$,

$$F(x) = F(2) + \int_2^x \frac{4}{t^3}\,\mathrm{d}t = \frac{2^2}{8} + \left[-\frac{2}{t^2}\right]_2^x = 1 - \frac{2}{x^2}.$$

Hence,

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{x^2}{8}, & 0 \leq x \leq 2, \\ 1 - \frac{2}{x^2}, & x > 2. \end{cases}$$



Consider $\mathbb{P}[X > a] = 0.005$, which is equivalent to $\mathbb{P}[X \leq a] = 0.995$. By inspection, $a > 2$. Hence,

$$\mathbb{P}[X \leq a] = 1 - \frac{2}{a^2} = 0.995 \implies a = 20.$$

* * * * *

Problem 6. The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} cx^2, & 0 \leq x \leq 2, \\ 2c(4-x), & 2 < x \leq 4, \\ 0, & \text{otherwise}, \end{cases}$$

where c is a constant.

(a) Show that $c = 0.15$.

(b) Find the mean of X.

(c) Find the lower quartile of X.

(d) Find the probability that a single observation of X lies between the lower quartile and the mean.

(e) Three independent observations of X are taken. Find the probability that one of the observations is greater than the mean and the other two are less than the median value of X.

Solution.

Part (a). Since the sum of probabilities is 1, we have

$$1 = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \int_0^2 cx^2\,\mathrm{d}x + \int_2^4 2c(4-x)\,\mathrm{d}x = c\left[\frac{x^3}{3}\right]_0^2 + 2c\left[4x - \frac{x^2}{2}\right]_2^4 = \frac{20}{3}c,$$

whence $c = 3/20 = 0.15$.

Part (b). We have

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x = \int_0^2 \frac{3}{20}x^3\,\mathrm{d}x + \int_2^4 \frac{3}{10}x(4-x)\,\mathrm{d}x = 2.2.$$

Part (c). Let F be the cdf of X. Clearly, $F(x) = 0$ for $x < 0$. For $0 \leq x \leq 2$,

$$F(x) = F(0) + \int_0^x \frac{3}{20}t^2\,\mathrm{d}t = \frac{3}{20}\left[\frac{t^3}{3}\right]_0^x = \frac{x^3}{20}.$$

For $2 < x \leq 4$,

$$F(x) = F(2) + \int_2^x \frac{3}{10}(4-t)\,\mathrm{d}t = \frac{2^3}{20} + \frac{3}{10}\left[4t - \frac{t^2}{2}\right]_2^x = \frac{-3x^2 + 24x - 28}{20}.$$

Finally, $F(x) = 1$ for $x > 4$. Thus,

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{20}x^3, & 0 \leq x \leq 2, \\ \frac{1}{20}(-3x^2 + 24x - 28), & 2 < x \leq 4, \\ 1, & x > 4. \end{cases}$$

Let l be the lower quartile of X. Then $\mathbb{P}[X < l] = \frac{1}{4}$. By inspection, $0 \leq l \leq 2$. Hence,

$$\mathbb{P}[X < l] = \frac{l^3}{20} = \frac{1}{4} \implies l = \sqrt[3]{5}.$$

Part (d). We have

$$\mathbb{P}\left[\sqrt[3]{5} < X < 2.2\right] = F(2.2) - F\left(\sqrt[3]{5}\right) = 0.264 \text{ (3 s.f.)}.$$

Part (e). By definition, the probability that an observation of X is less than the median value is $1/2$. Hence, the required probability is simply

$${}^3C_1\, \mathbb{P}[X > 2.2] \left(\frac{1}{2}\right)^2 = {}^3C_1 \left[1 - \mathbb{P}[X < 2.2]\right] \left(\frac{1}{2}\right)^2 = 0.365 \text{ (3 s.f.)}.$$

* * * * *

Problem 7. The cumulative distribution function of a continuous random variable X is given by

$$F(x) = \begin{cases} 0, & x < -2, \\ k\left(4x - \frac{x^3}{3} + \frac{16}{3}\right), & -2 \leq x \leq 2, \\ 1, & x > 2. \end{cases}$$

Find

(a) the value of k;

(b) the probability density function for X;

(c) the mean and variance of X.

Solution.

Part (a). Since F is continuous, we have $F(2) = F(2^+)$, i.e.

$$k\left(4(2) - \frac{2^3}{3} + \frac{16}{3}\right) = 1 \implies k = \frac{3}{32}.$$

Part (b). Note that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{3}{32}\left(4x - \frac{x^3}{3} + \frac{16}{3}\right)\right] = \frac{3}{32}\left(4 - x^2\right).$$

Hence,

$$f(x) = \begin{cases} \frac{3}{32}(4 - x^2), & -2 \leq x \leq 2, \\ 0, & \text{otherwise}. \end{cases}$$

Part (c). Observe that $xf(x)$ is an odd function (symmetric about $x = 0$). Hence, $\mathbb{E}[X] = 0$. Thus,

$$\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \int_{-2}^{2} \frac{3}{32} x^2 \left(4 - x^2\right)\,\mathrm{d}x = 0.8.$$

Problem 8. A continuous random variable X has distribution function given by

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{12}x^2, & 0 \leq x \leq 3, \\ -\frac{1}{4}x^2 + 2x - 3, & 3 < x \leq 4, \\ 1, & x > 4. \end{cases}$$

(a) Show that the median is $\sqrt{6}$.

(b) Find the corresponding probability density function, and show that the mean is $\frac{7}{3}$.

(c) Find the value of k such that $\mathbb{P}[X < k] = \mathbb{P}[3 < X < 4]$.

Solution.

Part (a). Let m be the median of X. By inspection, $0 \leq m \leq 3$. Thus,

$$\mathbb{P}[X < m] = \frac{1}{12}m^2 = \frac{1}{2} \implies m = \sqrt{6}.$$

Part (b). Differentiating F, we get

$$f(x) = \begin{cases} \frac{x}{6}, & 0 \leq x \leq 3, \\ -\frac{x}{2} + 2, & 3 < x \leq 4, \\ 0, & \text{otherwise.} \end{cases}$$

Hence,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) \, \mathrm{d}x = \int_0^3 \frac{x^2}{6} \, \mathrm{d}x + \int_3^4 x \left(-\frac{x}{2} + 2 \right) \, \mathrm{d}x = \left[\frac{x^3}{18} \right]_0^3 + \left[-\frac{x^3}{6} + x^2 \right]_3^4 = \frac{7}{3}.$$

Part (c). Note that

$$\mathbb{P}[3 < X < 4] = F(4) - F(3) = 1 - \frac{3^2}{12} = \frac{1}{4}.$$

Hence, by inspection, $0 \leq k \leq 3$. Thus,

$$\mathbb{P}[X < k] = \frac{k^2}{12} = \frac{1}{4} = \mathbb{P}[3 < X < 4] \implies k = \sqrt{3}.$$

* * * * *

Problem 9. The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{k}{(x+1)^4}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

where k is a constant.

(a) Show that $k = 3$, and find the cumulative distribution function. Find also the value of x such that $\mathbb{P}[X \leq x] = 7/8$.

(b) Find $\mathbb{E}[X + 1]$, and deduce that $\mathbb{E}[X] = 1/2$.

(c) By considering $\mathrm{Var}[X + 1]$, or otherwise, find $\mathrm{Var}[X]$.

Solution.

Part (a). Note that

$$\int \frac{1}{(x+1)^4}\,\mathrm{d}x = -\frac{1}{3(x+1)^3}.$$

Since the probabilities sum to 1,

$$1 = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = k\int_0^{\infty} \frac{1}{(x+1)^4}\,\mathrm{d}x = k\left[-\frac{1}{3(x+1)^3}\right]_0^{\infty} = \frac{k}{3} \implies k = 3.$$

Let F be the cdf of X. Clearly, $F(x) = 0$ for $x < 0$. For $x \geq 0$, we have

$$F(x) = F(0) + \int_0^x \frac{3}{(t+1)^4}\,\mathrm{d}t = 0 + 3\left[-\frac{1}{3(t+1)^3}\right]_0^x = 1 - \frac{1}{(x+1)^3}.$$

Thus,

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - \frac{1}{(x+1)^3}, & x \geq 0. \end{cases}$$

Consider $\mathbb{P}[X \leq x] = 7/8$:

$$\mathbb{P}[X \leq x] = 1 - \frac{1}{(x+1)^3} = \frac{7}{8} \implies x = 1.$$

Part (b). We have

$$\mathbb{E}[X+1] = \int_{-\infty}^{\infty} (x+1)f(x)\,\mathrm{d}x = 3\int_0^{\infty} \frac{1}{(x+1)^3}\,\mathrm{d}x = 3\left[-\frac{1}{2(x+1)^2}\right]_0^{\infty} = \frac{3}{2}.$$

Thus,

$$\mathbb{E}[X] = \mathbb{E}[(X+1)-1] = \mathbb{E}[X+1] - \mathbb{E}[1] = \frac{3}{2} - 1 = \frac{1}{2}.$$

Part (c). Consider $\mathbb{E}\big[(X+1)^2\big]$:

$$\mathbb{E}\big[(X+1)^2\big] = \int_{-\infty}^{\infty} (x+1)^2 f(x)\,\mathrm{d}x = 3\int_0^{\infty} \frac{1}{(x+1)^2}\,\mathrm{d}x = 3\left[-\frac{1}{x+1}\right]_0^{\infty} = 3.$$

Thus,

$$\mathrm{Var}[X] = \mathrm{Var}[X+1] = \mathbb{E}\big[(X+1)^2\big] - \mathbb{E}[X+1]^2 = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4}.$$

* * * * *

Problem 10. The probability that a randomly chosen flight from Stanston Airport is delayed by more than x hours is $\frac{1}{100}(x-10)^2$, $x \in \mathbb{R}$, $0 \leq x \leq 10$. No flights leave early, and none is delayed for more than 10 hours. The delay, in hours, for a randomly chosen flight is denoted by X.

(a) Find the median, m, of X, correct to three significant figures.

(b) Find the cumulative distribution function, F, of X and sketch the graph of F.

(c) Find the probability distribution function, f, of X and sketch the graph of f.

(d) Show that $\mathbb{E}[X] = 10/3$.

A random sample of 2 flights is taken. Find the probability that both flights are delayed by more than m hours, where m is the median of X.

Solution.

Part (a). Note that

$$\mathbb{P}[X > x] = \begin{cases} \frac{1}{100}(x-10)^2, & 0 \leq x \leq 10, \\ 0, & \text{otherwise.} \end{cases}$$

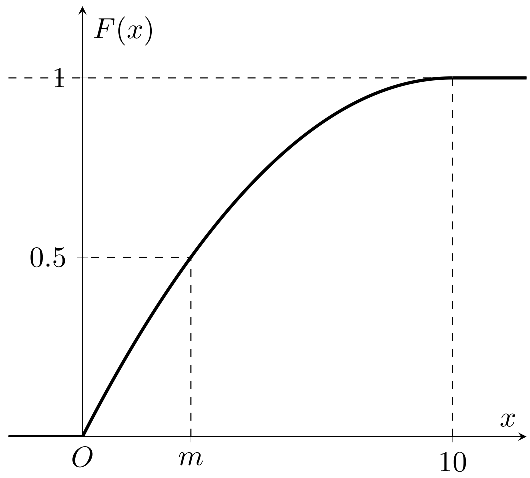
Thus,

$$\frac{1}{2} = \mathbb{P}[X > m] = \frac{1}{100}(m-10)^2 \implies m = 2.93 \text{ (3 s.f.)}.$$

Note that we reject $m = 17.1$ since $0 \leq m \leq 10$.

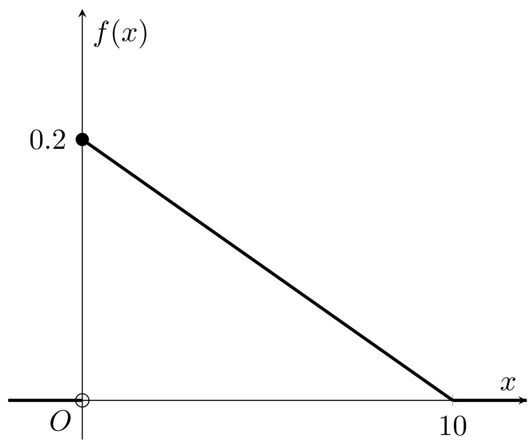
Part (b). We have

$$F(x) = \mathbb{P}[X \leq x] = 1 - \mathbb{P}[X > x] = \begin{cases} 0, & x < 0, \\ 1 - \frac{1}{100}(x-10)^2, & 0 \leq x \leq 10, \\ 1, & x > 10. \end{cases}$$



Part (c). Differentiating F, we get

$$f(x) = \begin{cases} -\frac{1}{50}(x-10), & 0 \leq x \leq 10, \\ 0, & \text{otherwise.} \end{cases}$$



Part (d). We have

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x = -\frac{1}{50}\int_0^{10} x(x-10)\,\mathrm{d}x = -\frac{1}{50}\left[\frac{x^3}{3} - 5x^2\right]_0^{10} = \frac{10}{3}.$$

The required probability is

$$\mathbb{P}[X > m]^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}.$$

* * * * *

Problem 11. The continuous random variable X has a probability density function

$$f(x) = \begin{cases} \frac{2}{\pi}, & 0 \le x \le \frac{\pi}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

The random variable Y is defined by $Y = \cos X$. Find $\mathbb{E}[Y]$, and show that $\mathrm{Var}[Y] = \frac{1}{2} - \frac{4}{\pi^2}$. Find the median of Y.

Solution. We have

$$\mathbb{E}[Y] = \mathbb{E}[\cos X] = \frac{2}{\pi}\int_0^{\pi/2} \cos(x)\;\mathrm{d}x = \frac{2}{\pi}\left[\sin x\right]_0^{\pi/2} = \frac{2}{\pi}.$$

Similarly,

$$\begin{aligned} \mathbb{E}\left[Y^2\right] = \mathbb{E}\left[\cos^2 X\right] &= \frac{2}{\pi}\int_0^{\pi/2} \cos^2(x)\;\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi/2} \frac{1+\cos(2x)}{2}\,\mathrm{d}x \\ &= \frac{1}{\pi}\left[x + \frac{\sin 2x}{2}\right]_0^{\pi/2} = \frac{1}{2}. \end{aligned}$$

Thus,

$$\mathrm{Var}[Y] = \mathbb{E}\left[Y^2\right] - \mathbb{E}[Y]^2 = \frac{1}{2} - \left(\frac{2}{\pi}\right)^2 = \frac{1}{2} - \frac{4}{\pi^2}.$$

Let m be the median of Y. We have

$$\frac{1}{2} = \mathbb{P}[Y < m] = \mathbb{P}[\cos x < m] = \mathbb{P}[x > \arccos m] = 1 - \frac{2}{\pi}\arccos m,$$

whence $m = \cos(\pi/4) = 1/\sqrt{2}$.

* * * * *

Problem 12. A random variable X has probability density function

$$f(x) = \begin{cases} \frac{2}{x^2}, & 1 \le x \le 2, \\ 0, & \text{otherwise,} \end{cases}$$

and the random variable Y is defined by $Y = 4X^3$. Find

(a) the mean and variance of Y;

(b) $\mathbb{P}[10 < Y < 20]$;

(c) the median of Y.

Solution.

Part (a). We have

$$\mathbb{E}[Y] = \mathbb{E}\big[4X^3\big] = \int_1^2 8x \, \mathrm{d}x = \big[4x^2\big]_1^2 = 12$$

and

$$\mathbb{E}\big[Y^2\big] = \mathbb{E}\big[16X^6\big] = \int_1^2 32x^4 \, \mathrm{d}x = \left[\frac{32}{5}x^5\right]_1^2 = \frac{992}{5}.$$

Thus,

$$\mathrm{Var}[Y] = \mathbb{E}\big[Y^2\big] - \mathbb{E}[Y]^2 = \frac{992}{5} - 12^2 = \frac{272}{5}.$$

Part (b). Note that

$$F(x) = \begin{cases} 0, & x < 1, \\ 2 - \frac{2}{x}, & 1 \leq x \leq 2, \\ 1, & x > 2. \end{cases}$$

Thus,

$$\begin{aligned} \mathbb{P}[10 < Y < 20] &= \mathbb{P}\big[10 < 4X^3 < 20\big] = \mathbb{P}\Big[\sqrt[3]{5/2} < X < \sqrt[3]{5}\Big] \\ &= F\Big(\sqrt[3]{5}\Big) - F\Big(\sqrt[3]{5/2}\Big) = 0.304 \text{ (3 s.f.)}. \end{aligned}$$

Part (c). Let m be the median of Y. We have

$$\frac{1}{2} = \mathbb{P}[Y < m] = \mathbb{P}\big[4X^3 < m\big] = \mathbb{P}\Big[X < \sqrt[3]{m/4}\Big] = 2 - \frac{2}{\sqrt[3]{m/4}}.$$

Solving, we get $m = 9.48$ (3 s.f.).

Assignment A15A

Problem 1. The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} a, & 0 \leq x \leq 1, \\ b, & 1 < x \leq 3, \\ 0, & \text{otherwise}, \end{cases}$$

where a and b are constants.

(a) If the mean of X is $5/4$, find a and b.

(b) Find the exact value of k such that $\mathbb{P}[X \leq k] = 5/8$.

Ten independent observations of X are taken, and the random variable R is the number of observations such that $X < 1/2$. Find $\mathbb{P}[R > 4]$.

Solution.

Part (a). Since the probabilities must sum to 1,

$$1 = \int_{-\infty}^{\infty} f(x) \, \mathrm{d}x = \int_0^1 a \, \mathrm{d}x + \int_1^3 b \, \mathrm{d}x = a + 2b. \tag{1}$$

Since the mean of X is $5/4$,

$$\frac{5}{4} = \mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) \, \mathrm{d}x = \int_0^1 ax \, \mathrm{d}x + \int_1^3 bx \, \mathrm{d}x = \left[\frac{ax^2}{2}\right]_0^1 + \left[\frac{bx^2}{2}\right]_1^3 = \frac{1}{2}a + 4b. \tag{2}$$

Solving (1) and (2) simultaneously, we get $a = 1/2$ and $b = 1/4$.

Part (b). Let $F(x)$ be the cdf of X. For $x < 0$, we have $F(x) = 0$. For $0 \leq x \leq 1$,

$$F(x) = F(0) + \int_0^x f(t) \, \mathrm{d}t = 0 + \int_0^x \frac{1}{2} \, \mathrm{d}t = \frac{1}{2}x.$$

For $1 < x \leq 3$,

$$F(x) = F(1) + \int_1^x f(t) \, \mathrm{d}t = \frac{1}{2} + \int_1^x \frac{1}{4} \, \mathrm{d}t = \frac{1}{4} + \frac{1}{4}x.$$

For $x > 3$, we have $F(x) = 1$. Thus,

$$F(x) = \begin{cases} 0, & 0 < x, \\ \frac{1}{2}x, & 0 \leq x \leq 1, \\ \frac{1}{4} + \frac{1}{4}x, & 1 < x \leq 3, \\ 1, & x > 3. \end{cases}$$

Note that $F(1) = 1/2 < 5/8$. Thus, $k \in (1, 3]$. Hence,

$$\frac{5}{8} = \mathbb{P}[X \leq k] = \frac{1}{4} + \frac{1}{4}k \implies k = \frac{3}{2}.$$

Since $\mathbb{P}[X < 1/2] = F(1/2) = 1/4$, it follows that $R \sim \mathrm{B}(10, 1/4)$. Using G.C., $\mathbb{P}[R > 4] = 0.0781$ (3 s.f.).

* * * * *

Problem 2. The random variable X is the distance, in metres, that an inexperienced tightrope walker has moved along a given tightrope before falling off. It is given that

$$\mathbb{P}[X > x] = 1 - \frac{1}{64}x^3, \quad 0 \leq x \leq 4.$$

(a) Show that $\mathbb{E}[X] = 3$.

(b) Find σ, the standard deviation of X.

(c) Show that $\mathbb{P}[|X-3| < \sigma] = \frac{69}{80}\sqrt{\frac{3}{5}}$.

Solution.

Part (a). Let $F(x)$ be the cdf of X. Then

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{64}x^3, & 0 \le x \le 4, \\ 1, & x > 4. \end{cases}$$

Let $f(x)$ be the pdf of X. Then

$$f(x) = F'(x) = \begin{cases} \frac{3}{64}x^2, & 0 \le x \le 4, \\ 0, & \text{otherwise}. \end{cases}$$

Thus,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_0^4 \frac{3}{64}x^3\,\mathrm{d}x = \frac{3}{64}\left[\frac{x^4}{4}\right]_0^4 = 3.$$

Part (b). Note that

$$\mathbb{E}\big[X^2\big] = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \int_0^4 \frac{3}{64}x^4\,\mathrm{d}x = \frac{3}{64}\left[\frac{x^5}{5}\right]_0^4 = \frac{48}{5}.$$

Thus,

$$\sigma^2 = \mathrm{Var}[X] = \mathbb{E}\big[X^2\big] - \mathbb{E}[X]^2 = \frac{48}{5} - 3^2 = \frac{3}{5} \implies \sigma = \sqrt{\frac{3}{5}}.$$

Part (c). We have

$$\begin{aligned} \mathbb{P}[|X-3| < \sigma] &= \mathbb{P}[3-\sigma < X < 3+\sigma] = F(3+\sigma) - F(3-\sigma) \\ &= \frac{1}{64}(3+\sigma)^3 - \frac{1}{64}(3-\sigma)^3 = \frac{1}{32}\sigma\left(27+\sigma^2\right) = \frac{69}{80}\sqrt{\frac{3}{5}}. \end{aligned}$$

* * * * *

Problem 3. In the triangle PQR, $PR = 6$ cm, $QR = 10$ cm and $\angle PRQ = y$ radians, where y is uniformly distributed on the interval from 0 to $\pi/2$. The area of triangle PQR is A cm units^2. Find the probability density function of A.

Solution. Since $Y \sim \mathrm{U}(0, \pi/2)$, its cdf is given by

$$\Phi_Y(y) = \begin{cases} 0, & y < 0, \\ \frac{2}{\pi}y, & 0 \le y \le \frac{\pi}{2}, \\ 1, & y > \frac{\pi}{2}. \end{cases}$$

Since

$$A = \frac{1}{2}(PR)(QR)\sin PRQ = \frac{1}{2}(6)(10)\sin y = 30\sin y,$$

we have

$$F_A(a) = \mathbb{P}[A < a] = \mathbb{P}[30\sin y < a] = \mathbb{P}\Big[y < \arcsin\frac{a}{30}\Big].$$

Hence,

$$F_A(a) = \begin{cases} 0, & \arcsin\frac{a}{30} < 0, \\ \frac{2}{\pi}\arcsin\frac{a}{30}, & 0 \le \arcsin\frac{a}{30} \le \frac{\pi}{2}, \\ 1, & \arcsin\frac{a}{30} > \frac{\pi}{2}. \end{cases} = \begin{cases} 0, & a < 0 \\ \frac{2}{\pi}\arcsin\frac{a}{30}, & 0 \le a \le 30, \\ 1, & a > 30. \end{cases}$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}a}\frac{2}{\pi}\arcsin\frac{a}{30} = \frac{2}{\pi}\frac{1}{\sqrt{1-(a/30)^2}}\frac{1}{30} = \frac{2}{\pi\sqrt{30^2-a^2}},$$

the pdf of A is given by

$$f_A(a) = F_A'(a) = \begin{cases} \frac{2}{\pi\sqrt{30^2-a^2}}, & 0 \le a < 30, \\ 0, & \text{otherwise.} \end{cases}$$

A15B Special Continuous Random Variables

Tutorial A15B

Problem 1. A continuous random variable X has a uniform distribution over the interval $[0, n]$. Write down $\mathbb{E}[X]$ and show that $\mathrm{Var}[X] = \frac{1}{12}n^2$. Denoting the expectation and standard deviation by μ and σ respectively, evaluate $\mathbb{P}[|X - \mu| < \sigma]$.

Solution. We have $X \sim \mathrm{U}(0, n)$. Hence, $\mathbb{E}[X] = n/2$. Note that X has pdf

$$f(x) = \begin{cases} \frac{1}{n}, & 0 \leq x \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

Thus,

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f(x) \, \mathrm{d}x = \int_0^{\infty} \frac{x^2}{n} \, \mathrm{d}x = \left[\frac{x^3}{3n}\right]_0^n = \frac{n^2}{3},$$

whence

$$\mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{n^2}{3} - \left(\frac{n}{2}\right)^2 = \frac{n^2}{12}.$$

Thus,

$$\sigma = \sqrt{\mathrm{Var}[X]} = \frac{n}{2\sqrt{3}}.$$

Note that X has cdf

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{n}, & 0 \leq x \leq n, \\ 1, & x > n. \end{cases}$$

Thus,

$$\mathbb{P}[|X - \mu| < \sigma] = \mathbb{P}[\mu - \sigma < X < \mu + \sigma] = \frac{\mu + \sigma}{n} - \frac{\mu - \sigma}{n} = \frac{2\sigma}{n} = \frac{1}{\sqrt{3}}.$$

* * * * *

Problem 2. The continuous random variable X has probability density function defined by

$$f(x) = \begin{cases} k\mathrm{e}^{-\lambda x}, & x \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

where k and λ are positive constants.

(a) Show that $k = \lambda$.

(b) Show that $\mathbb{E}[X] = 1/\lambda$.

(c) Find $\mathrm{Var}[X]$.

(d) Find the median of X.

The random variable X represents the lifetime in hours of a particular brand of torch battery. Show that the probability that a particular battery lasts at least twice as long as the mean lifetime is e^{-2}. Find, to three decimal places, the probability that the lifetime of a particular battery lies between the median and mean lifetimes.

Solution.

Part (a). Since probabilities sum to 1,

$$1 = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \int_0^{\infty} k\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \left[\frac{-k\mathrm{e}^{-\lambda x}}{\lambda}\right]_0^{\infty} = \frac{k}{\lambda} \implies k = \lambda.$$

Part (b). We have

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x = \int_0^{\infty} \lambda x\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \left[-x\mathrm{e}^{-\lambda x} - \frac{1}{\lambda}\mathrm{e}^{-\lambda x}\right]_0^{\infty} = \frac{1}{\lambda}.$$

Part (c). We have

$$\begin{aligned}\mathbb{E}\left[X^2\right] &= \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \int_0^{\infty} \lambda x^2\mathrm{e}^{-\lambda x}\,\mathrm{d}x \\ &= \left[-x^2\mathrm{e}^{-\lambda x}\right]_0^{\infty} + \frac{2}{\lambda}\int_0^{\infty} \lambda x\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{2}{\lambda}\,\mathbb{E}[X] = \frac{2}{\lambda^2}.\end{aligned}$$

Thus,

$$\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

Part (d). Let F be the cdf of X. For $x \geq 0$,

$$F(x) = \int_0^x f(t)\,\mathrm{d}t = \int_0^x \lambda\mathrm{e}^{-\lambda t}\,\mathrm{d}t = \left[-\mathrm{e}^{-\lambda t}\right]_0^x = 1 - \mathrm{e}^{-\lambda x}.$$

Let m be the median. Then

$$\frac{1}{2} = F(m) = 1 - \mathrm{e}^{-\lambda m} \implies m = \frac{\ln 2}{\lambda}.$$

* * * * *

Problem 3. Given that $X \sim \mathrm{N}(21.5, 7)$, evaluate

(a) $\mathbb{P}[18.7 < X \leq 24.5]$

(b) $\mathbb{P}[X < 21.5]$

(c) $\mathbb{P}[X \geq 23]$

(d) $\mathbb{P}[|X - 21.5| < 6]$

(e) $\mathbb{P}[|X - 21.5| > 4.5]$

Solution.

Part (a). $\mathbb{P}[18.7 < X \leq 24.5] = 0.727$ (3 s.f.).

Part (b). $\mathbb{P}[X < 21.5] = 0.5$.

Part (c). $\mathbb{P}[X \geq 23] = 1 - \mathbb{P}[X < 23] = 0.285$ (3 s.f.).

Part (d). $\mathbb{P}[|X - 21.5| < 6] = \mathbb{P}[15.5 < X < 27.5] = 0.977$ (3 s.f.).

Part (e). $\mathbb{P}[|X - 21.5| > 4.5] = 1 - \mathbb{P}[17 < X < 26] = 0.0890$ (3 s.f.).

Problem 4. If $X \sim \mathrm{N}(33, 10)$, find the value (or range of values) of a such that

(a) $\mathbb{P}[X < a] = 0.14$

(b) $\mathbb{P}[X > a] = 0.5$

(c) $\mathbb{P}[X \geq a] > 0.2$

Find the value (or range of values) of c such that

(a) $\mathbb{P}[Z < c] = 0.86$

(b) $\mathbb{P}[|Z| > c] = 0.05$

(c) $\mathbb{P}[|Z| \leq c] < 0.7$

Solution. Using G.C.,
Part (a). $\mathbb{P}[X < a] = 0.14 \implies a = 29.6$.
Part (b). $\mathbb{P}[X > a] = 0.5 \implies a = 33$.
Part (c). $\mathbb{P}[X \geq a] > 0.2 \implies a < 35.7$.
Part (d). $\mathbb{P}[Z < c] = 0.86 \implies c = 1.08$.
Part (e). $\mathbb{P}[|Z| > c] = 0.05 \implies \mathbb{P}[Z > c] = 0.025 \implies c = 1.96$.
Part (f). $\mathbb{P}[|Z| \leq c] < 0.7 \implies \mathbb{P}[0 \leq Z \leq c] < 0.35 \implies \mathbb{P}[Z \leq c] < 0.85 \implies c < 1.04$.

* * * * *

Problem 5. The continuous random variable X has a uniform distribution over the interval $0 \leq x \leq 1$. Write down $\mathbb{E}[X]$ and $\mathrm{Var}[X]$. The random variable Y is defined by $Y = \mathrm{e}^{-X}$. By considering $\mathbb{P}[Y \leq y]$, obtain the cumulative density function of Y. Hence, find $\mathbb{E}[Y]$ and $\mathrm{Var}[Y]$, leaving your answers correct to 2 decimal places.

Solution. Since $X \mathrm{U}(0, 1)$, we clearly have $\mathbb{E}[X] = 0.5$ and $\mathrm{Var}[X] = \frac{1}{12}$.

Since

$$F_X(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x \leq 1, \\ 1, & x > 1, \end{cases}$$

we have

$$\begin{aligned} F_Y(y) = \mathbb{P}[Y \leq y] &= \mathbb{P}\left[\mathrm{e}^{-X} \leq y\right] = \mathbb{P}[X \geq -\ln y] = 1 - \mathbb{P}[X \leq -\ln y] \\ &= \begin{cases} 1, & -\ln y < 0, \\ 1 + \ln y, & 0 \leq -\ln y \leq 1, \\ 0, & -\ln y > 1, \end{cases} = \begin{cases} 1, & y > 1, \\ 1 + \ln y, & \mathrm{e}^{-1} \leq y \leq 1, \\ 0, & y < \mathrm{e}^{-1}. \end{cases} \end{aligned}$$

Differentiating, we obtain the pdf of y:

$$f_Y(y) = \begin{cases} \frac{1}{y}, & \mathrm{e}^{-1} \leq y \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Thus,

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y) \,\mathrm{d}y = \int_{\mathrm{e}^{-1}}^{1} 1 \,\mathrm{d}y = 0.63212 = 0.63 \text{ (2 d.p.)}.$$

Also, we have

$$\mathbb{E}\big[Y^2\big] = \int_{-\infty}^{\infty} y^2 f_Y(y)\,\mathrm{d}y = \int_{\mathrm{e}^{-1}}^{1} y\,\mathrm{d}y = 0.43233.$$

Thus,

$$\operatorname{Var}[Y] = \mathbb{E}\big[Y^2\big] - \mathbb{E}[Y]^2 = 0.63212 - 0.43233^2 = 0.03 \text{ (2 d.p.)}.$$

* * * * *

Problem 6. A semicircular arc, with centre O and radius a, is drawn with AB as diameter. A point Q is taken at random on this arc, such that the angle $BOQ = \theta$ has the rectangular distribution between 0 and π. N is the point on the line segment AB such that QN is perpendicular to AB.

(a) Calculate, in terms of a, the mean and standard deviation of the length of QN.

(b) Find the probability that QN is longer than $a/2$.

Solution.

Part (a). Note that θ has pdf

$$f(\theta) = \begin{cases} \frac{1}{\pi}, & 0 \le \theta \le \pi, \\ 0, & \text{otherwise.} \end{cases}$$

Since $QN = a\sin\theta$, we have

$$\mathbb{E}[QN] = \int_0^\pi (a\sin\theta)\left(\frac{1}{\pi}\right)\mathrm{d}\theta = \frac{a}{\pi}\left[-\cos\theta\right]_0^\pi = \frac{2a}{\pi}.$$

Further,

$$\mathbb{E}\big[QN^2\big] = \int_0^\pi (a\sin\theta)^2\left(\frac{1}{\pi}\right)\mathrm{d}\theta = \frac{a^2}{\pi}\int_0^\pi \frac{1-\cos 2\theta}{2}\,\mathrm{d}t = \frac{a^2}{2\pi}\left[\theta - \frac{\sin 2\theta}{2}\right]_0^\pi = \frac{a^2}{2}.$$

Thus,

$$\operatorname{Var}[QN] = \mathbb{E}\big[QN^2\big] - \mathbb{E}[QN]^2 = \frac{a^2}{2} - \left(\frac{2a}{\pi}\right)^2 = \frac{a^2}{\pi^2}\frac{\pi^2-8}{2}.$$

The standard deviation σ of QN is thus

$$\sigma = \sqrt{\operatorname{Var}[QN]} = \frac{a}{\pi}\sqrt{\frac{\pi^2-8}{2}}.$$

Part (b). Observe that if $QN > a/2$, then $\sin\theta > 1/2$, whence $\pi/6 < \theta < 5\pi/6$. Since θ is uniformly distributed, we have

$$\mathbb{P}\left[QN > \frac{a}{2}\right] = \frac{5\pi/6 - \pi/6}{\pi} = \frac{2}{3}.$$

* * * * *

Problem 7. The object distance U and the image distance V for a concave mirror are related to the focal distance f by the formula

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f},$$

where f is a constant. U is a random variable uniformly distributed over the interval $(2f, 3f)$. Show that V is distributed with probability density function

$$\frac{f}{(v-f)^2}$$

and state the range of corresponding values for V. Obtain the mean and median of V.

Solution. Since $U \sim \mathrm{U}(2f, 3f)$, we have $u - f > 0$ and

$$F_U(u) = \begin{cases} 0, & u < 2f, \\ \frac{u}{f} - 2, & 2f \leq u \leq 3f, \\ 1, & u > 3f. \end{cases}$$

Note also that

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \implies v = \frac{uf}{u-f}.$$

Thus,

$$\begin{aligned} F_V(v) = \mathbb{P}[V < v] &= \mathbb{P}\left[\frac{uf}{u-f} < v\right] = \mathbb{P}\left[u > \frac{vf}{v-f}\right] = 1 - \mathbb{P}\left[u < \frac{vf}{v-f}\right] \\ &= \begin{cases} 1, & \frac{vf}{v-f} < 2f, \\ 3 - \frac{v}{v-f}, & 2f \leq \frac{vf}{v-f} \leq 3f, \\ 0, & \frac{vf}{v-f} > 3f, \end{cases} = \begin{cases} 1, & v > 2f, \\ 3 - \frac{v}{v-f}, & \frac{3}{2}f \leq v \leq 2f, \\ 0, & v < \frac{3}{2}f. \end{cases} \end{aligned}$$

Differentiating, we obtain the pdf of V:

$$f_V(v) = F_V'(v) = \begin{cases} \frac{f}{(v-f)^2}, & \frac{3}{2}f \leq v \leq 2f, \\ 0, & \text{otherwise.} \end{cases}$$

The range of V is hence $[\frac{3}{2}f, 2f]$.

We have

$$\mathbb{E}[V] = \int_{-\infty}^{\infty} v f_V(v) \, \mathrm{d}v = \int_{3f/2}^{2f} \frac{vf}{(v-f)^2} \, \mathrm{d}v.$$

Consider the substitution $w = v - f$. The integral transforms as

$$\mathbb{E}[V] = \int_{f/2}^{f} \frac{(w+f)f}{w^2} \, \mathrm{d}w = \int_{f/2}^{f} \left(\frac{f}{w} + \frac{f^2}{w^2}\right) \mathrm{d}w = \left[f \ln w - \frac{f^2}{w}\right]_{f/2}^{f} = f + f \ln 2.$$

Let m be the median. We have

$$\frac{1}{2} = F_V(m) = 3 - \frac{m}{m-f} \implies m = \frac{5f}{3}.$$

* * * * *

Problem 8. The lifetime, T hours, of a certain kind of lamp has probability density function

$$f(t) = \begin{cases} \frac{1}{a}\mathrm{e}^{-t/b}, & t \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

where a and b are positive constants. Show that $a = b$.

Given that 42.6% of lamps have a lifetime longer than 2000 hours, calculate the common value of a and b, correct to 3 significant figures.

Find the (cumulative) distribution function of T and hence prove that

$$\mathbb{P}[T > t + c \mid T > c] = \mathbb{P}[T > t],$$

where $t \geq 0$ and c is a positive constant.

Three of the lamps are fitted in a laboratory. One of the lamps is turned on end is still working 200 hours later. At this time the other two lamps are turned on. Calculate the probability that after a further 480 hours

(a) all three lamps are working,

(b) just one lamp is working.

Solution. Since probabilities sum to 1,

$$1 = \int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = \int_0^{\infty} \frac{1}{a}\mathrm{e}^{-t/b}\,\mathrm{d}t = \left[-\frac{b}{a}\mathrm{e}^{-t/b}\right]_0^{\infty} = \frac{b}{a} \implies a = b.$$

Note that the cdf of T is given by

$$F(t) = \begin{cases} 0, & t < 0, \\ 1 - \mathrm{e}^{-t/a}, & t \geq 0. \end{cases}$$

We are given that $\mathbb{P}[T > 2000] = 0.426$. Thus,

$$0.574 = \mathbb{P}[T \leq 2000] = 1 - \mathrm{e}^{-2000/a} \implies a = 2340 \text{ (3 s.f.)}.$$

Note that $\mathbb{P}[T > t] = \mathrm{e}^{-t/a}$. Thus,

$$\begin{aligned} \mathbb{P}[T > t + c \mid T > c] &= \frac{\mathbb{P}[T > t + c \text{ and } T > c]}{\mathbb{P}[T > c]} \\ &= \frac{\mathbb{P}[T > t + c]}{\mathbb{P}[T > c]} = \frac{\mathrm{e}^{-(t+c)/a}}{\mathrm{e}^{-c/a}} = \mathrm{e}^{-t/a} = \mathbb{P}[T > t] \end{aligned}$$

Part (a). The probability that all three lamps are working is given by

$$[\mathbb{P}[T > 480]]^3 = \left(\mathrm{e}^{-480/2340}\right)^3 = 0.540 \text{ (3 s.f.)}.$$

Part (b). The probability that only one lamp is working is

$$^3C_1\, \mathbb{P}[T > 480]\,[1 - \mathbb{P}[T < 480]]^2 = 3\mathrm{e}^{-480/2340}\left(1 - \mathrm{e}^{-480/2340}\right)^2 = 0.0840.$$

* * * * *

Problem 9. The working life, T, in hours, of a drill used in tunnelling machinery is a random variable with probability density function defined as

$$f(t) = \begin{cases} \mu \mathrm{e}^{-\mu t}, & t > 0, \\ 0, & \text{otherwise}, \end{cases}$$

where μ is a positive constant.

(a) If the mean life is 20 hours, show that $\mu = 0.05$.

(b) Drilling is planned to take place continuously for one six-hour shift each day. If a new drill is used for each shift, what is the probability that it will fail during the shift?

(c) How long should the shift be to yield a probability of 0.8 for the drill not to fail?

(d) If a drill fails while in use, half an hours' drilling time is lost while it is repaired, but if it fails during the last hour of the shift, drilling is abandoned for the day. The cost of any time loss during shifts is at a rate of \$10,000 per hour. Find the expected cost per six-hour shift of lost drilling time. (Ignore the possibility of a replacement drill also failing during the shift.)

Solution.

Part (a). We have

$$20 = \mathbb{E}[T] = \int_{-\infty}^{\infty} tf(t)\,\mathrm{d}t = \int_{0}^{\infty} \mu t\mathrm{e}^{-\mu t}\,\mathrm{d}t = \left[-t\mathrm{e}^{-\mu t} - \frac{1}{\mu}\mathrm{e}^{-\mu t}\right]_0^{\infty} = \frac{1}{\mu} \implies \mu = 0.05.$$

Part (b). Note that the cdf of T is given by

$$F(t) = \begin{cases} 0, & t \leq 0, \\ 1 - \mathrm{e}^{-t/20}, & t > 0. \end{cases}$$

Thus, the probability that the drill fails during the six-hour shift is

$$\mathbb{P}[T < 6] = 1 - \mathrm{e}^{-6/20} = 0.259 \text{ (3 s.f.)}.$$

Part (c). Let t be the required time. Then

$$\mathbb{P}[T \geq t] = 0.8 \implies \mathbb{P}[T < t] = 0.2 \implies 1 - \mathrm{e}^{-t/20} = 0.2 \implies t = 4.46.$$

Thus, the shift should be 4.46 hours long.

Part (d). Let W_1 be the time wasted (measured in hours) in the first 5 hours of the shift, and let W_2 be the time wasted (measured in hours) in the last hour of the shift.

We have

$$\mathbb{E}[W_1] = \frac{1}{2}\,\mathbb{P}[T < 5] = \frac{1}{2}\left(1 - \mathrm{e}^{-5/20}\right) = 0.11059961 \text{ (8 d.p.)}.$$

Now consider W_2. Let $t \in [5, 6]$ be the total number of hours elapsed. If the drill fails at time t, then the rest of the day is wasted, i.e. $6 - t$ hours are wasted. This gives

$$\mathbb{E}[W_2] = \int_5^6 (6 - t)\,\mathbb{P}[T < t]\ \mathrm{d}t = \int_5^6 (6 - t)\frac{\mathrm{e}^{-t/20}}{20}\,\mathrm{d}t = 0.01914954 \text{ (8 d.p.)}.$$

Thus, the total expected time wasted is

$$\mathbb{E}[W_1] + \mathbb{E}[W_2] = 0.11059961 + 0.01914954 = 0.129749 \text{ (6 s.f.)}.$$

The expected cost due to wasted time is hence $10000 \cdot 0.129749 = \$1297.49$.

* * * * *

Problem 10. The lifetime, T years, before a particular type of washing machine breaks down may be taken to have the probability density function f given by

$$f(t) = \begin{cases} at\mathrm{e}^{-bt}, & t > 0, \\ 0, & \text{otherwise}, \end{cases}$$

where a and b are positive constants. It may be assumed that, if n is a positive integer,

$$\int_0^{\infty} t^n \mathrm{e}^{-bt}\,\mathrm{d}t = \frac{n!}{b^{n+1}}.$$

(a) Records show that the mean of T is 1.5. Show that $b = 4/3$ and find the value of a.

(b) Find $\mathrm{Var}[T]$.

(c) Calculate $\mathbb{P}[T < 1.5]$. State, giving a reason, whether this value indicates that the median of T is smaller than the mean of T or greater than the mean of T.

Solution.

Part (a). Since probabilities sum to 1,

$$1 = \int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = \int_0^{\infty} ate^{-bt}\,\mathrm{d}t = a\frac{1!}{b^{1+1}} = \frac{a}{b^2} \implies a = b^2. \tag{1}$$

Since the mean of T is 1.5,

$$1.5 = \int_{-\infty}^{\infty} tf(t)\,\mathrm{d}t = \int_0^{\infty} at^2\mathrm{e}^{-bt}\,\mathrm{d}t = a\frac{2!}{b^{2+1}} = \frac{2a}{b^3} \implies a = \frac{3}{4}b^3. \tag{2}$$

Solving (1) and (2) simultaneously, we get $a = 16/9$ and $b = 4/3$.

Part (b). Note that

$$\mathbb{E}\left[T^2\right] = \int_{-\infty}^{\infty} t^2 f(t)\,\mathrm{d}t = \int_0^{\infty} at^3\mathrm{e}^{-bt}\,\mathrm{d}t = a\frac{3!}{b^{3+1}} = \frac{6a}{b^4} = \frac{6(16/9)}{(4/3)^4} = \frac{27}{8}.$$

Thus,

$$\mathrm{Var}[T] = \mathbb{E}\left[T^2\right] - \mathbb{E}[T]^2 = \frac{27}{8} - 1.5^2 = \frac{9}{8}.$$

Part (c). We have

$$\mathbb{P}[T < 1.5] = \int_{-\infty}^{1.5} f(t)\,\mathrm{d}t = \int_0^{1.5} \frac{16}{9}t\mathrm{e}^{-4t/3}\,\mathrm{d}t = 0.594.$$

Thus, $\mathbb{P}[T < \mu] = 0.594 > 0.5 = \mathbb{P}[T < m]$, whence the median m is smaller than the mean μ.

* * * * *

Problem 11. X and Y are continuous random variables having independent normal distributions. The means of X and Y are 10 and 12 respectively, and the standard deviations are 2 and 3 respectively. Find

(a) $\mathbb{P}[Y < 10]$,

(b) $\mathbb{P}[Y < X]$,

(c) $\mathbb{P}[4X + 5Y > 90]$,

(d) the value of a such that $\mathbb{P}[X_1 + X_2 > a] = 1/4$, where X_1 and X_2 are independent observations of X.

Solution. Note that $X \sim \mathrm{N}(10, 4)$ and $Y \sim \mathrm{N}(12, 9)$.

Part (a). $\mathbb{P}[Y \leq 10] = 0.252$ (3 s.f.).

Part (b). Note that $Y - X \sim \mathrm{N}(12 - 10, 4 + 9) = \mathrm{N}(10, 13)$. Thus,

$$\mathbb{P}[Y < X] = \mathbb{P}[Y - X < 0] = 0.290 \text{ (3 s.f.)}.$$

Part (c). Note that $4X + 5Y \sim \mathrm{N}\big(4(10) + 5(12), 4^2(4) + 5^2(9)\big) = \mathrm{N}(100, 289)$. Thus,

$$\mathbb{P}[4X + 5Y > 90] = 0.722 \text{ (3 s.f.)}.$$

Part (d). Note that $X_1 + X_2 \sim \mathrm{N}(2(10), 2(4)) = \mathrm{N}(20, 8)$. Thus,

$$\mathbb{P}[X_1 + X_2 > a] = \frac{1}{4} \implies a = 21.9 \text{ (3 s.f.)}.$$

Problem 12. The weights of Sunny brand oranges are normally distributed with mean μ grams and standard deviation σ grams respectively. An inspection of a shipment of Sunny brand oranges shows that 37% of the oranges have weights exceeding 379 grams and 40% of the oranges have weights between 366 grams and 379 grams. Find μ and σ.

Three oranges are selected at random. Find the probability that one orange has weight exceeding 379 grams and two oranges have weights between 366 grams and 379 grams.

Solution. Let W g be the weight of a Sunny brand orange. We are given

$$\mathbb{P}[W > 379] = 0.37 \implies \mathbb{P}[W \leq 379] = 1 - 0.37 = 0.63.$$

Normalizing this, we get

$$z = \frac{x-\mu}{\sigma} \implies 0.33185 = \frac{379-\mu}{\sigma} \implies \mu + 0.33185\sigma = 379. \tag{1}$$

We are also given

$$\mathbb{P}[366 < W < 379] = 0.40 \implies \mathbb{P}[W \leq 366] = 1 - \mathbb{P}[W > 366] = 1 - (0.40 + 0.37) = 0.23.$$

Normalizing this, we get

$$z = \frac{x-\mu}{\sigma} \implies -0.73885 = \frac{366-\mu}{\sigma} \implies \mu - 0.73885\sigma = 366. \tag{2}$$

Solving (1) and (2) simultaneously, we get $\mu = 375$ (3 s.f.) and $\sigma = 12.1$ (3 s.f.).

The required probability is given by

$$^3C_1\, \mathbb{P}[W > 379]\, [\mathbb{P}[366 < W < 379]]^2 = 3(0.37)(0.40)^2 = 0.178 \text{ (3 s.f.)}.$$

* * * * *

Problem 13. A shopper buys two kinds of vegetables in a shop. The mass of potatoes and the mass of onions bought are modelled as having independent normal distributions with the following means and standard deviations.

	Mean	Standard deviation
Mass of potatoes	3 kg	0.2 kg
Mass of onions	1 kg	0.05 kg

The price of potatoes is 50 cents a kilogram and the price of onions is \$1.20 a kilogram.

(a) Find the mean of the total cost of the vegetables and show that the standard deviation of the total cost is \$0.117, correct to 3 significant figures.

(b) Find the probability that the total cost of the vegetables lies between \$2.50 and \$2.80.

Solution.

Part (a). Let the random variables P kg and O kg be the mass of potatoes and onions respectively. We have $P \sim \mathrm{N}\big(3, 0.2^2\big)$ and $O \sim \mathrm{N}\big(1, 0.05^2\big)$.

Let the random variable $T = 0.50P + 1.20O$ be the total cost of the vegetables. Then

$$T \sim \mathrm{N}\big(0.50(3) + 1.20(1), 0.50^2\left(0.20^2\right) + 1.20^2\left(0.05^2\right)\big) = \mathrm{N}(2.7, 0.0136).$$

Thus, the mean of the total cost of the vegetables is \$2.7 and the standard deviation is $\sqrt{0.0136} = \$0.117$.

Part (b). $\mathbb{P}[2.50 < T < 2.80] = 0.761$ (3 s.f.).

* * * * *

Problem 14. It is given that $X \sim \mathrm{N}\big(\mu, \sigma^2\big)$ and $\mathbb{P}[X < 1] = \mathbb{P}[X > 9]$. Write down the value of μ. It is also given that $2\,\mathbb{P}[X < 2] = \mathbb{P}[X < 8]$. Find σ.

Three observations of X are taken. Determine the probability that two will be more than 7 and the other will be between 3 and 5 inclusive.

Solution. Clearly, $\mu = 5$. Using G.C., $\sigma = 6.96$ (3 s.f.).

The required probability is

$$^3C_1\,[\mathbb{P}[X > 7]]^2\,\mathbb{P}[3 \leq X \leq 5] = 0.0508 \text{ (3 s.f.)}.$$

* * * * *

Problem 15. The thickness in cm of a mechanics textbook is a random variable with the distribution $\mathrm{N}\big(2.5, 0.1^2\big)$.

(a) The mean thickness of n randomly chosen mechanics textbooks is denoted by $\overline{M}$ cm. Given that $\mathbb{P}\big[\overline{M} > 2.53\big] = 0.0668$, find the value of n.

The thickness in cm of a statistics textbook is a random variable with the distribution $\mathrm{N}\big(2.0, 0.08^2\big)$.

(b) Calculate the probability that 21 mechanics textbooks and 24 statistics textbooks will fit onto a bookshelf of length 1 m. State clearly the mean and variance of any normal distribution you use in your calculation.

(c) Calculate the probability that the total thickness of 4 statistics textbooks is less than three times the thickness of 1 mechanics textbook. State clearly the mean and variance of any normal distribution you use in your calculation.

(d) State an assumption needed for your calculations in parts (ii) and (iii).

Solution.

Part (a). Note that

$$\overline{M} = \frac{M_1 + M_2 + \cdots + M_n}{n} \sim \mathrm{N}\left(2.5, \frac{0.1^2}{n}\right).$$

Using G.C.,

$$\mathbb{P}\big[\overline{M} > 2.53\big] = 0.0668 \implies n = 25.$$

Part (b). Let the random variable T cm be the total length of 21 mechanics and 24 statistics textbooks. Then

$$T \sim \mathrm{N}\big(21(2.5) + 24(2.0), 21\left(0.1^2\right) + 24\left(0.08^2\right)\big) = \mathrm{N}(100.5, 0.3636).$$

Thus, the probability that the textbooks will fit on the shelf is

$$\mathbb{P}[T < 100] = 0.203 \text{ (3 s.f.)}.$$

Part (c). Let the random variable D cm be the difference between the total thickness of 4 statistics textbooks and 3 times the thickness of 1 mechanics textbook. Then

$$D = S_1 + \cdots + S_4 - 3M = \mathrm{N}\big(4(2.0) - 3(2.5), 4\left(0.08^2\right) + 3^2\left(0.1^2\right)\big) = \mathrm{N}(0.5, 0.1156).$$

The required probability is thus

$$\mathbb{P}[D < 0] = 0.0707 \text{ (3 s.f.)}.$$

Part (d). The thickness of the mechanics and statistics textbooks are independent of each other.

* * * * *

Problem 16. The lengths of metal rods in a box are normally distributed with mean 1.3 m and variance 0.7 m^2.

(a) Find the greatest length l for which the probability that a randomly chosen metal rod is shorter than l m is less than 0.3.

(b) Eight metal rods are chosen at random. Determine the probability that at least three are longer than 1.4 m.

(c) A random sample of 100 metal rods is selected. Find the expected number of metal rods with lengths that are between 1.2 m and 1.6 m.

Solution.

Part (a). Let the random variable L be the length of a metal rod. Then $L \sim \mathrm{N}(1.3, 0.7)$. Hence, if $\mathbb{P}[L < l] < 0.3$, then $l < 0.861$, whence the largest possible l is 0.861 m.

Part (b). Let the number of rods longer than 1.4 m be X. Note that $\mathbb{P}[L > 1.4] = 0.45243$. Thus, $X \sim \mathrm{B}(8, 0.45243)$, and

$$\mathbb{P}[X \geq 3] = 1 - \mathbb{P}[X \leq 2] = 0.784 \text{ (3 s.f.)}.$$

Part (c). Let the number of rods whose lengths are between 1.2 m and 1.6 m be Y. Note that $\mathbb{P}[1.2 < L < 1.6] = 0.18761$. Thus, $Y \sim \mathrm{B}(100, 0.18761)$, whence

$$\mathbb{E}[Y] = (100)(0.18761) = 18.8 \text{ (3 s.f.)}.$$

* * * * *

Problem 17. An ice-cream shop provides two types of paper cups, regular or large, for its customers. Each customer picks a cup according to his appetite and fills it with ice-cream of flavours of his choice. The mass of each cup together with its ice-cream content is measured at the cashier with a weighing machine, and the customer is charged at a rate of \$2 per 100g measured.

Let X and Y be the respective mass, in grams, of the regular and large cups with their ice-cream content. It is found that both X and Y independently follows a normal distribution, with parameters given in the table below:

	Mean	Standard Deviation
X	200	30
Y	350	60

A family of six, consisting of a couple and four boys, enter the ice-cream shop. The couple decides to share ice-cream in a large cup and each of the boys independently takes a regular cup of ice-cream.

(a) Find the probability that one of the boys pay more than \$5 for his regular cup of ice-cream and the other three pay less than \$4 each.

(b) Find the probability that the total cost of the children's regular cups of ice-cream exceeds twice that of the parents' large cup of ice-cream.

The mass, in grams, of an empty regular cup is known to follow a normal distribution with mean 30g and standard deviation 5g. The ice-cream content, in grams, in a regular cup also follows independently a normal distribution.

Find, with adequate justification, the variance of the ice-cream content in a regular cup.

Solution. We have $X \sim \mathrm{N}\big(200, 30^2\big)$ and $Y \sim \mathrm{N}\big(350, 60^2\big)$.

Part (a). Note that a price of \$5 corresponds to a mass of 250 g, while a price of \$4 corresponds to a mass of 200 g. The required probability is thus

$$^4C_1\,\mathbb{P}[X > 250]\,(\mathbb{P}[X < 200])^3 = 0.0239 \text{ (3 s.f.)}.$$

Part (b). Note that the difference in cost D has distribution

$$D = X_1 + \cdots + X_4 - 2Y \sim \mathrm{N}\big(4(200) - 2(350), 4\left(30^2\right) + 2^2\left(60^2\right)\big) = \mathrm{N}(100, 18000).$$

Thus, the required probability is

$$\mathbb{P}[D > 0] = 0.772 \text{ (3 s.f.)}.$$

The variance of the total mass is the sum of the variance of the mass of ice-cream and the variance of the mass of the empty cup. Thus, the variance of the mass of ice-cream is $30^2 - 5^2 = 875$ g.

Assignment A15B

Problem 1. P is a fixed point on the circumference of a circle, centre O, and radius r. Q is a point on the circumference such that $\angle POQ = \theta$, where θ is a random variable with a rectangular distribution in $[0, 2\pi)$. Find the mean and median values for the length of the shorter arc PQ. The length of the chord PQ is X. Find $\mathbb{E}[X]$ and show that $\operatorname{Var}[X] = 2r^2\left(1 - \frac{8}{\pi^2}\right)$. Find also $\mathbb{P}\left[X > r\sqrt{3}\right]$.

Solution. Let L be the length of the shorter arc PQ. Observe that

$$L = \begin{cases} r\theta, & 0 \le \theta < \pi, \\ r\left(2\pi - \theta\right), & \pi \le \theta < 2\pi. \end{cases}$$

Note also that θ has probability density function

$$f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi}, & 0 \le \theta < 2\pi, \\ 0, & \text{otherwise.} \end{cases}$$

Thus,

$$\mathbb{E}[L] = \int_0^\pi \frac{r\theta}{2\pi}\,\mathrm{d}\theta + \int_\pi^{2\pi} \frac{r\left(2\pi - \theta\right)}{2\pi}\,\mathrm{d}\theta = \frac{r}{2\pi}\left[\frac{\theta^2}{2}\right]_0^\pi + \frac{r}{2\pi}\left[2\pi\theta - \frac{\theta^2}{2}\right]_0^\pi = \frac{\pi r}{2}.$$

Now consider the cdf of L:

$$\begin{aligned} F_L(l) &= \mathbb{P}[L < l] = \mathbb{P}[r\theta < l] + \mathbb{P}[r\left(2\pi - \theta\right) < l] \\ &= \mathbb{P}[\theta < l/r] + \mathbb{P}[\theta > 2\pi - l/r] = \frac{l}{2\pi r} + \frac{l}{2\pi r} = \frac{l}{\pi r}. \end{aligned}$$

Let m be the median of L. Then

$$F_L(m) = \frac{1}{2} \implies \frac{m}{\pi r} = \frac{1}{2} \implies m = \frac{\pi r}{2}.$$

Thus, both the mean and median values of the shorter arc PQ are $\pi r/2$.

Note that $X = 2r\sin(\theta/2)$. Thus,

$$\mathbb{E}[X] = 2r\,\mathbb{E}\left[\sin\frac{\theta}{2}\right] = 2r\int_0^{2\pi} \sin\left(\frac{\theta}{2}\right)\frac{1}{2\pi}\,\mathrm{d}\theta = \frac{2r}{\pi}\left[-\cos\frac{\theta}{2}\right]_0^{2\pi} = \frac{4r}{\pi}.$$

We also have

$$\begin{aligned} \mathbb{E}\left[X^2\right] &= 4r^2\,\mathbb{E}\left[\sin^2\frac{\theta}{2}\right] = 4r^2\int_0^{2\pi} \sin^2\left(\frac{\theta}{2}\right)\frac{1}{2\pi}\,\mathrm{d}\theta \\ &= \frac{2r^2}{\pi}\int_0^{2\pi} \frac{1 - \cos\theta}{2}\,\mathrm{d}\theta = \frac{r^2}{\pi}\left[\theta - \sin\theta\right]_0^{2\pi} = 2r^2. \end{aligned}$$

Hence,

$$\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = 2r^2 - \left(\frac{4r}{\pi}\right)^2 = 2r^2\left(1 - \frac{8}{\pi^2}\right).$$

Observe that

$$2r\sin\frac{\theta}{2} = X > r\sqrt{3} \implies \sin\frac{\theta}{2} > \frac{\sqrt{3}}{2} \implies \theta \in \left(\frac{2\pi}{3}, \frac{4\pi}{3}\right).$$

Since θ is uniform on $[0, 2\pi)$,

$$\mathbb{P}\left[X > r\sqrt{3}\right] = \frac{4\pi/3 - 2\pi/3}{2\pi - 0} = \frac{1}{3}.$$

Problem 2. The continuous random variable X has the exponential distribution whose probability density function is given by

$$f(x) = \begin{cases} \mu e^{-\mu x}, & x \geq 0, \\ 0, & \text{otherwise}, \end{cases}$$

where μ is a positive constant. Television sets are hired out by a rental company. The time in months, X, between major repairs has the above exponential distribution with $\mu = 0.05$.

(a) Find, to 3 significant figures, the probability that a television set hired out by the company will not require a major repair for at least a two-year period.

(b) The company agreed to replace any sets for which the time between major repairs is less than M months, where M is a whole number. Given that the company does not want to have to replace more than one set in 5, find the set of possible values of M.

Solution.

Part (a).

$$\mathbb{P}[X > 24] = \int_{24}^{\infty} 0.05e^{-0.05x} \, dx = 0.301 \text{ (3 s.f.)}.$$

Part (b).

$$\mathbb{P}[X < M] \leq \frac{1}{5} \implies \mathbb{P}[X \geq M] \geq \frac{4}{5} \implies e^{-0.05M} \geq \frac{4}{5}.$$

Using G.C., $M \in \{1, 2, 3, 4\}$.

* * * * *

Problem 3. An examination is marked out of 100. It is taken by a large number of candidates. The mean mark, for all candidates, is 72.1, and the standard deviation is 15.2. Give a reason why a normal distribution, with this mean and standard deviation, would not give a good approximation to the distribution of marks.

Solution. Let X be the marks scored by a candidate. Then $X \sim \mathrm{N}(72.1, 15.2^2)$. Note that $\mathbb{P}[0 \leq X \leq 100] = 0.967$. Assuming no candidate receives a negative score, 3.32% of candidates are not accounted for under a normal distribution model.

* * * * *

Problem 4. The weights of boys in a certain age group are normally distributed, with mean 52 kg and standard deviation σ kg. The weights of girls in the same age group are normally distributed, with mean μ kg and standard deviation 5 kg. On average, 1 in 25 randomly chosen boys weighs less than 45 kg; and 2 in 25 randomly chosen girls weigh more than 49 kg.

(a) Find the values of μ and σ.

(b) Find the probability that the weight of two randomly chosen boys is more than thrice the weight of a randomly chosen girl.

(c) Find the probability that the mean weight of 10 girls chosen is less than 41 kg.

Solution. Let B kg be the weight of a boy, and let G kg be the weight of a girl. We have $B \sim \mathrm{N}(52, \sigma^2)$ and $G \sim \mathrm{N}(\mu, 5^2)$.

Part (a). Since $\mathbb{P}[B < 45] = 1/25$, using G.C., we obtain $\sigma = 4$. Since $\mathbb{P}[G > 49] = 2/25$, using G.C., we have $\mu = 42$.

Part (b). Let $T = B_1 + B_2 - 3G$. Then $T \sim \mathrm{N}\big(2(52) - 3(42), 2\left(4^2\right) + 3^2\left(5^2\right)\big) = \mathrm{N}(-22, 257)$. Thus, the required probability is

$$\mathbb{P}[T > 0] = 0.0850 \text{ (3 s.f.)}.$$

Part (c). Let $\overline{G} = \frac{1}{10}(G_1 + \cdots + G_{10})$. Then $\overline{G} \sim \mathrm{N}\big(42, 5^2/10\big)$. Hence, the required probability is

$$\mathbb{P}\big[\overline{G} < 41\big] = 0.264 \text{ (3 s.f.)}.$$

A16 Sampling

Tutorial A16

Problem 1. In a country, 75% of the population have height exceeding 1.50 m and 10% have height exceeding 1.90 m. Assuming a normal distribution of heights, show that the height exceeded by 20% of the population is 1.81 m, correct to 3 significant figures.

A random sample of 80 people is taken from the population. Find the probability that the sample mean exceeds 1.69 m.

Solution. Let the random variable H m be the height of a person. Let $H \sim \mathrm{N}(\mu, \sigma^2)$. We are given that $\mathbb{P}[H > 1.50] = 0.75$ and $\mathbb{P}[H > 1.90] = 0.10$. Standardizing,

$$\frac{1.50 - \mu}{\sigma} = -0.6745 \quad \text{and} \quad \frac{1.90 - \mu}{\sigma} = 1.2816.$$

Solving, we get $\mu = 1.6379$ and $\sigma = 0.2045$. Thus,

$$\mathbb{P}[H > h] = 0.20 \implies h = 1.81 \text{ (3 s.f.)}.$$

Let $\overline{H} = \frac{1}{80}(H_1 + \cdots + H_{80})$. Then $\overline{H} \sim \mathrm{N}\left(1.6379, \frac{1}{80}(0.2045)^2\right)$. Hence,

$$\mathbb{P}\left[\overline{H} > 1.69\right] = 0.0113 \text{ (3 s.f.)}.$$

* * * * *

Problem 2. A factory produces packets of peanuts. The mass of peanuts in a packet has mean 605 g and standard deviation 6 g. A sample of sixty packets is chosen. Find the probability that the mean mass of peanuts in a packet from this sample is between 600 g and 606 g. State the assumptions that you have made.

Solution. Let $\overline{M}$ g be the mean mass of a packet of peanuts in a sample. Assuming $\overline{M}$ follows a normal distribution (since the size of a sample, 60 packets, is large), we have $\overline{M} \sim \mathrm{N}(605, 6^2)$. Thus,

$$\mathbb{P}\left[600 < \overline{M} < 606\right] = 0.902 \text{ (3 s.f.)}.$$

* * * * *

Problem 3. A beekeeper sells jars of honey which are labelled, "Total weight: 300 grams". She takes a random sample of 10 filled jars and records the weight, x grams, of each filled jar. Her results are summarized below, with $\overline{x}$ denoting the sample mean.

$$\sum x = 3030, \quad \sum (x - \overline{x})^2 = 148.$$

Calculate unbiased estimates of the mean μ, and the variance σ^2, of the weight, X grams of a jar of honey.

Solution. We have

$$\overline{x} = \frac{1}{n}\sum x = \frac{3030}{10} = 303 \text{ g},$$

and

$$s^2 = \frac{1}{n-1}\sum (x-\overline{x})^2 = \frac{148}{10-1} = 16.4 \text{ g}^2.$$

* * * * *

Problem 4. A large multinational company has 100,000 employees based in several different countries. To celebrate the 90th anniversary of the founding of the company, the Chief Executive wishes to invite a representative sample of 90 employees to a party, to be held at the company's Headquarters in Singapore. Explain how simple random sampling could be carried out to choose the 90 employees.

Solution. Assign a unique number to each of the 100,000 employees. For each employee, place a corresponding numbered ball in a bag. Draw 90 balls from the bag, without replacement, at random. The numbers on the balls identify the chosen employees.

* * * * *

Problem 5. The mass of an abalone of a certain grade follows a normal distribution with mean 180 g and standard deviation 14.2 g.

(a) Find the probability that the mean mass of a sample of sixty abalones chosen at random differs from the population mean mass by more than 2g.

(b) This grade of abalones is priced at 450 dollars per kilogram. A customer orders five abalones. Find the probability that the customer ends up paying an average of more than 84 dollars per abalone.

Solution. Let the random variable $\overline{M}_i$ g be the mean mass of an abalone in a sample of i abalones. Note that $\overline{M}_i \sim \mathrm{N}\big(180, 14.2^2/n\big)$.

Part (a). $\mathbb{P}\big[\big|\overline{M}_{60} - 180\big| < 2\big] = \mathbb{P}\big[178 < \overline{M}_{60} < 182\big] = 0.725$ (3 s.f.).

Part (b). $\mathbb{P}\big[\frac{450}{1000}\overline{M}_5 > 84\big] = \mathbb{P}\big[\overline{M}_5 > \frac{560}{3}\big] = 0.147$ (3 s.f.).

* * * * *

Problem 6. The random variable X has the distribution $\mathrm{N}(1, 20)$.

(a) Given that $\mathbb{P}[X < a] = 2\,\mathbb{P}[X > a]$, find a.

(b) A random sample of n observations of X is taken. Given that the probability that the sample mean exceeds 1.5 is at most 0.01, find the possible values of n.

Solution.

Part (a). Using G.C., $a = 2.93$.

Part (b). Let $\overline{X} = \frac{1}{n}(X_1 + \ldots X_n)$. Then $\overline{X} \sim \mathrm{N}(1, 20/n)$. Consider $\mathbb{P}\big[\overline{X} > 1.5\big] \leq 0.01$. Using G.C., $n \geq 433$.

* * * * *

Problem 7. The random variable X has a Poisson distribution with mean 4. The random variable $\overline{X}$ is the mean of a random sample of 100 values of X. By using a suitable approximation, find $\mathbb{P}\big[\overline{X} < 3.5\big]$. The random variable Y has a binomial distribution with mean 4 and variance 3. The random variable $\overline{Y}$ is the mean of a random sample of 60 values of Y. By using a suitable approximation, find $\mathbb{P}\big[\overline{Y} - \overline{X} > 0.5\big]$.

Solution. Let $\overline{X} = \frac{1}{100}(X_1 + \cdots + X_{100})$. Since the sample size (100) is large, by the Central Limit Theorem, $\overline{X} \sim \mathrm{N}(4, 4/100)$. Thus, $\mathbb{P}\big[\overline{X} < 3.5\big] = 0.00621$.

Let $\overline{Y} = \frac{1}{60}(Y_1 + \cdots + Y_{60})$. Since the sample size (60) is large, by the Central Limit Theorem, $\overline{Y} \sim \mathrm{N}(4, 3/60)$. Thus, $\overline{Y} - \overline{X} \sim \mathrm{N}(4-4, 4/100 + 3/60) = \mathrm{N}(0, 0.09)$. Thus,

$$\mathbb{P}\big[\overline{Y} - \overline{X} > 0.5\big] = 0.0478 \text{ (3 s.f.)}.$$

* * * * *

Problem 8. The continuous random variable X has $\mathbb{E}[X] = 0$ and $\mathrm{Var}[X] = 4/5$. The random variable Y is defined by $Y = aX + b$, where a and b are positive constants. It is given that $\mathbb{E}[Y] = 50$ and $\mathrm{Var}[Y] = 80$. Find a and b.

A random sample consists of 160 independent observations of Y. Find an approximate value for the probability that the sample sum lies between 7840 and 8080.

Solution. Note that

$$50 = \mathbb{E}[Y] = \mathbb{E}[aX + b] = a\,\mathbb{E}[X] + b = b,$$

and

$$80 = \mathrm{Var}[Y] = \mathrm{Var}[aX + b] = a^2\,\mathrm{Var}[X] = a^2\left(\frac{4}{5}\right) \implies a^2 = 100 \implies a = 10.$$

Note that we reject $a = -10$ since a is positive.

Let $\Sigma = Y_1 + \cdots + Y_{160}$. Since the sample size (160) is large, $\Sigma \sim \mathrm{N}(160(50), 160(80))$. Thus,

$$\mathbb{P}[7840 < \Sigma < 8080] = 0.682 \text{ (3 s.f.)}.$$

* * * * *

Problem 9. The speeds of 120 randomly selected cars are measured as they pass a camera on a motorway. Denoting the speed by x km per hour, the results are summarized by

$$\sum(x - 100) = -221, \quad \sum(x - 100)^2 = 4708.$$

Suggest a reason why, in this context, the given data is summarized in terms of $(x - 100)$ rather than x.

Giving your answers correct to 2 places of decimals, find unbiased estimates of the population mean and variance.

If another sample of 50 cars is chosen, estimate the probability that mean speed of the 50 cars is at least 100 km per hour. State one assumption and one approximation used in obtaining this estimate.

Solution. The given data may be summarized in terms of $(x - 100)$ because the speed limit is 100 km/h.

Note that

$$\sum x = \sum(x - 100) + 100n = -221 + 100(120) = 11779$$

and

$$\begin{aligned}\sum x^2 &= \sum(x - 100)^2 + 200\sum x - 100^2 n \\ &= 4708 + 200(11779) - 100^2(120) = 1160508.\end{aligned}$$

Thus,

$$\overline{x} = \frac{1}{n}\sum x = \frac{11779}{120} = 98.16 \text{ (2 d.p.)}$$

and

$$s^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{1}{n}\left(\sum x\right)^2\right] = \frac{1}{120-1}\left[1160508 - \frac{11779^2}{120}\right] = 36.14 \text{ (2 d.p.)}.$$

Let $\overline{X}$ km/h be the mean speed of the 50 cars. Assuming that the speeds of the cars are independent, by the Central Limit Theorem, we can approximate $\overline{X}$ using a normal distribution: $\overline{X} \sim \mathrm{N}(98.16, 36.14/50)$. Hence,

$$\mathbb{P}\left[\overline{X} > 100\right] = 0.0152 \text{ (3 s.f.)}.$$

* * * * *

Problem 10. $100p\%$ of all insurance agents from a large insurance company, Avila, have an advanced diploma in insurance (ADI), where $p < 0.5$. A sample of 10 agents from Avila is obtained. It is given that the number of insurance agents with ADI in this sample can be modelled by a binomial distribution.

(a) It is given that the probability that 5 of the agents in this sample have an ADI each is 0.12294, correct to 5 decimal places. Show that p satisfies an equation of the form $p(1-p) = k$ for some real constant k to be determined, and hence find the value of p correct to 2 decimal places.

(b) Suppose instead that $p = 0.24$ and forty samples of 10 Avila insurance agents each are obtained. Find the probability that the average number of insurance agents with ADI of the forty samples is between 2.3 and 2.5.

(c) Explain, stating a reason, how increasing the number of samples of 10 Avila insurance agents each will affect your answer in part (b).

Solution.

Part (a). Let X be the number of insurance agents with an ADI in a sample. Then $X \sim \mathrm{B}(10, p)$. Since $\mathbb{P}[X = 5] = 0.12294$, we have

$$\binom{10}{5} p^5 (1-p)^5 = 0.12294 \implies p(1-p) = \sqrt[5]{\frac{0.12294}{\binom{10}{5}}} = 0.217600.$$

Thus, $k = 0.217600$. Solving for p, we get $p = 0.32$ or $p = 0.68$, which we reject since $p < 0.5$.

Part (b). Taking $p = 0.24$, we have

$$\mu = np = 2.4 \quad \text{and} \quad \sigma^2 = np(1-p) = 1.824.$$

Let $\overline{X} = \frac{1}{40}(X_1 + \cdots + X_{40})$. Since the sample size (40) is large, $\overline{X} \sim \mathrm{N}(2.4, 1.824/40)$. Hence,

$$\mathbb{P}\left[2.3 < \overline{X} < 2.5\right] = 0.360 \text{ (3 s.f.)}.$$

Part (c). As the number of samples increases, the variance of $\overline{X}$ will decrease. The distribution of $\overline{X}$ becomes more concentrated around 2.4, hence $\mathbb{P}\left[2.3 < \overline{X} < 2.5\right]$ will tend to 1.

* * * * *

Problem 11. In a certain country there are 100 professional football clubs, arranged in 4 divisions. There are 22 clubs in Division One, 24 in Division Two, 26 in Division Three and 28 in Division Four.

(a) Alice wishes to find out about approaches to training by clubs in Division One, so she sends a questionnaire to the 22 clubs in Division One. Explain whether these 22 clubs form a sample or a population.

(b) Dilip wishes to investigate the facilities for supporters at the football clubs, but does not want to obtain the detailed information necessary from all 100 clubs. Explain how he should carry out this investigation, and why he should do the investigation in this way.

(c) Find the number of different possible samples of 20 football clubs, with 5 clubs chosen from each division.

Solution.

Part (a). The 22 clubs form a population because she is studying the entire group relevant to her research question (training in Division One clubs).

Part (b). Let k be the number of clubs Dilip wishes to investigate. Assign each club a unique number. For each club, place a corresponding numbered ball in a bag. Draw k balls from the bag, without replacement, at random. The numbers on the balls identify the clubs that Dilip should investigate.

Part (c). The required number is

$$^{22}C_5\,^{24}C_5\,^{26}C_5\,^{28}C_5 = 7.24 \times 10^{18}.$$

Assignment A16

Problem 1. Alice receives "like" notifications from her Facebook friends at random, with an average of one "like" received over two days. Taking a year as 52 weeks, find the probabilities that in a year,

(a) there are not more than 2 weeks in which she receives 6 "like" notices in a week,

(b) the mean number of "like" notices received per week is at least 4 by the use of a suitable approximation.

Solution. Let L be the number of "likes" received in a week. Then $L \sim \text{Po}\left(\frac{7}{2}\right)$.

Part (a). Note that $\mathbb{P}[L = 6] = 0.077098$ (5 s.f.). Let W be the number of weeks in a year in which Alice receives 6 "likes". Then $W \sim \text{B}(52, 0.077098)$. The required probability is hence $\mathbb{P}[W \leq 2] = 0.225$ (3 s.f.).

Part (b). Let $\overline{L} = \frac{1}{52}(L_1 + \cdots + L_{52})$. Since the sample size (52) is large, by the Central Limit Theorem, $\overline{L} \sim \text{N}\left(\frac{7}{2}, \frac{1}{52}\frac{7}{2}\right)$ approximately. The required probability is hence $\mathbb{P}\left[\overline{L} \geq 4\right] = 0.0270$ (3 s.f.).

* * * * *

Problem 2. The mass of doughnuts produced in a doughnut factory is found to have mean 150 g and standard deviation 60 g.

(a) Find the probability that the mean mass of a random sample of 200 doughnuts is between 145 g and 160 g. Give a reason why it is not necessary to assume that the mass of doughnuts is normally distributed.

(b) Find the least value of n such that the probability that the mean mass of a sample of n doughnuts is greater than 140 g is greater than 0.8.

Solution. Let M g be the mass of a doughnut.

Part (a). Let $\overline{M}_k = \frac{1}{200}(M_1 + \cdots + M_k)$. Since the sample size (200) is large, by the Central Limit Theorem, $\overline{M}_{200} \sim \text{N}\left(150, 60^2/200\right)$ approximately. Hence, the required probability is

$$\mathbb{P}\left[145 < \overline{M}_{200} < 160\right] = 0.871 \text{ (3 s.f.)}.$$

The Central Limit Theorem applies to any distribution, so long as n is large. Hence, M need not be normally distributed.

Part (b). Suppose n is large. Then by the Central Limit Theorem, $\overline{M}_n \sim \text{N}\left(150, 60^2/n\right)$ approximately. Consider $\mathbb{P}\left[\overline{M}_n > 140\right] > 0.8$. Using G.C., the least n is 26, which is large.

* * * * *

Problem 3. A pharmaceutical company created a new drug to treat a particular illness. The patients experienced weight loss due to the side effects of the new drug. A random sample of 50 individuals was selected and the weight loss by each individual, x kg, is recorded and summarized as follows:

$$\sum (x - 5) = 120 \quad \text{and} \quad \sum (x - 5)^2 = 2500.$$

(a) Describe how the company can obtain the random sample of 50 individuals.

(b) Calculate unbiased estimates of μ, the population mean, and σ^2, the population variance.

(c) Estimate the probability that the mean weight loss by fifty randomly chosen patients who took the new drug is greater than 7 kg.

Solution.

Part (a). Assign each patient a unique positive integer. Using a random number generator, obtain 50 distinct positive integers. The patients assigned to these numbers are then sampled.

Part (b). Note that

$$\sum (x-5) = 120 \implies \sum x = 120 + 50(5) = 370$$

and

$$\sum (x-5)^2 = \sum \left(x^2 - 10x + 25\right) = 2500 \implies \sum x^2 = 4950.$$

Hence,

$$\overline{x} = \frac{1}{n}\sum x = 7.4 \quad \text{and} \quad s^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{1}{n}\left(\sum x\right)^2\right] = 45.152 \text{ (5 s.f.)}.$$

Let $\overline{W}$ kg be the mean weight loss by 50 randomly chosen patients. Since the sample size (50) is large, by the Central Limit Theorem, $\overline{W} \sim \mathrm{N}(7.4, 45.152)$. Hence, the required probability is $\mathbb{P}\left[\overline{W} \geq 7\right] = 0.524$ (3 s.f.).

A17 Confidence Intervals

Tutorial A17

Problem 1. The weights of 4-month-old pigs are known to be normally distributed with standard deviation 4 kg. A new diet is suggested and a sample of 25 pigs given this new diet have an average weight of 30.42 kg. Determine a 99% confidence interval for the mean weight of 4-month-old pigs that are fed this diet.

Solution. Let X kg be the weight of a pig. We are given that $\overline{x} = 30.42$. Hence, $\overline{X} \sim \mathrm{N}\big(30.42, 4^2/25\big)$. Using G.C., a 99% confidence interval for μ is $(28.4, 32.5)$.

* * * * *

Problem 2. A firm produces a type of car types called Standard. A random sample of 150 Standard types is examined, and the lifetimes (in thousands of kilometres) are summarized by

$$\sum x = 2850, \quad \sum (x - \overline{x})^2 = 1931.04.$$

(a) Obtain the unbiased estimates of the mean and variance of the lifetimes of Standard tyres.

(b) Calculate a 97% confidence interval for the mean lifetime of Standard tyres.

Solution.

Part (a). We have

$$\overline{x} = \frac{1}{n} \sum x = \frac{1}{150}(2850) = 19$$

and

$$s^2 = \frac{1}{n-1} \sum (x - \overline{x})^2 = \frac{1}{150-1}(1931.04) = 12.96.$$

Part (b). By the Central Limit Theorem, $\overline{X} \sim \mathrm{N}(19, 12.96)$. Using G.C., a 97% confidence interval for μ is $(18.4, 19.6)$.

* * * * *

Problem 3. The 95% confidence interval for the mean length of life, in hours, of a particular brand of light bulb is $(1023.3, 1101.7)$. It is known that standard deviation of the length of life in the brand of light bulb is σ. This interval is based on results from a random sample of 36 light bulbs. Find a 99% confidence interval for the mean length of life of this brand of light bulb, assuming the length of life is normally distributed.

Solution. Note that

$$\overline{x} = \frac{1107.7 - 1023.3}{2} + 1023.3 = 1065.5.$$

Hence,

$$\overline{x} + z_{0.975} \frac{\sigma}{\sqrt{n}} = 1107.7 \implies \sigma = \frac{(1101.7 - \overline{x})\sqrt{n}}{z_{0.975}} = 110.8.$$

Thus, using G.C., a 99% confidence interval for μ is $(1020, 1110)$.

Problem 4. A coin is chosen at random from a population of recently produced coins. The discrete random variable X is the age, in years, of the coin. The population mean of X is denoted by μ, the population standard deviation is denoted by σ, and the population proportion for which $X \leq 1$ is denoted by p. A random sample of 120 independent observations of X was taken, and the results can be summarized as follows.

Age (x)	0	1	2	3	4	5
Frequency (f)	14	26	24	23	17	19

(a) Calculate unbiased estimates of μ, σ^2 and p.

(b) Find a symmetric 95% confidence interval for μ.

(c) It is desired to find a symmetric 95% confidence interval for μ, of width 0.2, using a random sample of n coins. Estimate the smallest possible value for n.

(d) Find a 95% confidence interval for p.

Solution.

Part (a). Using G.C., we have $\overline{x} = 2.5$, $s^2 = 2.6387$ and $p_s = (14+26)/120 = 1/3$.

Part (b). By the Central Limit Theorem, $\overline{X} \sim \mathrm{N}(2.5, 2.6387/n)$ approximately. Taking $n = 120$, using G.C., a 95% confidence interval for μ is $(2.21, 2.79)$.

Part (c). We require

$$z_{0.975} \frac{s}{\sqrt{n}} \leq \frac{0.2}{2} = 0.1 \implies n \geq \left(\frac{z_{0.975}\,s}{0.1}\right)^2 = 1013.6.$$

Thus, the least n is 1014.

Part (d). Using G.C., a 95% confidence interval for p is $(0.249, 0.418)$.

* * * * *

Problem 5.

(a) $\overline{X}$ is the mean of a large random sample of size n_1 from a population with mean μ_1 and variance σ_1^2. $\overline{Y}$ is the mean of a large random sample of size n_2 from a population with mean μ_2 and variance σ_2^2. State the sampling distribution of $\left(\overline{Y} - \overline{X}\right)$, giving its mean and variance.

(b) Buildrite and Constructall are two building firms. The amount, X thousand dollars, paid to Buildrite by each 100 randomly chosen customers is summarized by $\sum x = 160$, $\sum x^2 = 265$.

   (i) Find an approximate 99.8% confidence interval for the mean amount paid per customer to Buildrite.

   The amount paid to Constructall by each customer was Y thousand dollars. Based on a random sample of 200 customers, unbiased estimates of the mean and variance of Y were 1.8 and 0.3216 respectively.

   (ii) Find, to the nearest dollar, an approximate 90% confidence interval for the value by which the mean amount paid per customer to Constructall exceeds that paid to Buildrite.

Solution.

Part (a). By the Central Limit Theorem, $\overline{X} \sim \mathrm{N}\left(\mu_1, \sigma_1^2/n_1\right)$ and $\overline{Y} \sim \mathrm{N}\left(\mu_2, \sigma_2^2/n_2\right)$, it follows that $\overline{Y} - \overline{X} \sim \mathrm{N}\left(\mu_2 - \mu_1, \sigma_1^2/n_1 + \sigma_2^2/n_2\right)$.

Part (b).

Part (b)(i). We have

$$\overline{x} = \frac{1}{n}\sum x = \frac{1}{100}(160) = 1.6$$

and

$$\sigma_x^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{1}{n}\left(\sum x\right)^2\right] = \frac{1}{100-1}\left[265 - \frac{1}{100}(160)^2\right] = 0.0909 \text{ (3 s.f.)}.$$

Thus, by the Central Limit Theorem, $\overline{X} \sim \mathrm{N}(1.6, 0.0909)$ approximately. Hence, using G.C., a 99.8% confidence interval for μ_x is $(1.507, 1.693)$. In dollars, this is $(1507, 1693)$.

Part (b)(ii). By the Central Limit Theorem, $\overline{Y} \sim \mathrm{N}(1.8, 0.3216/200)$. Hence,

$$\overline{Y} - \overline{X} \sim \mathrm{N}\left(1.8 - 1.6, \frac{0.0909}{100} + \frac{0.3216}{200}\right) = \mathrm{N}(0.2, 0.002517).$$

Using G.C., a 90% confidence interval for $\mu_y - \mu_x$ is $(0.117 - 0.283)$. In dollars, this is $(117, 283)$.

* * * * *

Problem 6. The speed at which a baseball is thrown, x km/h, is measured at the instant that it leaves the pitcher's hand. To join a particular baseball club, a pitcher has to be able to throw balls at 140 km/h. The results for 10 randomly chosen through by a young pitcher on a cool day are summarized by

$$\sum (x - 128) = 7.9, \quad \sum (x - 128)^2 = 338.4.$$

Assuming that these results are observations from a normal distribution, obtain unbiased estimates of the mean and variance of this distribution, and obtain a symmetric 99.5% confidence interval for the mean, and explain in context what it means. Can the young pitcher throw balls at 140 km/h on average?

Solution. Note that

$$\sum (x - 128) = 7.9 \implies \sum x = 1287.9$$

and

$$\sum (x - 128)^2 = \sum \left(x^2 - 256x + 128^2\right) = 338.4 \implies \sum x^2 = 166200.8.$$

Thus,

$$\overline{x} = \frac{1}{n}\sum x = 128.79 \quad \text{and} \quad s^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{1}{n}\left(\sum x\right)^2\right] = 36.907.$$

Since X follows a normal distribution and the sample size is small,

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim \mathrm{t}(n-1).$$

Using G.C., a symmetric 99.5% confidence interval for μ is $(122, 136)$. This means that we are 99.5% confident that the interval 122 – 136 km/h contains the average speed of a baseball thrown by the young pitcher. Hence, on average, the young pitcher cannot throw balls at 140 km/h.

* * * * *

Problem 7. Ten students independently performed an experiment to estimate the value of π. Their results were:

3.12, 3.16, 2.94, 3.33, 3.00, 3.11, 3.50, 2.81, 3.02, 3.10.

(a) Calculate the unbiased estimates of the population mean and variance.

(b) Stating any necessary assumption, calculate a 95% confidence interval for π based on these data, giving your answer to two decimal places.

(c) Estimate the minimum number of results that would be needed if it is required that the width of the resulting 95% confidence interval be at most 0.02.

Solution.
Part (a). Using G.C., $\overline{x} = 3.109$ and $s^2 = 0.038032$.
Part (b). Assuming that X follows a normal distribution, we have

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim \mathrm{t}(n-1).$$

Using G.C., a 95% confidence interval for π is $(2.97, 3.25)$.
Part (c). We require

$$t_{0.975}\frac{s}{\sqrt{n}} \leq \frac{0.02}{2} = 0.01 \implies n \geq \left(\frac{t_{0.975}s}{0.01}\right)^2 = 1946.2.$$

Thus, the least n needed is 1947.

* * * * *

Problem 8.

(a) In a market research survey 25 people out of a random sample of 100 from a certain area said that they used a particular brand of soap. Find a 97% confidence interval for the proportion of people in the area who use this brand of soap.

(b) A research lab published an article about this brand of soap, reporting it contains ingredients that is beneficial to one's health. A new survey was conducted, and a 97% confidence interval was found to be $(0.450, 0.620)$. Comment, with reference to the confidence interval computed in (a), whether the proportion of people in the area who use this brand of soap has changed after the research article was published.

Solution.
Part (a). We have $p_s = 25/100 = 1/4$. Using G.C., a 97% confidence interval for p is $(0.156, 0.344)$.
Part (b). Since $(0.156, 0.344) \cap (0.450, 0.620) = \varnothing$, there is sufficient evidence at a 97% confidence level that the mean has changed.

* * * * *

Problem 9. Based on previous records, it is known that p, the proportion of workers supporting the Thunder party, is about 40% in the last election. For this coming election, a market research organization intends to interview a random sample of n voters, and wishes to ensure that the probability is about 0.8 that its sample estimate of the proportion of Thunder voters lies within two percentage points of the sample percentage. Assuming that all voters interviewed do reveal which party they support, what is the least sample size the organization should take?

Solution. For large n, by the Central Limit Theorem, $P_s \sim \mathrm{N}(p_s, p_s(1-p_s)/n)$. Given $p_s = 0.40$, for an 80% confidence interval with error less than 0.02, we require

$$z_{0.9}\sqrt{\frac{p_s(1-p_s)}{n}} \leq 0.02 \implies n \geq \left(\frac{z_{0.9}\sqrt{p_s(1-p_s)}}{0.02}\right)^2 = 985.4.$$

Hence, the least n needed is 986.

Assignment A17

Problem 1. The proportion of letters sent by first-class post which are delivered on the next working day after they are posted is p. In order to obtain an estimate of p, 1000 letters were posted at randomly chosen times and places, and their times of arrivals were recorded. It was found that 900 were delivered on the next working day after posting. Calculate a 99.5% confidence interval for p.

Explain briefly what a 99.5% confidence interval means in this context.

Subsequently, it is proposed to conduct a larger trial to obtain a more precise estimate of p. Estimate the least number of letters to be posted in order for the value of p to be determined to within ± 0.005 with 99.9% confidence.

Solution. By the Central Limit Theorem, $P_s \sim \mathrm{N}(p, p(1-p)/n)$ approximately. From the given sample, we have $p_s = 900/1000$ and $n = 1000$. Using G.C., a 99.5% confidence interval for p is $(0.87337, 0.92663)$.

We can say at a 99.5% confidence interval that the interval $(0.87337, 0.92663)$ contains the proportion of letters delivered on the next working day.

For the width to be within ± 0.005, we must have

$$z_{0.9995}\sqrt{\frac{p(1-p)}{n}} \leq 0.005.$$

Using our estimate $p_s = 900/1000$, by G.C., we have $n \geq 38979.2$. Thus, at least 38980 letters should be posted.

* * * * *

Problem 2.

(a) A research firm conducted a survey to determine the mean amount students spent on drinks during a week. A sample of 60 students revealed that

$$\sum(x-18) = 388, \quad \sum(x-18)^2 = 2550,$$

where x is the amount a student spent on drinks during a week.

(i) Find the unbiased estimates of the population mean and variance, correct to 3 decimal places.

(ii) Estimate the sample size if it is intended that the resulting 94% confidence interval for the population mean should have a width of 0.18.

(iii) Based on the sample size found in (a)(ii), if a confidence interval has width greater than 0.18, should the confidence level be higher or lower than 94%? Justify your answer.

(b) A random sample of 50 Year 6 H2 Mathematics preliminary examination scripts are marked, and the passing rate is 60%. The 95% confidence interval for the passing rate of all Year 6 H2 Mathematics candidates sitting for the preliminary examination is $(a\%, b\%)$. Calculate the values of a and b.

Asked to explain the meaning of this interval, a student states that "95% of the Year 6 classes has a passing rate between $a\%$ and $b\%$ in their H2 Mathematics preliminary examination." Is this statement correct? State your reason.

Solution.

Part (a).

Part (a)(i). We have

$$\overline{x} = \frac{1}{n}\sum x = \frac{1}{n}\sum(x-18)+18 = 24.467 \text{ (3 d.p.)}$$

and

$$s^2 = \frac{1}{n-1}\left[\sum(x-18)^2 - \frac{1}{n}\left(\sum(x-18)\right)^2\right] = 0.694 \text{ (3 d.p.)}.$$

Part (a)(ii). Since the sample size (60) is large, by the Central Limit Theorem,

$$\overline{X} \sim \mathrm{N}\left(\overline{x}, s^2\right) = \mathrm{N}(24.467, 0.694)\text{approximately}.$$

For the width to be 0.18, we must have

$$z_{0.97}\sqrt{\frac{s^2}{n}} \leq \frac{0.18}{2} = 0.09.$$

Using G.C., we have $n \geq 303.0796$. Thus, the sample size should be approximately 304.

Part (a)(iii). The higher the confidence level, the longer the width. Hence, if the confidence interval has a width greater than 0.18, the confidence level should be larger than 94%.

Part (b). By the Central Limit Theorem, $P_s \sim \mathrm{N}(p, p(1-p)/n)$ approximately. From the sample, $p_s = 0.60$ and $n = 50$. Using G.C., a 95% confidence interval for p is $(0.464, 0.736)$. Hence, $a = 46.4$ and $b = 73.6$.

His statement is incorrect. The confidence interval does not say anything about the distribution of passes within a class.

A18A Hypothesis Testing I - Single Value

Tutorial A18A

Problem 1. For the scenario described below, set up the null and alternative hypotheses required in testing the various claims for (a), (b) and (c). Write the concluding statement for the following cases: (i) H_0 is rejected, or (ii) H_0 is not rejected.

The Health Ministry claimed that the average weight of babies born last year is 3 kg. Test at 10% significance level whether

(a) the average weight of babies born last year differs from the claim.

(b) the Health Ministry has overstated the average weight.

(c) the Health Ministry has understated the average weight.

Solution.

Part (a). Let H_0: $\mu = 3$, H_1: $\mu \neq 3$.

Part (a)(i). We reject H_0 and conclude there is sufficient evidence at the 10% significance level that there is a change in the weight of babies born last year.

Part (a)(ii). We do not reject H_0 and conclude there is insufficient evidence at the 10% significance level that there is a change in the weight of babies born last year.

Part (b). Let H_0: $\mu = 3$, H_1: $\mu < 3$.

Part (b)(i). We reject H_0 and conclude there is sufficient evidence at the 10% significance level that there is a decrease in the weight of babies born last year.

Part (b)(ii). We do not reject H_0 and conclude there is insufficient evidence at the 10% significance level that there is decrease in the weight of babies born last year.

Part (c). Let H_0: $\mu = 3$, H_1: $\mu > 3$.

Part (c)(i). We reject H_0 and conclude there is sufficient evidence at the 10% significance level that there is an increase in the weight of babies born last year.

Part (c)(ii). We do not reject H_0 and conclude there is insufficient evidence at the 10% significance level that there is an increase in the weight of babies born last year.

* * * * *

Problem 2. A random variable X is known to have a normal distribution with variance 36. The mean of the distribution of X is denoted by μ. A random sample of 50 observations of X has mean 23.8. Test, at 2% significance level, the null hypothesis $\mu = 22$ against the alternative hypothesis $\mu > 22$.

Solution. Let H_0: $\mu = 22$, H_1: $\mu > 2$. We perform a 1-tail test at 2% significance level. Under H_0, $\overline{X} \sim \mathrm{N}(22, 36/50)$. From the sample, $\overline{x} = 23.8$. Using G.C., the p-value is 0.0169, which is less than the significance level of 2%. Thus, we reject H_0 and conclude there is sufficient evidence at 2% significance level that $\mu > 22$.

* * * * *

Problem 3. A random sample of 10 observations of a normal variable X has mean $\overline{x}$, where

$$\overline{x} = 4.344 \quad \text{and} \quad \sum (x - \overline{x})^2 = 0.8022.$$

Carry out a 2-tail test, at the 5% level of significance, to test whether the mean of X is 4.58. State your null and alternative hypotheses clearly.

Solution. Let H_0: $\mu = 4.58$, H_1: $m \neq 4.58$. We perform a 2-tail test at 5% significance level. Under H_0,

$$\frac{\overline{X} - \mu}{S/\sqrt{10}} \sim \mathrm{t}(9).$$

From the sample, $\overline{x} = 4.344$ and

$$s^2 = \frac{1}{9}\sum (x - \overline{x})^2 = 0.089133 \text{ (5 s.f.)}.$$

Using G.C., the p-value is 0.0339, which is less than the significance level of 5%. Thus, we reject H_0 and conclude there is sufficient evidence at 5% significance level that $\mu \neq 4.58$.

* * * * *

Problem 4. The mean of a normally distributed random variable X is denoted by μ, and it is given that the population variance is 15. A sample of 50 random observations of X is taken, and the results are summarized by $\sum x = 527.1$.

(a) Carry out a 2-tail test of the null hypothesis $\mu = 9.5$, at the 5% significance level.

(b) Carry out an appropriate 1-tail test of the null hypothesis $\mu = 11.5$ at the 5% significance level. State your alternative hypothesis clearly, with explanation.

Solution.

Part (a). Let H_0: $\mu = 9.5$, H_1: $\mu \neq 9.5$. We perform a 2-tail test at 5% significance level. Under H_0, $\overline{X} \sim \mathrm{N}(9.5, 15/50)$. From sample,

$$\overline{x} = \frac{1}{50}\sum x = 10.542 \text{ (5 s.f.)}.$$

Using G.C., the p-value is 0.0571, which is greater than the significance level of 5%. Thus, we do not reject H_0 and conclude there is insufficient evidence at 5% significance level that $\mu \neq 9.5$.

Part (b). Let H_0: $\mu = 11.5$. Since $\overline{x} = 10.542 < 11.5$, the alternative hypothesis $\mathrm{H}_1$is $\mu < 11.5$. We perform a 1-tail test at 5% significance level. Under H_0, $\overline{X} \sim \mathrm{N}(11.5, 15/50)$. Using G.C., the p-value is 0.0401, which is less than the significance level of 5%. Thus, we reject H_0 and conclude there is sufficient evidence at 5% significance level that $\mu < 11.5$.

* * * * *

Problem 5. 'Brilliant' fireworks are intended to burn for 40 seconds. A random sample of 50 'Brilliant' fireworks is taken. Each firework in the sample is ignited and the burning time, x seconds, is measured. The results are summarized by

$$\sum(x - 40) = -27, \quad \sum(x - 40)^2 = 167.$$

(a) Test, at the 5% level of significance, whether the mean burning time of 'Brilliant' fireworks differs from 40 seconds.

(b) Suggest a reason why, in this context, the given data is summarized in terms of $(x - 40)$ rather than x.

(c) State, with a reason, whether, in using the above test, it is necessary to assume that the burning times of 'Brilliant' fireworks have a normal distribution.

(d) State what you understand by the expression 'at the 5% level of significance' in the context of the question.

(e) Explain what is meant by critical region of the test conducted in (c) in the context of the question. State the critical region.

Solution.

Part (a). Let H_0: $\mu = 40$, H_1: $\mu \neq 40$. We perform a 2-tail test at 5% significance level. Under H_0, by the Central Limit Theorem, $\overline{X} \sim N\big(40, \sigma^2/50\big)$ approximately. From the sample,

$$\overline{x} = \frac{1}{50} \sum (x - 40) + 40 = 39.46$$

and

$$s^2 = \frac{1}{49}\left[\sum (x-40)^2 - \frac{1}{50}\left(\sum (x-40)\right)^2\right] = 3.1106.$$

Thus, $\overline{X} \sim N(40, 3.1106/50)$ approximately. Using G.C., the p-value is 0.0303, which is less than the significance level of 5%. Thus, we reject H_0 and conclude that there is sufficient evidence at 5% significance level that the mean burning time of 'Brilliant' fireworks differs from 40 seconds.

Part (b). The claimed mean burning time is 40 seconds.

Part (c). No. Because the sample size (50) is large, the Central Limit Theorem ensures that $\overline{X}$ approximately follows a normal distribution, regardless of the distribution of X.

Part (d). There is a 5% chance of rejecting H_0 given that the mean burning time of 'Brilliant' fireworks is actually 40 seconds.

Part (e). The critical region is the set of values of $\overline{x}$ that leads to rejecting H_0. Using G.C., the critical region is $\overline{x} < 39.138$ or $\overline{x} > 40.862$.

* * * * *

Problem 6. A coin is chosen at random from a population of recently produced coins. The discrete random variable X is the age, in years, of the coin. The population mean of X is denoted by μ, and the population standard deviation is denoted by σ. A random sample of 150 independent observations of X was taken, and the results can be summarized as follows.

Age (x)	0	1	2	3	4	5
Frequency (f)	24	36	31	23	17	19

(a) Explain what is meant by random sample in context of the question.

(b) Calculate unbiased estimates of μ and σ^2.

(c) What do you understand by the term unbiased estimate?

(d) A one-tail test to test $\mu = 2$ against $\mu > 2$ is carried out. Find the smallest significance level of the test at which the claim $\mu > 2$ is supported.

Solution.

Part (a). Each recently-produced coin has an equal and independent chance of being selected for observation.

Part (b). Using G.C., $\overline{x} = 2.2$ and $s^2 = 1.614^2 = 2.60$.

Part (c). An unbiased estimate is an estimate of a population parameter such that the expected value of the estimator is equal to the true value of the parameter.

Part (d). Let H_0: $\mu = 2$, H_1: $\mu > 2$. Under H_0, by the Central Limit Theorem, $\overline{X} \sim \mathrm{N}(2, 2.60/150)$ approximately. Using G.C., the p-value is 0.0646. Thus, 6.46% is the smallest significance level that results in the rejection of H_0.

* * * * *

Problem 7. A random sample of 90 batteries, used in a particular model of mobile phone, is tested and the 'standby-time', x hours which is normally distributed, is measured. The results are summarized by

$$\sum x = 3040.8 \quad \text{and} \quad \sum x^2 = 115773.66.$$

Test, at the 1% significance level, whether the mean standby-time is less than 36.0 hours.

In a test at the 5% significance level, it is found that there is significant evidence that the population mean talk-time is less than 5 hours.

Using only this information, and giving a reason in each case, state whether each of the following statements is (i) necessarily true, (ii) necessarily false, or (iii) neither necessary true nor necessarily false.

(a) There is significant evidence at the 10% significance level that the population mean talk-time is less than 5 hours.

(b) There is significant evidence at the 5% significance level that the population mean talk-time is not 5 hours.

The manufacturer changed the production method of the batteries. It took a sample of 100 batteries, and obtained a 95% confidence interval for the mean standby-time of $(36.2, 37.4)$.

(c) Without further computation, explain if the mean standby-time of the batteries have changed from 36.0 hours.

Solution. Let H_0: $\mu = 36.0$, H_1: $\mu < 36.0$. We perform a 1-tail test at 1% significance level. From the sample,

$$\overline{x} = \frac{1}{90} \sum x = 33.787 \text{ (5 s.f.)}$$

and

$$s^2 = \frac{1}{89} \left[\sum x^2 - \frac{1}{90} \left(\sum x \right)^2 \right] = 146.46 \text{ (5 s.f.)}.$$

Under H_0, $\overline{X} \sim \mathrm{N}(36.0, 146.46/90)$ approximately. Using G.C., the p-value is 0.0414, which is greater than the significance level of 1%. Thus, we do not reject H_0 and conclude there is insufficient evidence at 1% significance level that the mean standby-time is less than 36.0 hours.

Part (a). It is necessarily true. The higher the significance level, the larger the critical region.

Part (b). It is neither necessarily true nor necessarily false. If the p-value is less than 2.5%, H_0 would be rejected under a 2-tail test. However, if the p-value is between 2.5% and 5%, H_0 would not be rejected.

Part (c). Since the 95% confidence interval $(36.2, 37.4)$ does not contain $\mu = 36.0$, under a 2-tail test at 5% significance level, we can reject H_0. Thus, there is sufficient evidence at a 5% significance level that the mean standby-time of the batteries has changed from 36.0 hours.

Problem 8. In a factory, the time in minutes for an employee to install an electronic component is a normally distributed continuous random variable T. The standard deviation of T is 5.0 and under ordinary conditions, the expected value of T is 38.0. After background music is introduced into the factory, a sample of n components is taken and the mean time for randomly chosen employees to install them is found to be $\bar{t}$ minutes. A test is carried out, at the 5% significance level, to determine whether the mean time taken to install a component has been reduced.

(a) State appropriate hypotheses for the test, defining any symbols you use.

(b) Given that $n = 50$, state the set of values of $\bar{t}$ for which the result of the test would be to reject the null hypothesis.

(c) It is given instead that $\bar{t} = 37.1$ and the result of the test is that the null hypothesis is not rejected. Obtain an inequality involving n, and hence find the set of values that n can take.

Solution.

Part (a). Let $\mu = \mathbb{E}[T]$. The hypotheses are H_0: $\mu = 38.0$, H_1: $\mu < 38.0$.

Part (b). We perform a 1-tail test at 5% significance level. Under H_0, $\overline{T} \sim \mathrm{N}\big(38.0, 5.0^2/50\big)$. Normalizing,

$$\frac{\overline{T} - 38.0}{5.0/\sqrt{50}} \sim \mathrm{N}(0, 1).$$

For the null hypotheses to be rejected, we must have

$$\frac{\bar{t} - 38.0}{5.0/\sqrt{50}} \leq z_{0.05}.$$

Thus, $\bar{t} < 36.8$. Further, $t > 0$, so $0 < \bar{t} < 36.8$.

Part (c). We perform a 1-tail test at 5% significance level. Under H_0, $\overline{T} \sim \mathrm{N}\big(38.0, 5.0^2/n\big)$. Normalizing,

$$\frac{\overline{T} - 38.0}{5.0/\sqrt{n}} \sim \mathrm{N}(0, 1).$$

For the null hypotheses to be rejected, we must have

$$\frac{37.1 - 38.0}{5.0/\sqrt{n}} \leq z_{0.05}.$$

Thus, $1 \leq n \leq 83$. The set of values that n can take is thus $\{n \in \mathbb{Z} : 1 \leq n \leq 83\}$.

* * * * *

Problem 9. A motoring magazine editor believes that the figures quoted by car manufacturers for distances travelled per litre of fuel are too high. He carries out a survey into this by asking for information by readers. For a certain model of car, 8 readers reply with the following data, measured in km per litre.

14.0 12.5 11.0 11.0 12.5 12.6 15.6 13.2.

(a) Calculate unbiased estimates of the population mean and variance.

The manufacturer claims that this model of car will travel 13.8 km per litre on average.

(b) Stating two assumptions, carry out a t-test of the magazine editor's belief at the 5% significance level.

(c) Explain the meaning of p-value in the context of the question.

Solution.

Part (a). Using G.C., $\overline{x} = 12.8$, $s^2 = 1.518458^2 = 2.3057$.

Part (b). Let the random variable X be the distance travelled per litre, measured in km. We assume that X is normally distributed, and that the information provided by the readers are truthful. Let H_0: $\mu = 13.8$, H_1: $\mu < 13.8$. We perform a 1-tail test at 5% significance level. Under H_0,

$$\frac{\overline{X} - 13.8}{2.3057/\sqrt{8}} \sim \text{t}(7).$$

Using G.C., the p-value is 0.0524, which is greater than the significance level of 5%. Thus, we do not reject H_0 and conclude there is insufficient evidence at 5% significance level that the model of car travels less than 13.8 km per litre.

Part (c). There is a 5.24% chance of obtaining a sample mean less than 12.8.

* * * * *

Problem 10. A company supplies sugar in small packets. The mass of sugar in one packet is denoted by X grams. The masses of a random sample of 9 packets are summarized by

$$\sum x = 86.4 \quad \text{and} \quad \sum x^2 = 835.82.$$

(a) Calculate unbiased estimates of the mean and variance of X.

The mean mass of sugar in a packet is claimed to be 10 grams. The company directors want to know whether the sample indicates that this claim is over-stated.

(b) Stating a necessary assumption, carry out a t-test at the 5% significance level. Explain why the Central Limit Theorem does not apply in this context.

(c) Suppose now that the population variance of X is known, and the assumption made in part (b) is still valid. What change would there be in carrying out the test? You do not have to carry out the test.

Solution.

Part (a). We have

$$\overline{x} = \frac{1}{9}\sum x = 9.6$$

and

$$s^2 = \frac{1}{8}\left[\sum x^2 - \frac{1}{9}\left(\sum x\right)^2\right] = 0.81.$$

Part (b). Assume that X is normally distributed. Let H_0: $\mu = 10$ and H_1: $\mu < 10$. We perform a 1-tail test at 5% significance level. Under H_0,

$$\frac{\overline{X} - 10}{\sqrt{0.81/9}} \sim \text{t}(8).$$

Using G.C., the p-value is 0.110, which is greater than the significance level of 5%. Thus, we do not reject H_0 and conclude there is insufficient evidence at 5% significance level that the mean mass of sugar per packet is less than 10g.

The Central Limit Theorem does not apply here as the sample size, 10, is too small.

Part (c). If the population variance of X is known, we will use a z-test instead of a t-test to calculate the p-value.

* * * * *

Problem 11. The number of minutes that the 0815 bus arrives late at my local bus stop has a normal distribution; the mean number of minutes the bus is late has been 4.3. A new company takes over the service, claiming the punctuality will be improved. After the new company takes over, a random sample of 10 days is taken and the number of minutes that the bus is late is recorded. The sample mean is $\bar{t}$ minutes and the sample variance is k^2 minutes^2. A test is to be carried out at the 10% level of significance to determine whether the mean number of minutes late has been reduced.

(a) State appropriate hypothesis for the test, defining any symbols that you use.

(b) Given that $k^2 = 3.2$, find the set of values of $\bar{t}$ for which the result of the test would be that the null hypothesis is not rejected.

(c) Given instead that $\bar{t} = 4.0$, find the set of values of k^2 for which the result of the test would be to reject the null hypothesis.

Solution.

Part (a). Let T be the number of minutes that the bus is late. Let $\mu = \mathbb{E}[T]$. The hypotheses are H_0: $\mu = 4.3$, H_1: $\mu < 4.3$.

Part (b). We perform a 1-tail test at 10% significance level. Note that the sample variance s^2 is given by

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

Under H_0,

$$\frac{\overline{T} - 4.3}{k/3} \sim \mathrm{t}(9).$$

To not reject H_0, we require

$$\frac{\bar{t} - 4.3}{k/3} > t_{0.10}.$$

Solving, we get $\bar{t} > 3.48$. Thus, the set of values that $\bar{t}$ can take on is

$$\left\{\bar{t} \in \mathbb{R} : \bar{t} > 3.48\right\}.$$

Part (c). We perform a 1-tail test at 10% significance level. Under H_0,

$$\frac{\overline{T} - 4.3}{k/3} \sim \mathrm{t}(9).$$

To reject H_0, we require

$$\frac{4.0 - 4.3}{k/3} \leq t_{0.10}.$$

Solving, we get $0 \leq k \leq 0.651$. Thus, the set of values that k^2 can take on is

$$\left\{k^2 \in \mathbb{R} : 0 < k^2 \leq 0.423\right\}.$$

Assignment A18A

Problem 1. The manufacturers of an electric water heater claim that their heaters will heat 500 litres of water from a temperature of 10°C to a temperature of 35°C in, on average no longer than 12 minutes. In order to test their claim, 14 randomly chosen heaters are bought and the times (x minutes) to heat 500 litres of water from 10°C to 35°C are measured. Correct to 1 decimal place, the results are as follows:

13.2, 12.2, 11.4, 14.5, 11.6, 12.9, 12.4, 10.3, 12.3, 11.8, 11.0, 13.0, 12.1, 12.6.

Stating an assumption necessary for validity, test the manufacturers' claim at a 10% significance level.

Solution. Let X min be the amount of time taken by a heater to heat 500 litres of water from 10°C to 35°C. Let μ min be the mean time taken. Let H_0: $\mu = 12$ and H_1: $\mu > 12$. We perform a one-tail test at 10% significance level. Assuming X is normally distributed, under H_0,

$$\frac{\overline{X} - 12}{S/\sqrt{14}} \sim \mathrm{t}(13).$$

From the sample, $\overline{x} = 12.246$ min and $s = 1.0315$ min. Using G.C., the p-value is 0.204, which is greater than the 10% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence at a 10% significance level that the mean amount of time taken is greater than 12 min.

* * * * *

Problem 2. The mass of vegetables in a randomly chosen bag has a normal distribution. The mass of the contents of a bag is supposed to be 10 kg. A random sample of 80 bags is taken and the mass of the contents of each bag, x grams, is measured. The data is summarized by

$$\sum (x - 10000) = -2510 \quad \text{and} \quad \sum (x - 10000)^2 = 2010203.$$

Test, at the 5% significance level, whether the mean mass of the contents of a bag is less than 10 kg. Explain, in the context of the questions, the meaning of 'at the 5% significance level', and the meaning of 'p-value'.

Solution. Let X g be the mass of vegetables in a bag, and let μ g be the mean mass of vegetables in a bag. Let H_0: $\mu = 10000$ and H_1: $\mu < 10000$. We perform a one-tail test at 5% significance level. From sample,

$$\overline{x} = 10000 + \frac{1}{80}(-2510) = 9968.63 \text{ g}$$

and

$$s^2 = \frac{1}{80-1}\left[2010203 - \frac{1}{80}(-2510)^2\right] = 24449 \text{ g}^2.$$

Under H_0,

$$\overline{X} \sim \mathrm{N}\left(10000, \frac{24449}{80}\right) \text{ approximately.}$$

Using G.C., the p-value is 0.0364, which is less than the 5% significance level. Thus, we reject H_0 and conclude there is sufficient evidence at a 5% significance level that the mean mass of vegetables in a bag is less than 10 kg.

'at the 5% significance level' means there is a 5% chance that we reject H_0 when the mean mass of vegetables in a bag is actually 10 kg. 'p-value' is the lowest significance level needed to reject H_0.

A18B Hypothesis Testing - Two Values

Tutorial A18B

Problem 1. A machine assesses the life of a ball-point pen by measuring the length of a continuous line drawn using the pen. A random sample of 80 pens of brand A have a total writing length of 96.84 km. A random sample of 75 pens of brand B have a total writing length of 93.75 km. Assuming that the standard deviation of the writing length of a single pen is 0.25 km for both brands, test at the 5% level, whether the mean writing lengths of the two brands differ significantly.

Solution. Let X_A km and X_B km be the total writing length of Brand A and Brand B pens respectively. Let $\mu_A = \mathbb{E}[X_A]$ and $\mu_B = \mathbb{E}[X_B]$.

Let H_0: $\mu_A - \mu_B = 0$ and H_1: $\mu_A - \mu_B \neq 0$. We perform a two-tail two-sample z-test at 5% significance level. Under H_0,

$$\overline{X}_A - \overline{X}_B \sim \mathrm{N}\left(0, 0.25^2\left(\frac{1}{80} + \frac{1}{75}\right)\right).$$

From the sample,

$$\overline{x}_A = \frac{96.84}{80} = 1.2105 \quad \text{and} \quad \overline{x}_B = \frac{93.75}{75} = 1.25.$$

Using G.C., the p-value is 0.326, which is greater than the 5% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence at the 5% significance level that the mean writing lengths of the two brands differ significantly.

* * * * *

Problem 2. I have two alternative routes to work. The times taken on the 8 randomly chosen occasions that I use route 1 are summarized by $\sum x = 182$ and $\sum x^2 = 4202$, while the times taken on the 12 randomly chosen occasions that I take route 2 are summarized by $\sum y = 238$ and $\sum y^2 = 5108$, with time being measured in minutes. Determine whether there is significant evidence, at the 5% level, of a difference in the mean times taken on the two routes. State any assumptions needed.

Solution. Let H_0: $\mu_X - \mu_Y = 0$ and H_1: $\mu_X - \mu_Y \neq 0$. We perform a two-tail two-sample t-test at 5% significance level. Assuming that X and Y are normally distributed and have a common variance, under H_0,

$$\frac{\overline{X} - \overline{Y}}{s_p\sqrt{1/8 + 1/12}} \sim \mathrm{t}(8 + 12 - 2) = \mathrm{t}(18).$$

From the sample,

$$\overline{x} = \frac{1}{8}(182) = 22.75, \quad \overline{y} = \frac{1}{12}(238) = 19.833.$$

Also,

$$s_X^2 = \frac{1}{8-1}\left[4202 - \frac{1}{8}(182)^2\right] = 8.7857, \quad s_Y^2 = \frac{1}{12-1}\left[5108 - \frac{1}{12}(238)^2\right] = 35.242.$$

Thus, the pooled variance is

$$s_p^2 = \frac{(8-1)(8.7857) + (12-1)(35.242)}{8+12-2} = 24.953.$$

Using G.C., the p-value is 0.217, which is greater than the 5% significance level. Thus, we do not reject H_0 and conclude that there is insufficient evidence at the 5% significance level that there is a difference in the mean times taken on the two routes.

* * * * *

Problem 3. In an experiment, twelve pairs of plants were positioned close to each other in various different locations in a large greenhouse. One plant in each pair was given fertilizer in April and the other in May. The yields are given below.

Pair	1	2	3	4	5	6	7	8	9	10	11	12
April	344	307	339	256	398	267	256	407	335	381	300	388
May	315	289	317	277	363	258	283	385	269	355	275	363

Test at 5% significance level whether the mean yield in April is at least 30 more than the mean yield in May. State any assumptions needed.

Solution. Let D = April yield − May yield. Let H_0: $\mu_D = 30$ and H_1: $\mu_D > 30$. We perform a one-tail paired-sample t-test at 5% significance level. Assuming that D is normally distributed, under H_0,

$$\frac{\overline{D} - 30}{S_D/\sqrt{12}} \sim \mathrm{t}(11).$$

From the sample, $\overline{d} = 19.083$ and $s_D = 24.347$. Using G.C. the p-value is 0.926, which is greater than the 5% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence at 5% significance level that the mean yield in April is at least 30 more than the mean yield in May.

* * * * *

Problem 4. A school teacher decides to test the effectiveness of using a computer-based lesson to teach trigonometry. The teacher selects and pairs students of equal ability. One student from each pair is randomly chosen and assigned to a control group that receives the standard lesson, while the other student in the pair is then assigned to the experimental group that receives the computer-based lesson. On completion of the course, students in both groups sat for the same test to evaluate their learning outcomes. The test consists of multiple-choice questions, and is administered and marked online. The marks are given in the table below.

Pair	1	2	3	4	5	6	7	8	9	10
Control	70	65	70	79	72	60	53	50	72	91
Experiment	89	80	72	72	91	65	60	65	70	88

(a) Find the largest value of significance level, α, at which it could not be rejected that there is no difference in the two methods. State any assumption(s) necessary for the test to be valid.

(b) At the 10% significance level it is to be concluded that the experimental group scores higher than the control group by more than k marks. Find the values of k.

Solution.

Part (a). Let $D =$ experiment score $-$ control score. Let H_0: $\mu_D = 0$ and H_1: $\mu_D \neq 0$. We perform a two-tail paired-sample t-test. Assuming that D is normally distributed, under H_0, our test statistic is

$$\frac{\overline{D}}{S_D/\sqrt{10}} \sim \mathrm{t}(9).$$

From the sample, $\overline{d} = 7$ and $s_D = 9.5568$. Using G.C., the p-value is 0.0457. Thus, the largest value of α to not reject H_0 is 0.0457.

Part (b). Let H_0: $\mu_D = k$ and H_1: $\mu_D > k$. We perform a one-tail paired-sample t-test at the 10% significance level. Assuming that D is normally distributed, under H_0, our test statistic is

$$\frac{\overline{D} - k}{S_D/\sqrt{10}} \sim \mathrm{t}(9).$$

The observed test statistic is thus

$$\frac{7-k}{9.5568/\sqrt{10}} = 2.3163 - 0.33089k.$$

For H_0 to be rejected, we require

$$2.3163 - 0.33089k > t_{0.9} \implies k < \frac{2.3163 - t_{0.9}}{0.33089} = 2.82.$$

* * * * *

Problem 5. Two cyclists cycled to work every weekday and recorded the times (in minutes) that they took each day.

The records for one randomly chosen week is as follows, where Cyclist 1 took x minutes and Cyclist 2 took y minutes:

$$\sum (x - \overline{x})^2 = 90.5, \quad \sum y = 160.6, \quad \sum y^2 = 5244.8.$$

Assuming that the variances of the times they took each day are the same, find an unbiased estimate for the common variance.

During another randomly chosen week, the times (in minutes) that they took is recorded as follows:

Cyclist 1 (x)	30.1	21.9	34.1	33.8	c
Cyclist 2 (y)	31.7	30.5	40.1	30.2	28.1

The missing data is denoted by c.

Based on the second set of data, the hypothesis that the mean time of the two cyclists does not differ is not rejected at 5% significance level. Find the range of values of c. State any assumptions needed for your test to be valid.

Solution. Their individual sample variances are given by

$$s_X^2 = \frac{1}{n-1} \sum (x - \overline{x})^2 = \frac{1}{5-1}(90.5) = 22.625$$

and

$$s_Y^2 = \frac{1}{n-1}\left[\sum y^2 - \frac{1}{n}\left(\sum y\right)^2\right] = \frac{1}{5-1}\left[5244.8 - \frac{1}{5}(160.6)^2\right] = 21.582.$$

Thus, the pooled variance is

$$s_p^2 = \frac{s_X^2 + s_Y^2}{2} = \frac{22.625 + 21.582}{2} = 22.104.$$

Let $\mathrm{H_0}$: $\mu_X - \mu_Y = 0$ and $\mathrm{H_1}$: $\mu_X - \mu_Y \neq 0$. We perform a two-tail two-sample t-test at a 5% significance level. Assuming the pooled variance is the same, and that X and Y are normally distributed, under $\mathrm{H_0}$, our test statistic is

$$T = \frac{\overline{X} - \overline{Y}}{S_p\sqrt{1/5 + 1/5}} \sim \mathrm{t}(5 + 5 - 2) = \mathrm{t}(8).$$

From the sample,

$$\overline{x} - \overline{y} = \frac{119.9 + c}{5} - 32.12 = \frac{c}{5} - 8.14.$$

Our observed test statistic is thus

$$t = \frac{\overline{x} - \overline{y}}{s_p\sqrt{1/5 + 1/5}} = \frac{c/5 - 8.14}{\sqrt{22.104}\sqrt{2/5}} = 0.067261c - 2.7375.$$

For the null hypothesis to not be rejected, we must have

$$t_{0.025} < 0.067261c - 2.7375 < t_{0.975}.$$

Solving for c, we get

$$6.42 = \frac{t_{0.025} + 2.7375}{0.067261} < c < \frac{t_{0.975} + 2.7375}{0.067261} = 74.9.$$

* * * * *

Problem 6. Due to the differences in environment, the masses of a certain species of small amounts are believed to be greater in Region A than in Region B. It is known that the masses in both regions are normally distributed, with masses in Region A having a standard deviation of 0.004 kg and masses in Region B having a standard deviation of 0.09 kg. To test the theory, random samples are taken: 60 animals from Region A had a mean of 3.03 kg and 50 animals from Region B had a mean mass of 3.00 kg. Test at the 1% significance level, whether the animals of this species in Region A have a greater mean mass than those in Region B.

Will your conclusion be affected if we do not have the information that the masses of the animals are normally distributed? Explain your answer.

Solution. Let X kg and Y kg represent the mass of the animal in Regions A and B respectively. Let $\mathrm{H_0}$: $\mu_x - \mu_y = 0$ and $\mathrm{H_1}$: $\mu_x - \mu_y > 0$. We perform a one-tail two-sample z-test at 1% significance level. Under $\mathrm{H_0}$,

$$\overline{X} - \overline{Y} \sim \mathrm{N}\left(0, \frac{0.04^2}{60} + \frac{0.09^2}{50}\right).$$

From the sample, $\overline{x} = 3.03$ and $\overline{y} = 3.00$. Using G.C., the p-value is 0.0145, which is greater than the 1% significance level. Thus, we do not reject $\mathrm{H_0}$ and conclude there is insufficient evidence at 1% significance level that the animals of this species in Region A have a greater mean mass than those in Region B.

The conclusion will not be affected. This is because the sizes of both samples are large, so the test statistics will be approximately the same by the Central Limit Theorem.

Assignment A18B

Problem 1. A lorry is transporting a large number of red apples. As it passes over a bump in the road 10 apples fall off its back. The masses (in g) of the fallen apples are summarized by

$$\sum(x - 100) = 23.7, \quad \sum(x - 100)^2 = 1374.86.$$

(a) Treating the fallen apples as being a random sample, determine a symmetric 99% confidence interval for the mean mass of a red apple, stating any assumptions that you have made.

(b) A two-tailed test of the null hypothesis H_0: $\mu_x = 116$, where μ_x is the population mean mass of red apples, is to be carried out at the 5% level of significance. Using the confidence interval obtained in (i), state the conclusion of the test.

On its return down the road the lorry is carrying a large number of green apples. When it passes over the bump 15 green apples fall off its back and are collected. The masses (in g) of these apples are summarized by

$$\sum(y - 110) = -73.2, \quad \sum(y - 110)^2 = 2114.33.$$

(c) Assuming that the distribution of the masses of green apples has the same variance as that for the red apples, and that the fallen apples constitute a random sample, obtain a pooled estimate of the common variance.

(d) Carry out an appropriate test on the data, using a 10% significance level, on whether the two distributions have the same mean. State any further assumptions needed for the test to be valid.

Solution.

Part (a). Let X g be the mass of a red apple. Let $\mu_X = \mathbb{E}[X]$. From the sample,

$$\overline{x} = 100 + \frac{23.7}{10} = 102.37 \quad \text{and} \quad s_X^2 = \frac{1}{10-1}\left(1374.86 - \frac{23.7^2}{10}\right) = 146.52.$$

Assuming that X is normally distributed, a 99% confidence interval is $(89.93, 114.81)$.

Part (b). Since $116 \notin (89.93, 114.81)$, at a 1% significance level, we can reject H_0. Thus, at a 5% significance level, we also reject H_0.

Part (c). Let Y g be the mass of a green apple. Let $\mu_Y = \mathbb{E}[Y]$. From the sample,

$$\overline{y} = 110 + \frac{-73.2}{15} = 105.12 \quad \text{and} \quad s_Y^2 = \frac{1}{15-1}\left(2114.33 - \frac{(-73.2)^2}{15}\right) = 125.51.$$

Let the pooled variance be s_p^2. Then

$$s_p^2 = \frac{(10-1)s_X^2 + (15-1)s_Y^2}{10+15-2} = 133.73.$$

Part (d). Let H_0: $\mu_X - \mu_Y = 0$ and H_1: $\mu_X - \mu_Y \neq 0$. Assuming that X and Y are independently and normally distributed, we perform a 2-tail 2-sample t-test at the 10% significance level. Under H_0,

$$\frac{\overline{X} - \overline{Y}}{s_p\sqrt{\frac{1}{10} + \frac{1}{15}}} \sim \mathrm{t}(10 + 15 - 2) = \mathrm{t}(23).$$

Using G.C., the p-value is 0.566, which is greater than our 10% significance level. We thus do not reject H_0 and conclude there is insufficient evidence at the 10% significance level that the two distributions have different means.

* * * * *

Problem 2. The thickness of plaque (measured in mm) in the carotid artery of 10 randomly selected patients with mild atherosclerotic disease were measured. Two measurements are taken: thickness before treatment with vitamin E (x) and after two years of taking vitamin E daily (y). The readings are given below.

X	0.66	0.72	0.85	0.62	0.54	0.63	0.64	0.67	0.73	0.68
Y	0.60	0.65	0.79	0.63	0.59	0.55	0.64	0.70	0.68	0.64

The medical research team is interested on whether taking vitamin E daily for two years will reduce the thickness of plaque.

Explain which test the medical research team should use. Carry out a suitable test at the 5% level of significance, stating the necessary assumptions.

Solution. They should use a 1-tail paired-sample t-test since the observations are taken from the same patients.

Let $D = Y - X$. We perform a 1-tail paired-sample t-test at the 5% significance level, assuming that D is normally distributed. Our hypotheses are H_0: $\mu_D = 0$ and H_1: $\mu_D < 0$.

From the sample, $\overline{d} = -0.027$ and $s_D = 0.0457$. Under H_0,

$$\frac{\overline{D}}{s_D/\sqrt{10}} \sim \mathrm{t}(9).$$

Using G.C., the p-value is 0.0473, which is less than our significance level of 5%. Thus, we reject H_0 and conclude there is sufficient evidence at the 5% significance level that taking vitamin E daily for two years will reduce thickness of plaque.

* * * * *

Problem 3. A supermarket get its supplies of potatoes from two different suppliers, A and B. The weights, x and y grams of random samples of potatoes from supplier A and B respectively are summarized as follows.

$$\begin{array}{lll} \text{Supplier A:} \quad \text{Sample size } n_1 = 100, & \sum x = 10313, & \sum x^2 = 1072660. \\ \text{Supplier B:} \quad \text{Sample size } n_2 = 85, & \sum y = 8982, & \sum y^2 = 956540. \end{array}$$

(a) Test, at the 5% level of significance, whether the potatoes supplied by the two suppliers have the same mean weight.

(b) State 2 changes you would need to make to the test statistic if both sample sizes are small.

Solution.

Part (a). We perform a 2-tail 2-sample z-test at the 5% significance level. Our hypotheses are H_0: $\mu_X - \mu_Y = 0$ and H_1: $\mu_X - \mu_Y \neq 0$. From the sample,

$$\overline{x} = \frac{10313}{100} = 103.13 \quad \text{and} \quad s_X^2 = \frac{1}{100-1}\left(1072660 - \frac{10313^2}{100}\right) = 91.720,$$

and

$$\overline{y} = \frac{8982}{85} = 105.67 \quad \text{and} \quad s_Y^2 = \frac{1}{85-1}\left(956540 - \frac{8982^2}{85}\right) = 88.176.$$

Under H_0,

$$\overline{X} - \overline{Y} \sim \mathrm{N}\left(0, \frac{91.720}{100} + \frac{88.176}{85}\right) \text{ approximately.}$$

Using G.C., the p-value is 0.0692, which is greater than our significance level of 5%. Thus, we do not reject $\mathrm{H}_0$and conclude there is insufficient evidence at the 5% significance level that the potatoes supplied by the two suppliers have different mean weight.

Part (b). The test statistic would follow a t-distribution, and would use a pooled variance.

A18C Hypothesis Testing - Goodness of Fit and Independence

Tutorial A18C

Problem 1. The positioning of nests in shrubs was noted for a species of birds.

Nest position	N	NE	E	SE	S	SW	W	NW
Number of nests	65	73	67	51	47	45	45	48

Using a 5% significance level, test the hypothesis that the birds have no directional preference in positioning their nests.

Solution. Let H_0: data consistent with (discrete) uniform distribution, and H_1: data inconsistent with uniform distribution. We take a 5% level of significance.

Under H_0, the observed and expected frequencies are

Nest position	N	NE	E	SE	S	SW	W	NW
O_i	65	73	67	51	47	45	45	48
E_i	55.125	55.125	55.125	55.125	55.125	55.125	55.125	55.125

Our test statistic is $\sum(O_i - E_i)^2/E_i \sim \chi^2_{8-1} = \chi^2_7$, so our p-value is 0.0228, which is less than our significance level of 5%. Thus, we reject H_0 and conclude there is sufficient evidence at the 5% level of significance that the birds have a directional preference in positioning their nests.

* * * * *

Problem 2. When a tetrahedral dice is thrown, the number landing face down counts as the score. Four such dice are thrown 200 times and the number of fours obtained are shown.

Number of fours	0	1	2	3	4
Frequency	20	47	83	41	9

(a) Use a χ^2 test at the 5% level to test whether the dice are fair, that is, a $\mathrm{B}(4, 1/4)$ model is appropriate.

(b) Use the data to estimate a value for p, the probability that the score from a tetrahedral die is 4. Test at the 5% level whether a $\mathrm{B}(4, p)$ model is appropriate.

Solution.

Part (a). Let H_0: data is consistent with $\mathrm{B}(4, 1/4)$, and H_1: data is inconsistent with $\mathrm{B}(4, 1/4)$. We take a 5% level of significance.

Under H_0, the observed and expected frequencies are

Number of fours	0	1	2	3	4
O_i	20	47	83	41	9
E_i	63.281	84.375	42.187	9.375	0.7812

Since the expected frequency in the last columns is less than 5, we group the last two columns together. Our test statistic is hence $\sum(O_i - E_i)^2/E_i \sim \chi^2_{4-1} = \chi^2_3$, giving a p-value of 3.60×10^{-52}, which is less than our significance level of 5%. Thus, we reject H_0 and conclude there is sufficient evidence at the 5% level of significance that a $\mathrm{B}(4, 1/4)$ model is not appropriate.

Part (b). Note that

$$\widehat{p} = \frac{\overline{x}}{n} = \frac{1.86}{4} = 0.465.$$

Let H_0: data is consistent with $\mathrm{B}(4, 0.465)$, and H_1: data is inconsistent with $\mathrm{B}(4, 0.465)$. We take a 5% level of significance.

Under H_0, the observed and expected frequencies are

Number of fours	0	1	2	3	4
O_i	20	47	83	41	9
E_i	16.385	56.964	74.267	43.033	9.3507

Since we estimated $\widehat{p}$ using $\overline{x}$, our test statistic is $\sum(O_i - E_i)^2/E_i \sim \chi^2_{5-2} = \chi^2_3$, giving a p-value of 0.299, which is larger than our significance level of 5%. Thus, we do not reject H_0 and conclude there is sufficient evidence at the 5% level of significance that a $\mathrm{B}(4, p)$ model is appropriate.

* * * * *

Problem 3. The accident rates, taken over a twelve-month period, for the workers in a particular company, classified by age, are given in the following table.

Age (years)	18-25	26-40	41-50	Over 50	Total
At least one accident	112	156	75	77	420
No accidents	175	267	179	228	849
Total	287	423	254	305	1269

Show that the data provides evidence, at the 0.1% significance level, that age and accident rate are not independent. Comment on the relation between age and accident rate.

Solution. Let H_0: age and accident rate are independent, and H_1: age and accident rate are dependent. We perform a χ^2 independence test at a 0.1% level of significance. Our test statistic is $\sum(O_i - E_i)^2/E_i \sim \chi^2_{(2-1)(4-1)} = \chi^2_3$. Under H_0, our p-vale is 6.31×10^{-4}, which is less than our significance level of 0.1%. Thus, we reject H_0 and conclude there is sufficient evidence, at the 0.1% level of significance, that age and accident rate are not independent.

From the data, the older workers are, the less likely they get into an accident, possibly because older workers have been working for a longer period of time, hence they have more experience when it comes to safety.

* * * * *

Problem 4. A random variable X is equally likely to take each integer value from 1 to n inclusive. In a random sample of N observations, the value of r is obtained O_r times for $r = 1, 2, \ldots, n$. Show that the calculated χ^2 statistics for these data can be expressed in the form

$$\frac{n}{N}\sum_{r=1}^{n} O_r^2 - N.$$

The table shows the monthly figures for road deaths occurring in a certain country over a 12-month period.

Month	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
Deaths	326	356	301	292	279	286	285	298	308	284	291	303

(a) Show that at the 5% significance level, these monthly figures conform to a uniform distribution.

(b) Partition the data in four groups of consecutive months, starting on December, March, June and September respectively. Show that there is evidence, at 5% significance level, of a seasonal variation and describe it. State what might cause the variation.

Solution. Note that

$$E_r = \mathbb{P}[X_1 = r] + \cdots + \mathbb{P}[X_N = r] = N\,\mathbb{P}[X = r] = \frac{N}{n}.$$

Hence,

$$\chi^2 = \sum_{r=1}^{n} \frac{(O_r - E_r)^2}{E_r} = \sum_{r=1}^{n} \frac{(O_r - N/n)^2}{N/n} = \frac{n}{N} \sum_{r=1}^{n} \left(O_r^2 - \frac{2N}{n} O_r + \frac{N^2}{n^2} \right).$$

Since $\sum O_r = N$, we have

$$\chi^2 = \frac{n}{N} \sum_{r=1}^{n} O_r^2 - \frac{n}{N} \cdot \frac{2N}{n} \cdot N + \frac{n}{N} \cdot n \cdot \frac{N^2}{n^2} = \frac{n}{N} \sum_{r=1}^{n} O_r^2 - N.$$

Part (a). Let H_0: data is consistent with uniform distribution, and H_1: data is inconsistent with uniform distribution. We take a 5% level of significance.

Our test statistic is $\sum (O_r - E_r)^2/E_r \sim \chi^2_{12-1} = \chi^2_{11}$. Under H_0, this evaluates to

$$\frac{n}{N} \sum_{r=1}^{n} O_r^2 - N = \frac{12}{3609}(1090553) - 3609 = 17.111,$$

which gives a p-value of 0.10464, which is greater than our 5% level of significance. Thus, we do not reject H_0 and conclude there is sufficient evidence at a 5% significance level that these monthly figures conform to a uniform distribution.

Part (b). Grouping by season, we get

Month	Dec – Feb	Mar – May	Jun – Aug	Sep – Nov
Deaths	983	857	891	878

Let H_0: data is consistent with uniform distribution, and H_1: data is inconsistent with uniform distribution. We take a 5% level of significance.

Our test statistic is $\sum (O_r - E_r)^2/E_r \sim \chi^2_{4-1} = \chi^2_3$. Under H_0, this evaluates to

$$\frac{n}{N} \sum_{r=1}^{n} O_r^2 - N = \frac{4}{3609}(3265503) - 3609 = 10.288,$$

which gives a p-value of 0.0358, which is less than our 5% level of significance. Thus, we reject H_0 and conclude there is sufficient evidence at a 5% significance level that there is a seasonal variation, which peaks during winter (Dec – Feb). This could be due to shorter days in winter, which leads to lower visibility and hence more car accidents, resulting in more deaths.

* * * * *

Problem 5.

(a) The random variable X has a normal distribution with mean 5 and standard deviation 3. The random variable Y is given by $Y = \left(\frac{X-5}{3}\right)^2$. State the distribution of Y, giving the value of any associated parameter. By considering the distribution of Y and X, use two methods to find $\mathbb{P}[Y < 1.35]$.

(b) A wine store wished to investigate whether there was an association between the sex of customers and their preference for red or white wine. During one week, 200 of the store's customers were questioned. The results are shown in the table.

	Red	White	No preference	Total
Male	59	37	26	122
Female	25	40	13	78
Total	84	77	39	200

Stating any necessary assumption, test at the 5% significance level whether there is an association between the sex of a customer and wine preference. Show in your working the contribution to the test statistic in each cell of the table.

Solution.

Part (a). Note that $(X-5)/3 \sim \mathrm{N}(0,1)$. Hence, $Y = Z^2 \sim \chi^2_1$.

Using the distribution of Y, we have $\mathbb{P}[Y < 1.35] = 0.755$. Alternatively, using the distribution of $(X-5)/3 = Z$, we have

$$\mathbb{P}[Y < 1.35] = \mathbb{P}\big[Z^2 < 1.35\big] = \mathbb{P}\Big[-\sqrt{1.35} < Z < \sqrt{1.35}\Big] = 0.755.$$

Part (b). Let H_0: sex of a customer and wine preference are independent, and H_1: sex of a customer and wine preference are dependent. We take a 5% level of significance. We assume that the data comes from a random sample and are representative of the population.

Under H_0, the expected frequencies are given by

	Red	White	No preference
Male	51.24	46.97	23.79
Female	32.76	30.03	15.21

Hence, the contributions are

	Red	White	No preference
Male	1.1752	2.1163	0.2053
Female	1.8381	3.3101	0.32111

Our test statistic is $\sum (O_i - E_i)^2/E_i \sim \chi^2_{(2-1)(3-1)} = \chi^2_2$, which gives a p-value of 0.0113, which is less than the 5% significance level. Thus, we reject H_0 and conclude there is sufficient evidence at a 5% significance level that there is an association between the sex of a customer and wine preference.

Problem 6. During the Second World War, records of the number of V-2 rockets launched by the Germans and their exact point of impact in South London were recorded. The rockets were launched from Germany and did not have sophisticated guidance systems. The particular part of the city studied was divided into 576 regions each having an area of 0.25 km^2. The table gives the number of regions experiencing x hits.

x	0	1	2	3	4	5	Total
Frequency	229	211	93	35	7	1	576

Give reasons why these data might be expected to fit a Poisson distribution. Test the above data at 1% significance level for a goodness of fit to a Poisson distribution with mean 0.95, listing the expected frequencies.

An important factory covered two neighbouring regions. It was estimated that two direct hits would cripple the factory. Find the probability that it was crippled.

Solution. Since the rockets do not have sophisticated guidance systems, the rocket strikes effectively form a Poisson process, with each region having an equal probability of being struck by each rocket.

Let H_0: data is consistent with Po(0.95), and H_1: data is inconsistent with Po(0.95). Under H_0, the expected frequencies are given by

x	0	1	2	3	4	5
O_i	229	211	93	35	7	1
E_i	222.76	211.62	100.52	31.832	7.5601	1.4364

Since the last column has an expected frequency of less than 5, we combine it with the second-last column. Our test statistic is hence $\sum(O_i - E_i)^2/E_i \sim \chi^2_{5-1} = \chi^2_4$, which gives a p-value of 0.884, which is greater than our level of significance of 1%. Thus, we do not reject H_0 and conclude there is sufficient evidence at the 1% level of significance that the data is consistent with a Poisson distribution with mean 0.95.

Note that $X_1 + X_2 \sim \mathrm{Po}(0.95 + 0.95) = \mathrm{Po}(1.9)$, so the desired probability is

$$\mathbb{P}[X_1 + X_2 \geq 2] = 1 - \mathbb{P}[X_1 + X_2 \leq 1] = 0.566.$$

* * * * *

Problem 7. A meteorologist conjectures that, at a certain location, the rainfall (x mm) on June 30th may be regarded as an observation from the exponential distribution with probability density function given by

$$f(x) = \lambda \mathrm{e}^{-\lambda x},$$

where $x \geq 0$. Show that $\mathbb{E}[X] = 1/\lambda$.

It is known that during the 25-year period 1992 to 2016, a total of 260 mm of rain fell on June 30th.

(a) Find an estimate for the value of λ. Hence, show that the probability that more than 20 mm of rain will fall at this location on June 30th, 2017 is 0.146.

The individual rainfall measurements on June 30th for the period 1961 to 1985 are summarized in the table below.

Rainfall (x mm)	$x \leq 4$	$4 < x \leq 9$	$9 < x \leq 16$	$x > 16$
Number of days	10	5	6	4

(b) Using your estimate of λ obtained above as the true value, test the goodness of fit of the exponential distribution to the data, using a 5% significance level.

Solution. We have

$$\mathbb{E}[X] = \int_0^\infty x f(x)\,\mathrm{d}x = \int_0^\infty x\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \left[-x\mathrm{e}^{-\lambda x} - \frac{1}{\lambda}\mathrm{e}^{-\lambda x}\right]_0^\infty = \frac{1}{\lambda}.$$

Part (a). Note that

$$\lambda = \frac{1}{\overline{x}} = \frac{1}{25/260} = \frac{5}{52}.$$

Hence, the desired probability is

$$\mathbb{P}[X > 20] = \mathrm{e}^{-\frac{5}{52}(20)} = 0.146.$$

Part (b). Let H_0: data is consistent with Exp(5/52), and H_1: data is inconsistent with Exp(5/52). We take a 5% level of significance.

Under H_0, the expected frequencies are

Rainfall (x mm)	$x \leq 4$	$4 < x \leq 9$	$9 < x \leq 16$	$x > 16$
O_i	10	5	6	4
E_i	7.9822	6.4956	5.1545	5.3678

Our test statistic is $\sum(O_i - E_i)^2/E_i \sim \chi^2_{4-1} = \chi^2_3$, which gives a p-value of 0.719, which is greater than the 5% significance level. Thus, we do not reject H_0 and conclude there is sufficient evidence at a 5% significance level that the data is consistent with Exp(5/52).

* * * * *

Problem 8. The following table shows the number of participants who received Gold, Silver and Bronze awards in a telematch.

	Gold	Silver	Bronze	Total
Male	50	s		120
Female				
Total	100	100		300

Copy and complete the missing entries in the observed frequency table. Construct in similar form, the expected frequency table for a χ^2 test that the gender of the participant is independent of the type of awards. State your hypotheses clearly. Show that $\chi^2 = \frac{1}{12}\left(s^2 - 70s + 1300\right)$. What is the range of s that leads to the rejection of the hypothesis at 2.5% of significance?

Solution. The complete observed frequency table is

	Gold	Silver	Bronze	Total
Male	50	s	$70 - s$	120
Female	50	$100 - s$	$30 + s$	180
Total	100	100	100	300

Let H_0: gender and type of award are independent, and H_1: gender and type of award are dependent. Under H_0, the expected frequency table is

	Gold	Silver	Bronze	Total
Male	40	40	40	120
Female	60	60	60	180
Total	100	100	100	300

Hence, the test statistic is

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{25}{6} + \left(\frac{1}{40} + \frac{1}{60}\right)\left((s-40)^2 + (30-s)^2\right) = \frac{s^2 - 70s + 1300}{12}.$$

To reject H_0 at a 2.5% significance level, we require $\chi^2 \geq 7.3778$, so $s \leq 31$ or $s \geq 39$. But both s and $70 - s$ must be positive integers, so we ultimately have $0 \leq s \leq 31$ or $39 \leq s \leq 70$.

* * * * *

Problem 9. A large international company employs university graduates having Class I, II or III degrees. After one year's employment, the performance of 250 of their graduates was assessed. It was found that 32 were graded A (high), 140 were graded B (average) and the rest were graded C (below average). Of these 250 graduates, 30 had Class I degrees, 150 had Class II degrees and the rest had Class III degrees.

(a) Assuming that performance grade and degree class are independent, draw up a table showing the expected frequencies of each performance grade for each degree class. Explain why, in this case, two rows, or two columns, must be combined in order for a χ^2 test of independence to be applied.

(b) Of the graduates with Class I degrees, 8 were graded A and 14 were graded B. Of those with Class II degrees, 21 were graded A and 90 were graded B. The rest were each graded C. Using the data for all 250 graduates, test, at the 5% significance level, whether performance grade and degree class are independent. State any assumptions necessary for the validity of your test.

Solution.

Part (a). Let H_0: performance grade and degree class are independent, and H_1: performance grade and degree class are dependent.

From the given data, we can construct the expected frequency table under H_0:

	Class I	Class II	Class III	Total
A	3.84	19.2	8.96	32
B	16.8	84	39.2	140
C	9.36	46.8	21.84	78
Total	30	150	70	250

Since the expected frequency of a Class I degree, performance grade A graduate is less than 5, two rows or two columns must be combined for a χ^2 test to be applied.

Part (b). Combining the rows corresponding to performance grades A and B, the expected frequency table becomes

	Class I	Class II	Class III	Total
A/B	20.64	103.2	48.16	172
C	9.36	46.8	21.84	78
Total	30	150	70	250

Further, from the given data, the observed frequency table is

	Class I	Class II	Class III	Total
A/B	22	111	39	172
C	8	39	31	78
Total	30	150	70	250

We take a 5% significance level and assume that the graduates were randomly hired and are thus representative of the population.

Our test statistic is $\sum(O_i - E_i)^2/E_i \sim \chi^2_{(2-1)(3-1)} = \chi^2_2$, which gives a p-value of 0.0206, which is less than our significance level of 5%. Thus, we reject H_0 and conclude there is sufficient evidence at a 5% significance level that performance grade and degree class are not independent.

Assignment A18C

Problem 1. For a long time, experts have been trying to explain the complex relations and interactions between leaders and other members in an organization. A study was carried out to investigate the relevance of a leader's gender in adopting a specific leadership style. A survey was given to a random sample of 79 people holding leadership positions from various organizations and institutions. Their responses to the questionnaire served to identify their dominant leadership style. The results are given in the following contingency table.

	Authoritarian	**Democratic**	**Laissez-faire**
Male	12	22	9
Female	20	13	3

Carry out a test for the independence of the two factors, gender and the dominant leadership style.

Discuss what the test indicates about the association, if any, between the two factors. You should refer to the p-value for your test and the largest two contributions to the test statistic.

Solution. Our hypotheses are H_0: gender and dominant leadership style are independent, H_1: gender and dominant leadership style are dependent. Under H_0, the expected frequencies are

	Authoritarian	**Democratic**	**Laissez-faire**
Male	17.418	19.051	6.5316
Female	14.582	15.949	5.4684

The test statistic is $\sum (O_i - E_i)^2/E_i \sim \chi^2_2$. Using G.C., the p-value is 0.034269. The individual contributions to the test statistic are

	Authoritarian	**Democratic**	**Laissez-faire**
Male	1.6853	0.4565	0.9328
Female	2.0131	0.5453	1.1142

The largest two contributions come from the authoritarian leadership style.

The small p-value indicates that the two factors are associated. Further, the large contributions from the authoritarian leadership style indicates that fewer people identify with the authoritarian leadership style than expected.

* * * * *

Problem 2.

(a) After the implementation of the Electronic Road Pricing Scheme, a survey was conducted where the number of vehicles passing the gantry point at Pan-Island Expressway in each period of 20 seconds was recorded. The results for 100 periods are as follows:

Number of vehicles in the period	0	1	2	3	4	5
Frequency	24	36	28	8	3	1

(i) Calculate the mean number of vehicles per period. Perform a goodness of fit test, at the 5% level of significance to determine whether the data could be a sample from a Poisson distribution.

(ii) How would the above test change if it were specified as a Po(1.33) distribution instead?

(b) A survey of 100 families, known to be regular television viewers, was undertaken. They were asked which of the two channels they watched most during an average week. A summary of their replies is given in the following table, together with the region in which they lived.

Region	North	South	East	West
Channel 5	10	17	$10+a$	$23-a$
Channel 8	5	8	$20-a$	$7+a$

To test the hypothesis that there is no association between the channel watched most and the region, show that the χ^2 statistic in terms of a can be simplified as

$$\chi^2 = \frac{10a^2 - 130a + 479}{36}.$$

Find the set of values of a that would result in the assumption not being rejected at the 5% level of significance.

Solution.

Part (a).

Part (a)(i). From the data, $\overline{x} = 1.33$. Our hypotheses are H_0: data consistent with Po(1.33) and H_1: data inconsistent with Po(1.33). We take a 5% level of significance.

Under H_0, the expected frequencies are

x	0	1	2	3	4	≥ 5
E_i	26.448	35.175	23.392	10.370	3.4481	1.1667

The last two columns have expected frequencies less than 5, so we combine the last three columns. The test statistic is $\sum(O_i - E_i)^2/E_i \sim \chi^2_2$. Using G.C., the p-value is 0.417, which is greater than our 5% significance level. Hence, we do not reject H_0 and conclude there is insufficient evidence at the 5% level that the data is inconsistent with Po(1.33).

Part (a)(ii). The degrees of freedom would be $4 - 1 = 3$, hence the test statistic would follow a χ^2_3 distribution instead.

Part (b). Our hypotheses are H_0: most watched channel and region are independent, H_1: most watched channel and region are dependent. Under H_0, the expected frequencies are

Region	North	South	East	West
Channel 5	9	15	18	18
Channel 8	6	10	12	12

The test statistic is

$$\begin{aligned}\sum \frac{(O_i - E_i)^2}{E_i} &= \frac{(10-9)^2}{9} + \frac{(5-6)^2}{6} + \frac{(17-15)^2}{15} + \frac{(8-10)^2}{10} \\ &\quad + \frac{(10+a-18)^2}{18} + \frac{(20-a-12)^2}{12} + \frac{(23-a-18)^2}{18} + \frac{(7+a-12)^2}{12} \\ &= \frac{17}{18} + \left(\frac{1}{18} + \frac{1}{12}\right)\left[(a-8)^2 + (a-5)^2\right] \\ &= \frac{10a^2 - 130a + 479}{36}.\end{aligned}$$

The critical value for a 5% level of significance is 7.815. Hence, to not reject H_0, we require

$$\frac{10a^2 - 130a + 479}{36} \leq 7.815 \implies 1.76 \leq a \leq 11.2.$$

Since a is an integer, the required range of values of a is $2 \leq a \leq 11$.

* * * * *

Problem 3. The proportions of blood types A, B, AB and O in the population of a country are p_1, p_2, p_3, p_4 respectively, where $\sum_{i=1}^{4} p_i = 1$. In order to test whether the population of a city in the country conforms to these figures, a random sample of size n is selected and the numbers of people with blood types A, B, AB and O are found to be a, b, c and d respectively.

(a) Show that the χ^2 statistic for a goodness of fit test simplifies to

$$\chi^2 = \frac{a^2}{np_1} + \frac{b^2}{np_2} + \frac{c^2}{np_3} + \frac{d^2}{np_4} - n.$$

(b) It is given that $p_1 = p_4$, $p_2 = 3p_3$ and the values of a, b, c, d are 26, 19, 10 and 45 respectively. Denoting the common value of p_1 and p_4 by p, show that

$$\chi^2 = \frac{2701}{100p} + \frac{661}{75\,(1-2p)} - 100.$$

Hence, find the value of p_0 of p for which this value of χ^2 is a minimum.

(c) Carry out the goodness of fit test at the 10% significance level, with $p = p_0$.

(d) State, giving your reason, the conclusion of your test for values of p other than p_0.

(e) Using the values of a, b, c and d in (b), construct a 95% confidence interval for the proportion of people in the country with blood type A.

(f) Discuss one advantage and one disadvantage of finding a 95% confidence interval instead of a 99% confidence interval.

Solution.

Part (a). The test statistic is

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \sum \left(\frac{O_i^2}{E_i} - 2O_i + E_i \right) = \sum \left(\frac{O_i^2}{E_i} \right) - 2n + n$$
$$= \frac{a^2}{np_1} + \frac{b^2}{np_2} + \frac{c^2}{np_3} + \frac{d^2}{np_4} - n.$$

Part (b). Note that $p_1 + p_2 + p_3 + p_4 = 1$. Hence,

$$p_2 = \frac{3\,(1-2p)}{4} \quad \text{and} \quad p_3 = \frac{1-2p}{4}.$$

The test statistic is thus

$$\chi^2 = \frac{26^2}{100p} + \frac{19^2}{100\left(\frac{3(1-2p)}{4}\right)} + \frac{10^2}{100\left(\frac{1-2p}{4}\right)} + \frac{45^2}{100p} - 100 = \frac{2701}{100p} + \frac{661}{75\,(1-2p)} - 100.$$

Using G.C., the minimum value of χ^2 is 6.47259, occurring when $p = p_0 = 0.35615$.

Part (c). Our hypotheses are H_0: data consistent with proportions p_1, p_2, p_3, p_4, and H_1: data inconsistent with proportions p_1, p_2, p_3, p_4. We take a level of significance of 10%. The test statistic is $\sum(O_i - E_i)^2/E_i \sim \chi_3^2$. From (c), the value of the test statistic is 6.47259, but the critical value for a 10% significance level is 6.2514. Hence, we reject H_0 and conclude there is sufficient evidence at the 10% level that the data is inconsistent with the proportions p_1, p_2, p_3, p_4.

Part (d). Because χ^2 already attains a minimum at $p = p_0$, the value of the test statistic for all other values of p will remain greater than the critical value of 6.2514m hence leading to a rejection of H_0 at the 10% significance level.

Part (e). A 95% confidence interval is $(0.17403, 0.34597)$.

Part (f). An advantage is that the interval will be smaller and hence more precise. A disadvantage is that we must accept a lower level of confidence.

Part X

Group B

B1 Graphs and Transformations I

Tutorial B1A

Problem 1. Without using a calculator, sketch the following graphs and determine their symmetries.

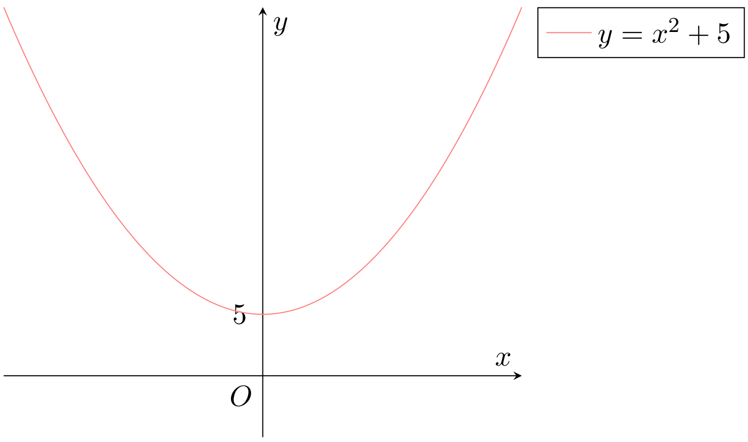
(a) $y = x^2 + 5$

(b) $y = 2x - x^3$

(c) $y = x^2 - 4x + 3$

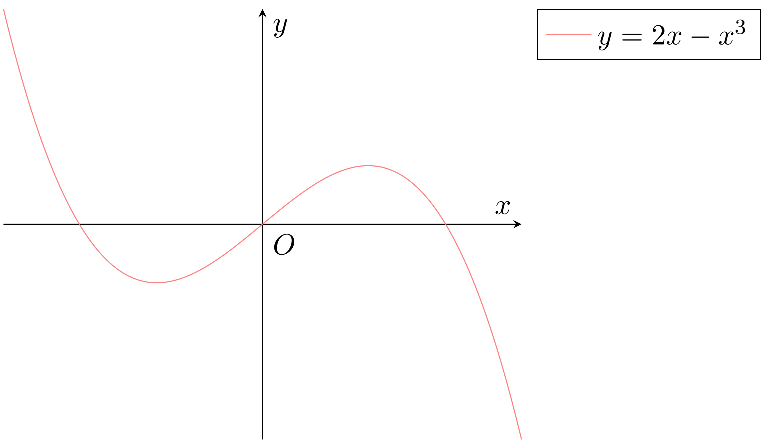
Solution.

Part (a).



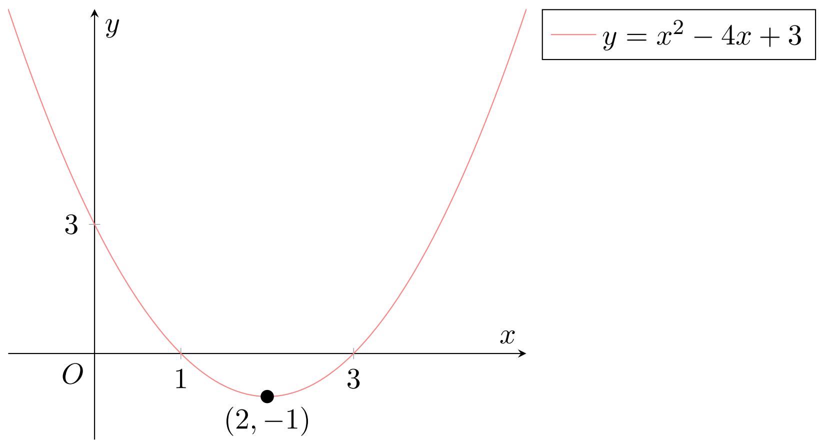
Symmetry: $x = 0$.

Part (b).



Symmetry: $(0, 0)$.

Part (c).



Symmetry: $x = 2$.

* * * * *

Problem 2. Sketch the following curves. Indicate using exact values, the equations of any asymptotes and the coordinates of any intersection with the axes.

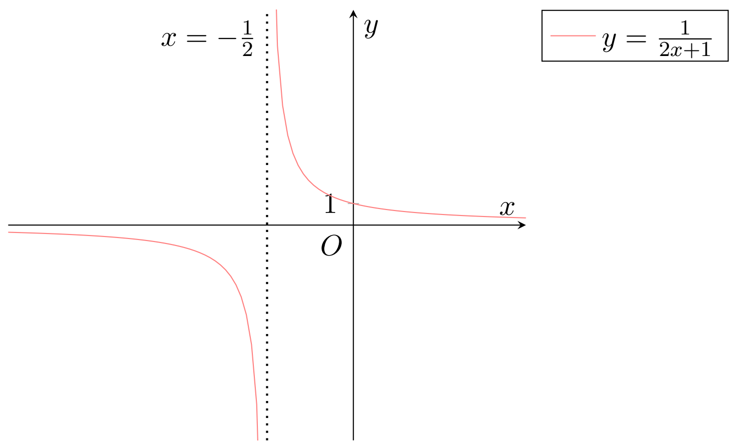
(a) $y = \frac{1}{2x+1}$

(b) $y = \frac{3x}{x-2}$

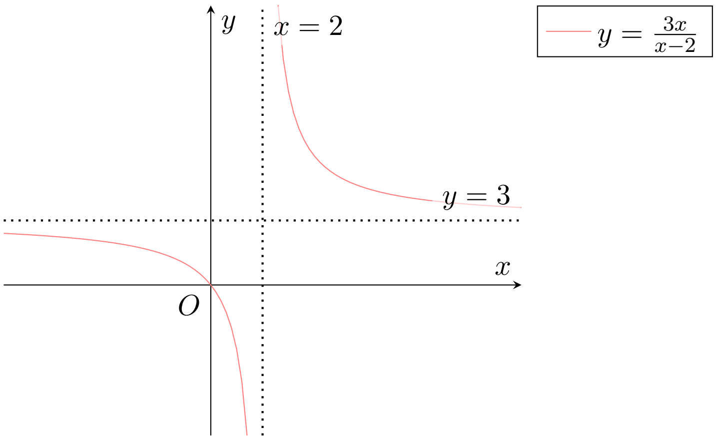
(c) $y = \frac{x^2+x-6}{2x-2}$

Solution.

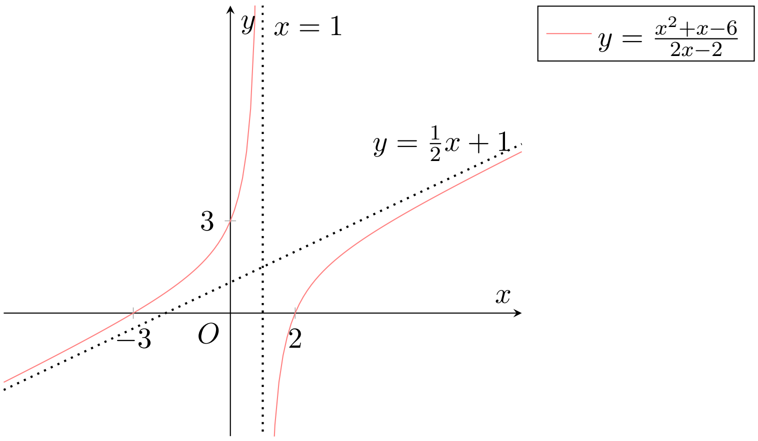
Part (a).



Part (b).



Part (c).



* * * * *

Problem 3. Sketch the following graphs

(a) $x^2 + 2x + 2y + 4 = 0$

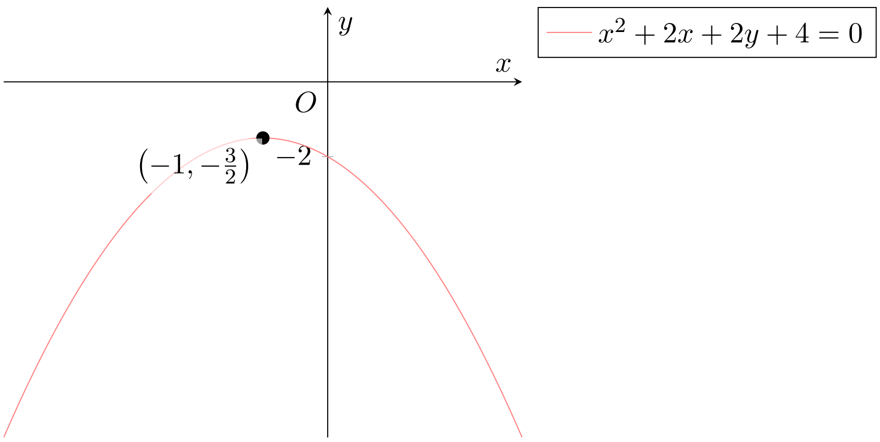
(b) $y^2 = x - 9$

(c) $y^2 = (x-2)^4 + 5$

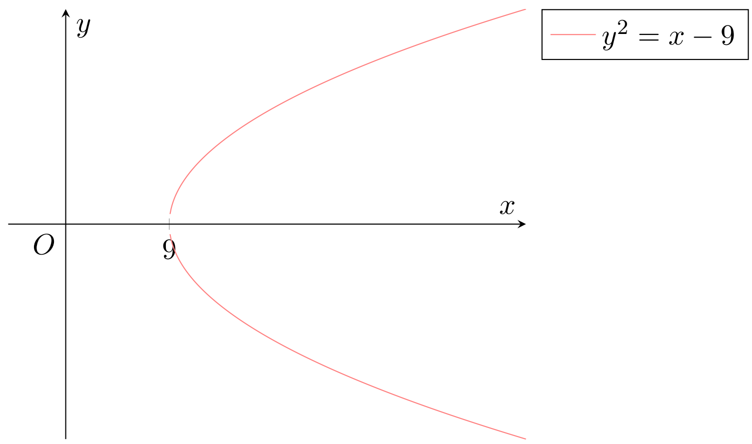
(d) $y = \tan\left(\frac{1}{2}x\right)$, $-2\pi \leq x \leq 2\pi$

Solution.

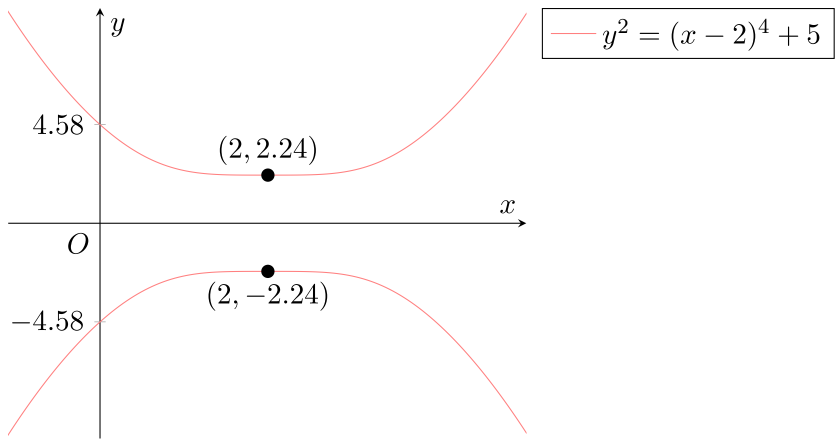
Part (a).



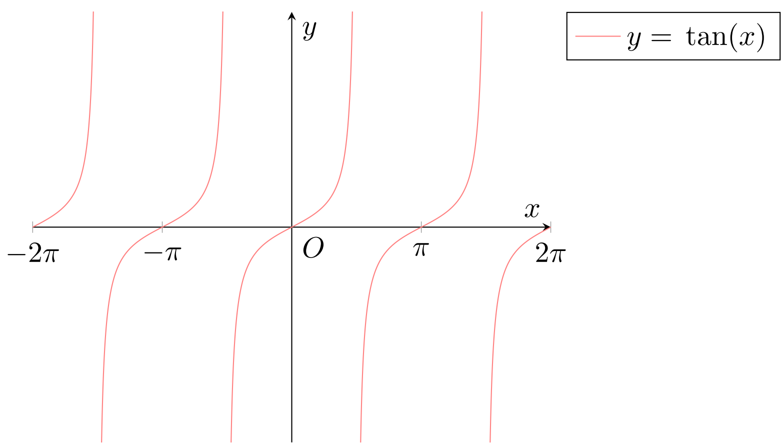
Part (b).



Part (c).



Part (d).



* * * * *

Problem 4. Sketch the following curves. Indicate using exact values, the equations of any asymptotes and the coordinates of any intersection with the axes.

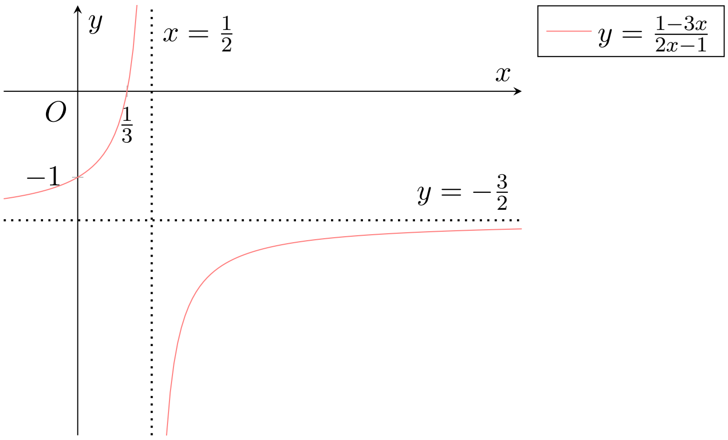
(a) $y = \frac{1-3x}{2x-1}$

(b) $y = \frac{ax}{x-a}$, $a < 0$

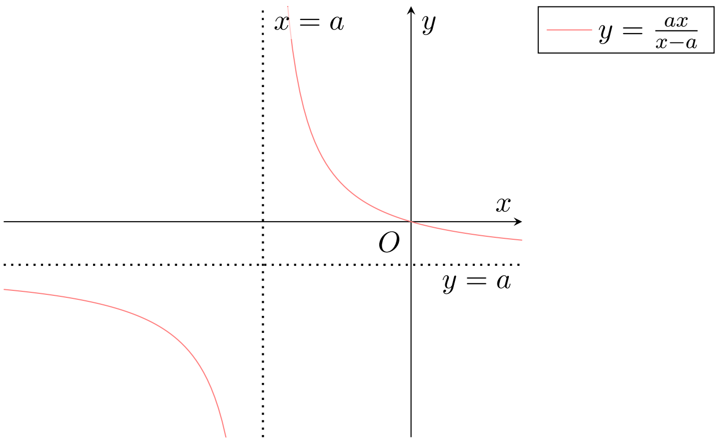
(c) $y = -\frac{b(x+3a)}{x+a}$, $a, b > 0$

Solution.

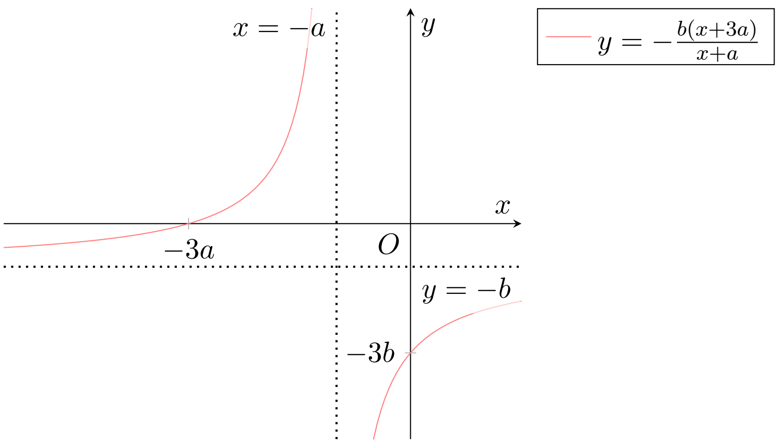
Part (a).



Part (b).



Part (c).



Problem 5. Sketch the following curves and find the coordinates of any turning points on the curves.

(a) $y = x + 2\sin x$, $0 \leq x \leq 2\pi$

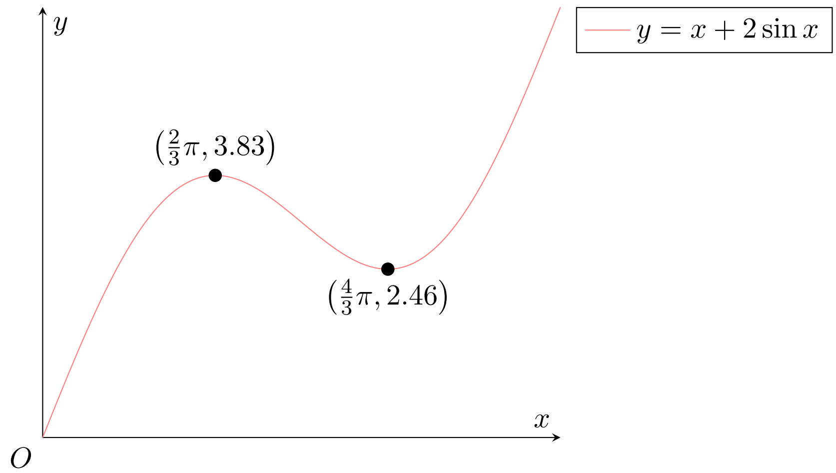
(b) $y = \frac{x}{\ln x}$, $x > 0$, $x \neq 1$

(c) $y = x\mathrm{e}^{-x}$

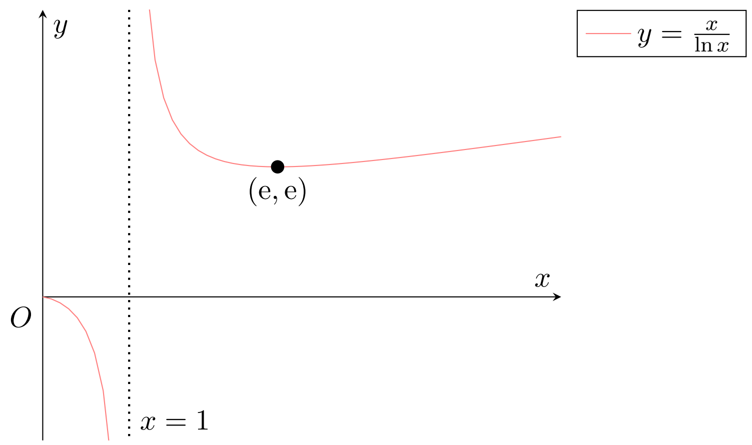
(d) $y = x\mathrm{e}^{-x^2}$

Solution.

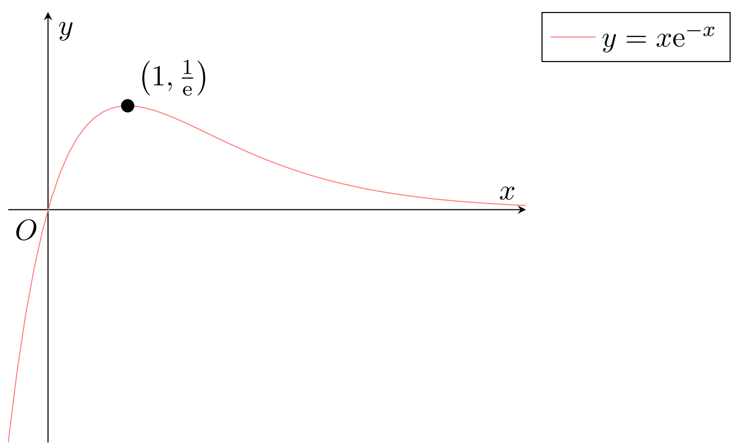
Part (a).



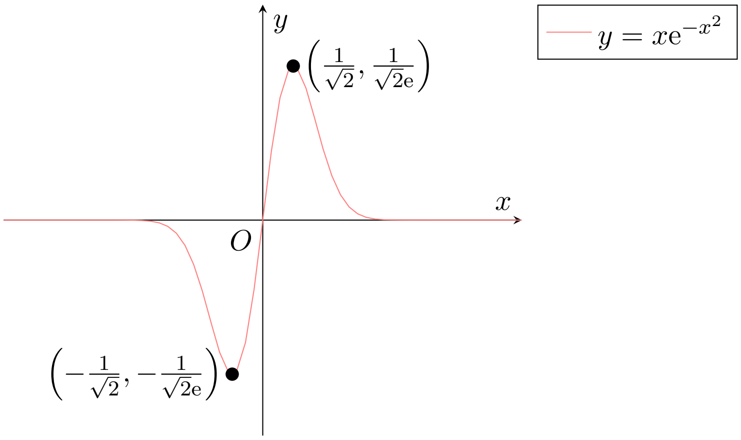
Part (b).



Part (c).



Part (d).



* * * * *

Problem 6. The equation of a curve C is $y = 1 + \frac{6}{x-3} - \frac{24}{x+3}$.

(a) Explain why $y = 1$ and $x = 3$ are asymptotes to the curve.

(b) Find the coordinates of the points where C meets the axes.

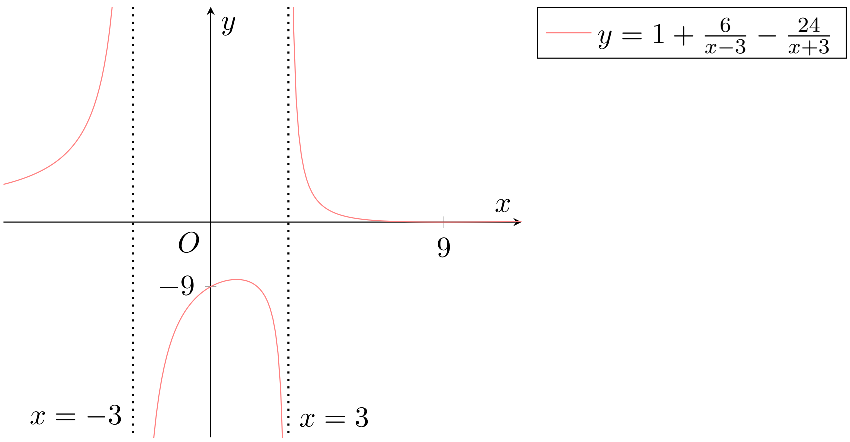
(c) Sketch C.

Solution.

Part (a). As $x \to \pm\infty$, $y \to 1$. Hence, $y = 1$ is an asymptote to C. As $x \to 3^{\pm}$, $y \to \pm\infty$. Hence, $x = 3$ is an asymptote to C.

Part (b). When $x = 0$, $y = -9$. When $y = 0$, $x = 9$. Hence, C meets the axes at $(0, -9)$ and $(9, 0)$.

Part (c).



* * * * *

Problem 7. The curve C has equation $y = \frac{ax^2+bx}{x+2}$, where $x \neq -2$. It is given that C has an asymptote $y = 1 - 2x$.

(a) Show (do not verify) that $a = -2$ and $b = -3$.

(b) Using an algebraic method, find the set of values that y can take.

(c) Sketch C, showing clearly the positions of any axial intercept(s), asymptote(s) and stationary point(s).

(d) Deduce that the equation $x^4 + 2x^3 + 2x^2 + 3x = 0$ has exactly one real non-zero root.

Solution.

Part (a).

$$y = \frac{ax^2 + bx}{x + 2} = \frac{(ax + b - 2a)(x + 2) - 2(b - 2a)}{x + 2} = ax + b - 2a - \frac{2(b - 2a)}{x + 2}.$$

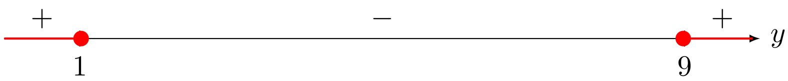
Since C has an asymptote $y = 1 - 2x$, we have $a = -2$ and $b - 2a = 1$, whence $b = -3$.

Part (b).

$$y = \frac{-2x^2 + -3x}{x + 2} \implies y(x + 2) = -2x^2 - 3x \implies 2x^2 + (3 + y)x + 2y = 0.$$

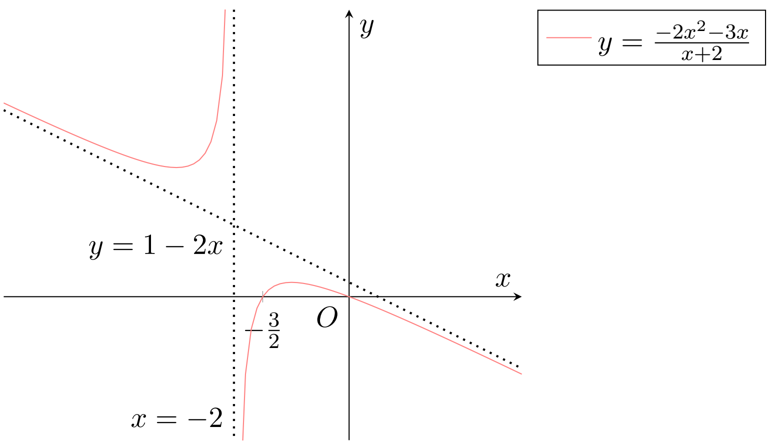
For all values that y can take on, there exists a solution to $2x^2 + (3 + y)x + 2y = 0$. Hence, $\Delta \geq 0$.

$$(3 + y)^2 - 4(2)(2y) \geq 0 \implies y^2 - 10y + 9 \geq 0 \implies (y - 1)(y - 9) \geq 0.$$



Thus, $\{y \in \mathbb{R} : y \leq 1 \text{ or } y \geq 9\}$.

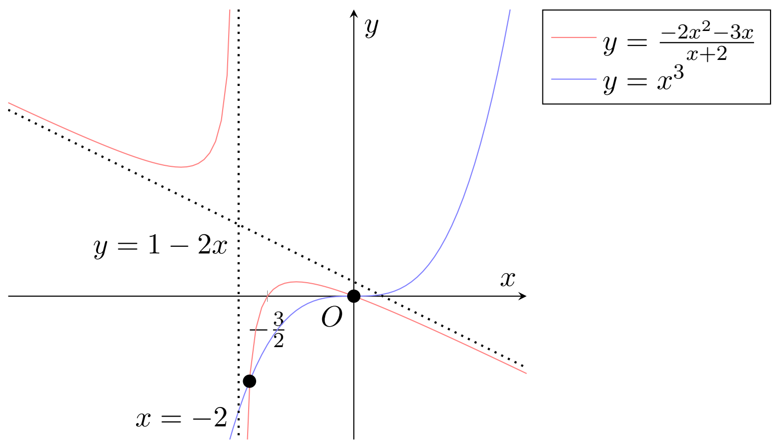
Part (c).



Part (d). Observe that

$$x^4 + 2x^3 + 2x^2 + 3x = 0 \implies x^3(x + 2) = -2x^2 - 3x \implies x^3 = \frac{-2x^2 - 3x}{x + 2}.$$

This motivates us to plot $y = x^3$ and $y = \frac{-2x^2-3x}{x+2}$ on the same graph.



We thus see that $y = x^3$ intersects $y = \frac{-2x^2-3x}{x+2}$ twice, with one intersection point being the origin. Thus, there is only one real non-zero root to $x^4 + 2x^3 + 2x^2 + 3x = 0$.

* * * * *

Problem 8. The curve C is defined by the equation $y = \frac{x}{x^2-5x+4}$.

(a) Write down the equations of the asymptotes.

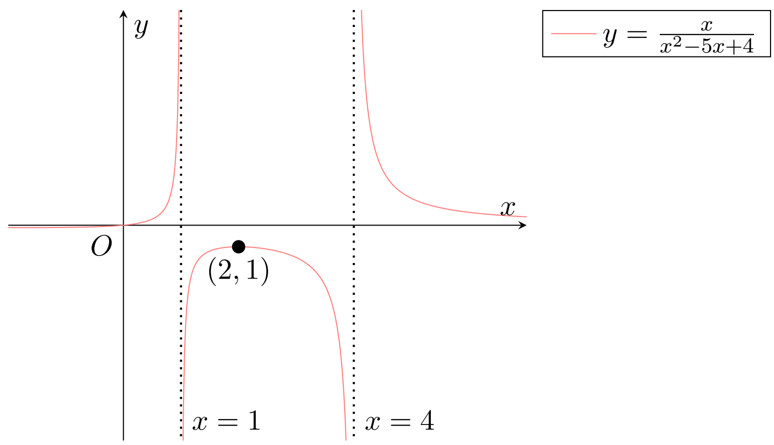
(b) Sketch C, indicating clearly the axial intercept(s), asymptote(s) and turning point(s).

(c) Find the positive value k such that the equation $\frac{x}{x^2-5x+4} = kx$ has exactly 2 distinct real roots.

Solution.

Part (a). As $x \to \pm\infty$, $y \to 0$. Hence, $y = 0$ is an asymptote. Observe that $x^2 - 5x + 4 = (x-1)(x-4)$. Hence, $x = 1$ and $x = 4$ are also asymptotes.

Part (b).



Part (c). Note that $x = 0$ is always a root of $\frac{x}{x^2-5x+4} = kx$. We thus aim to find the value of k such that $\frac{x}{x^2-5x+4} = kx$ has only one non-zero root.

We observe that if $k > 0$, $y = kx$ will intersect with $y = \frac{x}{x^2-5x+4}$ at least twice: before $x = 1$ and after $x = 4$. In order to have only one non-zero root, we must force the intersection point that comes before $x = 1$ to be at the origin $(0, 0)$. Hence, k is tangential to C at $(0, 0)$, thus giving $k = \mathrm{d}C/\mathrm{d}x|_{x=0}$.

$$k = \left.\frac{\mathrm{d}C}{\mathrm{d}x}\right|_{x=0} = \left.\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x}{x^2-5x+4}\right)\right|_{x=0} = \left.\frac{3x^2-10x+4}{(x^2-5x+4)^2}\right|_{x=0} = \frac{1}{4}.$$

Tutorial B1B

Problem 1. Without using a calculator, sketch the following graphs of conics.

(a) $y^2 - 4x = 12$

(b) $(x+1)^2 + y^2 = 4$

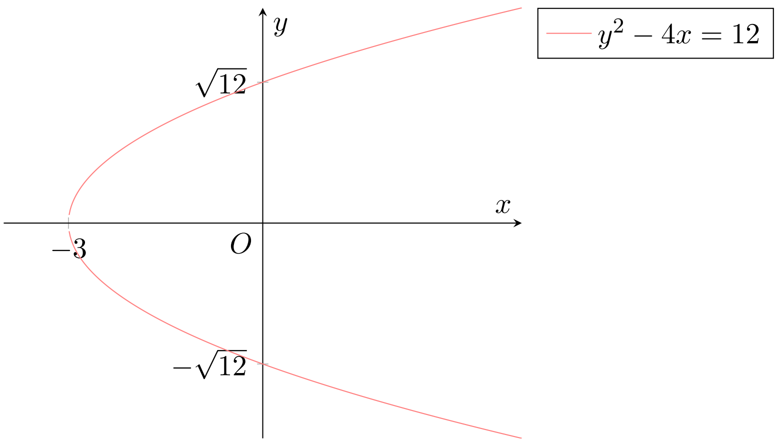
(c) $\frac{(x-3)^2}{9} + \frac{y^2}{2} = 1$

(d) $4x^2 + y^2 = 4$

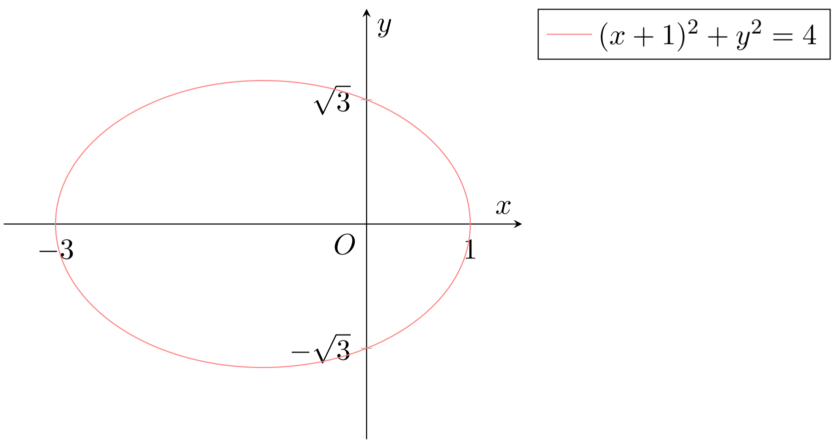
(e) $8y^2 - 2x^2 = 16$

Solution.

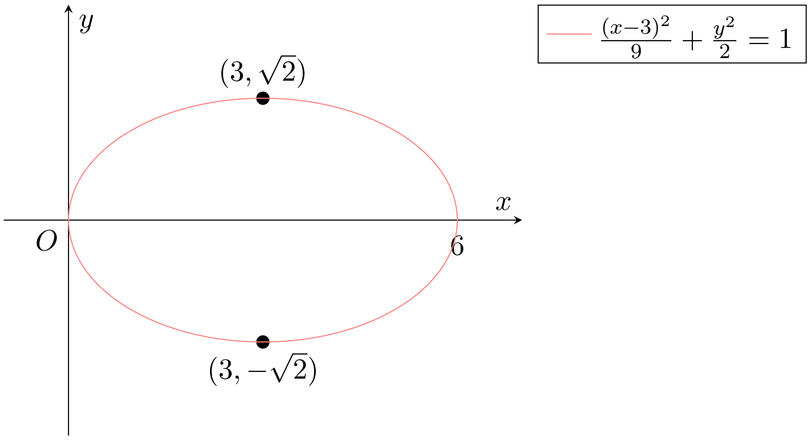
Part (a).



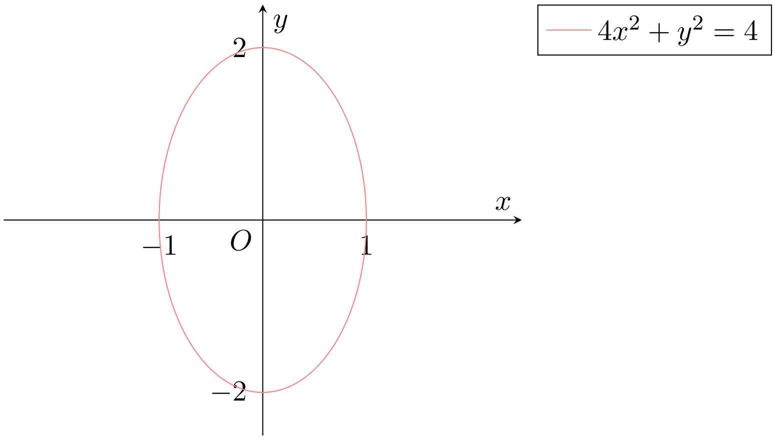
Part (b).



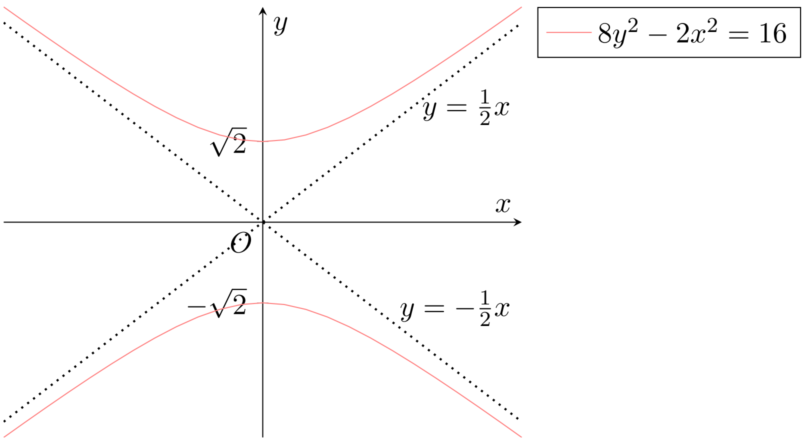
Part (c).



Part (d).



Part (e).



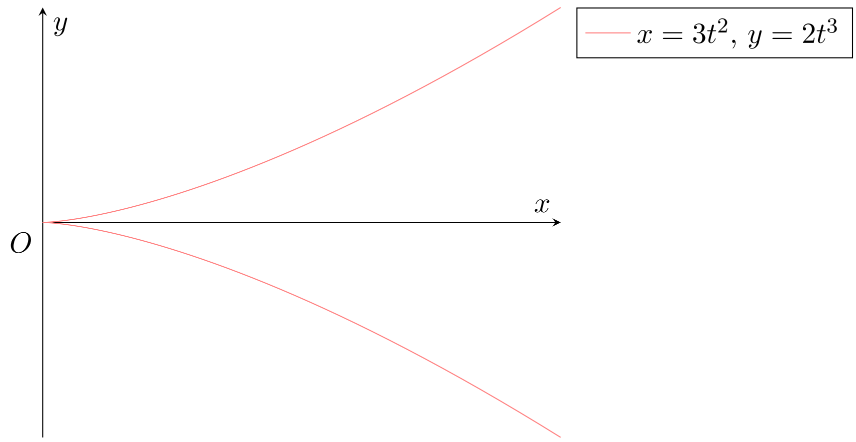
Problem 2. Sketch the curves defined by the following parametric equations, indicating the coordinates of any intersection with the axes.

(a) $x = 3t^2$, $y = 2t^3$

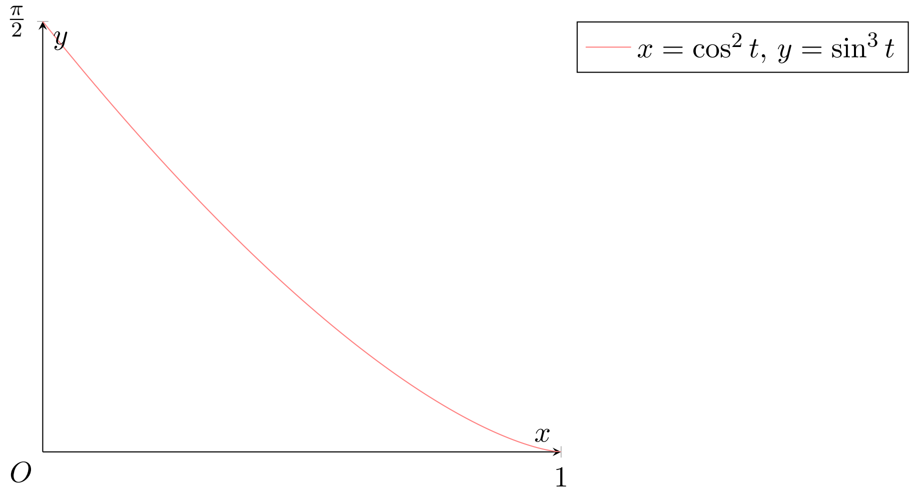
(b) $x = \cos^2 t$, $y = \sin^3 t$, $0 \le t \le \frac{\pi}{2}$

Solution.

Part (a).



Part (b).



Problem 3. Without using a calculator, sketch the following graphs of conics.

(a) $y^2 + 4y + x = 0$

(b) $x^2 + y^2 - 4x - 4y = 0$

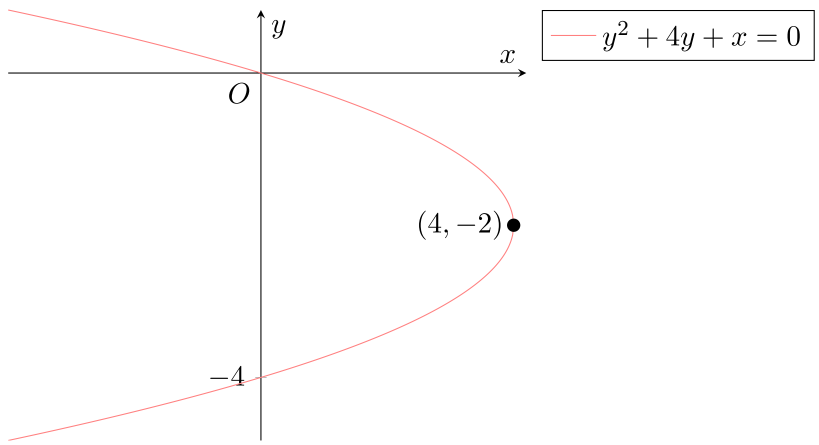
(c) $x^2 + 4y^2 - 2x - 24y + 33 = 0$

(d) $4x^2 - y^2 - 8x + 4y = 1$

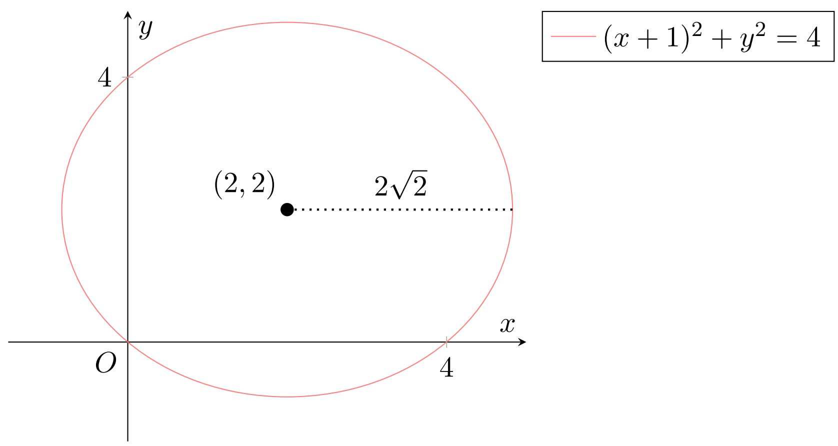
(e) $x = -\sqrt{17 - y^2}$

Solution.

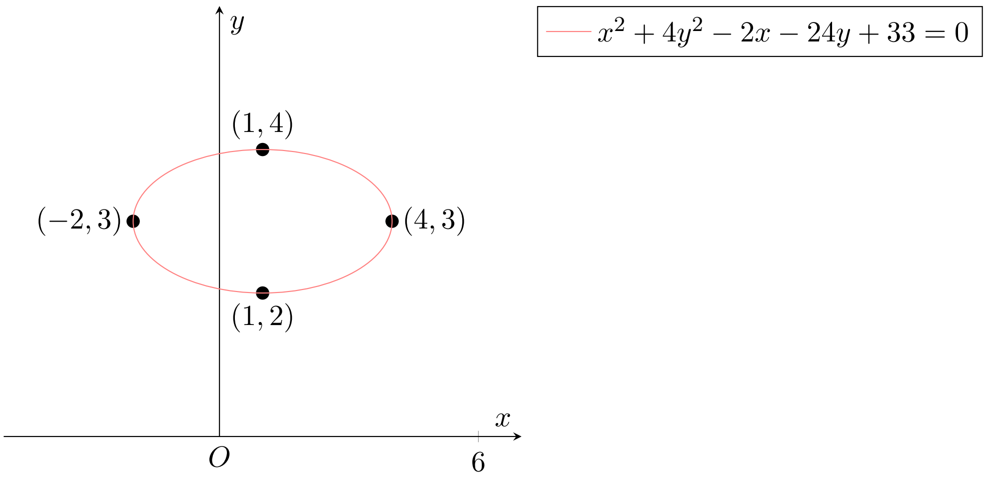
Part (a).



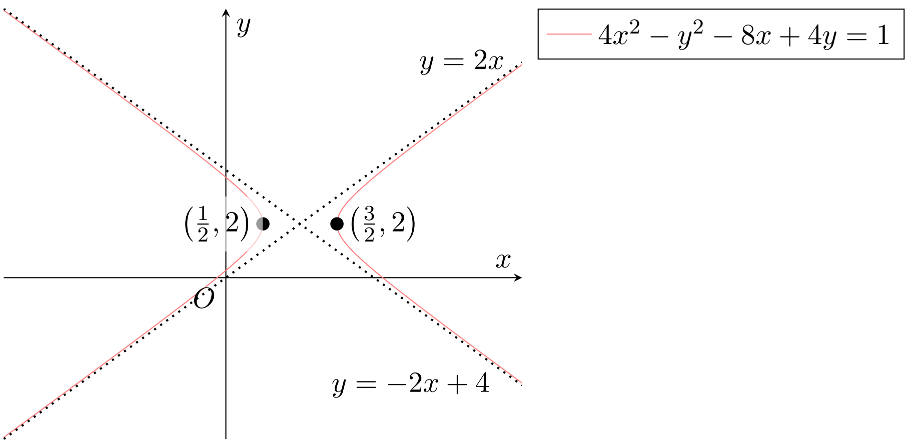
Part (b).



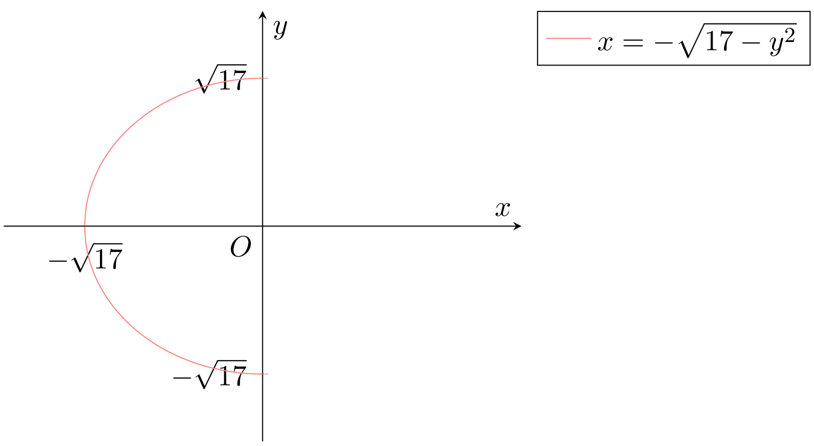
Part (c).



Part (d).



Part (e).



* * * * *

Problem 4. Sketch the curves defined by the following parametric equations. Find also their respective Cartesian equations.

(a) $x = 4t + 3$, $y = 16t^2 - 9$, $t \in \mathbb{R}$

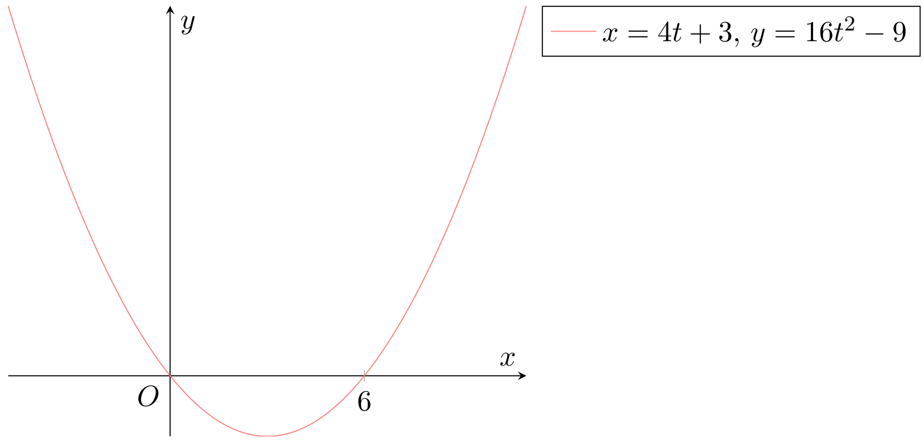
(b) $x = t^2$, $y = 2 \ln t$, $t \geq 1$

(c) $x = 1 + 2\cos\theta$, $y = 2\sin\theta - 1$, $0 \le \theta \le \frac{\pi}{2}$

(d) $x = t^2$, $y = \frac{2}{t}$, $t \neq 0$

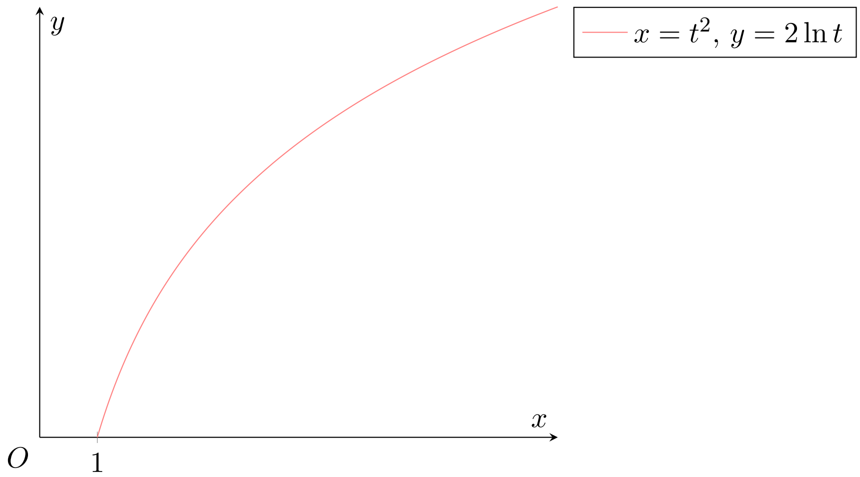
Solution.

Part (a).



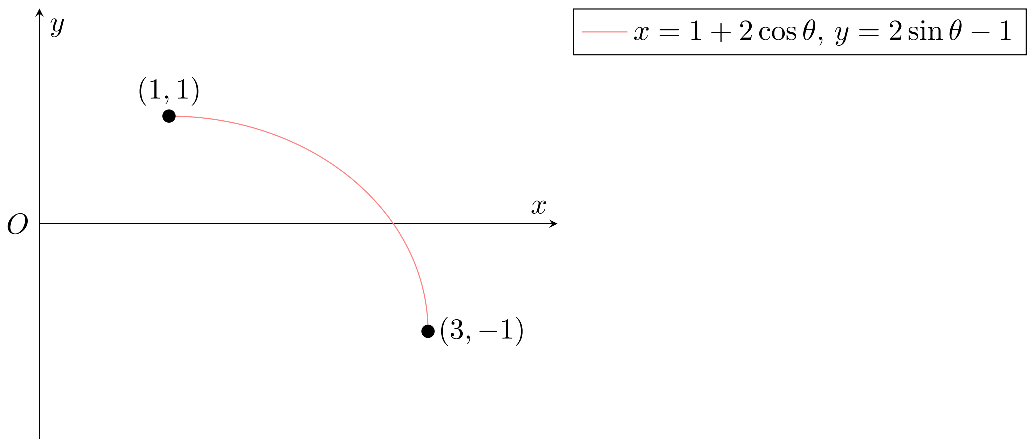
Since $x = 4t + 3$, we have $t = \frac{1}{4}(x - 3)$. Thus, $y = 16\left(\frac{1}{4}(x-3)\right)^2 - 9 = (x-3)^2 - 9$.

Part (b).



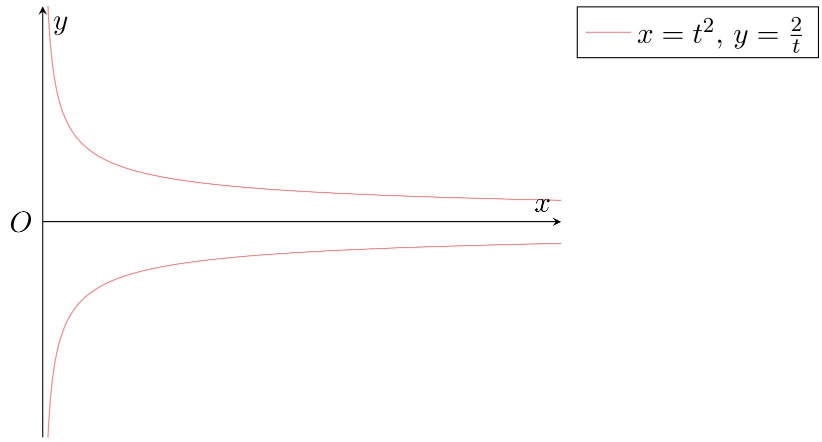
Since $x = t^2$ and $t \ge 1 > 0$, we have $t = \sqrt{x}$. Thus, $y = 2\ln(t) = 2\ln(\sqrt{x}) = \ln(x)$.

Part (c).



We have $2\cos\theta = x - 1$ and $2\sin\theta = y + 1$. Squaring both equations and adding them, we obtain $4\cos^2\theta + 4\sin^2\theta = (x-1)^2 + (y+1)^2$, which simplifies to $(x-1)^2 + (y+1)^2 = 4$.

Part (d).



Since $x = t^2$, we have $t = \pm\sqrt{x}$. Hence, $y = \pm\frac{2}{\sqrt{x}}$.

* * * * *

Problem 5. The curve C_1 has equation $y = \frac{x-2}{x+2}$. The curve C_2 has equation $\frac{x^2}{6} + \frac{y^2}{3} = 1$.

(a) Sketch C_1 and C_2 on the same diagram, stating the exact coordinates of any points of intersections with the axes and the equations of any asymptotes.

(b) Show algebraically that the x-coordinates of the points of intersection of C_1 and C_2 satisfy the equation $2(x-2)^2 = (x+2)^2\left(6 - x^2\right)$.

(c) Use your calculator to find these x-coordinates.

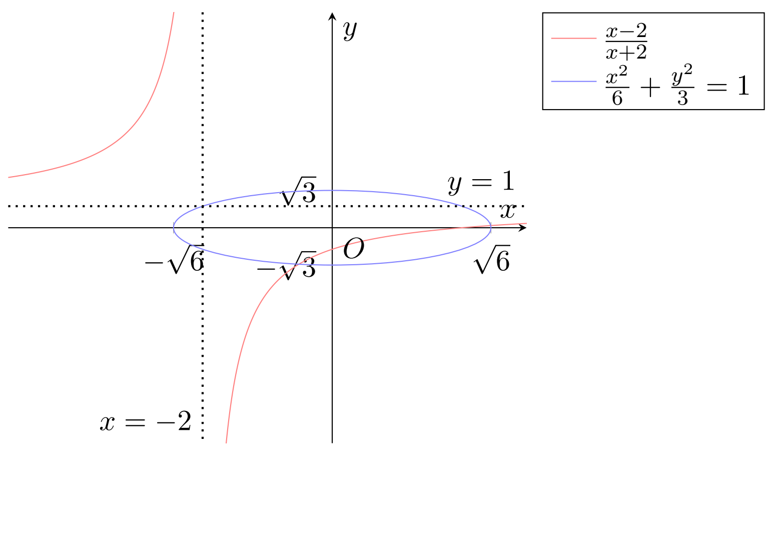
Another curve is defined parametrically by

$$x = \sqrt{6}\cos\theta,\ y = \sqrt{3}\sin\theta,\ -\pi \le \theta \le \pi.$$

(d) Find the Cartesian equation of this curve and hence determine the number of roots to the equation $\sqrt{3}\sin\theta = \frac{\sqrt{6}\cos\theta - 2}{\sqrt{6}\cos\theta + 2}$ for $-\pi \le \theta \le \pi$.

Solution.

Part (a).



Part (b). From C_1, we have $y(x+2) = x - 2$. Hence,

$$y^2(x+2)^2 = (x-2)^2.$$

From C_2, we have $x^2 + 2y^2 = 6$. Hence,

$$y^2 = \frac{6 - x^2}{2}.$$

Putting both equations together, we have

$$(x-2)^2 = \frac{\left(6 - x^2\right)(x+2)^2}{2} \implies 2(x-2)^2 = \left(6 - x^2\right)(x+2)^2.$$

Part (c). The x-coordinates are $x = -0.515$ or $x = 2.45$.

Part (d). Since $x = \sqrt{6}\cos\theta$ and $y = \sqrt{3}\sin\theta$, we have $x^2 = 6\cos^2\theta$ and $2y^2 = 6\sin^2\theta$. Adding both equations together, we have

$$x^2 + 2y^2 = 6\cos^2\theta + 6\sin^2\theta = 6 \implies \frac{x^2}{6} + \frac{y^2}{3} = 1.$$

This is the equation that gives C_1. We further observe that the equation $\sqrt{3}\sin\theta = \frac{\sqrt{6}\cos\theta - 2}{\sqrt{6}\cos\theta + 2}$ simplifies to $y = \frac{x-2}{x+2}$. This is the equation that gives C_2. Since there are two intersections between C_1 and C_2, there are thus two roots to the equation $\sqrt{3}\sin\theta = \frac{\sqrt{6}\cos\theta - 2}{\sqrt{6}\cos\theta + 2}$.

Self-Practice B1

Problem 1. The equations of the curves C_1 and C_2 are given by $y = \frac{2x+1}{x-3}$ and $3(x-1)^2 + 4y^2 = 12$ respectively. Sketch C_1 and C_2 on the same diagram, stating the exact coordinates of any points of intersection with the axes and the equations of any asymptotes.

* * * * *

Problem 2. The curve C has equation $y = \frac{x^2-4x}{x^2-9}$.

(a) Express y in the form $P + \frac{Q}{x-3} + \frac{R}{x+3}$, where P, Q and R are constants.

(b) Sketch C, showing clearly the asymptotes and the coordinates of the points of intersection with the axes.

* * * * *

Problem 3. The curve C has the equation $y = \frac{x^2+px-q}{x+r}$. It is given that C has a vertical asymptote at $x = -3$ and intersects the x-axis at $x = -2$ and $x = 1$.

(a) Determine the values of p, q and r.

(b) State the equation of other asymptote(s).

(c) Prove, using an algebraic method, that y cannot lie between two values which are to be determined.

(d) Hence, sketch C, labelling clearly the axial intercepts, asymptotes and the coordinates of any turning points.

(e) Deduce the number of roots of the equation $3x^4 + 3x^3 - 6x^2 - x - 3 = 0$.

* * * * *

Problem 4. The curve C has equation $y^2 = 5x^2 + 4$.

(a) Sketch C, indicating clearly the axial intercepts, the equations of the asymptotes and the coordinates of the stationary points.

(b) Hence by inserting a suitable graph, determine the range of values of h, where h is a positive constant, such that the equation $5x^2 + 4 = h^2\left(1 - x^2\right)$ has no real roots.

* * * * *

Problem 5. The curve C has equation $y = \frac{mx^2+2x+m}{x}$, where m is a non-zero constant.

(a) Find the range of values of m for C to cut the x-axis at two distinct points.

(b) For $m = \frac{1}{2}$, find the equations of the asymptotes of C.

(c) Hence, sketch the curve C for $m = \frac{1}{2}$: indicating clearly the asymptotes, any turning points and axial intercepts.

(d) By drawing a sketch of another suitable curve in the same diagram as your sketch of C, deduce the number of real roots of the equation $x^2 + 4x + 1 = -2xe^x$.

Problem 6. The curve C_1 has equation $\frac{(x-1)^2}{4} = \frac{y^2}{9} + 4$.

Sketch C_1, making clear the main relevant features, and state the set of values that x can take.

Another curve C_2 is defined by the parametric equations

$$x = \frac{2}{t^2+1}, \quad y = 3\sqrt{t}\ln t, \quad t > 1.$$

Use a non-graphical method to determine the set of possible values of x. Sketch the curve C_2, labelling all axial intercepts and asymptotes (if any) clearly.

Hence, **without solving the equation**, state the number of real roots to the equation

$$9\left(\frac{2}{t^2+1} - 1\right)^2 = 4\left(3\sqrt{t}\ln t\right)^2 + 144,$$

explaining your reason(s) clearly.

Given that $k > 0$, state the smallest integer value of k such that the equation

$$9\left(\frac{2}{t^2+1} + k - 1\right)^2 = 4\left(3\sqrt{t}\ln t\right)^2 + 144,$$

has exactly one real root which is positive.

* * * * *

Problem 7 (🌶).

(a) An ellipse of equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where $0 < b < a$, has two points called foci $F_1(-c, 0)$ and $F_2(c, 0)$. The definition of the ellipse is such that for every point P on the ellipse, the sum of the distance of P to F_1 and F_2 is always a constant k.

- (i) By considering a suitable point on the ellipse, determine the value of k in terms of a and/or b.
- (ii) By considering another suitable point on the ellipse, find c in terms of a and b.

(b) A hyperbola with equation $(y-h)^2 - 1 = \frac{1}{4}(x-k)^2$ has $y = \frac{1}{2}x + \frac{3}{2}$ as one of its asymptotes, and the point $(1, 3)$ is on the hyperbola. Find the values of h and k.

Assignment B1A

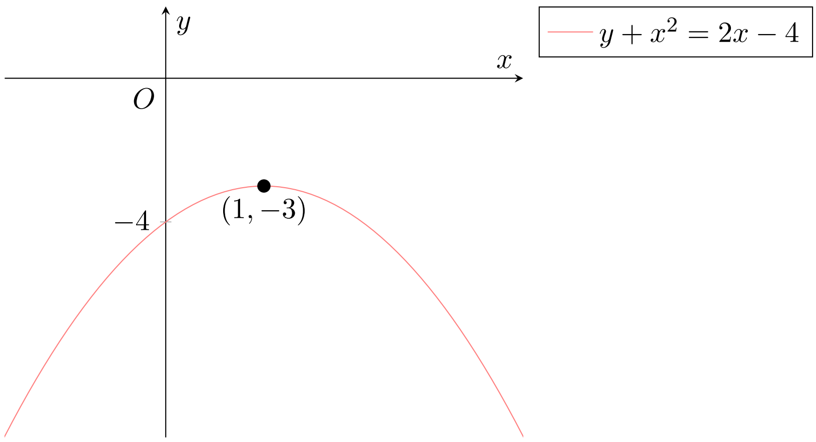
Problem 1. Sketch clearly labelled diagrams of each of the following curves, giving exact values of axial intercepts, stationary points and equations of asymptotes, if any.

(a) $y + x^2 = 2x - 4$

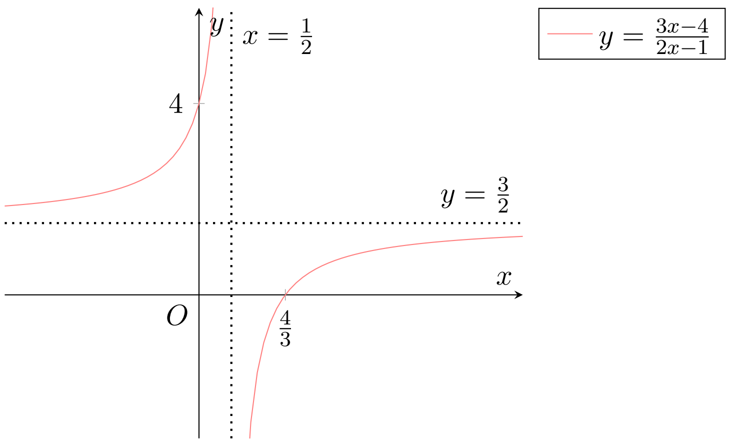
(b) $y = \frac{3x-4}{2x-1}$

Solution.

Part (a).



Part (b).



* * * * *

Problem 2. On separate diagrams, sketch the graphs of

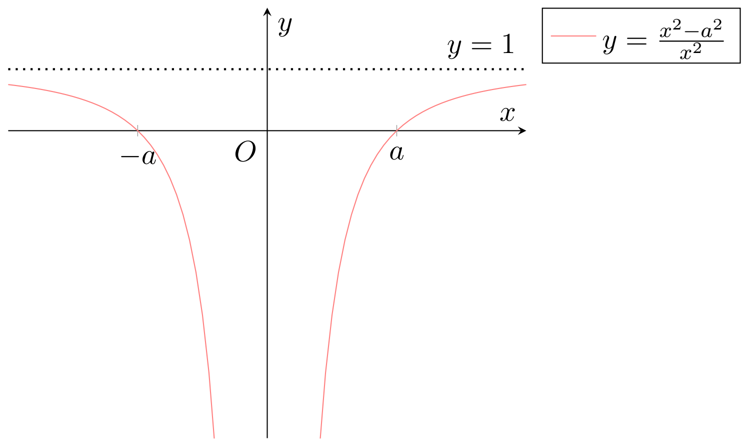
(a) $y = \frac{x^2-a^2}{x^2}$, $a > 0$

(b) $y = \frac{x-1}{2x(x+3)}$

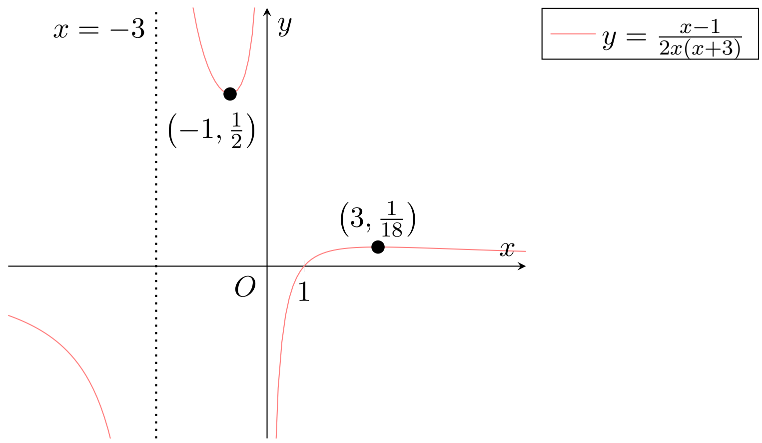
Indicate clearly the coordinates of axial intercepts, stationary points and equations of asymptotes, if any.

Solution.

Part (a).



Part (b).



* * * * *

Problem 3. The curve C has equation $y = \frac{ax^2+bx-2}{x+4}$, where a and b are constants. It is given that $y = 2x - 5$ is an asymptote of C.

(a) Find the values of a and b.

(b) Sketch C.

(c) Using an algebraic method, find the set of values that y cannot take.

(d) By drawing a sketch of another suitable curve in the same diagram as your sketch of C in part (b), deduce the number of distinct real roots of the equation $x^3+6x^2+3x-2 = 0$.

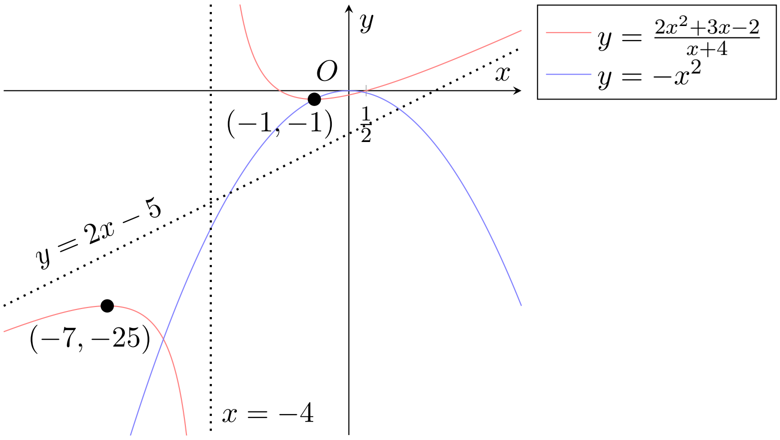
Solution.

Part (a). Since $y = 2x - 5$ is an asymptote of C, $\frac{ax^2+bx-2}{x+4}$ can be expressed in the form $2x - 5 + \frac{k}{x+4}$, where k is a constant.

$$\frac{ax^2+bx-2}{x+4} = 2x-5+\frac{k}{x+4} \implies ax^2+bx-2 = (2x-5)(x+4)+k = 2x^2+3x-20+k.$$

Comparing coefficients of x^2, x and constant terms, we have $a = 2$, $b = 3$ and $k = 18$.

Part (b).

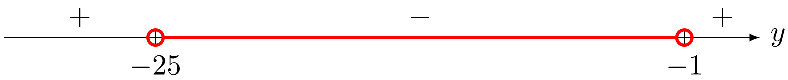


Part (c).

$$y = \frac{2x^2 + 3x - 2}{x + 4} \implies (x + 4)y = 2x^2 + 3x - 2 \implies 2x^2 + (3 - y)x - (2 + 4y) = 0.$$

For values that y cannot take on, there exist no solutions to $2x^2 + (3 - y)x - (2 + 4y) = 0$. Hence, $\Delta < 0$. Hence,

$$(3 - y)^2 - 4(2)\left(-(2 + 4y)\right) < 0 \implies y^2 + 26y + 25 < 0 \implies (y + 25)(y + 1) < 0.$$



Thus, the set of values that y cannot take is $\{y \in \mathbb{R} \colon -25 < y < -1\}$.

Part (d).

$$x^3 + 6x^2 + 3x - 2 = 0 \implies \frac{x^3 + 4x^2}{x + 4} + \frac{2x^2 + 3x - 2}{x + 4} = x^2 + \frac{2x^2 + 3x - 2}{x + 4} = 0$$
$$\implies \frac{2x^2 + 3x - 2}{x + 4} = -x^2.$$

Plotting $y = -x^2$ on the same digram, we see that there are 3 intersections between $y = x^2$ and C. Hence, there are 3 distinct real roots to $x^3 + 6x^2 + 3x - 2 = 0$.

Assignment B1B

Problem 1. Without using a calculator, sketch the graphs of the conics in parts (a), (b) and c.

(a) $3x^2 + 2y^2 = 6$

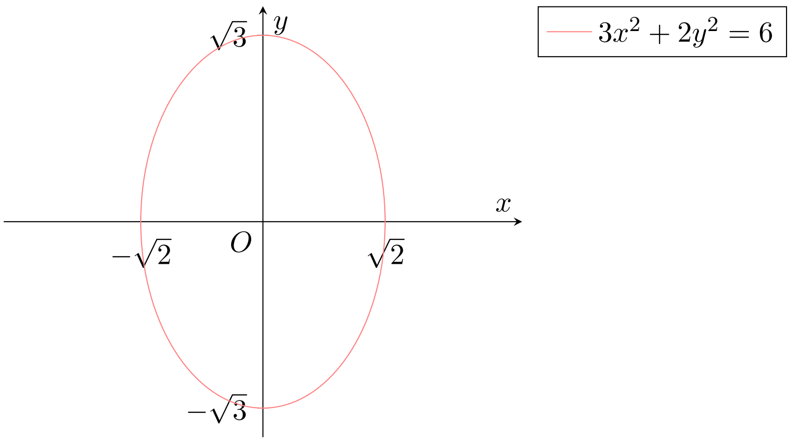
(b) $x^2 + y^2 + 4x - 2y - 20 = 0$

(c) $4(y-1)^2 - x^2 = 4$

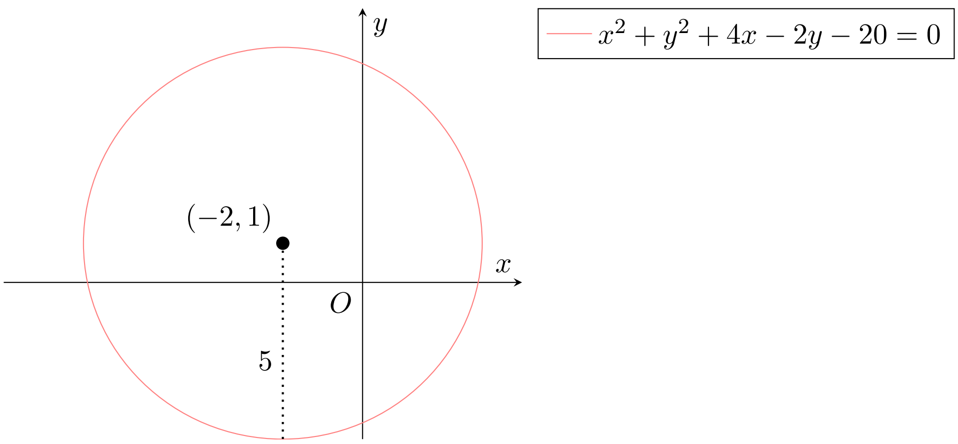
State a transformation that will transform the graph of (a) to a circle with centre $(0,0)$ and radius $\sqrt{3}$.

Solution.

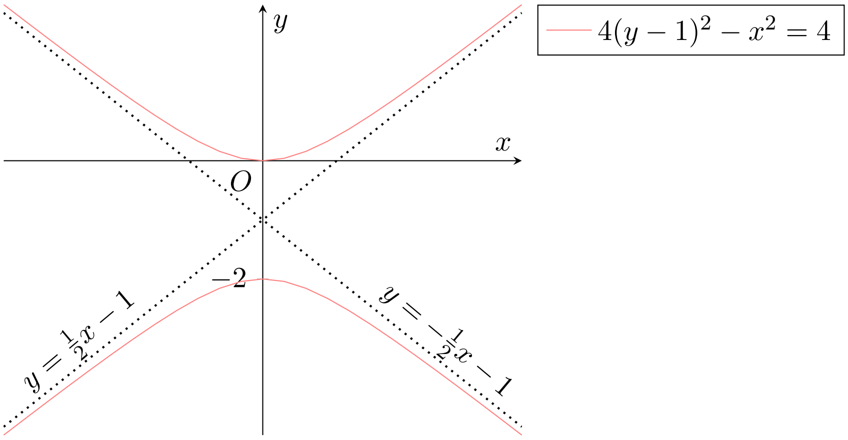
Part (a).



Part (b).



Part (c).



The transformation is $x \mapsto \sqrt{\frac{2}{3}}x$.

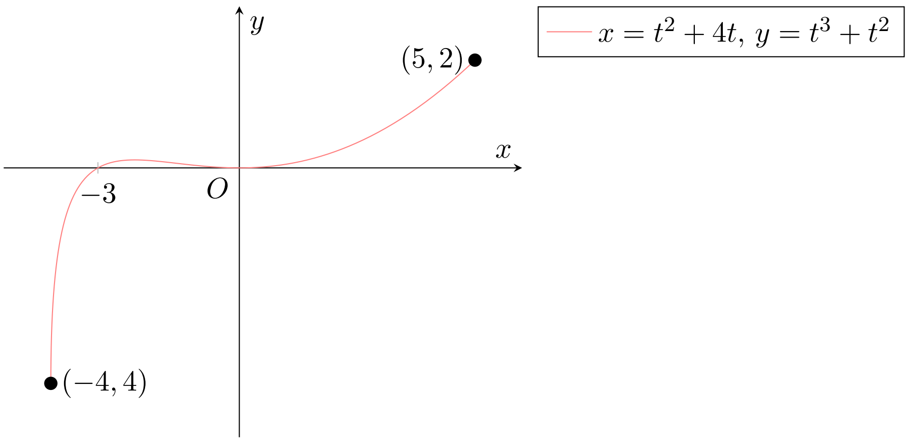
* * * * *

Problem 2. The curve C has parametric equations

$$x = t^2 + 4t,\ y = t^3 + t^2.$$

Sketch the curve for $-2 \leq t \leq 1$, stating the axial intercepts.

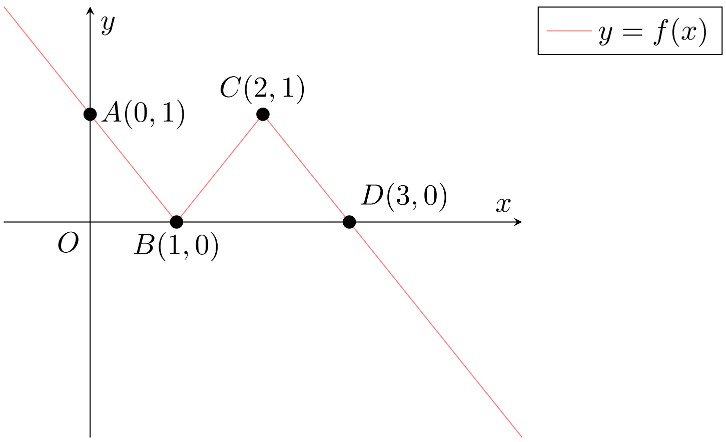
Solution.



B2 Graphs and Transformations II

Tutorial B2

Problem 1.



The graph of $y = f(x)$ is shown here. The points A, B, C and D have coordinates $(0,1)$, $(1,0)$, $(2,1)$ and $(3,0)$ respectively. Sketch, separately, the graphs of

(a) $y = f(2x)$

(b) $y = f(x+3)$

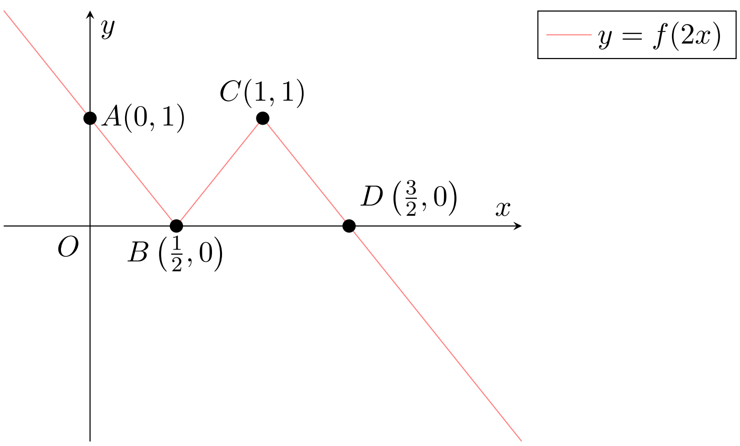
(c) $y = 1 - f(x)$

(d) $y = 3f\left(\frac{x}{2} - 1\right)$

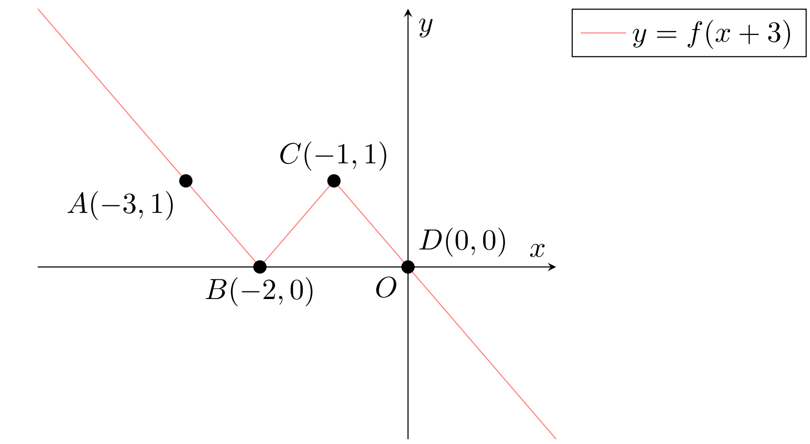
stating, in each case, the coordinates of the points corresponding to A, B, C and D.

Solution.

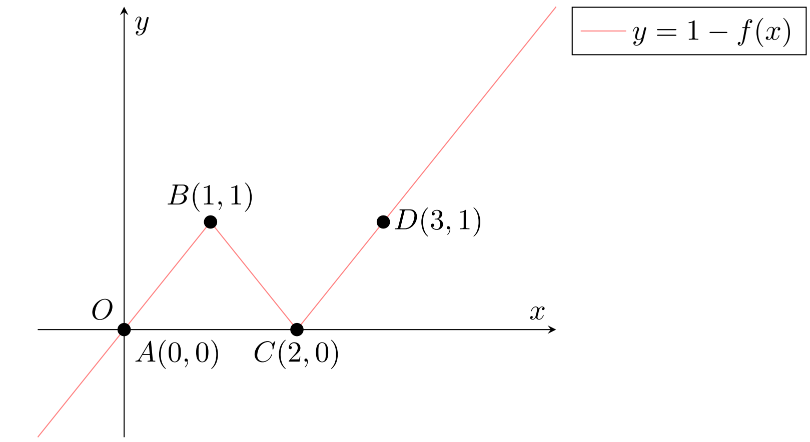
Part (a).



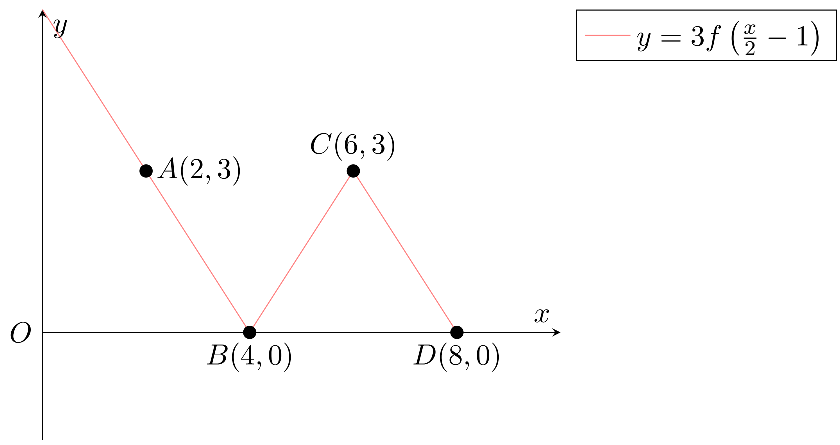
Part (b).



Part (c).



Part (d).



* * * * *

Problem 2. Sketch, on a single clear diagram, the graphs of

(a) $y = x^2$

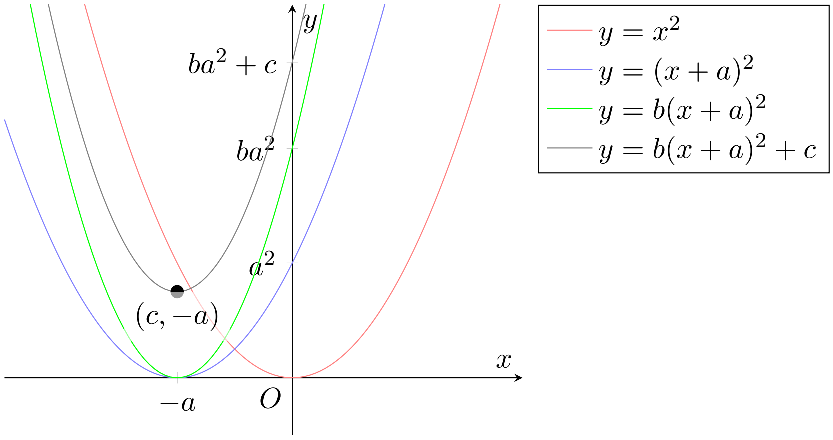
(b) $y = (x+a)^2$

(c) $y = b(x+a)^2$

(d) $y = b(x+a)^2 + c$

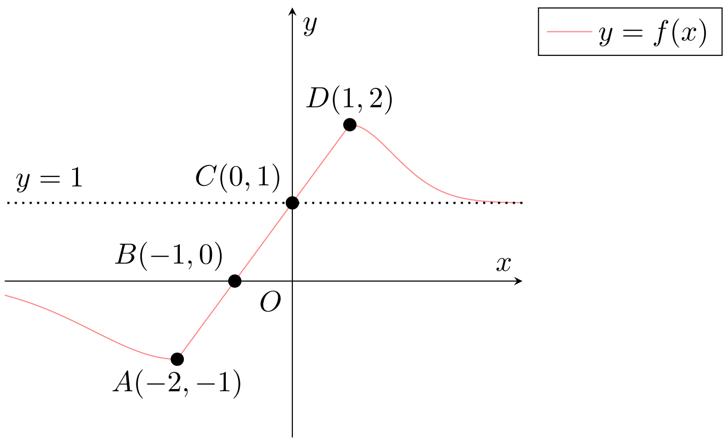
Assume constants $a > 0$, $c > 0$ and $b > 1$.

Solution.



* * * * *

Problem 3. The graph below has equation $y = f(x)$. It has asymptotes $y = 1$ and $y = 0$, a maximum point at $D(1, 2)$, a minimum point at $A(-2, -1)$, cuts the x-axis at $B(-1.0)$ and cuts the y-axis at $C(0, 1)$.



Sketch on separate diagrams the graphs of the following curves, labelling each curve clearly, indicating the horizontal asymptotes and showing the coordinates of the points corresponding to the points A, B, C and D.

(a) $y = f(x+1)$

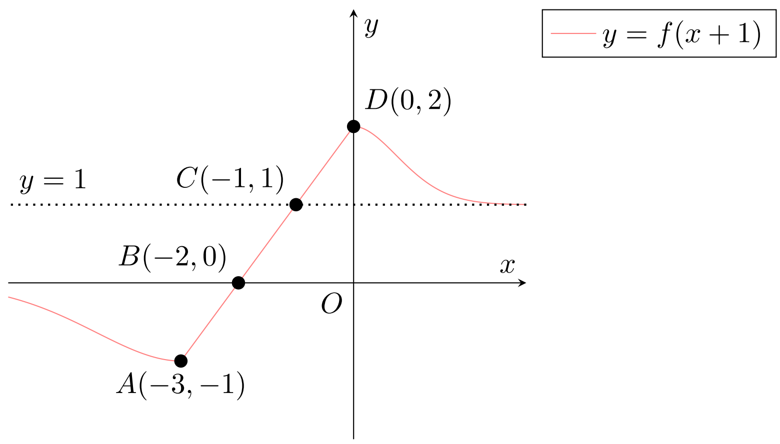
(b) $y = f\left(\frac{x}{2}\right)$

(c) $y = 2f(x) - 2$

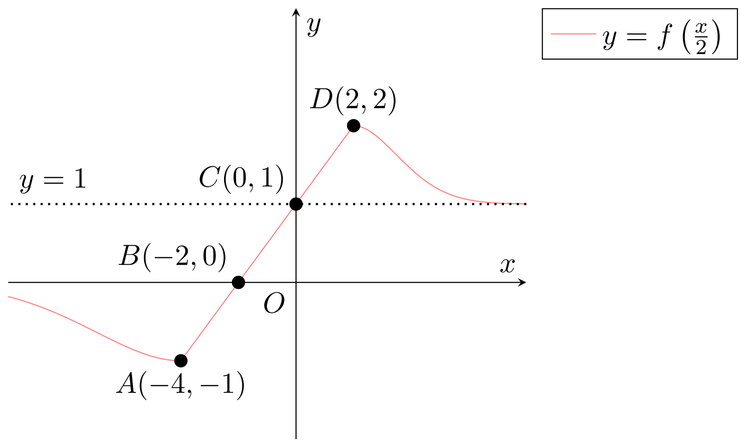
Find the number of solutions to the equation $f(x) = af(x)$ where $a \geq 2$.

Solution.

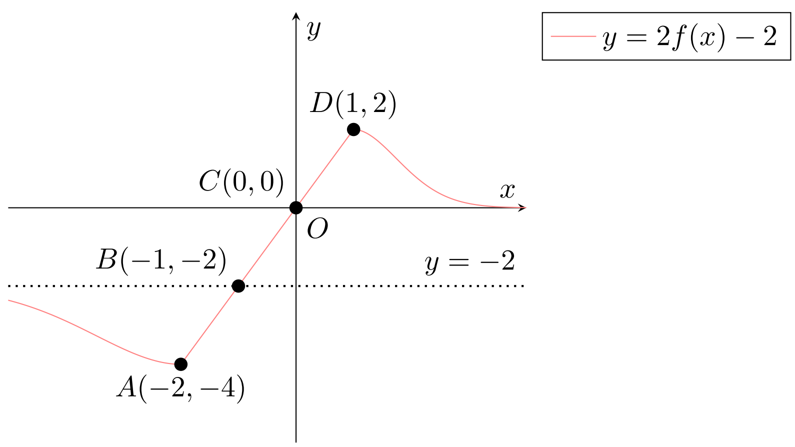
Part (a).



Part (b).



Part (c).



All points with a y-coordinate of 0 are invariant under the transformation $f(x) \mapsto af(x)$. Since there is only one such point $(B(-1,0))$, there is only 1 solution to the equation $f(x) = af(x)$, where $a \geq 2$.

Problem 4. The curve with equation $y = x^2$ is transformed by a translation of 2 units in the positive x-direction, followed by a stretch with scale factor $\frac{1}{2}$ parallel to the y-axis, followed by a translation of 6 units in the negative y-direction. Find the equation of the new curve in the form $y = f(x)$ and the exact coordinates of the points where this curve crosses the x- and y-axes.

Solution.

$$\begin{tikzcd} & y = x^3 \arrow[r, "x \mapsto x-2"] & y = (x-2)^3 \arrow[d, "y \mapsto 2y"] \\ 2(y+6) = (x-2)^3 & & 2y = (x-2)^3 \arrow[ll, "y \mapsto y+6"'] \end{tikzcd}$$

Hence, $y = \frac{1}{2}(x-2)^3 - 6$

When $x = 0$, $y = -10$. When $y = 0$, $x = 2 + \sqrt[3]{12}$. Thus, the curve crosses the x-axis at $(2 + \sqrt[3]{12}, 0)$ and the y-axis at $(0, -10)$.

* * * * *

Problem 5. Find the values of the constants A and B such that $\frac{x^2-4x}{(x-2)^2} = A + \frac{B}{(x-2)^2}$ for all values of x except $x = 2$.

Hence, state a sequence of transformations by which the graph of $y = \frac{x^2-4x}{(x-2)^2}$ may be obtained from the graph of $y = \frac{1}{x^2}$.

Solution.

$$\frac{x^2-4x}{(x-2)^2} = \frac{(x-2)^2-4}{(x-2)^2} = 1 + \frac{-4}{(x-2)^2}.$$

Thus, $A = 1$ and $B = -4$.

$$\begin{tikzcd} y = \dfrac{1}{x^2} \arrow[r, "x \mapsto x-2"] & y = \dfrac{1}{(x-2)^2} \arrow[r, "y \mapsto \frac{1}{4}y"] & y = \dfrac{4}{(x-2)^2} \arrow[d, "y \mapsto -y"] \\ & y = 1 + \dfrac{-4}{(x-2)^2} & y = \dfrac{-4}{(x-2)^2} \arrow[l, "y \mapsto y-1"'] \end{tikzcd}$$

1. Translate the curve 2 units in the positive x-direction.
2. Stretch the curve with a scale factor of 4 parallel to the y-axis.
3. Reflect the curve about the x-axis.
4. Translate the curve 1 unit in the positive y-direction.

* * * * *

Problem 6. The transformations A, B, C and D are given as follows:

- A: A reflection about the y-axis.
- B: A translation of 2 units in the positive x-direction.
- C: A scaling parallel to the y-axis by a factor of 3.

- D: A translation of 1 unit in the positive y-direction.

A curve undergoes the transformations A, B, C and D in succession, and the equation of the resulting curve is $y = 3\sqrt{2-x}+1$. Determine the equation of the curve before the transformations were effected.

Solution.

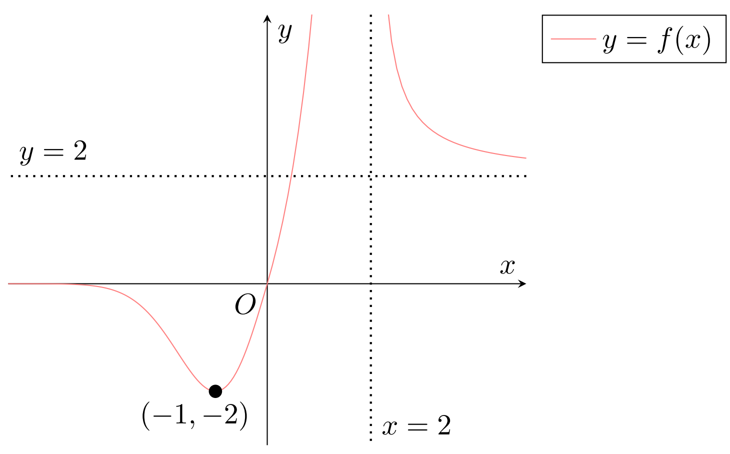
$$\begin{aligned} A\colon x &\mapsto -x && \Longrightarrow && A^{-1}\colon x \mapsto -x \\ B\colon x &\mapsto x-2 && \Longrightarrow && B^{-1}\colon x \mapsto x+2 \\ C\colon y &\mapsto \frac{1}{3}y && \Longrightarrow && C^{-1}\colon y \mapsto 3y \\ D\colon y &\mapsto y-1 && \Longrightarrow && D^{-1}\colon y \mapsto y+1 \end{aligned}$$

$$\begin{array}{c} y = 3\sqrt{2-x}+1 \\ \Big\downarrow D^{-1} \\ y+1 = 3\sqrt{2-x}+1 \\ \Big\downarrow C^{-1} \\ 3y+1 = 3\sqrt{2-x}+1 \\ \Big\downarrow B^{-1} \\ 3y+1 = 3\sqrt{2-(x+2)}+1 \\ \Big\downarrow A^{-1} \\ 3y+1 = 3\sqrt{2-(-x+2)}+1 \end{array}$$

Thus, the original curve has equation $y = \sqrt{x}$.

* * * * *

Problem 7.



The diagram shows the graph of $y = f(x)$. The curve passes through the origin and has minimum point $(-1, -2)$. The asymptotes are $x = 2$, $y = 0$ and $y = 2$.

Sketch, on separate diagrams, the graphs of

(a) $y = f(x-1)$

(b) $y = f(|x|)$

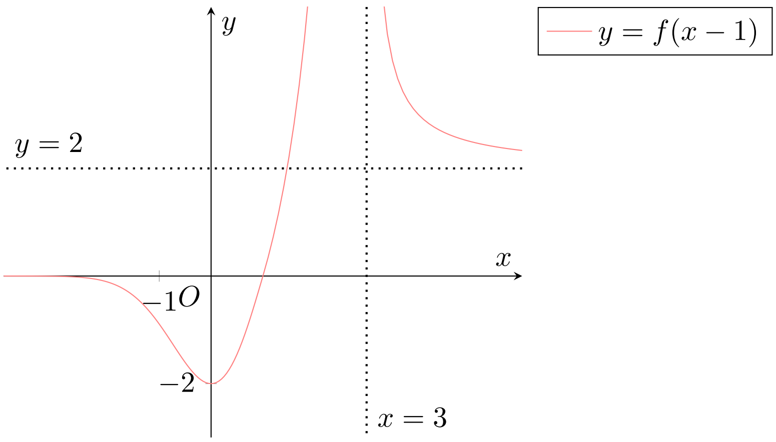
(c) $y = f(|x-1|)$

(d) $y = |f(x)|$

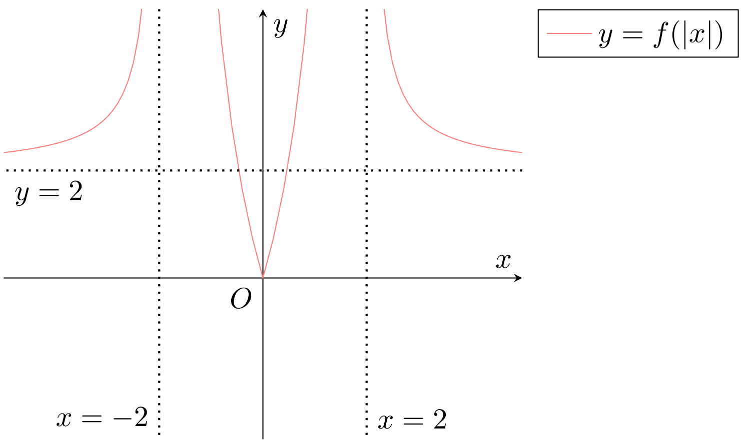
(e) $y = \frac{1}{f(x)}$

Solution.

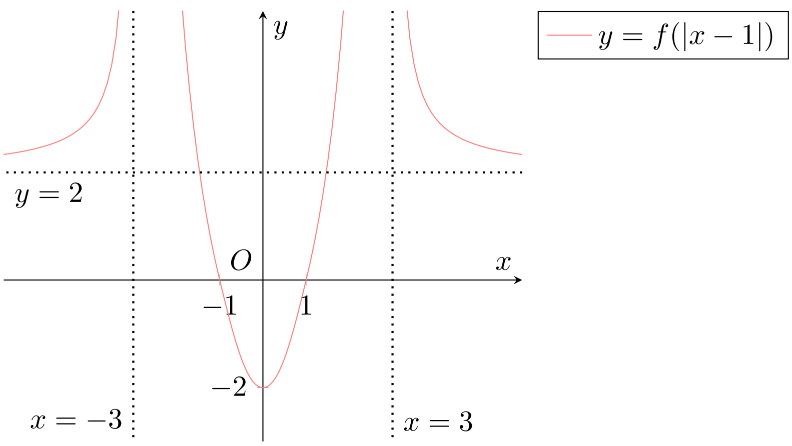
Part (a).



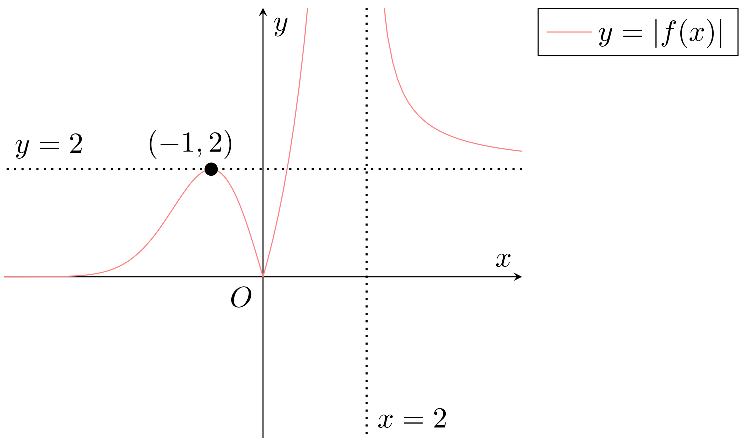
Part (b).



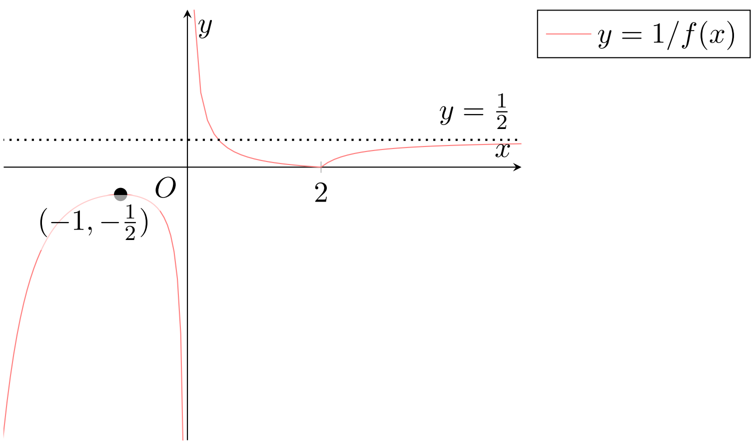
Part (c).



Part (d).

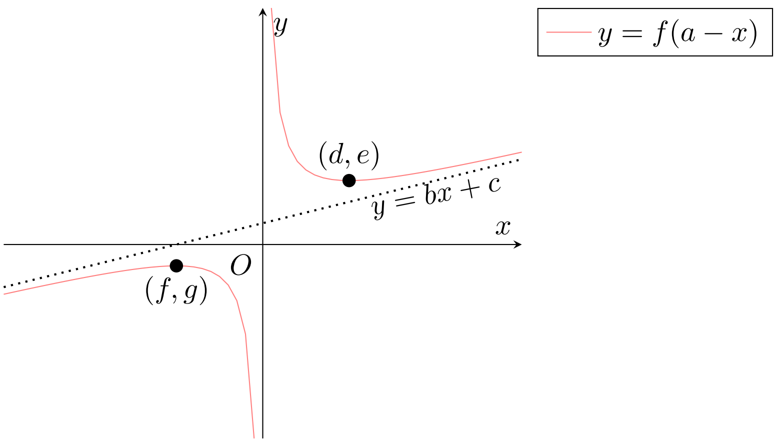


Part (e).



* * * * *

Problem 8.



The graph of $y = f(a-x)$ is shown in the figure, where $a > 0$. The curve has asymptotes $x = 0$, $y = bx + c$, a minimum point at (d, e) and a maximum point at (f, g).

Given $a > d$, sketch separately, the graphs of

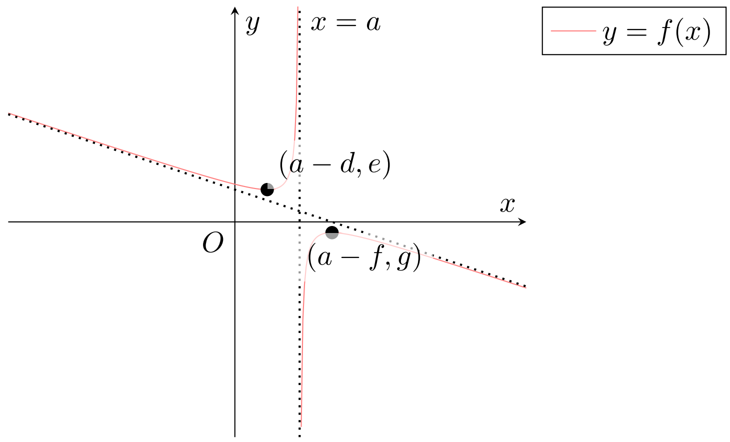
(a) $y = f(x)$

(b) $y = f(|x|)$

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

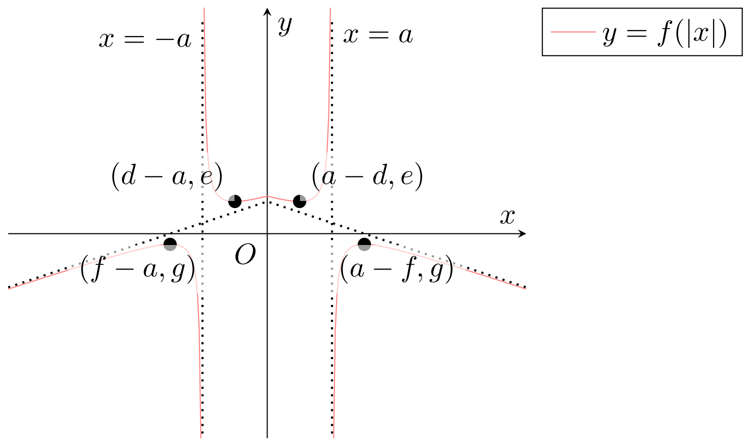
Solution.

Part (a).



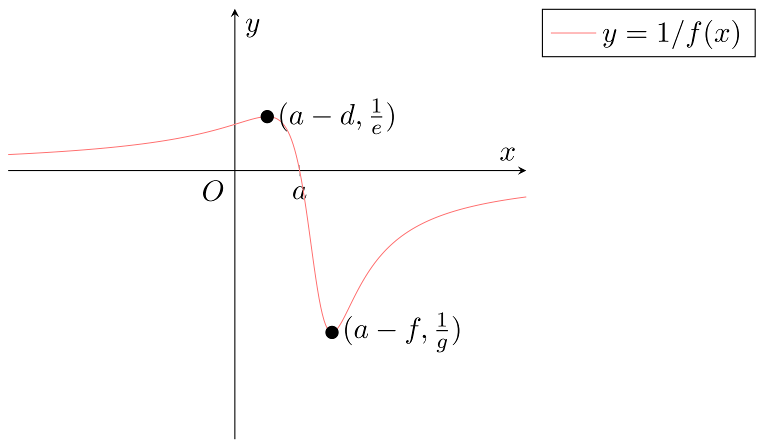
Equation of asymptote: $y = b(a - x) + c$

Part (b).



Equation of asymptotes: $y = b(a + x) + c$, $y = b(a - x) + c$

Part (c).



Problem 9. A curve C_1 is defined by the parametric equations

$$x = t(t+2),\ y = 2(t+1).$$

(a) Find the axial intercepts of the curve.

(b) Sketch C_1.

(c) A curve C_2 is defined by the parametric equations $x = t(t+2)$, $y = t+1$. Describe a geometrical transformation which maps C_1 to C_2. Hence, sketch the curve C_2 in the same diagram as C_1.

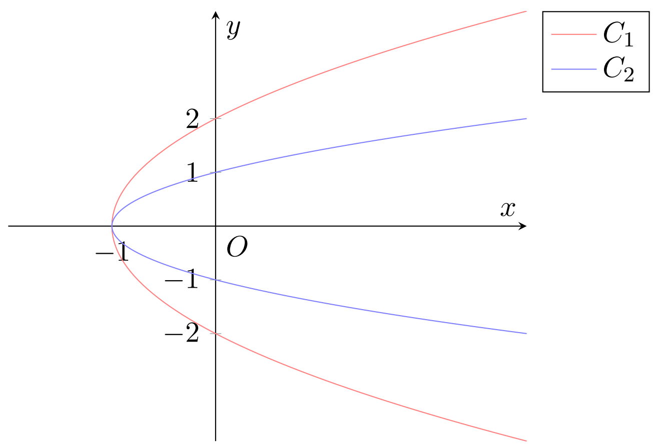
(d) Show that the Cartesian equation of the curve C_1 is given by $y^2 = 4(x+1)$.

Solution.

Part (a). Consider $x = 0$. Then $t(t+2) = 0$, whence $t = 0$ or $t = -2$. When $t = 0$, $y = 2$. When $t = -2$, $y = -2$. Hence, the curve intercepts the y-axis at $(0, 2)$ and $(0, -2)$.

Consider $y = 0$. Then $t = -1$, whence $x = -1$. Hence, the curve intercepts the x-axis at $(-1, 0)$.

Part (b).



Part (c). Scale by a factor of $\frac{1}{2}$ parallel to the y-axis.

Part (d).

$$y^2 = (2(t+1))^2 = 4(t^2+2t+1) = 4(t(t+1)+1) = 4(x+1).$$

Self-Practice B2

Problem 1. Show that the equation $y = \frac{2x+7}{x+2}$ can be written as $y = A + \frac{B}{x+2}$, where A and B are constants to be found. Hence, state a sequence of transformations which transform the graph of $y = \frac{1}{x}$ to the graph of $y = \frac{2x+7}{x+2}$.

Sketch the graph of $y = \frac{2x+7}{x+2}$, giving the equations of any asymptotes and the coordinates of any points of intersection with the x- and y-axes.

* * * * *

Problem 2. The diagram shows the curve with equation $y = f(x)$. The curve passes through the origin, and has asymptotes $x = a$ and $y = b$, where a and b are positive constants.

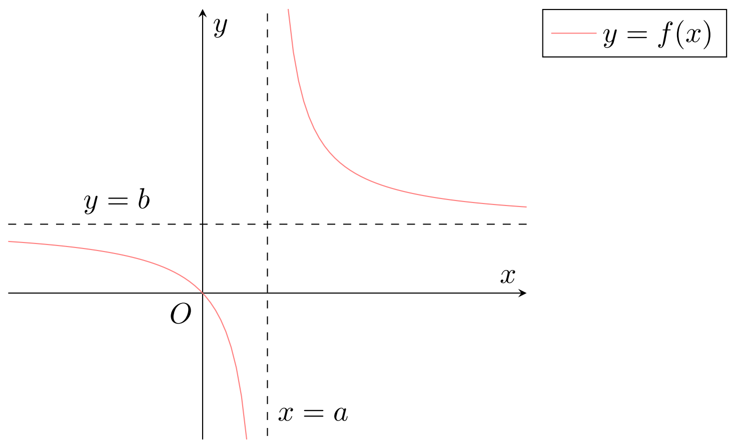


Figure 35.1

On separate diagrams, draw sketches of the graphs of

(a) $y = f(x+a) - b$,

(b) $y = 1/f(x)$,

showing clearly the axial intercepts and asymptotes (if any).

* * * * *

Problem 3. The curves C_1 and C_2 are given by the equations $x^2 + y^2 = 1$ and $x^2 - 2x + 9y^2 = a$ respectively, where a is a real constant. The curve C_2 cuts the x-axis at the origin O and is symmetrical about the line $x = b$, as shown in the diagram below.

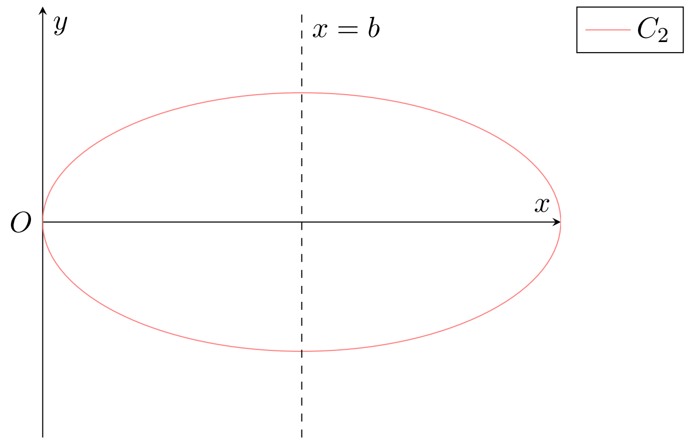


Figure 35.2

(a) Determine the values of a and b.

(b) Describe clearly a sequence of transformations that maps C_1 onto C_2.

* * * * *

Problem 4. It is given that the curve $y = f(x)$, where $f(x) = \frac{ax+b}{2x+c}$, where a, b, c are constants, has an asymptote $x = \frac{1}{2}$. The point A with coordinates $(2, \frac{5}{3})$ lies on the curve. The tangent to the curve at A has gradient $\frac{2}{9}$.

(a) Write down the value of c.

(b) Show that $a = 4$ and $b = -3$.

(c) Sketch the graph of $y = f(x)$, showing clearly all the asymptotes and the exact coordinates of the intersection with the axes.

(d) Describe a sequence of three transformations which transforms the graph $y = 2 + \frac{1}{x}$ to $y = f(x)$.

* * * * *

Problem 5. The curve whose equation is $\frac{(x-3)^2}{2^2} + \frac{y^2}{3^2} = 1$ undergoes, in succession, the following transformations:

A : A translation of magnitude 1 unit in the direction of the x-axis.

B : A reflection in the y-axis.

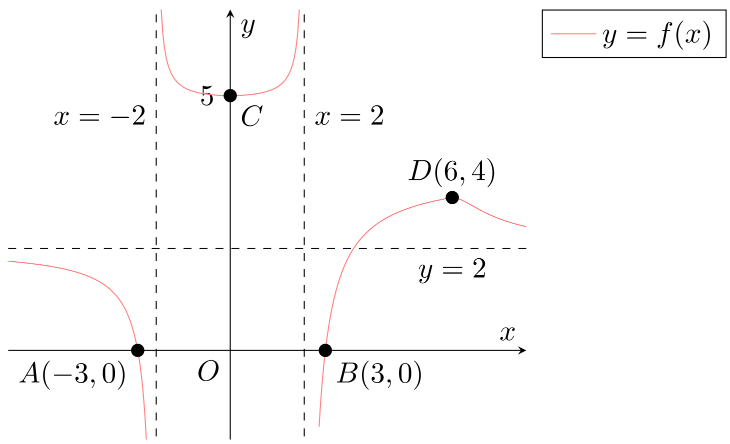
C : A scaling parallel to the y-axis by a scale factor of k.

(a) Find the equation of the resulting curve.

(b) State the value of k for which the resulting curve takes on the shape of a circle.

* * * * *

Problem 6. The diagram shows the graph of $y = f(x)$.



On separate diagrams, sketch the graphs of

(a) $y = f(4x + 3)$,

(b) $y = 1/f(x)$.

In each case, state the equations of any asymptotes and the coordinates of the points corresponding to A, B, C and D where appropriate.

* * * * *

Problem 7. A curve C_1 is defined parametrically by

$$x = \frac{2}{t-1}, \quad y = \frac{4}{t+1}, \quad t \neq \pm 1.$$

Sketch a clearly labelled diagram of C_1.

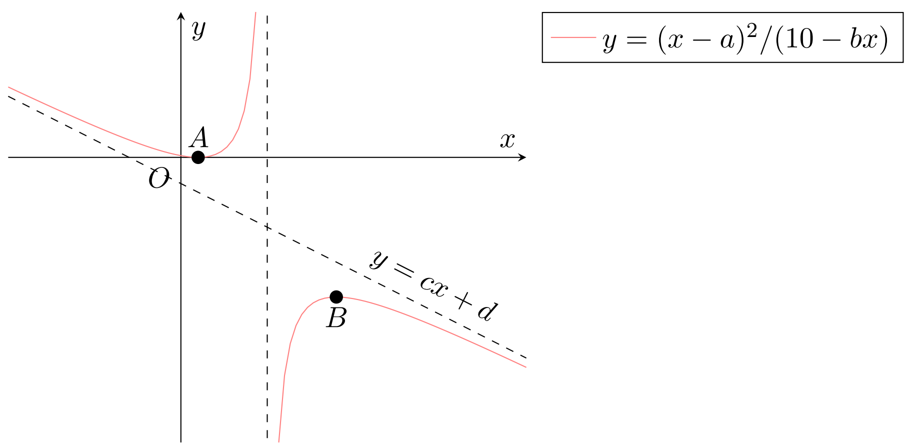
Describe a sequence of geometrical transformations which maps C_1 to C_2 defined by

$$x = \frac{1}{1-t}, \quad y = \frac{4}{t+1}, \quad t \neq \pm 1.$$

Sketch C_3, which is the reciprocal function of C_1, stating the equations of any asymptotes and any points of intersection with the axes.

* * * * *

Problem 8 (🌶). The curve of $y = \frac{x^2}{4-x}$ undergoes two transformations. The resulting curve whose equation is $y = \frac{(x-a)^2}{10-bx}$ has stationary points $A(1, 0)$ and $B(9, -8)$, and asymptotes $x = 5$ and $y = cx + d$, where a, b, c and d are constants.

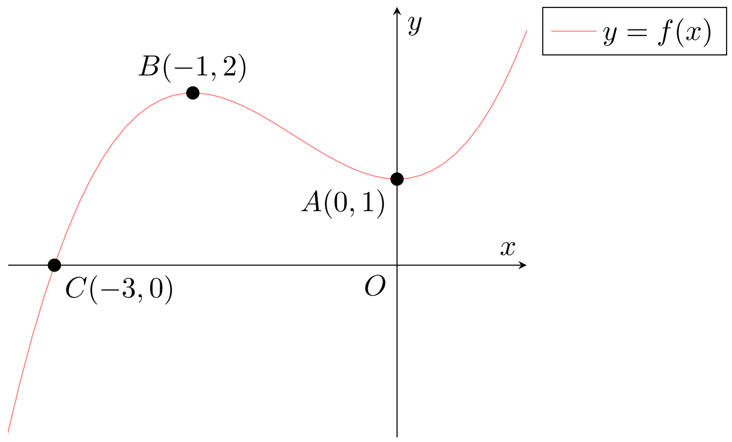


(a) Show that $a = 1$, and find the values of b, c and d.

(b) Describe the sequence of transformations undergone by the graph of $y = \frac{x^2}{4-x}$ to attain that of $y = \frac{(x-a)^2}{10-bx}$, where a and b are the values found in (a).

Assignment B2

Problem 1.



The diagram shows the graph of $y = f(x)$. The points A, B and C have coordinates $(0, 1)$, $(-1, 2)$ and $(-3, 0)$ respectively. Sketch, separately, the graphs of

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

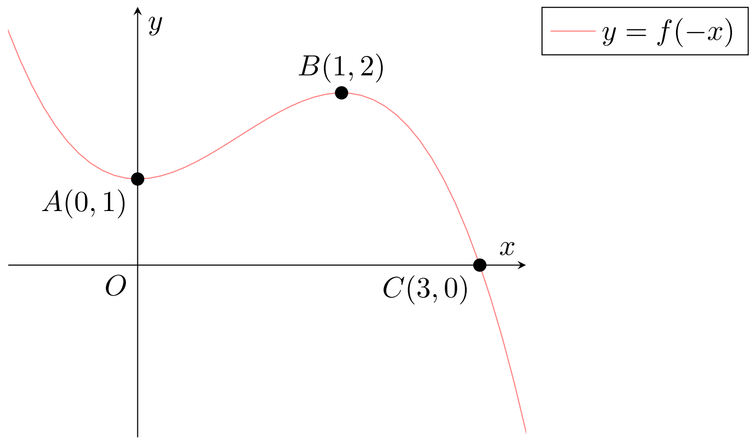
(b) $y = 3 - 2f(x)$

(c) $y = 3f\left(\frac{x}{2} + 1\right)$

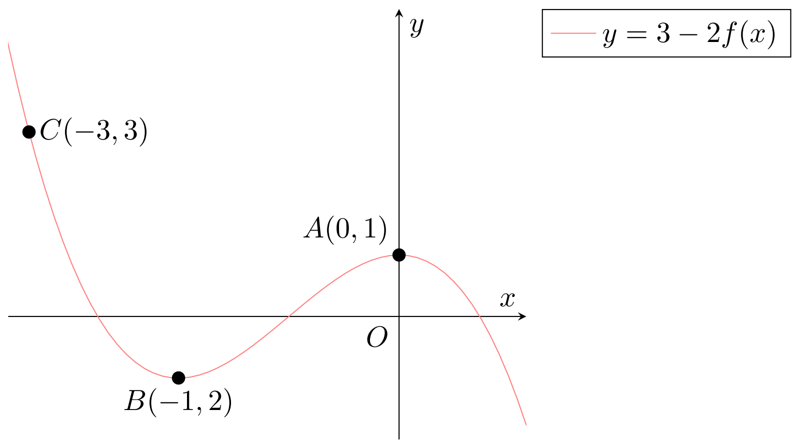
showing in each case the coordinates of the points corresponding to A, B and C.

Solution.

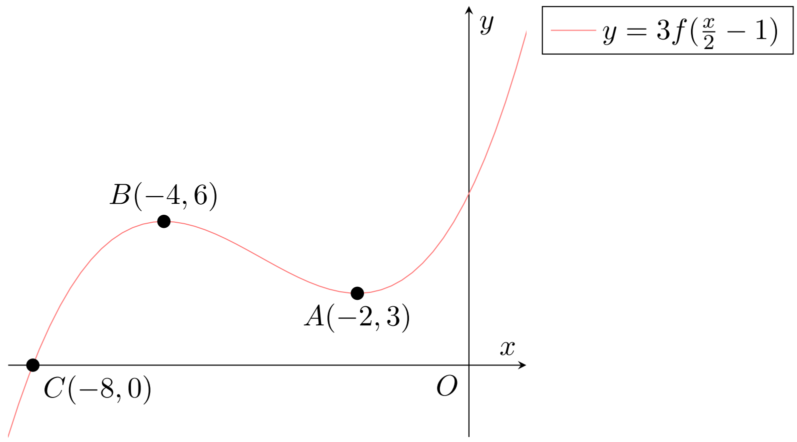
Part (a).



Part (b).

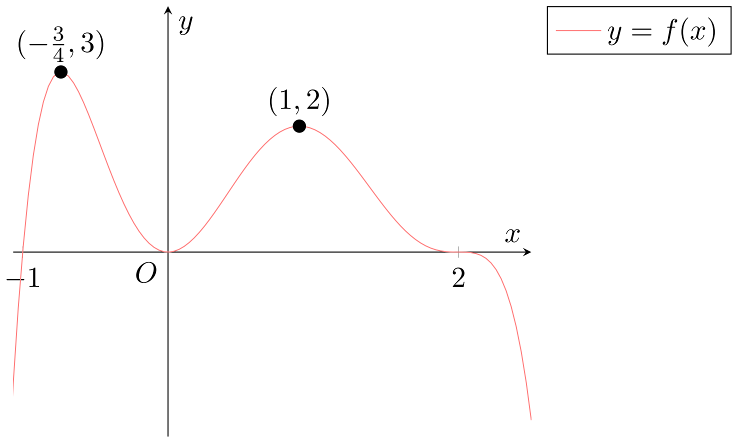


Part (c).



* * * * *

Problem 2.



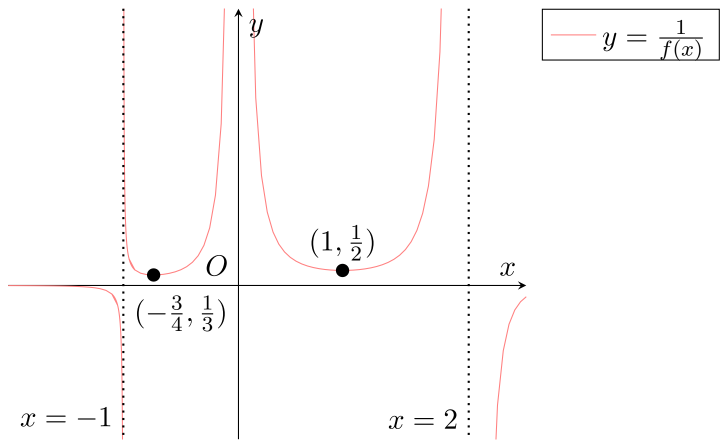
The curve shown is the graph of $y = f(x)$. The x-axis is a tangent at the origin and at $(2, 0)$. The curve has two maximum points at $\left(-\frac{3}{4}, 3\right)$ and $(1, 2)$. On two separate diagrams, sketch the graphs of the following equations. Show clearly the shapes of the graphs where they meet the x-axis and any asymptotes.

(a) $y = \frac{1}{f(x)}$, $x \neq -1, 0, 2$

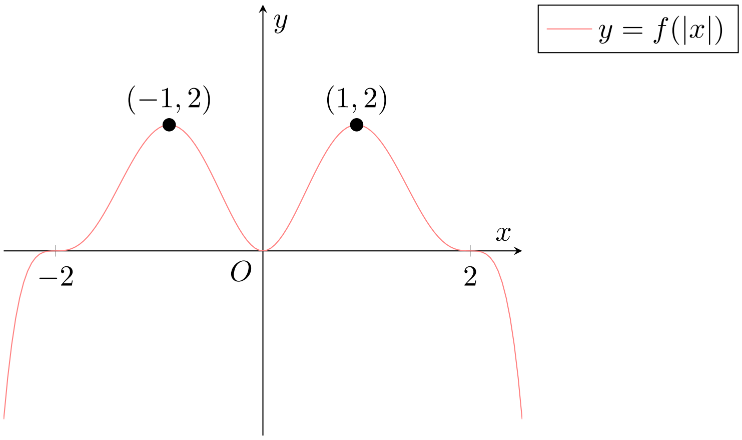
(b) $y = f(|x|)$

Solution.

Part (a).



Part (b).



* * * * *

Problem 3. A graph with equation $y = f(x)$ undergoes transformation A followed by transformation B where A and B are described as follows:

- A: a translation of 1 unit in the positive direction of the x-axis
- B: a scaling parallel to the x-axis by a factor $\frac{1}{2}$

The resulting equation is $y = 4x^2 - 4x + 1$. Find the equation $y = f(x)$.

Solution. Note that

$$\begin{aligned} A&\colon x \mapsto x - 1 &&\implies A^{-1}\colon x \mapsto x + 1 \\ B&\colon x \mapsto 2x &&\implies B^{-1}\colon x \mapsto \frac{1}{2}x. \end{aligned}$$

Hence,

$$y = 4x^2 - 4x + 1$$

$$\Big\downarrow B^{-1}$$

$$y = 4\left(\frac{1}{2}x\right)^2 - 4\left(\frac{1}{2}x\right) + 1$$

$$\Big\downarrow A^{-1}$$

$$y = 4\left[\frac{1}{2}(x+1)\right]^2 - 4\left[\frac{1}{2}(x+1)\right] + 1$$

Observe that y simplifies to

$$y = 4\left[\frac{1}{2}(x+1)\right]^2 - 4\left[\frac{1}{2}(x+1)\right] + 1 = (x+1)^2 - 2(x+1) + 1 = x^2 + 2x + 1 - 2x - 2 + 1 = x^2.$$

B3 Functions

Tutorial B3

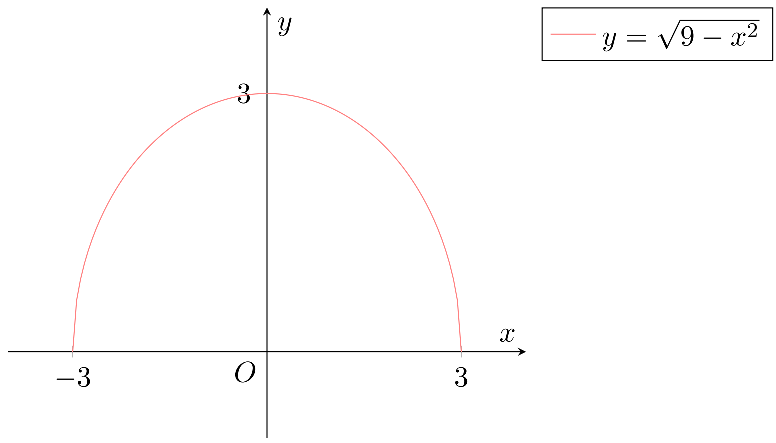
Problem 1. Sketch the following graphs and determine whether each graph represents a function for the given domain.

(a) $y = \sqrt{9 - x^2}$, $-3 \le x \le 3$

(b) $x = (y - 4)^2$, $y \in \mathbb{R}$

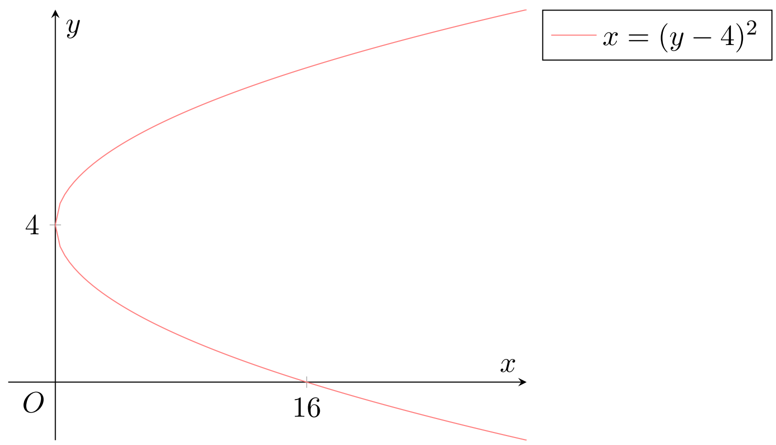
Solution.

Part (a).



$y = \sqrt{9 - x^2}$ passes the vertical line test for $-3 \le x \le 3$ and is hence a function.

Part (b).



$x = (y - 4)^2$ does not pass the vertical line test for $y \in \mathbb{R}$ and is hence not a function.

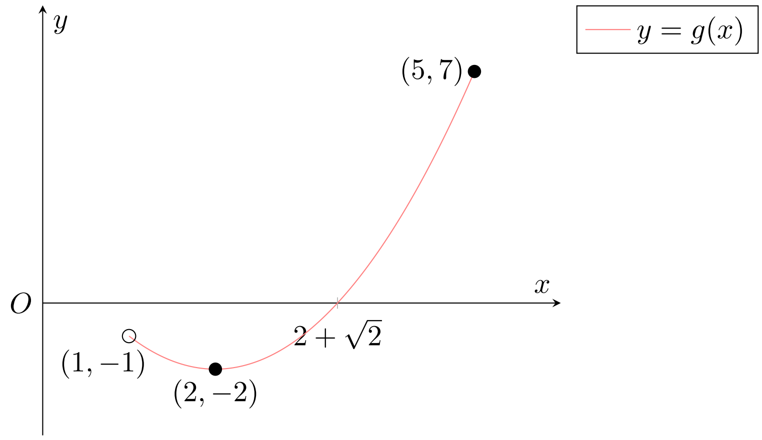
Problem 2. Sketch the graph and find the range for each the following functions.

(a) $g\colon x \mapsto x^2 - 4x + 2,\ 1 < x \leq 5$

(b) $h\colon x \mapsto |2x - 3|\,,\ x < 3$

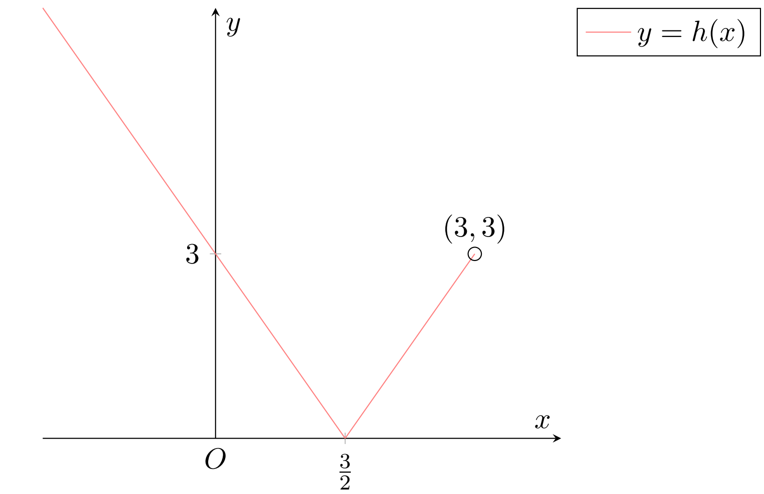
Solution.

Part (a).



From the graph, $R_g = [-2, 7)$.

Part (b).



From the graph, $R_h = [0, \infty)$.

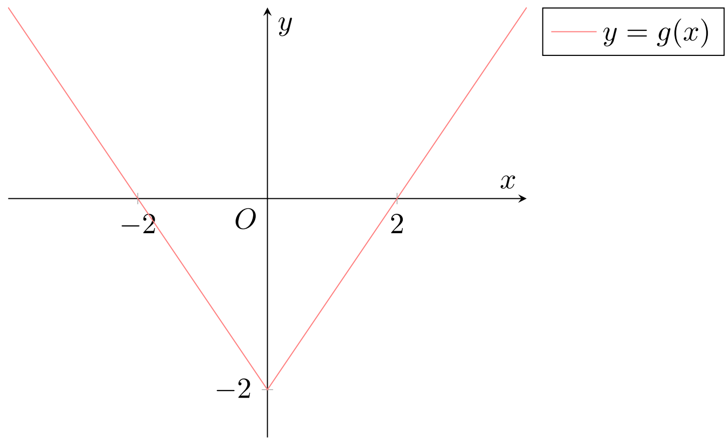
Problem 3. For each of the following functions, sketch its graph and determine if the function is one-one. If it is, find its inverse in a similar form.

(a) $g\colon x \mapsto |x| - 2$, $x \in \mathbb{R}$

(b) $h\colon x \mapsto x^2 + 2x + 5$, $x \leq -2$

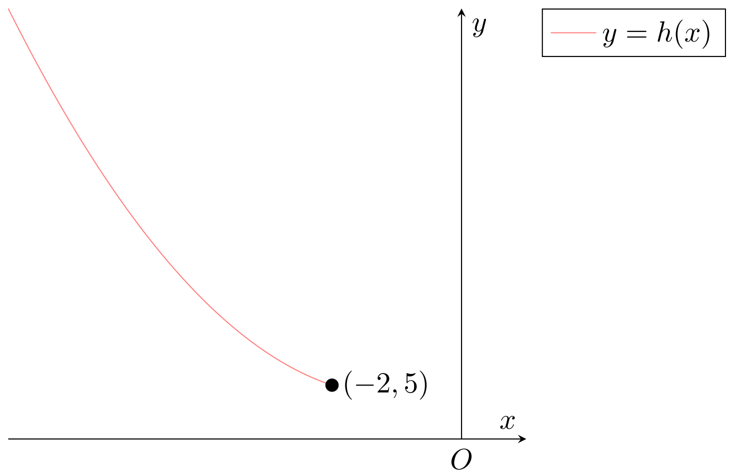
Solution.

Part (a).



$y = g(x)$ does not pass the horizontal line test. Hence, g is not one-one.

Part (b).



$y = h(x)$ passes the horizontal line test. Hence, h is one-one.
Note that $y = h(x) \implies x = h^{-1}(y)$. Now consider $y = h(x)$.

$$y = h(x) = x^2 + 2x + 5 = (x+1)^2 + 4 \implies x = -1 \pm \sqrt{y-4}.$$

Since $x \leq -2$, we reject $x = -1 + \sqrt{y-4}$. Note that $D_{h^{-1}} = R_h = [5, \infty)$. Hence,

$$h^{-1}\colon x \mapsto -1 - \sqrt{x-4},\ x \geq 5.$$

Problem 4. The function f is defined by

$$f\colon x \mapsto x + \frac{1}{x},\ x \neq 0.$$

(a) Sketch the graph of f and explain why f^{-1} does not exist.

(b) The function h is defined by $h\colon x \mapsto f(x)$, $x \in \mathbb{R}$, $x \geq \alpha$, where $\alpha \in \mathbb{R}^+$. Find the smallest value of α such that the inverse function of h exists.

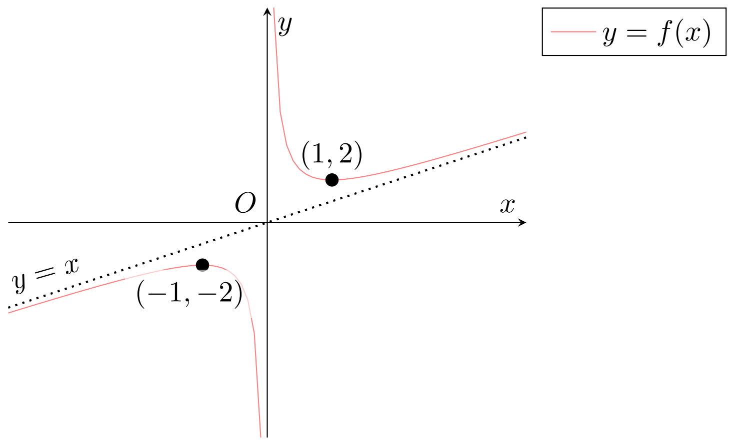
Using this value of α,

(c) State the range of h.

(d) Express h^{-1} in a similar form and sketch on a single diagram, the graphs of h and h^{-1}, showing clearly their geometrical relationship.

Solution.

Part (a).



$y = f(x)$ does not pass the horizontal line test. Hence, f is not one-one. Hence, f^{-1} does not exist.

Part (b). Consider $f'(x) = 0$ for $x > 0$.

$$f'(x) = 1 - \frac{1}{x^2} = 0 \implies x^2 = 1 \implies x = 1.$$

Note that we reject $x = -1$ since $x > 0$.

Looking at the graph of $y = f(x)$, we see that $f(x)$ achieves a minimum at $x = 1$. Hence, f is increasing for all $x \geq 1$. Thus, the smallest value of α is 1.

Part (c). Note $f(1) = 2$. Hence, from the graph, $R_h = [2, \infty)$.

Part (d). Note that $y = h(x) \implies x = h^{-1}(y)$. Now consider $y = h(x)$.

$$y = x + \frac{1}{x} \implies xy = x^2 + 1 \implies x^2 - yx + 1 = 0 \implies x = \frac{1}{2}\left(y \pm \sqrt{y^2 - 4}\right).$$

Note that $f(2) = \frac{5}{2}$. Since $2 = \frac{1}{2}\left(\frac{5}{2} + \sqrt{\left(\frac{5}{2}\right)^2 - 4}\right)$ and $2 \neq \frac{1}{2}\left(\frac{5}{2} - \sqrt{\left(\frac{5}{2}\right)^2 - 4}\right)$, we reject $x = \frac{1}{2}(y - \sqrt{y^2 - 4})$. Since $D_{f^{-1}} = R_f = [2, \infty)$, we thus have

$$h^{-1}\colon x \mapsto \frac{1}{2}\left(x + \sqrt{x^2 - 4}\right),\ x \geq 2.$$

Problem 5. The function f is defined as follows:

$$f\colon x \mapsto x^3 + x - 7,\ x \in \mathbb{R}.$$

(a) By using a graphical method or otherwise, show that the inverse of f exists.

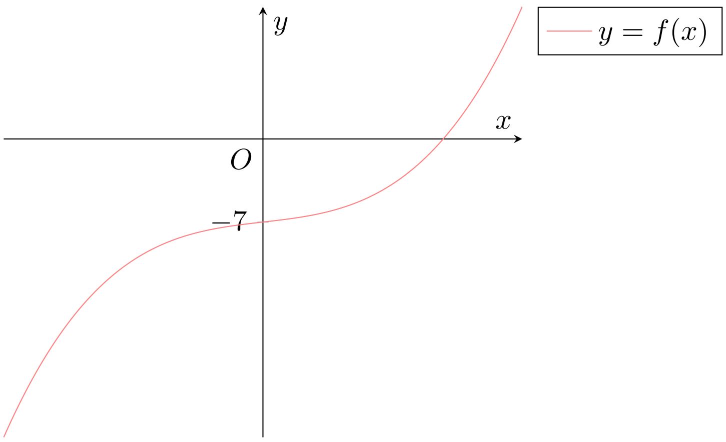
(b) Solve exactly the equation $f^{-1}(x) = 0$. Sketch the graph of f^{-1} together with the graph of f on the same diagram.

(c) Find, in exact form, the coordinates of the intersection point(s) of the graphs of f and f^{-1}.

(d) Given that the gradient of the tangent to the curve with equation $y = f^{-1}(x)$ is $\frac{1}{4}$ at the point with $x = p$, find the possible values of p.

Solution.

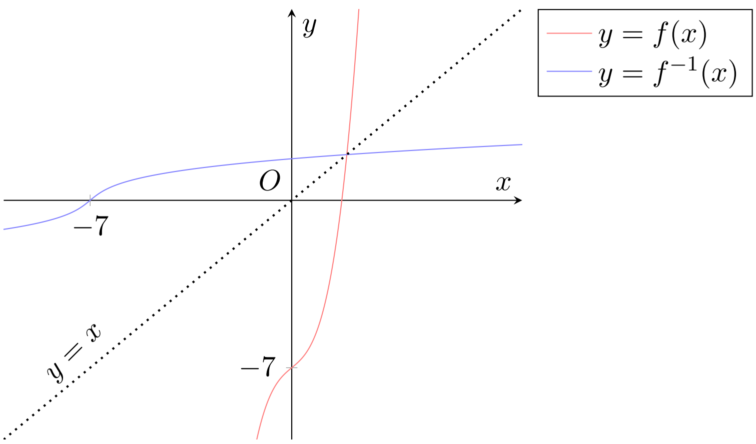
Part (a).



$y = f(x)$ passes the horizontal line test. Hence, f is one-one. Thus, f^{-1} exists.

Part (b). We have

$$f^{-1}(x) = 0 \implies x = f(0) = -7.$$



Part (c). Let (α, β) be the coordinates of the intersection between $f(x)$ and f^{-1}. From the graph, we see that $\alpha = \beta$, hence $f(\alpha) = \alpha$. Hence,

$$f(\alpha) = \alpha^3 + \alpha - 7 = \alpha \implies \alpha^3 = 7 \implies \alpha = \sqrt[3]{7}.$$

The coordinates are thus $\left(\sqrt[3]{7}, \sqrt[3]{7}\right)$.

Part (d). Note that

$$[f^{-1}(x)]' = \frac{1}{f'(f^{-1}(x))}.$$

Evaluating at $x = p$, we obtain

$$\frac{1}{4} = \left.\frac{1}{f'(f^{-1}(x))}\right|_{x=p} \implies f'(f^{-1}(x))\big|_{x=p} = 4.$$

Since $f'(x) = 3x^2 + 1$,

$$3f^{-1}(p)^2 + 1 = 4 \implies f^{-1}(p)^2 = 1 \implies f^{-1}(p) = \pm 1.$$

Case 1: $f^{-1}(p) = 1$. Then $p = f(1) = -5$.
Case 2: $f^{-1}(p) = -1$. Then $p = f(-1) = -9$.
Hence, $p = -5$ or $p = -9$.

* * * * *

Problem 6. The functions g and h are defined as follows:

$$g\colon x \mapsto \ln(x+2)\,, \qquad x \in (-1, 1)$$
$$h\colon x \mapsto x^2 - 2x - 1, \qquad x \in \mathbb{R}^+$$

(a) Sketch, on separate diagrams, the graphs of g and h.

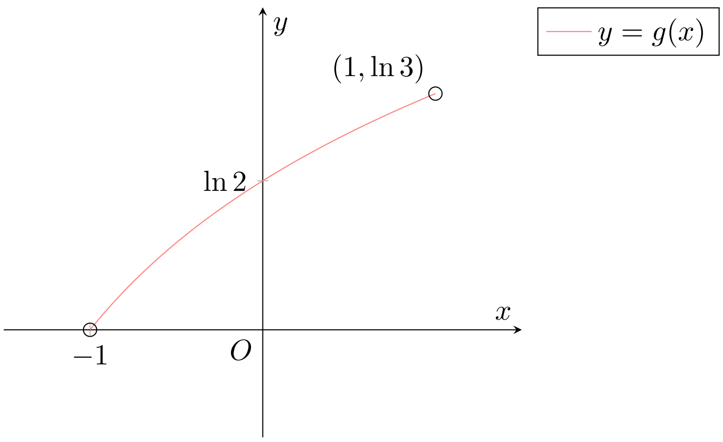
(b) Determine whether the composite function gh exists.

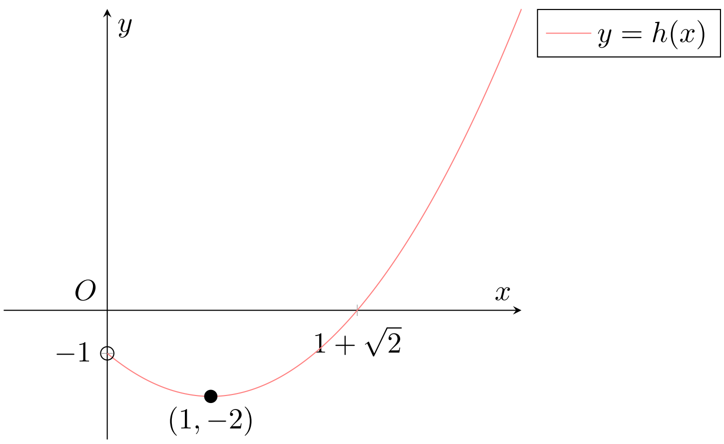
(c) Give the rule and domain of the composite function hg and find its range.

(d) The image of a under the composite function hg is -1.5. Find the value of a.

Solution.

Part (a).





Part (b). Observe that $R_h = [-2, \infty)$ and $D_g = (-1, 1)$. Hence, $R_h \nsubseteq D_g$. Thus, gh does not exist.

Part (c).

$$hg(x) = h(\ln(x+2)) = \ln(x+2)^2 - 2\ln(x+2) - 1.$$

Also note that $D_{hg} = D_g = (-1, 1)$. Hence,

$$hg\colon x \mapsto \ln(x+2)^2 - 2\ln(x+2) - 1,\ x \in (-1, 1).$$

Observe that h is decreasing on the interval $(0, 1]$ and increasing on the interval $[1, \infty)$. Note that $R_g = (0, \ln 3)$. Hence,

$$R_{hg} = [-2, \max\{h(0), h(\ln 3)\}) = [-2, -1).$$

Part (d). Note that $h(x) = (x-1)^2 - 2$. Hence, $h^{-1}(x) = 1 + \sqrt{x+2}$ (we reject $h^{-1}(x) = 1 - \sqrt{x+2}$ since $R_{h^{-1}} = D_h = \mathbb{R}^+$). Also note that $g^{-1} = \mathrm{e}^x - 2$. Thus,

$$\begin{aligned} hg(a) = -1.5 &\implies g(a) = h^{-1}(-1.5) = 1 + \sqrt{-1.5+2} = 1 + \frac{1}{\sqrt{2}} \\ &\implies a = g^{-1}\left(1 + \frac{1}{\sqrt{2}}\right) = \mathrm{e}^{1+\frac{1}{\sqrt{2}}} - 2. \end{aligned}$$

* * * * *

Problem 7. The functions f and g are defined as follows:

$$\begin{aligned} f\colon x &\mapsto 3 - x, && x \in \mathbb{R} \\ g\colon x &\mapsto \frac{4}{x}, && x \in \mathbb{R},\ x \neq 0 \end{aligned}$$

(a) Show that the composite function fg exists and express the definition of fg in a similar form. Find the range of fg.

(b) Find, in similar form, g^2 and g^3, and deduce g^{2017}.

(c) Find the set of values of x for which $g(x) = g^{-1}(x)$.

Solution.

Part (a). Note that $R_g = \mathbb{R} \setminus \{0\}$ and $D_g = \mathbb{R}$. Hence, $R_g \subseteq D_g$. Thus, fg exists.

$$fg(x) = f\left(\frac{4}{x}\right) = 3 - \frac{4}{x}.$$

Observe that $D_{fg} = D_g = \mathbb{R} \setminus \{0\}$. Thus,

$$fg\colon x \mapsto 3 - \frac{4}{x},\ x \in \mathbb{R} \setminus \{0\}.$$

Since $\frac{4}{x}$ can take on any value except 0, then $fg(x) = 3 - \frac{4}{x}$ can take on any value except 3. Thus,

$$R_{fg} = \mathbb{R} \setminus \{3\}.$$

Part (b). We have

$$g^2(x) = g\left(\frac{4}{x}\right) = \frac{4}{4/x} = x.$$

Hence,

$$g^2\colon x \mapsto x,\ x \in \mathbb{R} \setminus \{0\}\,.$$

We have

$$g^3(x) = g(g^2(x)) = g(x) = \frac{4}{x}.$$

Hence,

$$g^3\colon x \mapsto \frac{4}{x},\ x \in \mathbb{R} \setminus \{0\}\,.$$

Thus,

$$g^{2017} = g^{2016}(g(x)) = \left(g^2\right)^{1008} \circ g(x) = g(x) = \frac{4}{x}.$$

Hence,

$$g^{2017}\colon x \mapsto \frac{4}{x},\ x \in \mathbb{R} \setminus \{0\}\,.$$

Part (c). Note that $g(x) = g^{-1}(x) \implies g^2(x) = x$. From the definition of $g^2(x)$, we know that $g^2(x) = x$ for all x in D_{g^2}. Hence, the solution set is $\mathbb{R} \setminus \{0\}$.

* * * * *

Problem 8. The function f is defined by

$$f(x) = \begin{cases} 2x+1, & 0 \leq x < 2 \\ (x-4)^2+1, & 2 \leq x < 4. \end{cases}$$

It is further given that $f(x) = f(x+4)$ for all real values of x.

(a) Find the values of $f(1)$ and $f(5)$ and hence explain why f is not one-one.

(b) Sketch the graph of $y = f(x)$ for $-4 \leq x < 8$.

(c) Find the range of f for $-4 \leq x < 8$.

Solution.

Part (a). We have

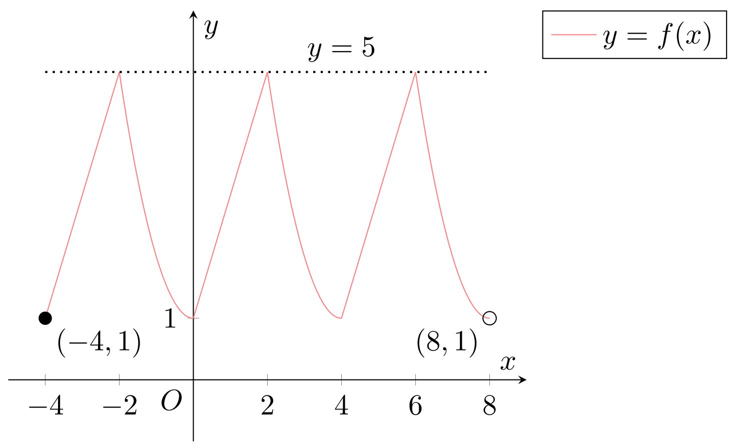
$$f(1) = 2(1) + 1 = 3$$

and

$$f(5) = f(1+4) = f(1) = 3.$$

Since $f(1) = f(5)$, but $1 \neq 5$, f is not one-one.

Part (b).



Part (c). From the graph, $R_f = [1, 5]$.

* * * * *

Problem 9.

(a) The function f is given by $f\colon x \mapsto 1 + \sqrt{x}$ for $x \in \mathbb{R}^+$.

  (i) Find $f^{-1}(x)$ and state the domain of f^{-1}.

  (ii) Find $f^2(x)$ and the range of f^2.

  (iii) Show that if $f^2(x) = x$ then $x^3 - 4x^2 + 4x - 1 = 0$. Hence, find the value of x for which $f^2(x) = x$. Explain why this value of x satisfies the equation $f(x) = f^{-1}(x)$.

(b) The function g, with domain the set of non-negative integers, is given by

$$g(n) = \begin{cases} 1, & n = 0 \\ 2 + g\left(\frac{1}{2}n\right), & n \text{ even} \\ 1 + g(n-1), & n \text{ odd} \end{cases}$$

  (i) Find $g(4)$, $g(7)$ and $g(12)$.

  (ii) Does g have an inverse? Justify your answer.

Solution.

Part (a).

Part (a)(i). Let $y = f(x)$. Then $x = f^{-1}(y)$.

$$y = f(x) = 1 + \sqrt{x} \implies \sqrt{x} = y - 1 \implies x = (y-1)^2.$$

Hence, $f^{-1}(x) = (x-1)^2$.

Observe that $D_{f^{-1}} = R_f = (1, \infty)$. Thus, $D_{f^{-1}} = (1, \infty)$.

Part (a)(ii). We have

$$f^2(x) = f(1 + \sqrt{x}) = 1 + \sqrt{1 + \sqrt{x}}.$$

Observe that $\sqrt{1 + \sqrt{x}} > 1$. Hence, $1 + \sqrt{1 + \sqrt{x}} > 1 + 1 = 2$, whence $R_{f^2} = (2, \infty)$.

Part (a)(iii). Note that $f^2(x) = x \implies 1 + \sqrt{1 + \sqrt{x}} = x$, whence x satisfies the recursion $1 + \sqrt{x} = x$. Hence,

$$1 + \sqrt{x} = x \implies \sqrt{x} = x - 1 \implies x = x^2 - 2x + 1 \implies x^2 - 3x + 1 = 0.$$

We can manipulate this to form the desired cubic equation:

$$0 = x\left(x^2 - 3x + 1\right) - \left(x^2 - 3x + 1\right) = x^3 - 4x^2 + 4x - 1.$$

Solving the initial quadratic equation yields $x = \frac{1}{2}\left(3 \pm \sqrt{5}\right)$. Observe that $\frac{3-\sqrt{5}}{2} < 2$ and $\frac{3+\sqrt{5}}{2} > 2$. Thus, the sole solution is $x = \frac{3+\sqrt{5}}{2}$.

Consider $f(x) = f^{-1}(x)$. Applying f on both sides of the equation, we have $f^2(x) = f(x)$. Since $x = \frac{3+\sqrt{5}}{2}$ satisfies $f^2(x) = f(x)$, it also satisfies $f(x) = f^{-1}(x)$.

Part (b).

Part (b)(i). We have

$$g(4) = 2 + g(2) = 2 + 2 + g(1) = 4 + 1 + g(0) = 5 + 1 = 6,$$

$$g(7) = 1 + g(6) = 1 + 2 + g(3) = 3 + 1 + g(2) = 4 + (g(4) - 2) = 2 + 6 = 8,$$

and

$$g(12) = 2 + g(6) = 2 + (g(7) - 1) = 1 + 8 = 9.$$

Part (b)(ii). Consider $g(5)$ and $g(6)$.

$$g(5) = 1 + g(4) = 1 + 6 = 7, \quad g(6) = g(7) - 1 = 8 - 1 = 7.$$

Since $g(5) = g(6)$, but $5 \neq 6$, g is not one-one. Hence, g^{-1} does not exist.

Self-Practice B3

Problem 1. Functions f and g are defined by

$$f : x \mapsto \frac{3x-2}{x+1}, \quad x \in \mathbb{R}, \quad x \neq -1,$$
$$g : x \mapsto 3x+4, \quad x \in \mathbb{R}.$$

(a) Find $f^{-1}(x)$ and state the domain and range of f^{-1}.

(b) Express in similar form, the functions g^2 and gf.

(c) Find the value of x for which $(gf)^{-1}(x) = 0$.

Solution.

Part (a). Let $y = f(x)$. Then

$$y = \frac{3x-2}{x+1} \implies xy + y = 3x - 2 \implies x = \frac{y+2}{3-y}.$$

Thus, $f^{-1}(x) = (x+2)/(3-x)$. Further, $D_{f^{-1}} = \mathbb{R} \setminus \{3\}$ and $R_{f^{-1}} = \mathbb{R} \setminus \{-1\}$.

Part (b). We have

$$g^2(x) = 3\,(3x+4) + 4 = 9x + 16$$

and

$$gf(x) = 3\left(\frac{3x-2}{x+1}\right) + 4 = \frac{13x-2}{x+1},$$

so

$$g^2 : x \mapsto 9x + 16, \quad x \in \mathbb{R},$$
$$gf : x \mapsto \frac{13x-2}{x+1}, \quad x \in \mathbb{R}, \quad x \neq -1.$$

Part (c). We have $(gf)^{-1}(x) = 0$, so $x = gf(0) = -2$.

* * * * *

Problem 2. The function f is defined by

$$f : x \mapsto 2 - (x-1)^2, \quad x \leq 1.$$

(a) Sketch the graphs of $y = f(x)$, $y = f^{-1}(x)$ and $y = f^{-1}f(x)$ on a single diagram.

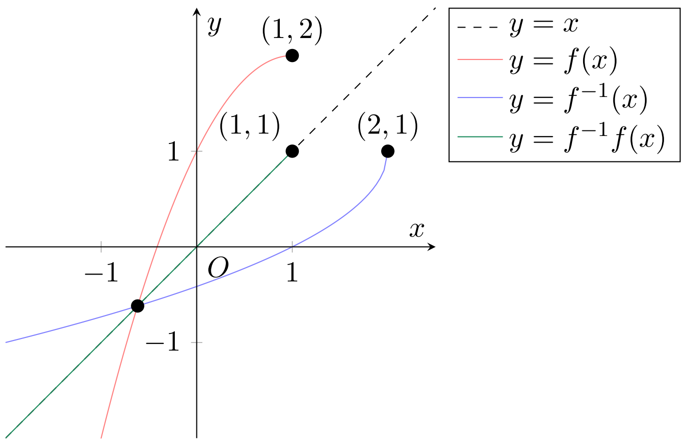
(b) If $f(\beta) = f^{-1}(\beta)$, find the values of the constant p and q such that

$$\beta^2 - p\beta + q = 0.$$

(c) Define f^{-1} in a similar form.

Solution.

Part (a).



Part (b). Graphically, we see that the intersection points of $y = f(\beta)$ and $y = f^{-1}(\beta)$ lie on the line $y = x$. It thus suffices to solve $f(\beta) = \beta$, from which we gather

$$2 - (\beta - 1)^2 = \beta \implies \beta^2 - \beta - 1 = 0,$$

so $p = 1$ and $q = -1$.

Part (c). Let $y = f(x)$. Then

$$y = 2 - (x - 1)^2 \implies (x - 1)^2 = 2 - y \implies x - 1 = \pm\sqrt{2 - y} \implies x = 1 \pm \sqrt{2 - y}.$$

Since $x \leq 1$, we take the negative branch, for

$$f^{-1} : x \mapsto 1 - \sqrt{2 - x}, \quad x \leq 2.$$

* * * * *

Problem 3. The function f is defined by $f : x \mapsto \cos \frac{\pi x}{2}$, $x \in \mathbb{R}$, $-2 < x \leq 0$.

(a) Sketch the graph of $y = f(x)$ indicating clearly the coordinates of all axial intercepts and end points.

(b) Show that f^{-1} exists, and find its rule and domain.

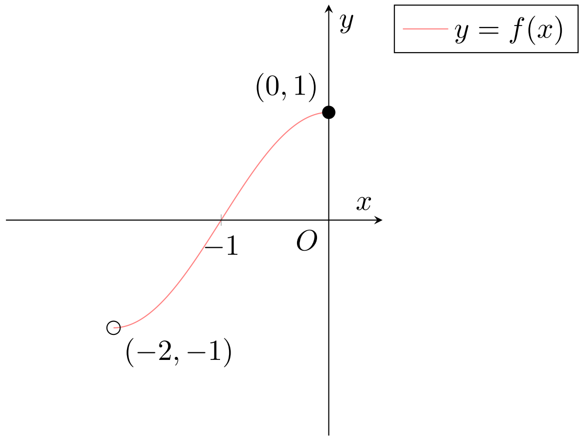
The function g is defined by $g : x \mapsto (2x + 1)^{2/3}$, $x \in \mathbb{R}$, $-2 < x \leq 2$.

(c) Find the set of values of x such that $g(x) \geq f(x)$

(d) Explain clearly why gf exists. Hence, find the range of gf.

Solution.

Part (a).



Part (b). For all constants k, the line $y = k$ and $y = f(x)$ has at most one intersection point. Hence, $y = f(x)$ passes the horizontal rule test, so it is one-one and thus invertible, i.e. f^{-1} exists.

Let $y = f(x)$. Then

$$y = \cos\frac{\pi x}{2} \implies \frac{\pi x}{2} = \arccos y \implies x = \frac{2}{\pi}\arccos y.$$

Note further that $D_{f^{-1}} = R_f = [-1, 1]$, so

$$f^{-1} : x \mapsto \frac{2}{\pi}\arccos x, \quad -1 \leq x \leq 1.$$

Part (c). Using G.C., we see that the solution set is

$$\{x \in \mathbb{R} \mid -2 < x \leq 0.673 \text{ or } x = 0\}.$$

Part (d). Since $R_f = (-1, 1]$ and $D_g(-2, 2]$, we have $R_f \subseteq D_g$, thus gf exists.

* * * * *

Problem 4. Functions f and$g$ are defined by:

$$f : x \mapsto x^2 + c, \quad x \leq 2,$$
$$g : x \mapsto 5 + \frac{3}{x}, \quad x \geq k,$$

where c, k are positive constants and $c > k$.

(a) Show that g^{-1} exists.

(b) Find g^{-1} in similar form, expressing its domain in terms of k.

(c) Determine whether each of the two functions, fg and gf, exists. Where it exists, express the composite function in similar form and state its range.

Solution.

Part (a). Let $x_1, x_2 \in D_g$ such that $g(x_1) = g(x_2)$. Then

$$5 + \frac{3}{x_1} = 5 + \frac{3}{x_2} \implies \frac{3}{x_1} = \frac{3}{x_2} \implies x_1 = x_2,$$

so $g(x)$ is one-one and thus invertible, i.e. g^{-1} exists.

Part (b). Let $y = g(x)$. Then

$$y = 5 + \frac{3}{x} \implies \frac{3}{x} = y - 5 \implies \frac{x}{3} = \frac{1}{y-5} \implies x = \frac{3}{y-5}.$$

Note also that for $x \geq k$,

$$5 < 5 + \frac{3}{x} \leq 5 + \frac{3}{k},$$

so

$$g^{-1} : x \mapsto \frac{3}{x-5}, \quad 5 < x \leq 5 + \frac{3}{k}.$$

Part (c). Note that $D_f = (-\infty, 2]$, $R_f = [c, \infty)$, $D_g = [k, \infty)$, and $R_g = (5, 5 + 3/k]$.

Since $R_g \not\subseteq D_f$, the composite function fg does not exist. However, because $c > k$, we have $R_f \subseteq D_g$, so gf exists and is given by

$$gf : x \mapsto 5 + \frac{3}{x^2 + c}, \quad x \leq 2.$$

Its range is $(5, 5 + 3/c]$.

* * * * *

Problem 5 (🌶). Functions f and g are defined such that

$$f : x \mapsto \arccos\left(x^2\right), \quad -1 \leq x \leq 1,$$
$$g : x \mapsto x^3 + 1, \quad x \in \mathbb{R}.$$

(a) Explain why the composite function fg does not exist.

The function h is defined such that $h(x) = g(x)$ and the domain of h is $a \leq x \leq 0$. It is given that $a = -5/4$.

(b) Find the range of fh in exact form.

(c) Determine all the possible value(s) of x that satisfies $g^{-1}(x^2) = 2$. Hence, explain why $h^{-1}(x^2) = 2$ has no solution.

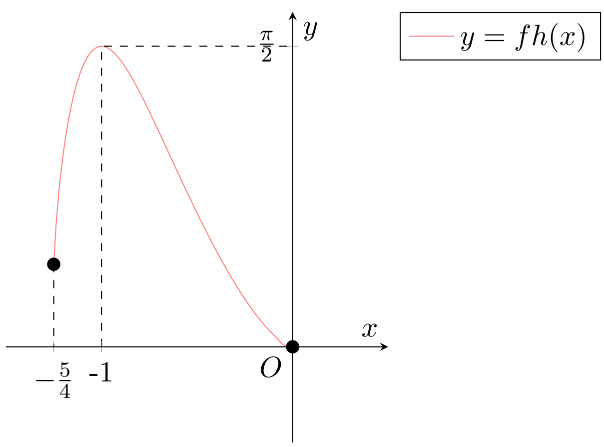
Solution.

Part (a). Note that $D_f = [-1, 1]$ and $R_g = \mathbb{R}$, hence $R_g \not\subseteq D_f$, so the function fg does not exist.

Part (b). Note that

$$fh(x) = \arccos\left(\left(x^3 + 1\right)^2\right),$$

where $-5/4 \leq x \leq 0$.



From the graph of $y = fh(x)$, we see that the maximum is attained when $x = -1$, where $y = \arccos 0 = \pi/2$. Meanwhile, the minimum is attained when $x = 0$, where $y = arccos 1 = 0$. Thus, $R_f h = [0, \pi/2]$.

Part (c). We have

$$g^{-1}(x^2) = 2 \implies x^2 = g(2) = 9 \implies x = \pm 3.$$

$h^{-1}(x^2) = 2$ has no solution since $2 \notin D_{h^{-1}} = R_h = [-5/4, 0]$.

Assignment B3

Problem 1. Functions f and g are defined as follows:

$$\begin{aligned} f\colon x &\mapsto (x-3)^2+6, && x\in\mathbb{R},\ x\leq 2\\ g\colon x &\mapsto \ln(x-2)\,, && x\in\mathbb{R},\ x>3 \end{aligned}$$

(a) Show that f^{-1} exists and define f^{-1} in a similar form.

(b) Sketch, on the same diagram, the graphs of f, f^{-1} and ff^{-1}.

(c) Find fg and gf if they exist, and find their ranges (where applicable).

Solution.

Part (a). Note that $f' = 2(x-3) < 0$ for all $x \leq 2$. Thus, f is strictly decreasing. Since f is also continuous, f is one-one. Thus, f^{-1} exists.

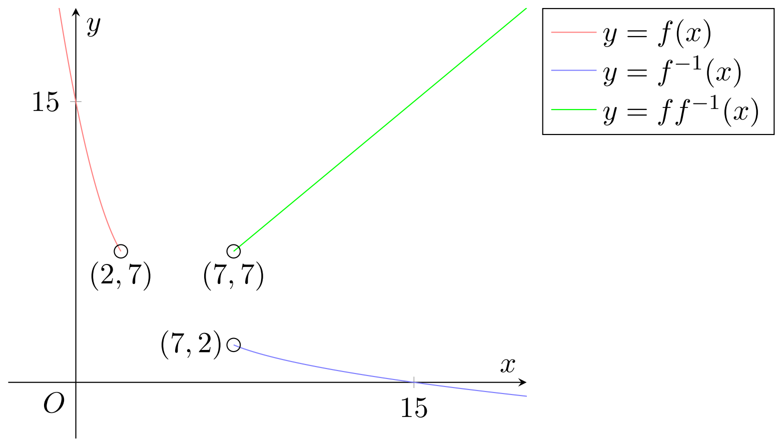
Let $y = f(x) \implies x = f^{-1}(y)$.

$$y = f(x) = (x-3)^2 + 6 \implies x = 3 \pm \sqrt{y-6}.$$

Since $x < 3$, we reject $x = 3+\sqrt{y-6}$. Lastly, observe that $D_{f^{-1}} = R_f = [f(2), \infty) = [7, \infty)$. Thus,

$$f^{-1}\colon x \mapsto 3-\sqrt{x-6},\ x\in\mathbb{R},\ x\geq 7.$$

Part (b).



Part (c). Note that $R_g = (0,\infty)$ and $D_f = (-\infty, 2]$. Hence, $R_g \nsubseteq D_f$. Thus, fg does not exist. Note that $R_f = [7,\infty)$ and $D_g = (3,\infty)$. Hence, $R_f \subseteq D_g$. Thus, gf exists.

Since $\ln x$ is a strictly increasing function, we have that g is also strictly increasing. Hence, $R_{gf} = [\ln(7-2)\,, \infty) = [\ln 5, \infty)$.

* * * * *

Problem 2. The function f is defined as follows:

$$f\colon x\mapsto \frac{1}{x^2-1}, \qquad x\in\mathbb{R},\ x\neq -1,\ x\neq 1.$$

(a) Sketch the graph of $y = f(x)$.

(b) If the domain of f is further restricted to $x \geq k$, state with a reason the least value of k for which the function f^{-1} exists.

In the rest of the question, the domain of f is $x \in \mathbb{R}$, $x \neq -1$, $x \neq 1$, as originally defined.

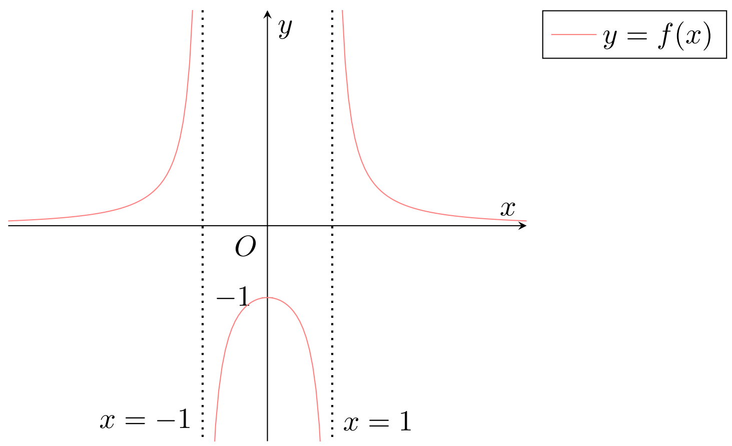
The function g is defined as follows:

$$g\colon x \mapsto \frac{1}{x-3}, \qquad x \in \mathbb{R},\ x \neq 2,\ x \neq 3, x \neq 4.$$

(c) Find the range of fg.

Solution.

Part (a).



Part (b). If the domain of f is further restricted to $x \geq 0$, f would pass the horizontal line test, whence f^{-1} would exist. Hence, $\min k = 0$.

Part (c). Observe that $R_g = \mathbb{R} \setminus \{g(2), g(4)\} = \mathbb{R} \setminus \{-1, 1\}$. Hence, $R_{fg} = R_f = \mathbb{R} \setminus (-1, 0\,]$.

* * * * *

Problem 3. The function f is defined by

$$f\colon x \mapsto \frac{x}{x^2-1}, \qquad x \in \mathbb{R},\ x \neq -1,\ x \neq 1.$$

(a) Explain why f does not have an inverse.

(b) The function f has an inverse if the domain is restricted to $x \leq k$. State the largest value of k.

The function g is defined by

$$g\colon x \mapsto \ln x - 1, \qquad x \in \mathbb{R},\ 0 < x < 1.$$

(c) Find an expression for $h(x)$ for each of the following cases:

(i) $gh(x) = x$

(ii) $hg(x) = x^2 + 1$

Solution.

Part (a). Observe that $f(1/2) = -2/3$ and $f(-2) = -2/3$. Hence, $f(1/2) = f(-2)$. Since $1/2 \neq -2$, f is not one-one. Thus, f does not have an inverse.

Part (b). Clearly, $\max k = 0$.

Part (c).

Part (c)(i). Note that $gh(x) = x \implies h(x) = g^{-1}(x)$. Hence, consider $y = g(x) \implies x = h(y)$.

$$y = g(x) = \ln x - 1 \implies \ln x = y + 1 \implies x = \mathrm{e}^{y+1}.$$

Hence, $h(x) = \mathrm{e}^{x+1}$.

Part (c)(ii). Let $h = h_2 \circ h_1$ such that $h_1 g(x) = x \implies h_1(x) = g^{-1}(x) \implies h_1(x) = \mathrm{e}^{x+1}$. Then

$$hg(x) = x^2 + 1 \implies h_2 h_1 g(x) = x^2 + 1 \implies h_2(x) = x^2 + 1.$$

Hence, $h(x) = h_2 h_1(x) = h_2(\mathrm{e}^{x+1}) = \left(\mathrm{e}^{x+1}\right)^2 + 1 = \mathrm{e}^{2x+2} + 1$.

B4 Differentiation

Tutorial B4

Problem 1. Evaluate the following limits.

(a) $\lim\limits_{x \to 5} (6x + 7)$

(b) $\lim\limits_{x \to 1} \frac{x^3-1}{1-x}$

(c) $\lim\limits_{x \to \infty} \frac{3x}{2x^2-5}$

Solution.

Part (a).

$$\lim_{x \to 5} (6x + 7) = 6(5) + 7 = 37.$$

Part (b).

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

Part (c).

$$\lim_{x \to \infty} \frac{3x}{2x^2 - 5} = \lim_{x \to \infty} \frac{3}{2x - 5/x}.$$

Note that as $x \to \infty$, $2x - \frac{5}{x} \to \infty$. Hence, $\lim\limits_{x \to \infty} \frac{1}{2x - 5/x} = 0$.

* * * * *

Problem 2. Differentiate the following with respect to x from first principles.

(a) $3x + 4$

(b) x^3

Solution.

Part (a).

$$\frac{\mathrm{d}}{\mathrm{d}x}(3x + 4) = \lim_{h \to 0} \frac{[3(x + h) + 4] - [3x + 4]}{h} = \lim_{h \to 0} \frac{3h}{h} = \lim_{h \to 0} 3 = 3.$$

Part (b).

$$\frac{\mathrm{d}}{\mathrm{d}x} x^3 = \lim_{h \to 0} \frac{(x + h)^3 - x^3}{h} = \lim_{h \to 0} \frac{3hx^2 + 3h^2x + h^3}{h} = \lim_{h \to 0} \left(3x^2 + 3hx + h^2\right) = 3x^2.$$

Problem 3. Differentiate each of the following with respect to x, simplifying your answer.

(a) $\left(x^2+4\right)^2\left(2x^3-1\right)$

(b) $\frac{x^2}{\sqrt{4-x^2}}$

(c) $\sqrt{1+\sqrt{x}}$

(d) $\left(\frac{x^3-1}{2x^3+1}\right)^4$

Solution.

Part (a).

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2+4\right)^2\left(2x^3-1\right)=\left(2x^3-1\right)\left[4x\left(x^2+4\right)\right]+\left(x^2+4\right)^2\left(6x^2\right)\\&=2x\left(x^2+4\right)\left[2\left(2x^3-1\right)+3x\left(x^2+4\right)\right]=2x\left(x^2+4\right)\left(7x^3+12x-2\right).\end{aligned}$$

Part (b).

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^2}{\sqrt{4-x^2}}=\frac{\sqrt{4-x^2}\,(2x)-x^2\left(\frac{-2x}{2\sqrt{4-x^2}}\right)}{4-x^2}=\frac{2x\left(4-x^2\right)+x^3}{(4-x^2)^{3/2}}=\frac{x\left(8-x^2\right)}{(4-x^2)^{3/2}}.$$

Part (c).

$$\frac{\mathrm{d}}{\mathrm{d}x}\sqrt{1+\sqrt{x}}=\frac{1}{2\sqrt{1+\sqrt{x}}}\cdot\frac{1}{2\sqrt{x}}=\frac{1}{4\sqrt{x(1+\sqrt{x})}}.$$

Part (d). Note that

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

Hence,

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^3-1}{2x^3+1}=\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{1}{2}-\frac{3}{2}\left(\frac{1}{2x^3+1}\right)\right]=-\frac{3}{2}\left[\frac{-6x^2}{(2x^3+1)^2}\right]=\frac{9x^2}{(2x^3+1)^2}.$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x^3-1}{2x^3+1}\right)^4=4\left(\frac{x^3-1}{2x^3+1}\right)^3\frac{9x^2}{(2x^3+1)^2}=\frac{36x^2\left(x^3-1\right)^3}{(2x^3+1)^5}.$$

* * * * *

Problem 4. Using a graphing calculator, find the derivative of $\frac{\mathrm{e}^{2x}}{x^2+1}$ when $x=1.5$.

Solution.

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^{2x}}{x^2+1}\right)\bigg|_{x=1.5}=6.66.$$

* * * * *

Problem 5. Find the derivative with respect to x of

(a) $\cos x^\circ$

(b) $\cot\left(1-2x^2\right)$

(c) $\tan^3(5x)$

(d) $\frac{\sec x}{1+\tan x}$

Solution.

Part (a).

$$\frac{\mathrm{d}}{\mathrm{d}x}\cos x^\circ = \frac{\mathrm{d}}{\mathrm{d}x}\cos\Big(\frac{\pi}{180}x\Big) = -\frac{\pi}{180}\sin\Big(\frac{\pi}{180}x\Big)\,.$$

Part (b).

$$\frac{\mathrm{d}}{\mathrm{d}x}\cot\big(1-2x^2\big) = 4x\csc\big(1-2x^2\big)\,.$$

Part (c).

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan^3(5x) = 15\tan^2(5x)\sec^2(5x).$$

Part (d). Note that

$$\frac{\sec x}{1+\tan x} = \frac{1}{\sin x+\cos x} = \frac{1}{\sqrt{2}\sin(x+\pi/4)} = \frac{1}{\sqrt{2}}\csc\Big(x+\frac{\pi}{4}\Big)\,.$$

Hence,

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\sec x}{1+\tan x} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{\sqrt{2}}\csc\Big(x-\frac{\pi}{4}\Big) = -\frac{1}{\sqrt{2}}\csc\Big(x+\frac{\pi}{4}\Big)\cot\Big(x+\frac{\pi}{4}\Big)\,.$$

* * * * *

Problem 6. Find the derivative with respect to x of

(a) $y = \mathrm{e}^{1+\sin 3x}$

(b) $y = x^2\mathrm{e}^{\frac{1}{x}}$

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

(d) $y = \frac{\ln(2x)}{x}$

(e) $y = \log_2(3x^4 - \mathrm{e}^x)$

(f) $y = 3^{\ln\sin x}$

(g) $y = a^{2\log_a x}$

(h) $y = \sqrt[3]{\frac{\mathrm{e}^x(x+1)}{x^2+1}}$, $x > 0$

Solution.

Part (a).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{1+\sin 3x} = 3\mathrm{e}^{1+\sin 3x}\cos(3x)\,.$$

Part (b).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}x^2\mathrm{e}^{1/x} = -x^2\left(-\frac{\mathrm{e}^{1/x}}{x^2}\right) + \mathrm{e}^{1/x}(2x) = \mathrm{e}^{1/x}(2x-1).$$

Part (c). Note that

$$y = \ln\left(\frac{1-x}{\sqrt{1+x^2}}\right) = \ln(1-x) - \frac{1}{2}\ln\left(1+x^2\right).$$

Hence,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\left[\ln(1-x) - \frac{1}{2}\ln\left(1+x^2\right)\right] = -\frac{1}{1-x} - \frac{x}{1+x^2} = -\frac{1+x}{(1-x)\left(1+x^2\right)}.$$

Part (d).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\ln(2x)}{x} = \frac{x\left(\frac{1}{x}\right) - \ln(2x)\,(1)}{x^2} = \frac{1-\ln(2x)}{x^2}.$$

Part (e). Note that

$$y = \log_2\left(3x^4 - \mathrm{e}^x\right) \implies 2^y = 3x^4 - \mathrm{e}^x.$$

Implicitly differentiating with respect to x,

$$2^y \ln 2 \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = 12x^3 - \mathrm{e}^x \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{12x^3 - \mathrm{e}^x}{2^y \ln 2} = \frac{12x^3 - \mathrm{e}^x}{\left(3x^4 - \mathrm{e}^x\right)\ln 2}.$$

Part (f). Note that

$$y = 3^{\ln\sin x} \implies \log_3 y = \frac{\ln y}{\ln 3} = \ln\sin x \implies \ln y = \ln 3 \ln\sin x.$$

Implicitly differentiating with respect to x,

$$\frac{1}{y}\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = \ln 3\left(\frac{\cos x}{\sin x}\right) = \ln 3 \cdot \cot x \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \ln 3 \cdot y\cot x = \ln 3 \cdot \cot(x)\cdot 3^{\ln\sin x}.$$

Part (g). Observe that

$$y = a^{2\log_a x} = a^{\log_a x^2} = x^2 \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}x^2 = 2x.$$

Part (h). Note that

$$y = \sqrt[3]{\frac{\mathrm{e}^x(x+1)}{x^2+1}} \implies \left(x^2+1\right)y^3 = \mathrm{e}^x(x+1).$$

Implicitly differentiating with respect to x,

$$\left(x^2+1\right)\left(3y^2\cdot\frac{\mathrm{d}y}{\mathrm{d}x}\right) + y^3\,(2x) = \mathrm{e}^x + (x+1)\mathrm{e}^x \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{e}^x(x+2) - 2xy^3}{3\left(x^2+1\right)y^2}.$$

Now observe that

$$\frac{\mathrm{e}^x(x+2)}{3\left(x^2+1\right)y^2} = \frac{\mathrm{e}^x(x+1)(x+2)}{3\left(x^2+1\right)(x+1)y^2} = \frac{y^3(x+2)}{3(x+1)y^2} = y\left(\frac{x+2}{3(x+1)}\right).$$

Thus,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y\left(\frac{x+2}{3(x+1)}\right) - y\left(\frac{2x}{3\left(x^2+1\right)}\right) = \frac{1}{3}\sqrt[3]{\frac{\mathrm{e}^x(x+1)}{x^2+1}}\left(\frac{x+2}{x+1} - \frac{2x}{x^2+1}\right).$$

Problem 7. Find the derivative with respect to x of

(a) $\arccos\left(\frac{x}{10}\right)$

(b) $\arctan\left(\frac{1}{1-x}\right)$

(c) $\arcsin(\tan x)$

Solution.

Part (a).

$$\frac{\mathrm{d}}{\mathrm{d}x}\arccos\frac{x}{10} = -\frac{1}{10\sqrt{1-\left(\frac{x}{10}\right)^2}} = -\frac{1}{\sqrt{100-x^2}}.$$

Part (b).

$$\frac{\mathrm{d}}{\mathrm{d}x}\arctan\left(\frac{1}{1-x}\right) = \frac{1}{1+\left(\frac{1}{1-x}\right)^2}\left(\frac{1}{(1-x)^2}\right) = \frac{1}{(1-x)^2+1}.$$

Part (c).

$$\frac{\mathrm{d}}{\mathrm{d}x}\arcsin(\tan x) = \frac{\sec^2 x}{1-\tan^2 x}.$$

* * * * *

Problem 8. Find an expression for $\mathrm{d}y/\mathrm{d}x$ in terms of x and y.

(a) $(y-x)^2 = 2a(y+x)$, where a is a constant

(b) $y^2 = \mathrm{e}^{2x}y + x\mathrm{e}^x$

(c) $y = x^y$

(d) $\sin x\cos y = \frac{1}{2}$

Solution.

Part (a). Implicitly differentiating with respect to x,

$$(y-x)\left(\frac{\mathrm{d}y}{\mathrm{d}x}-1\right) = a\left(\frac{\mathrm{d}y}{\mathrm{d}x}+1\right) \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{a+y-x}{y-x-a}.$$

Part (b). Implicitly differential with respect to x,

$$2y\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = \left(\mathrm{e}^{2x}\frac{\mathrm{d}y}{\mathrm{d}x}+2ye^{2x}\right)+(x\mathrm{e}^x+\mathrm{e}^x) \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{e}^x\,(2y\mathrm{e}^x+x+1)}{2y-\mathrm{e}^{2x}}.$$

Part (c). Note that

$$y = x^y \implies \ln y = y\ln x.$$

Implicitly differentiating with respect to x,

$$\frac{1}{y}\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y}{x}+\ln x\cdot\frac{\mathrm{d}y}{\mathrm{d}x} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y/x}{1/y-\ln x} = \frac{y^2}{x-xy\ln x}.$$

Part (d). Note that

$$\sin x\cos y = \frac{1}{2} \implies \cos y = \frac{1}{2}\csc x.$$

Implicitly differentiating with respect to x,

$$-\sin y \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{1}{2} \csc x \cot x \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\csc x \cot x}{2 \sin y}.$$

* * * * *

Problem 9. It is given that x and y satisfy the equation

$$\arctan x + \arctan y + \arctan(xy) = \frac{7}{12}\pi.$$

(a) Find the exact value of y when $x = 1$.

(b) Express $\frac{\mathrm{d}}{\mathrm{d}x} \arctan(xy)$ in terms of x, y and y'.

(c) Show that, when $x = 1$, $y' = -\frac{1}{3} - \frac{1}{2\sqrt{3}}$.

Solution.

Part (a). Substituting $x = 1$ into the given equation,

$$\frac{1}{4}\pi + 2 \arctan y = \frac{7\pi}{12} \implies \arctan y = \frac{\pi}{6} \implies y = \frac{1}{\sqrt{3}}.$$

Part (b).

$$\frac{\mathrm{d}}{\mathrm{d}x} \arctan(xy) = \frac{xy' + y}{1 + (xy)^2}.$$

Part (c). Differentiating the given equation with respect to x,

$$\frac{1}{1+x^2} + \frac{y'}{1+y^2} + \frac{xy'+y}{1+(xy)^2} = 0.$$

Substituting $x = 1$,

$$\frac{1}{2} + \frac{3y'}{4} + \frac{3}{4}\left(y' + \frac{1}{\sqrt{3}}\right) = 0 \implies y' = \frac{2}{3}\left(-\frac{3}{4\sqrt{3}} - \frac{1}{2}\right) = -\frac{1}{2\sqrt{3}} - \frac{1}{3}.$$

* * * * *

Problem 10. Find $\mathrm{d}y/\mathrm{d}x$ for

(a) $x = \frac{1}{1+t^2}$, $y = \frac{t}{1+t^2}$

(b) $x = \frac{1}{2}(\mathrm{e}^t - \mathrm{e}^{-t})$, $y = \frac{1}{2}(\mathrm{e}^t + \mathrm{e}^{-t})$

(c) $x = a \sec \theta$, $y = a \tan \theta$

(d) $x = \mathrm{e}^{3\theta} \cos(3\theta)$, $y = \mathrm{e}^{3\theta} \sin(3\theta)$

Solution.

Part (a). Observe that $y = xt$. Hence,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x\left(\frac{\mathrm{d}t}{\mathrm{d}x}\right) + t = x\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^{-1} + t = \frac{1}{1+t^2}\left(-\frac{2t}{(1+t^2)^2}\right)^{-1} + t = \frac{t^2-1}{2t}.$$

Part (b).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{\frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)}{\frac{1}{2}\left(\mathrm{e}^t + \mathrm{e}^{-t}\right)} = \frac{\mathrm{e}^t - \mathrm{e}^{-t}}{\mathrm{e}^t + \mathrm{e}^{-t}}.$$

Part (c). Recall that $\tan^2\theta + 1 = \sec^2\theta$. Hence, $y^2 + a^2 = x^2$. Implicitly differentiating with respect to x, we have

$$2y \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = 2x \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x}{y} = \frac{a\sec\theta}{a\tan\theta} = \csc\theta.$$

Part (d).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}(3\theta)}{\mathrm{d}x/\mathrm{d}(3\theta)} = \frac{\mathrm{e}^{3\theta}\cos 3\theta + \mathrm{e}^{3\theta}\sin 3\theta}{-\mathrm{e}^{3\theta}\sin 3\theta + \mathrm{e}^{3\theta}\cos 3\theta} = \frac{\cos 3\theta + \sin 3\theta}{\cos 3\theta - \sin 3\theta}.$$

* * * * *

Problem 11. A curve is defined by the parametric equation

$$x = 120t - 4t^2,\ y = 60t - 6t^2.$$

Find the value of $\mathrm{d}y/\mathrm{d}x$ at each of the points where the curve cross the x-axis.

Solution. The curve crosses the x-axis when $y = 0$:

$$y = 60t - 6t^2 = 6t(10 - t) = 0.$$

Hence, $t = 0$ or $t = 10$. Now, consider the derivative with respect to x of the curve.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{60 - 12t}{120 - 8t}.$$

Case 1: $t = 0$.

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{t=0} = \frac{60 - 12(0)}{120 - 8(0)} = \frac{1}{2}.$$

Case 2: $t = 10$.

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{t=10} = \frac{60 - 12(10)}{120 - 8(10)} = -\frac{3}{2}.$$

* * * * *

Problem 12. A curve has parametric equations $x = 2t - \ln(2t)$, $y = t^2 - \ln t^2$, where $t > 0$. Find the value of t at the point on the curve at which the gradient is 2.

Solution.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{2t - 2/t}{2 - 1/t} = \frac{2t^2 - 2}{2t - 1}.$$

Consider $\mathrm{d}y/\mathrm{d}x = 2$.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2t^2 - 2}{2t - 1} = 2 \implies \frac{t^2 - 1}{2t - 1} = 1 \implies t^2 - 1 = 2t - 1 \implies t^2 - 2t = t(t - 2) = 0.$$

Hence, $t = 0$ or $t = 2$. Since $t > 0$, we reject $t = 0$. Thus, $t = 2$.

Problem 13. If $y = \ln\left(\sin^3 2x\right)$, find $\frac{\mathrm{d}y}{\mathrm{d}x}$ and prove that $3\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 + 36 = 0$.

Solution.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{6\sin^2 2x\cos 2x}{\sin^3 2x} = 6\cot 2x.$$

Hence,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -12\csc^2 2x = -12\left(1+\cot^2 2x\right) = -12 - \frac{1}{3}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2.$$

Thus, we clearly have

$$3\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 + 36 = 0.$$

* * * * *

Problem 14. Given that $y = \mathrm{e}^{\arcsin(2x)}$, show that $\left(1-4x^2\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 4x\frac{\mathrm{d}y}{\mathrm{d}x} = 4y$. Differentiate this result further to obtain a differential equation for $\frac{\mathrm{d}^3y}{\mathrm{d}x^3}$.

Solution. Note that

$$y = \mathrm{e}^{\arcsin(2x)} \implies \ln y = \arcsin(2x)\,.$$

Implicitly differentiating with respect to x,

$$\frac{1}{y}\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2}{\sqrt{1-4x^2}} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2y}{\sqrt{1-4x^2}}.$$

Implicitly differentiating with respect to x once again,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\sqrt{1-4x^2}\left(2\cdot\frac{\mathrm{d}y}{\mathrm{d}x}\right) - 2y\left(\frac{-4x}{\sqrt{1-4x^2}}\right)}{1-4x^2}.$$

Now observe that

$$2\sqrt{1-4x^2}\cdot\frac{\mathrm{d}y}{\mathrm{d}x} + 4x\left(\frac{2y}{\sqrt{1-4x^2}}\right) = 4y + 4x\cdot\frac{\mathrm{d}y}{\mathrm{d}x}.$$

Hence,

$$\left(1-4x^2\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 4y + 4x\cdot\frac{\mathrm{d}y}{\mathrm{d}x} \implies \left(1-4x^2\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 4x\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = 4y.$$

Implicitly differentiating with respect to x once again,

$$\left(1-4x^2\right)\frac{\mathrm{d}^3y}{\mathrm{d}x^3} - 8x\cdot\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 4\left(x\cdot\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + \frac{\mathrm{d}y}{\mathrm{d}x}\right) = 4\cdot\frac{\mathrm{d}y}{\mathrm{d}x}.$$

Rearranging,

$$\left(1-4x^2\right)\frac{\mathrm{d}^3y}{\mathrm{d}x^3} - 12x\cdot\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 8\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

Self-Practice B4

Problem 1. Differentiate each of the following with respect to x, simplifying your answer.

(a) $\frac{x^2-2x}{(x+2)^2}$,

(b) $x\left(x^3+1\right)^{1/3}$,

(c) $\cot x \csc x$,

(d) $\left(\sin^3 x\right)(\sin 3x)$,

(e) $\arctan \sqrt{x}$,

(f) $\arcsin \sqrt{1-x^2}$,

(g) $y = \frac{\mathrm{e}^{2x}}{1+\mathrm{e}^x}$,

(h) $y = \ln \frac{4}{x^2}$,

(i) $y = 3^x$.

Solution.

Part (a). We have

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^2-2x}{(x+2)^2} = \frac{(x+2)^2(2x-2)-\left(x^2-2x\right)[2(x+2)]}{(x+2)^4} = \frac{2(3x-2)}{(x+2)^3}.$$

Part (b). We have

$$\frac{\mathrm{d}}{\mathrm{d}x}x\left(x^3+1\right)^{1/3} = \left(x^3+1\right)^{1/3} + x\left[\frac{1}{3}\left(x^3+1\right)^{-2/3}\cdot 3x^2\right] = \frac{2x^3+1}{(x^3+1)^{2/3}}.$$

Part (c). We have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\cot x\csc x &= \left(-\csc^2 x\right)(\csc x) + (\cot x)(-\csc x\cot x)\\ &= -\csc x\left(\csc^2 x + \cot^2 x\right)\\ &= \frac{1+\cos^2 x}{\sin^3 x}.\end{aligned}$$

Part (d). We have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\sin^3 x\sin 3x &= \left(\sin^3 x\right)(3\cos 3x) + \left(3\sin^2 x\cos x\right)(\sin 3x)\\ &= 3\sin^2 x(\sin x\cos 3x + \cos x\sin 3x)\\ &= 3\sin^2 x\sin 4x.\end{aligned}$$

Part (e). We have

$$\frac{\mathrm{d}}{\mathrm{d}x}\arctan\sqrt{x} = \frac{1}{1+(\sqrt{x})^2}\cdot\frac{1}{2\sqrt{x}} = \frac{1}{2\sqrt{x}(1+x)}.$$

Part (f). We have

$$\frac{\mathrm{d}}{\mathrm{d}x}\arctan\sqrt{1-x^2} = \frac{1}{\sqrt{1-\left(\sqrt{1-x^2}\right)^2}}\cdot\frac{-2x}{2\sqrt{1-x^2}} = -\frac{x}{|x|\sqrt{1-x^2}}.$$

Part (g). We have

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{(1+\mathrm{e}^x)\left(2\mathrm{e}^{2x}\right) - (\mathrm{e}^x)\left(\mathrm{e}^{2x}\right)}{(1+\mathrm{e}^x)^2} = \frac{\mathrm{e}^{2x}(2+\mathrm{e}^x)}{(1+\mathrm{e}^x)^2}.$$

Part (h). Note that $y = \ln\left(4/x^2\right) = \ln 4 - 2\ln x$. Hence,

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

Part (i). Taking logarithms, we get $\ln y = x \ln 3$. Implicitly differentiating with respect to x, we obtain

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \ln 3 \implies \frac{\mathrm{d}y}{\mathrm{d}x} = y\ln 3 = 3^x \ln 3.$$

* * * * *

Problem 2. Find an expression for $\frac{\mathrm{d}y}{\mathrm{d}x}$ in terms of x and y.

(a) $x^3 + y^3 + 3xy - 1 = 0$,

(b) $y^x = x$.

Solution.

Part (a). Implicitly differentiating with respect to x, we obtain

$$3x^2 + 3y^2\frac{\mathrm{d}y}{\mathrm{d}x} + 3x\frac{\mathrm{d}y}{\mathrm{d}x} + 3y = 0 \implies \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{x^2+y}{x+y^2}.$$

Part (b). Taking logarithms, we have $x\ln y = \ln x$. Implicitly differentiating with respect to x,

$$x\left(\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x}\right) + \ln y = \frac{1}{x} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \left(\frac{1}{x} - \ln y\right)\frac{y}{x} = \frac{y(1-\ln x)}{x^2}.$$

* * * * *

Problem 3. It is given that, at any point on the graph of $y = f(x)$, $\frac{\mathrm{d}y}{\mathrm{d}x} = \sqrt{1+y^3}$.

(a) Show that $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{3}{2}y^2$.

(b) Find the expressions for $\frac{\mathrm{d}^3y}{\mathrm{d}x^3}$ and $\frac{\mathrm{d}^4y}{\mathrm{d}x^4}$ in terms of y and $\frac{\mathrm{d}y}{\mathrm{d}x}$.

Solution.

Part (a). Differentiating with respect to x,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{3y^2}{2\sqrt{1+y^3}}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3y^2}{2\sqrt{1+y^3}}\left(\sqrt{1+y^3}\right) = \frac{3y^2}{2}.$$

Part (b). Differentiating once more with respect to x, we see that

$$\frac{\mathrm{d}^3y}{\mathrm{d}x^3} = \frac{3}{2}(2y)\frac{\mathrm{d}y}{\mathrm{d}x} = 3y\frac{\mathrm{d}y}{\mathrm{d}x}.$$

Differentiating again, we have

$$\frac{\mathrm{d}^4y}{\mathrm{d}x^4} = 3y\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 3\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 = 3y\left(\frac{3}{2}y^2\right) + 3\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 = \frac{9y^3}{2} + 3\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2.$$

Problem 4. Given that $y = \mathrm{e}^{\sqrt{1+x}}$, show that

$$4(1+x)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 2\frac{\mathrm{d}y}{\mathrm{d}x} = y.$$

Solution. Note that

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

Differentiating once more, we find that

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\left(2\sqrt{1+x}\right)y' - \left(1/\sqrt{1+x}y\right)}{4\left(1+x\right)}.$$

Rearranging, we get the desired result:

$$4\left(1+x\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 2\sqrt{1+x}\frac{\mathrm{d}y}{\mathrm{d}x} - \frac{y}{\sqrt{1+x}} = y - 2\frac{\mathrm{d}y}{\mathrm{d}x}.$$

* * * * *

Problem 5.

(a) A curve has parametric equations $x = t/(1+t)$, $y = \ln\cos t$, where $t \neq -1, \pi/2$. Find $\mathrm{d}y/\mathrm{d}x$ in terms of t.

(b) A curve has equation $\arcsin y + x\mathrm{e}^y = 3x$. Find $\mathrm{d}y/\mathrm{d}x$ in terms of x and y.

Solution.

Part (a). Note that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{(1+t)-t}{(1+t)^2} = \frac{1}{(1+t)^2} \quad \text{and} \quad \frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{\sin t}{\cos t} = -\tan t.$$

Thus,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{-\tan t}{(1+t)^{-2}} = -\left(1+t\right)^2\tan t.$$

Part (b). Implicitly differentiating with respect to x,

$$\frac{y'}{\sqrt{1-y^2}} + x\mathrm{e}^y y' + \mathrm{e}^y = 3 \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3-\mathrm{e}^y}{(1-y^2)^{-1/2} + x\mathrm{e}^y} = \frac{\left(3-\mathrm{e}^y\right)\sqrt{1-y^2}}{1+x\mathrm{e}^y\sqrt{1-y^2}}.$$

* * * * *

Problem 6.

(a) Differentiate $\frac{x-2x^3}{\ln x}$ with respect to x.

(b) Given that $0 < x < \frac{\pi}{2}$, show that $\frac{\mathrm{d}}{\mathrm{d}x}\arcsin(\cos x) = k$, where k is a real constant to be determined.

(c) Given that $\mathrm{e}^{xy} = \left(1+y^2\right)^2$, find $\frac{\mathrm{d}y}{\mathrm{d}x}$ in terms of x and y, simplifying your answer.

Solution.

Part (a). We have

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x-2x^3}{\ln x}=\frac{\ln x\left(1-6x^2\right)-(1/x)\left(x-2x^3\right)}{\ln^2 x}=\frac{\left(1-6x^2\right)\ln x-1+2x^2}{\ln^2 x}.$$

Part (b). We have

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

Part (c). Taking logarithms, we have

$$xy=2\ln\left(1+y^2\right).$$

Implicitly differentiating with respect to x,

$$xy'+y=2\left(\frac{2y\cdot y'}{1+y^2}\right)\implies\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{y}{\frac{4y}{1+y^2}-x}=\frac{y\left(1+y^2\right)}{4y-x\left(1+y^2\right)}.$$

* * * * *

Problem 7. It is given that $y=\ln\sin\left(\frac{\pi}{4}+x\right)$. Show that

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2+1=0.$$

Solution. It suffices to show that

$$\frac{\mathrm{d}^2y}{\mathrm{d}u^2}+\left(\frac{\mathrm{d}y}{\mathrm{d}u}\right)^2+1=0,$$

where $u=\pi/4+x$. Indeed, we have $y=\ln\sin u$, so

$$\frac{\mathrm{d}y}{\mathrm{d}u}=\frac{\cos u}{\sin u}=\cot u\quad\text{and}\quad\frac{\mathrm{d}^2y}{\mathrm{d}u^2}=-\csc^2 u,$$

so

$$\frac{\mathrm{d}^2y}{\mathrm{d}u^2}+\left(\frac{\mathrm{d}y}{\mathrm{d}u}\right)^2+1=-\csc^2 u+\cot^2 u+1=-\csc^2 u+\csc^2 u=0.$$

* * * * *

Problem 8.

(a) Differentiate the following with respect to x, giving your answers as single fractions.

  (i) $\ln\frac{x}{\sqrt{1-2x}}$,

  (ii) $\frac{1}{\arccos(x^2)}$.

(b) The variables x and y are related by

$$\mathrm{e}^{xy^2}=y\left(x^2+2\mathrm{e}^x\right).$$

Find the value of $\frac{\mathrm{d}y}{\mathrm{d}x}$ when $x=0$ and $y=\frac{1}{2}$.

Solution.

Part (a).

Part (a)(i). We have

$$\frac{\mathrm{d}}{\mathrm{d}x} \ln \frac{x}{\sqrt{1-2x}} = \frac{\sqrt{1-2x}}{x} \cdot \frac{\sqrt{1-2x} - x\left(-\frac{2}{2\sqrt{1-2x}}\right)}{1-2x} = \frac{1-x}{x\,(1-2x)}.$$

Part (a)(ii). We have

$$\frac{\mathrm{d}}{\mathrm{d}x} \frac{1}{\arccos(x^2)} = \left(-\frac{1}{[\arccos(x^2)]^2}\right)\left(-\frac{1}{\sqrt{1-(x^2)^2}}\right)(2x) = \frac{2x}{\sqrt{1-x^4}\,[\arccos(x^2)]^2}.$$

Part (b). Implicitly differentiating with respect to x, we get

$$\mathrm{e}^{xy^2}\left(y^2 + 2xy \cdot y'\right) = y'\left(x^2 + 2\mathrm{e}^x\right) + y\left(2x + 2\mathrm{e}^x\right).$$

At $(0, 1/2)$, we have

$$\frac{1}{4} = 2y' + 1 \implies y' = -\frac{3}{8}.$$

Assignment B4

Problem 1. Differentiate the following with respect to x.

(a) $\ln \frac{x^3}{\sqrt{1+x^2}}$

(b) $\arctan\left(\frac{x^2}{2}\right)$

(c) $e^{2x} \sec x$

Solution.

Part (a).

$$\ln \frac{x^3}{\sqrt{1+x^2}} = 3\ln x - \frac{1}{2}\ln\left(1+x^2\right) \implies \frac{d}{dx}\left(\ln \frac{x^3}{\sqrt{1+x^2}}\right) = \frac{3}{x} - \frac{x}{1+x^2}.$$

Part (b).

$$\frac{d}{dx}\arctan\left(\frac{x^2}{2}\right) = \frac{x}{1+x^4/4} = \frac{4x}{4+x^4}.$$

Part (c).

$$\frac{d}{dx}e^{2x}\sec x = e^{2x}\left(\sec x \tan x\right) + \sec x\left(2e^{2x}\right) = e^{2x}\sec x\left(\tan x + 2\right).$$

* * * * *

Problem 2. Find the gradient of the curve $x^3 + xy^2 = 5y$ at the point where $x = 1$ and $0 < y < 1$, leaving your answer to 3 significant figures.

Solution. Substituting $x = 1$ into the given equation,

$$y^2 - 5y + 1 = 0 \implies y = \frac{5 \pm \sqrt{21}}{2}.$$

Since $0 < y < 1$, we reject $y = \frac{1}{2}\left(5 + \sqrt{21}\right)$ and take $y = \frac{1}{2}\left(5 - \sqrt{21}\right) = 0.20871$ (5 s.f.).

Implicitly differentiating the given equation,

$$3x^2 + 2xy \cdot y' + y^2 = 5y' \implies y' = \frac{3x^2 - y^2}{5 - 2xy}.$$

Substituting $x = 1$ and $y = 0.20871$ into the above equation,

$$y' = \frac{3(1)^2 - (0.20871)^2}{2(1)(0.20871) - 5} = 0.664 \text{ (3 s.f.)}.$$

Hence, the gradient of the curve is 0.664.

* * * * *

Problem 3. A curve C has parametric equations

$$x = \sin^3\theta,\ y = 3\sin^2\theta\cos\theta, \qquad 0 \leq \theta \leq \frac{\pi}{2}.$$

Show that $dy/dx = a\cot\theta + b\tan\theta$, where a and b are values to be determined.

Solution. Note that

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = 3\sin^2\theta\cos\theta, \quad \frac{\mathrm{d}y}{\mathrm{d}\theta} = 3\left(2\sin\theta\cos^2\theta - \sin^3\theta\right).$$

Hence,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = \frac{3\left(2\sin\theta\cos^2\theta - \sin^3\theta\right)}{3\sin^2\theta\cos\theta} = \frac{2\cos\theta}{\sin\theta} - \frac{\sin\theta}{\cos\theta} = 2\cot\theta - \tan\theta.$$

Thus, $a = 2$ and $b = -1$.

B5 Applications of Differentiation

Tutorial B5A

Problem 1. The equation of a curve is $y = 2x^3 + 3x^2 + 6x + 4$. Find $\mathrm{d}y/\mathrm{d}x$ and hence show that y is increasing for all real values of x.

Solution.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 6x^2 + 6x + 6 = 6\left(x + \frac{1}{2}\right)^2 + \frac{18}{4}.$$

For all $x \in \mathbb{R}$, we have $\left(x + \frac{1}{2}\right)^2 \geq 0$. Hence, $\mathrm{d}y/\mathrm{d}x > 0$. Thus, y is increasing for all real values of x.

* * * * *

Problem 2. Find, by differentiation, the x-coordinates of all the stationary points on the curve $y = \frac{x^3}{(x+1)^2}$ stating, with reasons, the nature of each point.

Solution.

$$y = \frac{x^3}{(x+1)^2} \implies (x+1)^2 y = x^3 \implies y'(x+1)^2 + 2y(x+1) = 3x^2.$$

For stationary points, $y' = 0$. Thus,

$$2y(x+1) = \frac{2x^3}{x+1} = 3x^2 \implies 2x^3 - 3x^2(x+1) = x^2(-x-3) = 0.$$

Hence, $x = 0$ or $x = -3$.

x	0^-	0	0^+	$(-3)^-$	-3	$(-3)^+$
$\mathrm{d}y/\mathrm{d}x$	+ve	0	+ve	+ve	0	-ve

By the first derivative test, there is a stationary point of inflexion at $x = 0$ and a maximum point at $x = -3$.

* * * * *

Problem 3. Differentiate $f(x) = 8\sin(x/2) - \sin x - 4x$ with respect to x and deduce that $f(x) < 0$ for $x > 0$.

Solution.

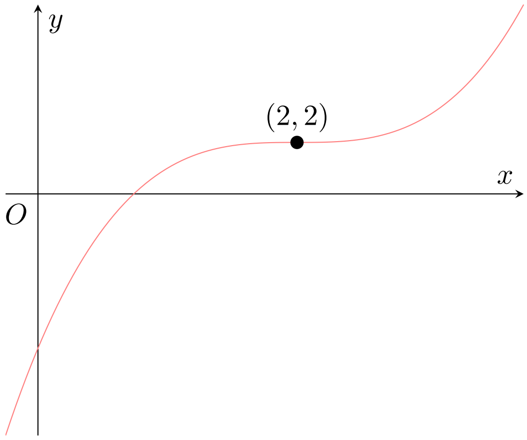
$$f'(x) = 4\cos\frac{x}{2} - \cos x - 4 = 4\cos\frac{x}{2} - \left(2\cos^2\frac{x}{2} - 1\right) - 4 = -2\left(\cos\frac{x}{2} - 1\right)^2 - 1.$$

Observe that for all $x \in \mathbb{R}$, $\left(\cos\frac{x}{2} - 1\right)^2 \geq 0$. Hence, $f'(x) < 0$ for all real values of x. Thus, $f(x)$ is strictly decreasing on $\mathbb{R}$.

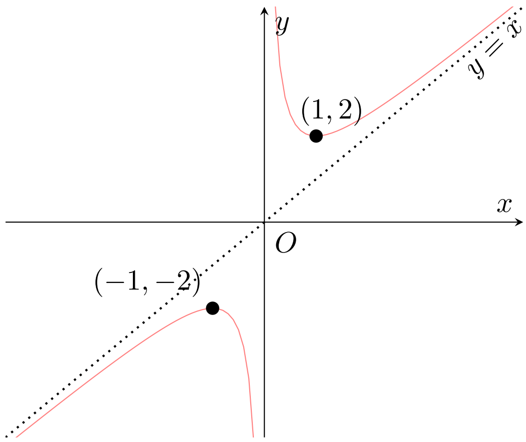
Note that $f(0) = 0$. Since $f(x)$ is strictly decreasing, for all $x > 0$, $f(x) < f(0) = 0$.

* * * * *

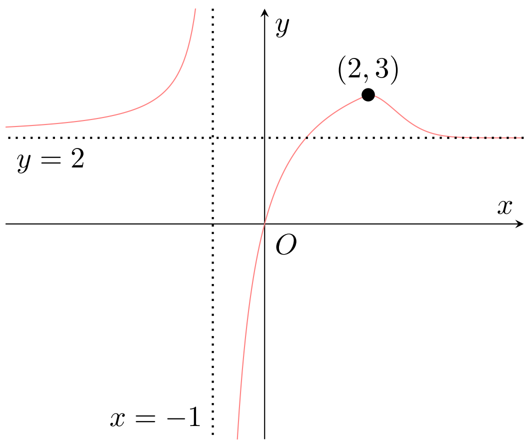
Problem 4. Sketch the graphs of the derivative functions for each of the graphs of the following functions below. In each graph, the point(s) labelled in coordinate form are stationary points.



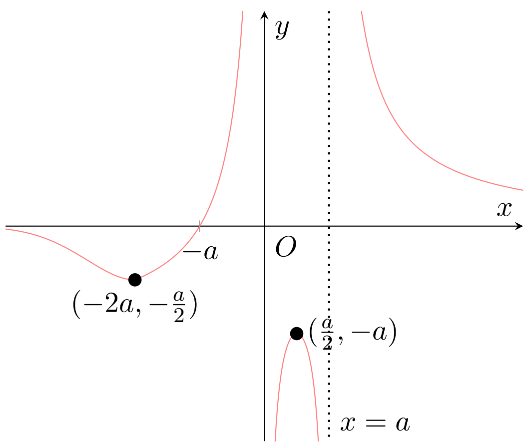
(a)



(b)



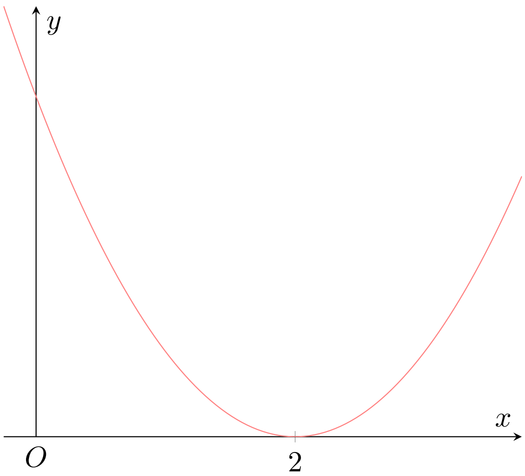
(c)



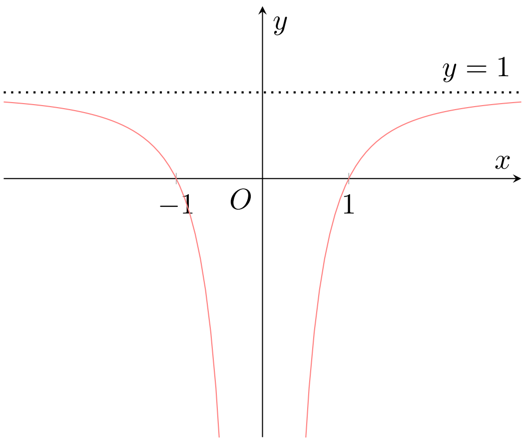
(d)

Solution.

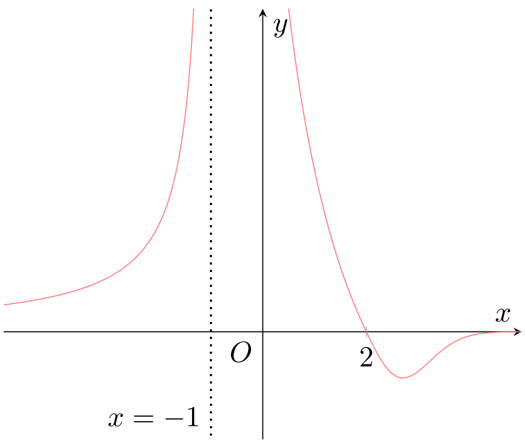
Part (a).



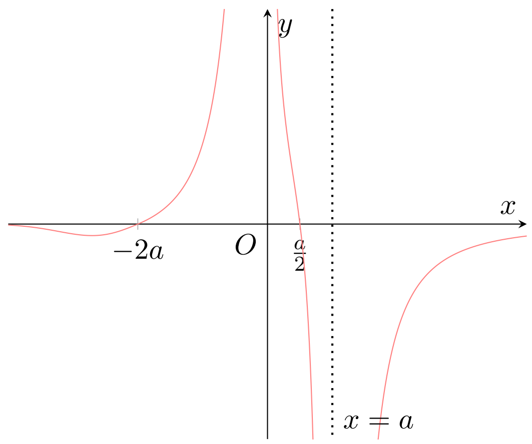
Part (b).



Part (c).



Part (d).



* * * * *

Problem 5.

(a) Given that $y = ax\sqrt{x+2}$ where $a > 0$, find $\mathrm{d}y/\mathrm{d}x$, expressing your answer as a single algebraic fraction. Hence, show that the curve $y = ax\sqrt{x+2}$ has only one turning point, and state its coordinates in exact form.

(b) Sketch the graph of $y = f'(x)$, where $f(x) = ax\sqrt{x+2}$, where $a > 0$.

Solution.

Part (a).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = a\left(\frac{x}{2\sqrt{x+2}} + \frac{2(x+2)}{2\sqrt{x+2}}\right) = \frac{a(3x+4)}{2\sqrt{x+2}}.$$

Consider the stationary points of $y = ax\sqrt{x+2}$. For stationary points, $\mathrm{d}y/\mathrm{d}x = 0$.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{a(3x+4)}{2\sqrt{x+2}} = 0 \implies a(3x+4) = 0.$$

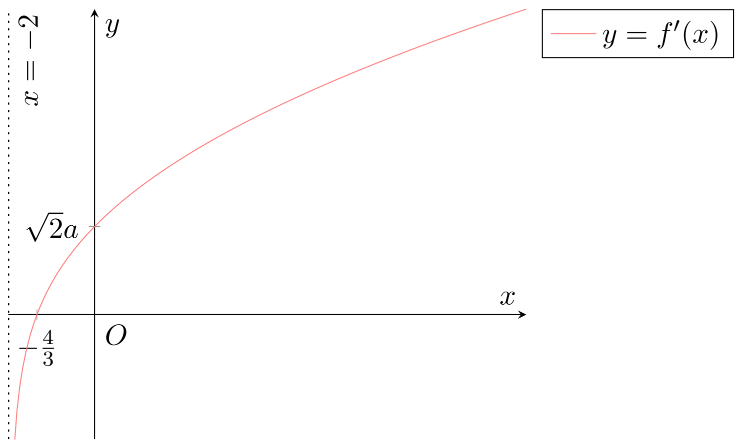
Since $a > 0$, we have $3x + 4 = 0$, whence $x = -4/3$.

x	$(-4/3)^-$	$-4/3$	$(-4/3)^+$
$\mathrm{d}y/\mathrm{d}x$	-ve	0	+ve

Hence, by the first derivative test, there is a turning point (minimum point) at $x = -4/3$. Thus, $y = ax\sqrt{x+2}$ has only one turning point.

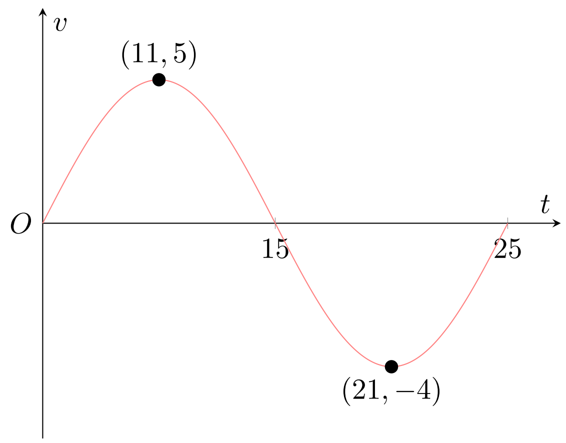
Substituting $x = -4/3$ into $y = ax\sqrt{x+2}$, we see that the coordinate of the turning point is $(-\frac{4}{3}, -\frac{4a}{3}\sqrt{\frac{2}{3}})$.

Part (b).



* * * * *

Problem 6. A particle P moves along the x-axis. Initially, P is at the origin O. At time t s, the velocity is v ms^{-1} and the acceleration is a ms^{-2}. Below is the velocity-time graph of the particle for $0 \leq t \leq 25$.



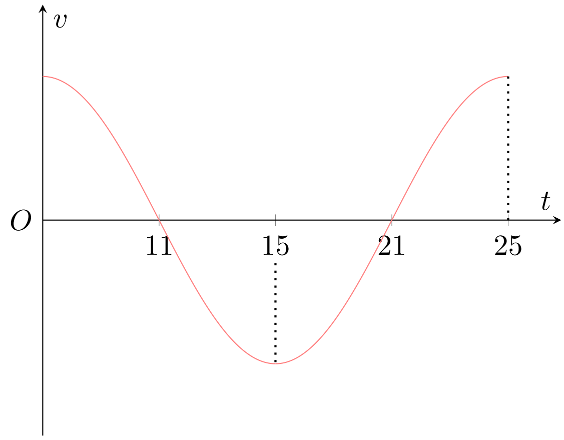
(a) Describe the motion of the particle for $0 \leq t \leq 25$.

(b) Sketch the acceleration-time graph of the particle P.

Solution.

Part (a). From $t = 0$ to $t = 11$, P speeds up and reaches a top speed of 5 ms^{-1}. From $t = 11$ to $t = 15$, P slows down. At $t = 15$, P is instantaneously at rest. From $t = 15$ to $t = 21$, P speeds up and moves in the opposite direction, reaching a top speed of 4 ms^{-1}. From $t = 21$ to $t = 25$, P slows down. At $t = 25$, P is instantaneously at rest.

Part (b).



* * * * *

Problem 7. The function f defined by $f(x) = \ln x - 2(x - 1/2)$, where $x \in \mathbb{R}, x > 0$. Find $f'(x)$ and show that the function is decreasing for $x > 1/2$. Hence, show that for $x > 1/2$, $2(x - 1/2) - \ln x > \ln 2$.

Solution. Observe that $f'(x) = 1/x - 2 < 0$ for $x > 1/2$. Thus, $f(x)$ is decreasing for all $x > 1/2$. Since $f(1/2) = -\ln 2$, it follows that

$$\left(\forall x > \frac{1}{2}\right): \quad -\ln 2 = f(1/2) > f(x) = \ln x - 2\left(x - \frac{1}{2}\right) \implies 2\left(x - \frac{1}{2}\right) - \ln x > \ln 2.$$

Tutorial B5B

Problem 1. The equation of a closed curve is $(x+2y)^2+3(x-y)^2=27$.

(a) Show, by differentiation, that the gradient at the point (x, y) on the curve may be expressed in the form $\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{y-4x}{7y-x}$.

(b) Find the equations of the tangents to the curve that are parallel to

  (i) the x-axis,

  (ii) the y-axis.

Solution.

Part (a). Implicitly differentiating the given equation,

$$(x+2y)\left(1+2y'\right)+3(x-y)(1-y')=(-x+7y)y'+4x-y=0 \implies y'=\frac{y-4x}{7y-x}.$$

Part (b).

Part (b)(i). When the tangent to the curve is parallel to the x-axis, $y'=0$, whence $y=4x$. Substituting $y=4x$ into the given equation,

$$(9x)^2+3(-3x)^2=27 \implies 108x^2=27 \implies x^2=\frac{1}{4} \implies x=\pm\frac{1}{2} \implies y=\pm 2.$$

The equations of the tangents are hence $y=\pm 2$.

Part (b)(ii). When the tangent to the curve is parallel to the y-axis, y' is undefined. Hence, $7y-x=0 \implies x=7y$. Substituting $x=7y$ into the given equation,

$$(9y)^2+3(6y)^2=27 \implies 189y^2=27 \implies y^2=\frac{1}{7} \implies y=\pm\frac{1}{\sqrt{7}} \implies x=\pm\sqrt{7}.$$

The equations of the tangents are hence $x=\pm\sqrt{7}$.

* * * * *

Problem 2. A piece of wire of length 8 cm is cut into two pieces, one of length x cm, the other of length $(8-x)$ cm. The piece of length x cm is bent to form a circle with circumference x cm. The other piece is bent to form a square with perimeter $(8-x)$ cm. Show that, as x varies, the sum of the areas enclosed by these two pieces of wire is a minimum when the radius of the circle is $\frac{4}{4+\pi}$ cm.

Solution. Let the radius of the circle be r cm. Then we have $x=2\pi r \implies r=x/2\pi$. Let the side length of the square be s cm. Then we have $8-x=4s \implies s=2-x/4$. Let the total area enclosed by the circle and the square be $A(x)$.

$$A(x)=\pi r^2+s^2=\pi\left(\frac{x}{2\pi}\right)^2+\left(2-\frac{x}{4}\right)^2=\left(\frac{1}{4\pi}+\frac{1}{16}\right)x^2-x+4.$$

Consider the stationary points of $A(x)$. For stationary points, $A'(x)=0$.

$$A'(x)=\left(\frac{1}{2\pi}+\frac{1}{8}\right)x-1=0 \implies x=\frac{1}{\frac{1}{2\pi}+\frac{1}{8}}=\frac{8\pi}{4+\pi}.$$

x	$\left(\frac{8\pi}{4+\pi}\right)^-$	$\frac{8\pi}{4+\pi}$	$\left(\frac{8\pi}{4+\pi}\right)^+$
$\mathrm{d}A/\mathrm{d}x$	$-$ve	0	$+$ve

Hence, by the first derivative test, the minimum value of $A(x)$ is achieved when $x = \frac{8\pi}{4+\pi}$, whence

$$r = \frac{1}{2\pi} \cdot \frac{8\pi}{4+\pi} = \frac{4}{4+\pi} \text{ cm.}$$

* * * * *

Problem 3. A spherical balloon is being inflated in such a way that its volume is increasing at a constant rate of 150 cm^3s^{-1}. At time t seconds, the radius of the balloon is r cm.

(a) Find $\mathrm{d}r/\mathrm{d}t$ when $r = 50$.

(b) Find the rate of increase of the surface area of the balloon when its radius is 50 cm.

Solution. Let the volume of the balloon be $V(r) = \frac{4}{3}\pi r^3$ cm^3.
Part (a). Note that $\frac{\mathrm{d}V}{\mathrm{d}t} = 150$ and $\frac{\mathrm{d}V}{\mathrm{d}r} = 4\pi r^2$.

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{\mathrm{d}r/\mathrm{d}V}{\mathrm{d}t/\mathrm{d}V} = \frac{\mathrm{d}V/\mathrm{d}t}{\mathrm{d}V/\mathrm{d}r} = \frac{150}{4\pi r^2} = \frac{75}{2\pi r^2}.$$

Evaluating $\frac{\mathrm{d}r}{\mathrm{d}t}$ at $r = 50$,

$$\left.\frac{\mathrm{d}r}{\mathrm{d}t}\right|_{r=50} = \frac{75}{2\pi \cdot 50^2} = \frac{3}{200\pi}.$$

Part (b). Let the surface area of the balloon be $A(r) = 4\pi r^2$. Observe that $\frac{\mathrm{d}A}{\mathrm{d}r} = 8\pi r$.

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\mathrm{d}A}{\mathrm{d}r} \cdot \frac{\mathrm{d}r}{\mathrm{d}t} \implies \left.\frac{\mathrm{d}A}{\mathrm{d}t}\right|_{r=50} = (8\pi \cdot 50)\left(\frac{3}{200\pi}\right) = 6.$$

Thus, the rate of increase of the surface area of the balloon when its radius is 50 cm is 6 cm/s.

* * * * *

Problem 4. A curve has parametric equations $x = 5\sec\theta$, $y = 3\tan\theta$, where $-\frac{1}{2}\pi < \theta < \frac{1}{2}\pi$. Find the exact coordinates of the point on the curve at which the normal is parallel to the line $y = x$.

Solution. Observe that $x^2 = 25\sec^2\theta$ and $\frac{25}{9}y^2 = 25\tan^2\theta$. Using the identity $\tan^2\theta + 1 = \sec^2\theta$, we get

$$\frac{25}{9}y^2 + 25 = x^2. \tag{$*$}$$

Implicitly differentiating with respect to x, we get

$$\frac{25}{9}y \cdot y' = x.$$

Since the normal is parallel to $y = x$, the tangent is parallel to $y = -x$, whence $y' = -1$. Thus,

$$y = -\frac{9}{25}x.$$

Substituting $y = -\frac{9}{25}x$ into $(*)$,

$$\frac{25}{9}\left(-\frac{9}{25}x\right)^2 + 25 = x^2 \implies \frac{16}{25}x^2 = 25 \implies \frac{4}{5}x = \pm 5 \implies x = \pm\frac{25}{4}.$$

Observe that for $-\pi/2 < \theta < \pi/2$, $x = 5\sec\theta \geq 5$. We thus take $x = 25/4$, whence $y = -9/4$. The coordinate of the required point is thus $(25/4, -9/4)$.

Problem 5. The parametric equations of a curve are

$$x = t^2,\ y = \frac{2}{t}.$$

(a) Find the equation of the tangent to the curve at the point $(p^2, 2/p)$, simplifying your answer.

(b) Hence, find the coordinates of the points Q and R where this tangent meets the x- and y-axes respectively.

(c) The point F is the mid-point of QR. Find a Cartesian equation of the curve traced by F as p varies.

Solution.

Part (a). Observe that $\mathrm{d}x/\mathrm{d}t = 2t$ and $\mathrm{d}y/\mathrm{d}t = -2/t^2$. Hence,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{-2/t^2}{2t} = -\frac{1}{t^3}.$$

Using the point-slope formula, the tangent to the curve at $(p^2, 2/p)$ is given by the equation

$$y - \frac{2}{p} = -\frac{1}{p^3}\left(x - p^2\right) \implies y = \frac{3}{p} - \frac{1}{p^3}x.$$

Part (b). Consider the case where $y = 0$:

$$0 = \frac{3}{p} - \frac{1}{p^3}x \implies x = 3p^2 \implies Q\left(3p^2, 0\right).$$

Consider the case where $x = 0$:

$$y = \frac{3}{p} \implies R\left(0, \frac{3}{p}\right).$$

Part (c). Note that

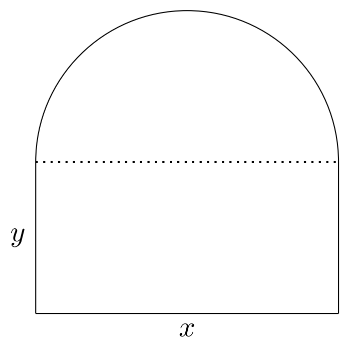
$$F = \left(\frac{3}{2}p^2, \frac{3}{2p}\right).$$

As p varies, F traces a curve given by the parametric equations $x = 3p^2/2$, $y = 3/2p$. Hence,

$$p^2 = \frac{2}{3}x = \frac{9}{4y^2} \implies y^2 = \frac{27}{8x}.$$

* * * * *

Problem 6.



A new flower-bed is being designed for a large garden. The flower-bed will occupy a rectangle x m by y m together with a semicircle of diameter x m, as shown in the diagram. A low wall will be built around the flowerbed. The time needed to build the wall will be 3 hours per metre for the straight parts and 9 hours per metre for the semicircular part. Given that a total time of 180 hours is taken to build the wall, find, using differentiation, the values of x and y which give a flower-bed of maximum area.

Solution. Observe that the length of the straight parts is $(2y + x)$ m, while the length of the semicircular part is $\frac{1}{2}\pi x$ m. Since a total time of 180 hours is taken to build the wall,

$$3(2y + x) + 9\left(\frac{1}{2}\pi x\right) = 180 \implies 4y + 2x + 3\pi x = 120 \implies x = \frac{120 - 4y}{2 + 3\pi}.$$

Differentiating with respect to y, we get $x' = -4/(2 + 3\pi)$. Let $A(y)$ be the total area enclosed by the garden, in m^2. Observe that

$$A(y) = xy + \frac{1}{2}\pi\left(\frac{x}{2}\right)^2 = xy + \frac{\pi}{8}x^2.$$

Consider the stationary points of $A(y)$. For stationary points, $A'(y) = 0$.

$$A'(y) = \big(x'y + x\big) + \frac{\pi}{4}x \cdot x' = 0.$$

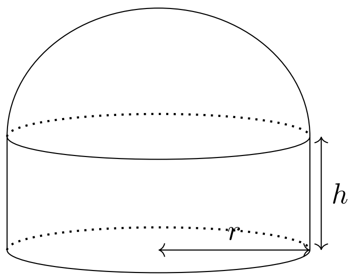
Substituting in our values of x and x', we get

$$\left[y\left(-\frac{4}{2+3\pi}\right) + \frac{120 - 4y}{2 + 3\pi}\right] + \left[\frac{\pi}{4}\left(\frac{120 - 4y}{2 + 3\pi}\right)\left(-\frac{4}{2+3\pi}\right)\right] = 0.$$

Using G.C., we get $y = 12.6$ (3 s.f.), whence $x = 6.09$ (3 s.f.).

* * * * *

Problem 7.



A model of a concert hall is made up of three parts.

- The roof is modelled by the curved surface of a hemisphere of radius r cm.
- The walls are modelled by the curved surface of a cylinder of radius r cm and height h cm.
- The floor is modelled by a circular disc of radius r cm.

The three parts are joined together as shown in the diagram. The model is made of material of negligible thickness.

(a) It is given that the volume of the model is a fixed value k cm^3, and the external surface area is a minimum. Use differentiation to find the values of r and h in terms of k. Simplify your answers.

(b) It is given instead that the volume of the model is 200 cm^3 and its external surface area is 180 cm^2. Show that there are two possible values of r. Given also that $r < h$, find the value of r and the value of h.

Solution.

Part (a). Let the volume of the model be V cm^3. Then

$$V = \frac{1}{2}\left(\frac{4}{3}\pi r^3\right) + \pi r^2 h = k \implies h = \frac{k}{\pi r^2} - \frac{2}{3}r. \tag{1}$$

Let the external surface area of the model be A cm^2. Then

$$A = \frac{4\pi r^2}{2} + 2\pi r h + \pi r^2 = 3\pi r^2 + 2\pi r\left(\frac{k}{\pi r^2} - \frac{2}{3}r\right) = \frac{5\pi}{3}r^2 + \frac{2k}{r}. \tag{2}$$

Consider the stationary points of A. For stationary points, $\mathrm{d}A/\mathrm{d}r = 0$.

$$\frac{\mathrm{d}A}{\mathrm{d}r} = \frac{10\pi}{3}r - \frac{2k}{r^2} = 0 \implies r^3 = \frac{3k}{5\pi} \implies r = \sqrt[3]{\frac{3k}{5\pi}}.$$

r	$\sqrt[3]{\frac{3k}{5\pi}}^-$	$\sqrt[3]{\frac{3k}{5\pi}}$	$\sqrt[3]{\frac{3k}{5\pi}}^+$
$\mathrm{d}A/\mathrm{d}r$	$-$ve	0	$+$ve

Hence, by the first derivative test, A is at a minimum when $r = \sqrt[3]{\frac{3k}{5\pi}}$.
Substituting $r^3 = \frac{3k}{5\pi}$ into (1),

$$\frac{2}{3}\pi\left(\frac{3k}{5\pi}\right) + \pi r^2 h = \frac{2}{5}k + \pi r^2 h = k \implies r^2 h = \frac{3k}{5\pi} = r^3 \implies h = r = \sqrt[3]{\frac{3k}{5\pi}}.$$

Part (b). From (2), we have

$$\frac{5\pi}{3}r^2 + \frac{2(200)}{r} = 180 \implies \pi r^3 - 108r + 240 = 0.$$

Let $f(r) = \pi r^3 - 108r + 240$. Consider the stationary points of $f(r)$. For stationary points, $f'(r) = 0$.

$$f'(r) = 3\pi r^2 - 108 = 0 \implies r^2 = \frac{36}{\pi} \implies r = \pm\frac{6}{\sqrt{\pi}}.$$

Since $f(r)$ is a cubic with two turning points, it follows that there is exactly one root in each of the following three intervals:

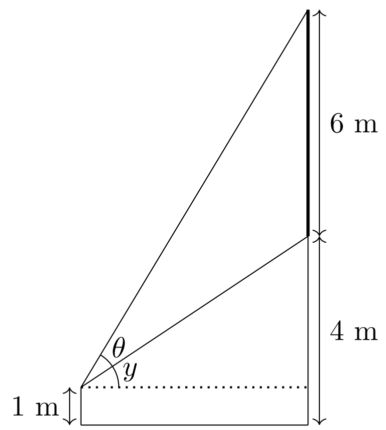
$$\left(-\infty, -\frac{6}{\sqrt{\pi}}\right), \quad \left(-\frac{6}{\sqrt{\pi}}, \frac{6}{\sqrt{\pi}}\right), \quad \left(\frac{6}{\sqrt{\pi}}, \infty\right).$$

We now show that the root in the interval $\left(-\frac{6}{\sqrt{\pi}}, \frac{6}{\sqrt{\pi}}\right)$ is positive. Since $f(r)$ has a positive leading coefficient, it must be decreasing in the interval $\left(-\frac{6}{\sqrt{\pi}}, \frac{6}{\sqrt{\pi}}\right)$. Since $f(0) = 240 > 0$, the root in said interval must be positive. Hence, $f(r) = 0$ has two positive roots. Using G.C., the roots are $r = 3.04$ and $r = 3.72$. From (1), we know that

$$h = \frac{200}{\pi r^2} - \frac{2}{3}r.$$

When $r = 3.04$, $h = 4.88 > r$. When $r = 3.72$, $h = 2.12 < r$. Thus, given that $r < h$, we have $r = 3.04$ and $h = 4.88$.

Problem 8.



A movie screen on a vertical wall is 6 m high and 4 m above the horizontal floor. A boy who is standing at x m away from the wall has eye level at 1 m above the floor as shown in the diagram.

The viewing angle of the boy at that position is θ and the angle of elevation of the bottom of the screen is y.

(a) Express y in terms of x.

(b) By expressing θ in terms of x or otherwise, find the stationary value of θ, giving your answers in exact form. Determine if the value is a maximum or minimum value, showing your working clearly.

Solution.

Part (a). Observe that $\tan y = 3/x$, whence $y = \arctan(3/x)$.

Part (b). Observe that $\tan(y+\theta) = 9/x$. Hence,

$$\tan(y+\theta) = \frac{\tan y + \tan\theta}{1 - \tan y \tan\theta} = \frac{3/x + \tan\theta}{1 - (3/x)\tan\theta} = \frac{3 + x\tan\theta}{x - 3\tan\theta} = \frac{9}{x} \implies \tan\theta = \frac{6x}{x^2+27}.$$

Hence,

$$\theta = \arctan\left(\frac{6x}{x^2+27}\right).$$

Differentiating with respect to x,

$$\frac{\mathrm{d}\theta}{\mathrm{d}x} = \frac{1}{1 + \left(\frac{6x}{x^2+27}\right)^2}\left[\frac{6\left(x^2+27\right) - 6x(2x)}{\left(x^2+27\right)^2}\right] = \frac{-6x^2+162}{36x^2 + \left(x^2+27\right)^2}.$$

For stationary points, $\mathrm{d}\theta/\mathrm{d}x = 0$. Hence,

$$-6x^2 + 162 = 0 \implies x^2 = 27 \implies x = \pm 3\sqrt{3}.$$

Since $x > 0$, we only take $x = 3\sqrt{3}$. Thus,

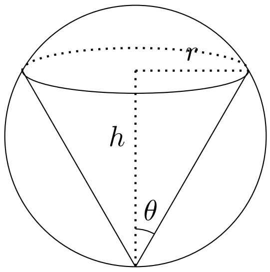
$$\theta = \arctan\left(\frac{6\left(3\sqrt{3}\right)}{27+27}\right) = \arctan\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6}.$$

x	$3\sqrt{3}^-$	$3\sqrt{3}$	$3\sqrt{3}^+$
$\mathrm{d}\theta/\mathrm{d}x$	+ve	0	−ve

Thus, by the first derivative test, $\theta = \frac{\pi}{6}$ is a maximum value.

* * * * *

Problem 9.



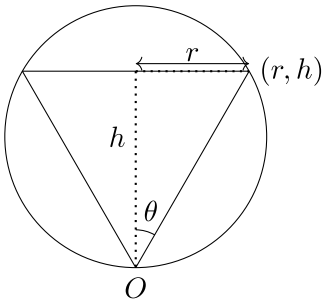
The diagram shows a right inverted cone of radius r, height h and semi-vertical angle θ, which is inscribed in a sphere of radius 1 unit.

Prove that $r^2 = 2h - h^2$.

(a) As r and h varies, determine the exact maximum volume of the cone.

(b) Show that $h = 2\cos^2\theta$. The volume of the cone is increasing at a rate of 6 unit^3/s when $h = \frac{3}{2}$. Determine the rate of change of θ at this instant, leaving your answer in an exact form.

Solution. Consider the following diagram of the cone and sphere.



Let the origin be the tip of the cone. Since the sphere has radius 1 unit, the circle is given by the equation $x^2 + (y-1)^2 = 1$. Since the point (r, h) lies on the circle,

$$r^2 + (h-1)^2 = 1 \implies r^2 = 2h - h^2. \tag{$*$}$$

Part (a). Implicitly differentiating $(*)$ with respect to r,

$$2r = 2 \cdot \frac{\mathrm{d}h}{\mathrm{d}r} - 2h \cdot \frac{\mathrm{d}h}{\mathrm{d}r} \implies \frac{\mathrm{d}h}{\mathrm{d}r} = \frac{r}{1-h}.$$

Let the volume of the cone be $V(r)$ units^3. Then

$$V(r) = \frac{1}{3}\pi r^2 h = \frac{1}{3}\pi\left(2h - h^2\right)h = \frac{1}{3}\pi\left(2h^2 - h^3\right).$$

Differentiating $V(r)$ with respect to r,

$$V'(r) = \frac{1}{3}\pi\left(4h \cdot \frac{\mathrm{d}h}{\mathrm{d}r} - 3h^2 \cdot \frac{\mathrm{d}h}{\mathrm{d}r}\right) = \frac{1}{3}\left(\frac{\pi r h}{1-h}\right)(4-3h).$$

Consider the stationary values of $V(r)$. For stationary values, $V'(r) = 0$, whence $h = 4/3$. Substituting this into $(*)$, we obtain

$$r^2 = 2\left(\frac{4}{3}\right) - \left(\frac{4}{3}\right)^2 = \frac{8}{9} \implies r = \sqrt{\frac{8}{9}}.$$

Note that we reject $r = -\sqrt{8/9}$ as $r > 0$.

r	$\sqrt{8/9}^-$	$\sqrt{8/9}$	$\sqrt{8/9}^+$
$V'(r)$	+ve	0	−ve

Hence, the maximum volume is achieved when $r = \sqrt{8/9}$. Note that

$$V\left(\sqrt{\frac{8}{9}}\right) = \frac{1}{3}\pi\left(\frac{8}{9}\right)\left(\frac{4}{3}\right) = \frac{32}{81}\pi.$$

The maximum volume of the cone is hence $32\pi/81$ units^3.

Part (b). From the diagram, we have

$$\cos\theta = \frac{h}{\sqrt{r^2+h^2}} \implies 2\cos^2\theta = \frac{2h^2}{r^2+h^2} = \frac{2h^2}{2h-h^2+h^2} = h.$$

Observe that

$$V = \frac{\pi}{3}\left(2h^2 - h^3\right) = \frac{\pi}{3}\left(8\cos^4\theta - 8\cos^6\theta\right) = \frac{8\pi}{3}\left(\cos^4\theta - \cos^6\theta\right).$$

Differentiating with respect to θ, we get

$$\frac{\mathrm{d}V}{\mathrm{d}\theta} = \frac{8\pi}{3}\left(-4\cos^3\theta\sin\theta + 6\cos^5\theta\sin\theta\right) = \frac{16\pi}{3}\cos^3\theta\sin\theta\left(-2+3\cos^2\theta\right).$$

Since $2\cos^2\theta = h = 3/2$, we clearly have $\theta = \pi/6$. Thus,

$$\left.\frac{\mathrm{d}V}{\mathrm{d}\theta}\right|_{h=3/2} = \frac{16\pi}{3}\cos^3\frac{\pi}{6}\sin\frac{\pi}{6}\left(-2+3\cos^2\frac{\pi}{6}\right) = \frac{\sqrt{3}\pi}{4}.$$

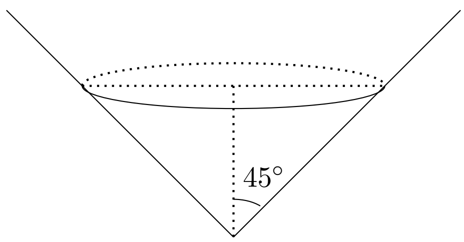
Hence,

$$\left.\frac{\mathrm{d}\theta}{\mathrm{d}t}\right|_{h=3/2} = \left.\left(\frac{\mathrm{d}\theta}{\mathrm{d}V}\cdot\frac{\mathrm{d}V}{\mathrm{d}t}\right)\right|_{h=3/2} = \frac{6}{\sqrt{3}\pi/4} = \frac{8\sqrt{3}}{\pi}.$$

θ is thus increasing at a rate of $8\sqrt{3}/\pi$ radians per second when $h = \frac{3}{2}$.

* * * * *

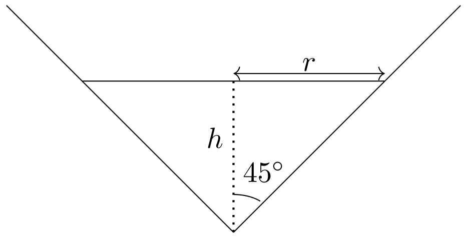
Problem 10.



A hollow cone of semi-vertical angle 45° is held with its axis vertical and vertex downwards. At the beginning of an experiment, it is filled with 390 cm^3 of liquid. The liquid runs out through a small hole at the vertex at a constant rate of 2 cm^3/s.

Find the rate at which the depth of the liquid is decreasing 3 minutes after the start of the experiment.

Solution. Consider the following diagram.



Let the volume of liquid be $V = \frac{1}{3}\pi r^2 h$ cm^3. From the diagram, we have $r = h$. Thus,

$$V = \frac{1}{3}\pi h^3.$$

Differentiating V with respect to h,

$$\frac{\mathrm{d}V}{\mathrm{d}h} = \frac{1}{3}\pi \cdot 3h^2 = \pi h^2.$$

Let t be the time since the start of the experiment in seconds. Consider $\mathrm{d}h/\mathrm{d}t$.

$$\frac{\mathrm{d}h}{\mathrm{d}t} = \frac{\mathrm{d}h}{\mathrm{d}V} \cdot \frac{\mathrm{d}V}{\mathrm{d}t} = \left(\frac{\mathrm{d}h}{\mathrm{d}V}\right)^{-1} \frac{\mathrm{d}V}{\mathrm{d}t} = \frac{-2}{\pi h^2}.$$

When $t = 180$, there is $390 - 180(2) = 30$ cm^3 of liquid left in the cone. Thus,

$$V = \frac{1}{3}\pi h^3 = 30 \implies h^3 = \frac{90}{\pi} \implies h = \sqrt[3]{\frac{90}{\pi}}.$$

Evaluating $\mathrm{d}h/\mathrm{d}t$ at $t = 180$,

$$\left.\frac{\mathrm{d}h}{\mathrm{d}t}\right|_{t=180} = \frac{-2}{\pi\left(\sqrt[3]{\frac{90}{\pi}}\right)^2} = -0.0680 \text{ (3 s.f.)}.$$

Thus, the depth of the liquid is decreasing at a rate of 0.0680 cm/s 3 minutes after the start of the experiment.

* * * * *

Problem 11. A particle is projected from the origin O, and it moves freely under gravity in the x-y plane. At time t s after projection, the particle is at the point (x, y) where $x = 30t$ and $y = 20t - 5t^2$, with x and y measured in metres.

(a) Given that the particle passes through two points A and B which are at a distance 15 m above the x-axis, find the time taken for the particle to travel from A to B. Find also the distance AB.

(b) It is known that the particle always travels in a direction tangential to its path. Show that, when $x = 10$, the particle is travelling at an angle of $\arctan(5/9)$ above the horizontal.

The speed of the particle is given by $\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}$. Find the speed of the particle when $x = 10$.

(c) Show that the equation of trajectory is $y = \frac{2}{3}x - \frac{1}{180}x^2$.

Solution.

Part (a). Consider $y = 15$.

$$y = 20t - 5t^2 = 15 \implies t^2 - 4t + 3 = (t-1)(t-3) = 0.$$

Hence, $t = 1$ or $t = 3$. Thus, the particle takes $3 - 1 = 2$ seconds to travel from A to B.

Note that $x = 30$ when $t = 1$, and $x = 90$ when $t = 3$. Hence, $A(30, 15)$ and $B(90, 15)$, whence $AB = 60$ m.

Part (b). Note that $\mathrm{d}x/\mathrm{d}t = 30$ and $\mathrm{d}y/\mathrm{d}t = 20 - 10t$. Thus,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{20 - 10t}{30} = \frac{2-t}{3}.$$

When $x = 10$, $t = 1/3$. Evaluating $\frac{\mathrm{d}y}{\mathrm{d}x}$ at $t = 1/3$,

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{t=\frac{1}{3}} = \frac{2 - 1/3}{3} = \frac{5}{9}.$$

Hence, the line tangent to the curve at $x = 10$ has gradient $5/9$. Thus, the particle is travelling at an angle of $\arctan(5/9)$ above the horizontal when $x = 10$.

Note that

$$\left.\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\right|_{t=\frac{1}{3}} = \sqrt{30^2 + \left(20 - \frac{10}{3}\right)^2} = 34.3 \text{ (3 s.f.)}.$$

Hence, the particle is travelling at a speed of 34.3 m/s when $x = 10$.

Part (c). Note that $t = x/30$. Hence,

$$y = 20t - 5t^2 = 20\left(\frac{x}{30}\right) - 5\left(\frac{x}{30}\right)^2 = \frac{2}{3}x - \frac{1}{180}x^2.$$

Self-Practice B5A

Problem 1. It is given that $f(x) = \frac{x^2-2x}{e^x}$.

Find the range of values of x for which the curve $y = f(x)$ is concave upward. Hence, sketch the graph of $y = f(x)$, indicating clearly the equations of any asymptotes and the coordinates of any stationary points and any intersections with the axes.

* * * * *

Problem 2. It is given that x and y satisfy the equation

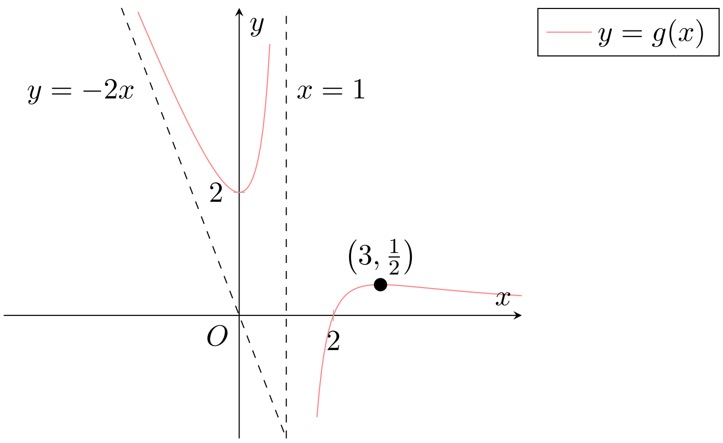
$$y^4 - \ln\frac{y^2}{4} = x^4 - 6x^2, \quad y > 0.$$

(a) Show that $\frac{dy}{dx} = \frac{2xy\left(x^2-3\right)}{2y^4-1}$.

(b) Hence, obtain the possible exact value(s) of $\frac{dy}{dx}$ when $y = 2$.

* * * * *

Problem 3. The diagram below shows the graph of $y = g(x)$. The graph has a minimum point at $(0, 2)$ and a maximum point at $(3, \frac{1}{2})$. The equations of the asymptotes are $x = 1$, $y = 0$ and $y = -2x$.



(a) State the interval(s) on which g is

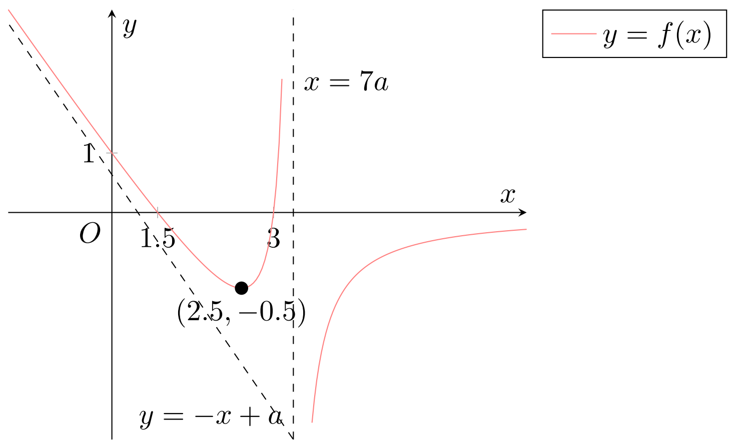
  (i) increasing;

  (ii) increasing and concave upward.

(b) Sketch $y = g'(x)$, showing clearly the equations of the asymptotes and the coordinates of the turning points and axial intercepts, where applicable.

* * * * *

Problem 4. The diagram below shows the graph of $y = f(x)$. It cuts the axes at the points $(0, 1)$, $(1.5, 0)$ and $(3, 0)$. It has a minimum point at $(2.5, -0.5)$. The horizontal, vertical and oblique asymptotes are $y = 0$, $x = 7a$ and $y = -x + a$ respectively, where a is a positive constant.



On separate diagrams, sketch the graphs of

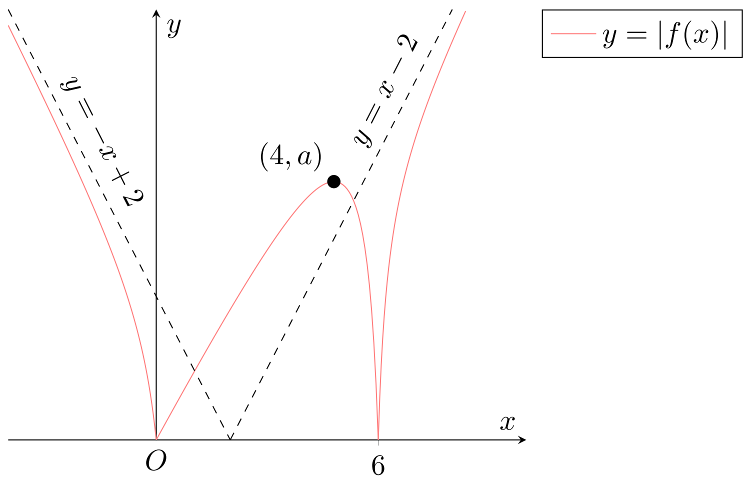
(a) $y = \frac{1}{f(x)}$,

(b) $y = f'(x)$,

showing clearly the axial intercepts, the stationary points and the equations of the asymptotes where applicable.

* * * * *

Problem 5 (🌶). The graph of $y = |f(x)|$ is shown in the diagram, with a maximum point $(4, a)$, and $x = 0$ and $x = 6$ are tangents to both graphs.



It is given that the graph of the continuous function f has **only** one oblique asymptote, and that $f'(1) > 0$ and $f'(7) < 0$.

Sketch the graph of $y = f'(x)$, showing clearly the stationary point(s), the asymptote(s) and the intercept(s), if any.

Self-Practice B5B

Problem 1. Find the coordinates of the points on the curve $3x^2 + xy + y^2 = 33$ at which the tangent is parallel to the x-axis.

$* * * * *$

Problem 2. Given the equation $x^{1/2} + y^{1/2} = k^{1/2}$, where k is a constant,

(a) show that the equation of the tangent at the point $(p,\ q)$ is given by

$$y = -\sqrt{\frac{q}{p}}x + q + \sqrt{pq}.$$

(b) Hence or otherwise, prove that the sum of the x and y-intercepts of any tangent line to the curve $x^{1/2} + y^{1/2} = k^{1/2}$ is constant and equal to k.

$* * * * *$

Problem 3. A curve C is defined by the parametric equations $x = t^2(t+1)$, $y = 4t - 5$, $t \geq 0$.

(a) Find the equation of the tangent to the curve C at the point where $y = -5$.

(b) Find the equation of the normal to the curve C when $t = 2$ and hence show that this normal does not intersect the curve C again.

$* * * * *$

Problem 4. A curve C has parametric equations

$$x = t + \frac{1}{t}, \quad y = t - \frac{1}{t}.$$

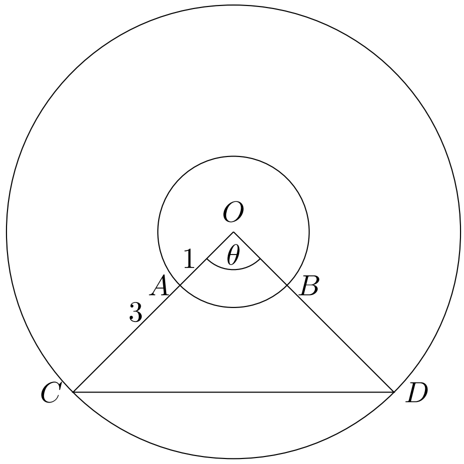
(a) The point P on the curve has parameter p. Show that the equation of the tangent at P is $\left(p^2+1\right)x - \left(p^2-1\right)y = 4p$.

(b) The tangent at P meets the line $y = x$ at the point A and the line $y = -x$ at the point B. Show that the area of triangle OAB is independent of p, where O is the origin.

(c) Find a Cartesian equation of C. Sketch C, giving the coordinates of any points where C crosses the x- and y-axes and the equations of any asymptotics.

$* * * * *$

Problem 5. The diagram shows two circles, of radii 1 and 3, each with centre O. The angle between the lines OAC and OBD is θ radians. The region R is bounded by the minor arc AB and the lines AC, CD and DB.



(a) Find the area of R.

(b) Find the value of θ for which the area of R is greatest.

(c) Find the greatest value of θ which ensures that the whole line segment CD lies between the two circles.

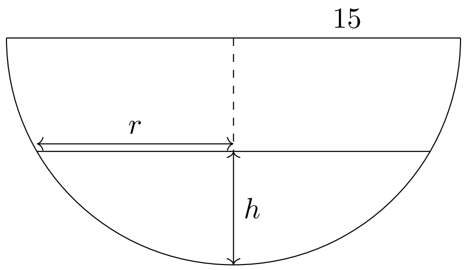
* * * * *

Problem 6. A company manufactures closed hollow cylindrical cans of cross-sectional radius r cm^2 and height h cm. A can is made of two different materials. Its top and base cost 0.09 cents per cm^2 and its curved surface costs 0.06 cents per cm^2 to manufacture.

Show that the radius of the cheapest can of volume 300 cm^3 is $\sqrt[3]{a/\pi}$, where a is a constant to be determined.

* * * * *

Problem 7. A hemispherical goldfish tank with radius 15 cm (as shown in the figure above) was initially filled with water. The tank has a defect and water is leaking at a constant rate of 20 cm^3 per min. The volume of water in the tank is given by $V = \frac{\pi}{3}\left(45h^2 - h^3\right)$ where h is the depth of water at the centre of the tank in cm. Show that r, the radius of the water surface in cm, is given by $r = \sqrt{30h - h^2}$.



Given that the minimum depth of water needed for the goldfish to survive is 5 cm, find, at this instant,

(a) the rate of change of the depth of water, and

(b) the rate of decrease of the radius of the water surface.

* * * * *

Problem 8. A circular cylinder is expanding in such a way that, at time t seconds, the height of the cylinder is y cm and the area of the cross-section is $\frac{1}{3}y^2$ cm^2. At the instant when $y = 3$ cm, the height is increasing at a rate of 0.5 cm/s. Find the rate of increase, at this instant, of:

(a) the area of the cross-section of the cylinder,

(b) the volume of the cylinder.

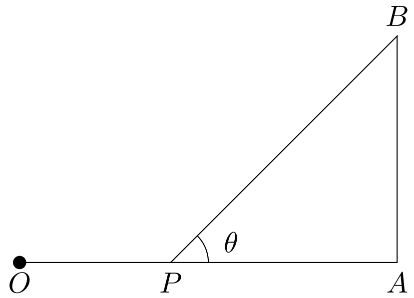
* * * * *

Problem 9. Two variables u and v are connected by the relation $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$, where f is a constant.

Given that u and v both vary with time, t, find an equation connecting $\frac{\mathrm{d}u}{\mathrm{d}t}$, $\frac{\mathrm{d}v}{\mathrm{d}t}$, u and v. Given also that u is decreasing at a constant rate of 2 cm per second and that $f = 10$ cm, calculate the rate of increase of v when $u = 50$ cm.

* * * * *

Problem 10. In the diagram, O and A are fixed points 1000 m apart on horizontal ground. The point B is vertically above A, and represents a balloon which is ascending at a steady rate of 2 ms^{-1}. The balloon is being observed from a moving point P on the line OA.



At time $t = 0$, the balloon is at A and the observer is at O. The observation point P moves towards A with steady speed 6 ms^{-1}. At time t, the angle APB is θ radians.

Show that

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{500}{t^2 + (500 - 3t)^2}.$$

* * * * *

Problem 11 (🌶). The normal to the rectangular hyperbola $xy = c^2$ at the point $P(cp, c/p)$, $p > 0$, meets the curve again at the point Q.

(a) Determine the coordinates of Q.

(b) Prove that $PQ^2 = 3OP^2 + OQ^2$.

Assignment B5A

Problem 1.

(a) Show, algebraically, that the derivative of the function

$$\ln(1+x) - \frac{2x}{x+2}$$

is never negative.

(b) Hence, show that $\ln(1+x) \geq \frac{2x}{x+2}$ when $x \geq 0$.

Solution. Let

$$f(x) = \ln(1+x) - \frac{2x}{x+2} = \ln(1+x) - 2 + \frac{4}{x+2}.$$

Part (a).

$$f'(x) = \frac{1}{1+x} - \frac{4}{(x+2)^2} = \frac{x^2}{(1+x)(x+2)^2}.$$

Given that $\ln(1+x)$ is defined, it must be that $1+x > 0$. We also know that $x^2 \geq 0$ and $(x+2)^2 \geq 0$. Hence, $f'(x) \geq 0$ for all x in the domain of f and is thus never negative.

Part (b). Note that $f(0) = 0$. Since $f'(x) \geq 0$ for all $x \geq 0$,

$$\ln(1+x) - \frac{2x}{x+2} = f(x) \geq f(0) = 0 \implies \ln(1+x) \geq \frac{2x}{x+2}.$$

* * * * *

Problem 2. The equation of a curve is $y = ax^2 - 2bx + c$, where a, b and c are constants, with $a > 0$.

(a) Using differentiation, find the coordinates of the turning point on the curve, in terms of a, b and c. State whether it is a maximum point or a minimum point.

(b) Given that the turning point of the curve lies on the line $y = x$, find an expression for c in terms of a and b. Show that in this case, whatever the value of b, $c \geq -1/4a$.

(c) Find the numerical values of a, b and c when the curve passes through the point $(0, 6)$ and has a turning point at $(2, 2)$.

Solution.

Part (a). For stationary points, $\mathrm{d}y/\mathrm{d}x = 0$. Hence,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2ax - 2b = 0 \implies x = \frac{b}{a} \implies y = a\left(\frac{b}{a}\right)^2 - 2b\left(\frac{b}{a}\right) + c = -\frac{b^2}{a} + c.$$

Since $a > 0$, the graph of y is concave upwards. Thus, there is a maximum point at $\left(\frac{b}{a}, -\frac{b^2}{a} + c\right)$.

Part (b). Since the turning point $\left(\frac{b}{a}, -\frac{b^2}{a} + c\right)$ lies on the line $y = x$,

$$\frac{b}{a} = -\frac{b^2}{a} + c \implies c = \frac{b + b^2}{a} = \frac{(b+1/2)^2 - 1/4}{a}.$$

Since $(b + 1/2)^2 \geq 0$, it follows that $c \geq -1/4a$.

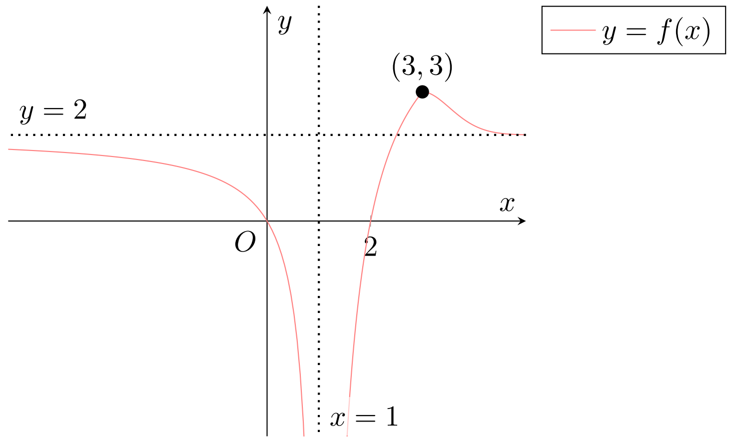
Part (c). Since the curve passes through $(0, 6)$, it is obvious to see that $c = 6$. Furthermore, since the curve has a turning point at $(2, 2)$, we know that $\frac{b}{a} = 2$ and $-\frac{b^2}{a} + c = 2$. Hence,

$$-\frac{b^2}{a} = 2 - c = -4 \implies b\left(\frac{b}{a}\right) = 4 \implies b = 2 \implies a = 1.$$

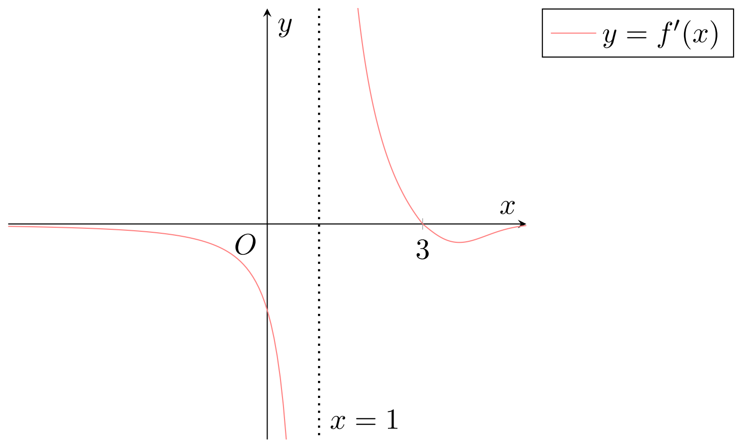
Thus, $a = 1$, $b = 2$, and $c = 6$.

* * * * *

Problem 3. The diagram below shows the graph of $y = f(x)$. Sketch the graph of $y = f'(x)$.



Solution.



Problem 4. The curve C has equation

$$x - y = (x + y)^2.$$

It is given that C has only one turning point.

(a) Show that $1 + \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2}{2x+2y+1}$.

(b) Hence, or otherwise, show that $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -\left(1 + \frac{\mathrm{d}y}{\mathrm{d}x}\right)^2$.

(c) Hence, state, with a reason, whether the turning point is a maximum or a minimum.

Solution.

Part (a). Implicitly differentiating the given equation,

$$1 - \frac{\mathrm{d}y}{\mathrm{d}x} = 2(x+y)\left(1 + \frac{\mathrm{d}y}{\mathrm{d}x}\right) \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1 - (2x+2y)}{2x+2y+1} \implies 1 + \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2}{2x+2y+1}.$$

Part (b). Implicitly differentiating the above equation,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -\frac{2\left(2 + 2 \cdot \frac{\mathrm{d}y}{\mathrm{d}x}\right)}{(2x+2y+1)^2} = -\left(\frac{2}{2x+2y+1}\right)^2\left(1 + \frac{\mathrm{d}y}{\mathrm{d}x}\right) = -\left(1 + \frac{\mathrm{d}y}{\mathrm{d}x}\right)^3.$$

Part (c). For turning points, $\mathrm{d}y/\mathrm{d}x = 0$. Hence, $\mathrm{d}^2y/\mathrm{d}x^2 = -1 < 0$. Thus, the turning point is a maximum.

Assignment B5B

Problem 1. Sketch the curve with parametric equations

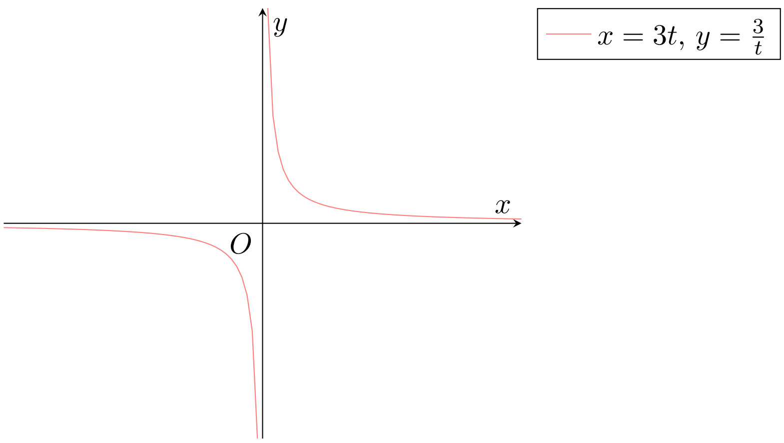
$$x = 3t,\ y = \frac{3}{t}.$$

The point P on the curve has parameter $t = 2$. The normal at P meets the curve again at the point Q.

(a) Show that the normal at P has equation $2y = 8x - 45$.

(b) Find the value of t at Q.

Solution.



Part (a). Consider $\mathrm{d}y/\mathrm{d}x$.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{-3/t^2}{3} = -\frac{1}{t^2}.$$

Hence, the tangent to the curve has gradient $-1/t^2$, whence the normal to the curve has gradient $\frac{-1}{-1/t^2} = t^2$. Thus, the normal to the curve at P has gradient $2^2 = 4$. Note that P has coordinates $(6, 3/2)$. Using the point-slope formula, the normal at P has equation

$$y - \frac{3}{2} = 4(x-6) \implies 2y = 8(x-6) + 3 = 8x - 45.$$

Part (b). Observe that

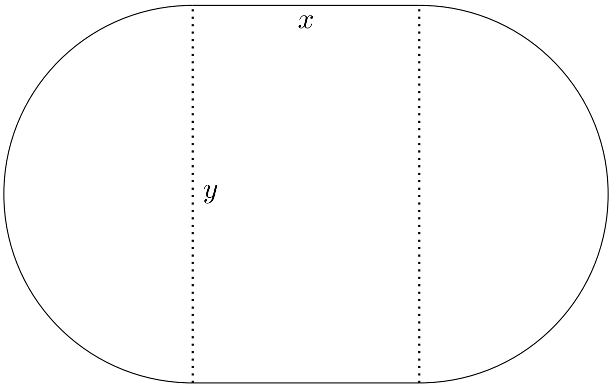
$$x = 3t \implies t = \frac{x}{3} \implies y = \frac{3}{x/3} = \frac{9}{x}.$$

Substituting $y = 9/x$ into the equation of the normal at P,

$$2\left(\frac{9}{x}\right) = 8x - 45 \implies 8x^2 - 45x - 18 = (x-6)(8x+3) = 0.$$

Hence, the x-coordinate of Q is $-3/8$ (note that we reject $x = 6$ since that corresponds to P). Thus, $t = -1/8$ at Q.

Problem 2.



A pond with a constant depth of 1 m is being designed for a park. The pond comprises a rectangle x m by y m and two semicircles of diameter y m, as shown in the diagram. The cost to build a boundary around the pond is \$30 per metre for straight parts and \$60 per metre for the semicircular parts. Given that the budget for the boundary is fixed at \$6000, using differentiation or otherwise, find in terms of π, the exact values of x and y which give the pond a maximum volume.

Solution. Observe that the total length of the straight parts is $2x$ m and the total length of the semicircular parts is πy m. Hence,

$$30(2x) + 60(\pi y) = 6000 \implies x + \pi y = 100 \implies x = 100 - \pi y.$$

Let $V(y)$ m^3 be the volume of the pond.

$$V(y) = \pi\left(\frac{y}{2}\right)^2 + xy = \frac{\pi}{4}y^2 + (100 - \pi y)\,y = -\frac{3\pi}{4}y^2 + 100y.$$

Consider the stationary points of $V(y)$. For stationary points, $V'(y) = 0$.

$$V'(y) = -\frac{3\pi}{2}y + 100 = 0 \implies y = \frac{200}{3\pi}.$$

y	$\left(\frac{200}{3\pi}\right)^-$	$\frac{200}{3\pi}$	$\left(\frac{200}{3\pi}\right)^+$
$V'(y)$	+ve	0	−ve

By the first derivative test, the maximum volume of the pond is achieved when $y = 200/3\pi$. Thus, $x = 100 - \pi y = 100/3$.

* * * * *

Problem 3. A circular cylinder is expanding in such a way that, at time t seconds, the length of the cylinder is $20x$ cm and the area of the cross-section is x cm^2. Given that, when $x = 5$, the area of the cross-section is increasing at a rate of 0.025 cm^{2}s^{-1}, find the rate of increase at this instant of

(a) the length of the cylinder,

(b) the volume of the cylinder,

(c) the radius of the cylinder.

Solution. Let $A = x$ cm^2 be the cross-sectional area of the cylinder. Then

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

and

$$\left.\frac{\mathrm{d}A}{\mathrm{d}t}\right|_{x=5} = 0.025.$$

Part (a). Let $L = 20x$ cm be the length of the cylinder. Then

$$\frac{\mathrm{d}L}{\mathrm{d}t} = 20 \cdot \frac{\mathrm{d}x}{\mathrm{d}t} \implies \left.\frac{\mathrm{d}L}{\mathrm{d}t}\right|_{x=5} = 20\,(0.025) = 0.5.$$

Thus, the length of the cylinder is increasing at a rate of 0.5 cm/s.

Part (b). Let $V = AL = 20x^2$ cm^3 be the volume of the cylinder. Then

$$\frac{\mathrm{d}V}{\mathrm{d}t} = 40x \cdot \frac{\mathrm{d}x}{\mathrm{d}t} \implies \left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{x=5} = 40(5)(0.025) = 5.$$

Thus, the volume of the cylinder is increasing at a rate of 5 cm^3/s.

Part (c). Let R cm be the radius of the cylinder. Observe that

$$\pi R^2 = A = x \implies R = \sqrt{\frac{x}{\pi}} = \frac{\sqrt{x}}{\sqrt{\pi}}.$$

Hence,

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \frac{1}{\sqrt{\pi}} \cdot \frac{1}{2\sqrt{x}} \cdot \frac{\mathrm{d}x}{\mathrm{d}t} \implies \left.\frac{\mathrm{d}R}{\mathrm{d}t}\right|_{x=5} = \frac{1}{\sqrt{\pi}}\left(\frac{1}{2\sqrt{5}}\right)(0.025) = 0.00315 \text{ (3 s.f.)}.$$

Thus, the radius of the cylinder is increasing at a rate of 0.00315 cm/s.

* * * * *

Problem 4. The curve C has equation $2^{-y} = x$. The point A on C has x-coordinate a where $a > 0$. Show that $\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{1}{a \ln 2}$ at A and find the equation of the tangent to C at A. Hence, find the equation of the tangent to C which passes through the origin.

The straight line $y = mx$ intersects C at 2 distinct points. Write down the range of values of m.

Solution. Observe that

$$2^{-y} = x \implies y = -\log_2 x = -\frac{\ln x}{\ln 2} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{1}{x \ln 2}.$$

At A, $x = a$. Hence,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{1}{a \ln 2}.$$

Also, we clearly have $A(a, -\ln a/\ln 2)$. Using the point-slope formula, the tangent to C at A has equation

$$y - \left(-\frac{\ln a}{\ln 2}\right) = -\frac{1}{a \ln 2}(x - a) \implies y = -\frac{x}{a \ln 2} + \frac{1 - \ln a}{\ln 2}.$$

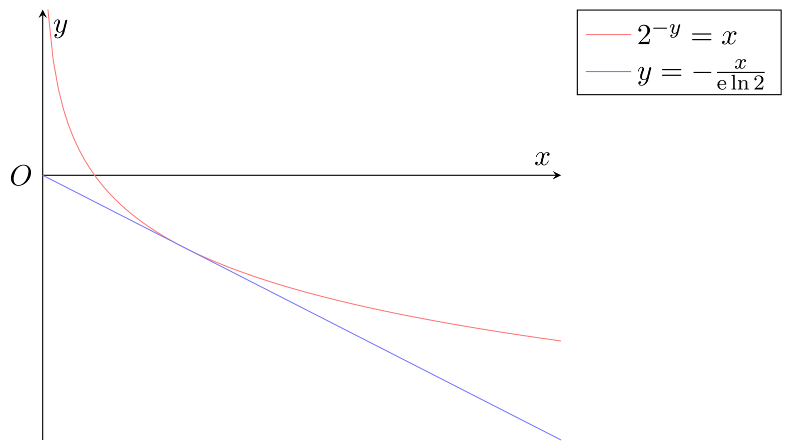
Consider the straight line $y = mx$ that is tangent to C and passes through the origin.

$$0 = -\frac{0}{a \ln 2} + \frac{1 - \ln a}{\ln 2} \implies 1 - \ln a = 0 \implies a = \mathrm{e}.$$

Hence, the equation of the tangent to C that passes through the origin is

$$y = -\frac{x}{\mathrm{e} \ln 2}.$$

Consider the graph of $2^{-y} = x$.



Hence, $m \in (-1/\mathrm{e}\ln 2, 0)$.

B6 MacLaurin Series

Tutorial B6

Problem 1.

(a) Given that $f(x) = \mathrm{e}^{\cos x}$, find $f(0)$, $f'(0)$ and $f''(0)$. Hence, write down the first two non-zero terms in the MacLaurin series for $f(x)$. Give the coefficients in terms of e.

(b) Given that $g(x) = \tan\left(2x + \frac{1}{4}\pi\right)$, find $g(0)$, $g'(0)$ and $g''(0)$. Hence, find the first three terms in the MacLaurin series of $g(x)$.

Solution.

Part (a). Note that

$$f'(x) = -\mathrm{e}^{\cos x} \sin x = -f(x) \sin x \implies f''(x) = -f(x) \cos x - f'(x) \sin x.$$

Evaluating $f(x)$, $f'(x)$ and $f''(x)$ at 0,

$$f(0) = \mathrm{e}, \quad f'(0) = 0, \quad f''(0) = -\mathrm{e}.$$

Hence,

$$f(x) = \frac{\mathrm{e}}{0!} + \frac{0}{1!}x + \frac{-\mathrm{e}}{2!}x^2 = \mathrm{e} - \frac{\mathrm{e}}{2}x^2 + \cdots.$$

Part (b). Note that

$$g'(x) = 2\sec^2\left(2x + \frac{\pi}{4}\right) = 2\left(1 + \tan^2\left(2x + \frac{\pi}{4}\right)\right) = 2 + 2g^2(x) \implies g''(x) = 4g(x)g'(x).$$

Evaluating $g(x)$, $g'(x)$ and $g''(x)$ at 0,

$$g(x) = 1, \quad g'(x) = 4, \quad g''(x) = 16.$$

Hence,

$$g(x) = \frac{1}{0!} + \frac{4}{1!}x + \frac{16}{2!}x^2 + \cdots = 1 + 4x + 8x^2 + \cdots.$$

* * * * *

Problem 2. Find the first three non-zero terms of the MacLaurin series for the following in ascending powers of x. In each case, find the range of values of x for which the series is valid.

(a) $\frac{(1+3x)^4}{\sqrt{1+2x}}$

(b) $\frac{\sin 2x}{2+3x}$

Solution.

Part (a). Observe that

$$(1+3x)^4 = 1 + 4(3x) + 6(3x)^2 + \cdots = 1 + 12x + 54x^2 + \cdots$$

and

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

Thus,

$$\begin{aligned} y = \frac{(1+3x)^4}{\sqrt{1+2x}} &= \left(1 + 12x + 54x^2 + \cdots\right)\left(1 - x + \frac{3}{2}x^2 + \cdots\right) \\ &= \left(1 - x + \frac{3}{2}x^2\right) + \left(12x - 12x^2\right) + \left(54x^2\right) + \cdots = 1 + 11x + \frac{87}{2}x^2 + \cdots . \end{aligned}$$

Note that the series is valid only when

$$|2x| < 1 \implies -\frac{1}{2} < x < \frac{1}{2}.$$

Part (b). Note that

$$\sin 2x = 2x - \frac{(2x)^3}{3!} + \cdots = 2x - \frac{4}{3}x^3 + \cdots$$

and

$$\begin{aligned} \frac{1}{2+3x} = \frac{1}{2}\left(1 + \frac{3x}{2}\right)^{-1} &= \frac{1}{2}\left[1 - \frac{3x}{2} + \left(\frac{3x}{2}\right)^2 - \left(\frac{3x}{2}\right)^3 + \cdots\right] \\ &= \frac{1}{2} - \frac{3}{4}x + \frac{9}{8}x^2 - \frac{27}{16}x^3 + \cdots . \end{aligned}$$

Thus,

$$\begin{aligned} \frac{\sin 2x}{2+3x} &= \left(2x - \frac{4}{3}x^3 + \cdots\right)\left(\frac{1}{2} - \frac{3}{4}x + \frac{9}{8}x^2 - \frac{27}{16}x^3 + \cdots\right) \\ &= \left(x - \frac{3}{2}x^2 + \frac{9}{4}x^3\right) + \left(-\frac{2}{3}x^3\right) + \cdots = x - \frac{3}{2}x^2 + \frac{19}{12}x^3 + \cdots . \end{aligned}$$

The series is only valid when

$$\left|\frac{3}{2}x\right| < 1 \implies -\frac{2}{3} < x < \frac{2}{3}.$$

* * * * *

Problem 3. Find the MacLaurin series of $\ln(1+\cos x)$, up to and including the term in x^4.

Solution. Let $y = \ln(1+\cos x)$. Then

$$y = \ln(1+\cos x) \implies \mathrm{e}^y = 1 + \cos x.$$

Implicitly differentiating repeatedly with respect to x,

$$\mathrm{e}^y y' = -\sin x \implies \mathrm{e}^y\left[\left(y'\right)^2 + y''\right] = -\cos x \implies \mathrm{e}^y\left[\left(y'\right)^3 + 3y'y'' + y'''\right] = \sin x$$
$$\implies \mathrm{e}^y\left[\left(y'\right)^4 + 3\left(y''\right)^2 + 6\left(y'\right)^2 y'' + 4y'y''' + y^{(4)}\right] = \cos x.$$

Evaluating the above at $x = 0$, we get

$$y(0) = \ln 2, \quad y'(0) = 0, \quad y''(0) = -\frac{1}{2}, \quad y'''(0) = 0, \quad y^{(4)}(0) = -\frac{1}{4}.$$

Thus,

$$\ln(1+\cos x) = \ln 2 + \frac{-1/2}{2!}x^2 + \frac{-1/4}{4!}x^4 + \cdots = \ln 2 - \frac{1}{4}x^2 - \frac{1}{96}x^4 + \cdots.$$

* * * * *

Problem 4.

(a) Find the first three terms of the MacLaurin series for $\mathrm{e}^x(1+\sin 2x)$.

(b) It is given that the first two terms of this series are equal to the first two terms in the series expansion, in ascending powers of x, of $\left(1+\frac{4}{3}x\right)^n$. Find n and show that the third terms in each of these series are equal.

Solution.

Part (a). Observe that

$$\mathrm{e}^x = 1 + x + \frac{x^2}{2} + \cdots$$

and

$$1 + \sin 2x = 1 + 2x + \cdots.$$

Hence,

$$\mathrm{e}^x\left(1+\sin 2x\right) = \left(1 + x + \frac{x^2}{2} + \cdots\right)\left(1 + 2x + \cdots\right)$$
$$= (1+2x) + \left(x + 2x^2\right) + \left(\frac{x^2}{2}\right) + \cdots = 1 + 3x + \frac{5}{2}x^2 + \cdots.$$

Part (b). Note that

$$\left(1+\frac{4}{3}x\right)^n = 1 + n\left(\frac{4}{3}x\right) + \frac{n(n-1)}{2}\left(\frac{4}{3}x\right)^2 + \cdots = 1 + \frac{4n}{3}x + \frac{8n(n-1)}{9}x^2\cdots.$$

Comparing the second terms of both series, we get

$$\frac{4n}{3} = 3 \implies n = \frac{9}{4}.$$

Thus, the third term of $(1+\frac{4}{3}x)^n$ is

$$\frac{8(\frac{9}{4})(\frac{9}{4}-1)}{9}x^2 = \frac{5}{2}x^2.$$

Hence, the third terms in each of these series are equal.

Problem 5.

(a) Show that the first three non-zero terms in the expansion of $\left(\frac{8}{x^3}-1\right)^{1/3}$ in ascending powers of x are in the form $\frac{a}{x}+bx^2+cx^5$, where a, b and c are constants to be determined.

(b) By putting $x=\frac{2}{3}$ in your result, obtain an approximation for $\sqrt[3]{26}$ in the form of a fraction in its lowest terms.

A student put $x=6$ into the expansion to obtain an approximation of $\sqrt[3]{26}$. Comment on the suitability of this choice of x for the approximation of $\sqrt[3]{26}$.

Solution.

Part (a). We have

$$\begin{aligned}\left(\frac{8}{x^3}-1\right)^{\frac{1}{3}} = \frac{2}{x}\left(1-\frac{x^3}{8}\right)^{\frac{1}{3}} &= \frac{2}{x}\left[1+\frac{1}{3}\left(-\frac{x^3}{8}\right)+\frac{\left(\frac{1}{3}\right)\left(\frac{1}{3}-1\right)}{2}\left(-\frac{x^3}{8}\right)^2+\cdots\right]\\ &= \frac{2}{x}\left(1-\frac{x^3}{24}-\frac{x^6}{576}+\cdots\right)=\frac{2}{x}-\frac{x^2}{12}-\frac{x^5}{288}+\cdots.\end{aligned}$$

Part (b). Evaluating the above equation at $x=2/3$,

$$\sqrt[3]{26}\approx\left(\frac{8}{(2/3)^3}-1\right)^{1/3}=\frac{2}{2/3}-\frac{(2/3)^2}{12}-\frac{(2/3)^5}{288}=\frac{6479}{2187}.$$

Observe that the validity range for the series is

$$\left|-\frac{x^3}{8}\right|<1 \implies -2<x<2.$$

Since 6 is outside this range, it is not an appropriate choice.

* * * * *

Problem 6. Let $f(x)=\mathrm{e}^x\sin x$.

(a) Sketch the graph of $y=f(x)$ for $-3\le x\le 3$.

(b) Find the series expansion of $f(x)$ in ascending powers of x, up to and including the term in x^3.

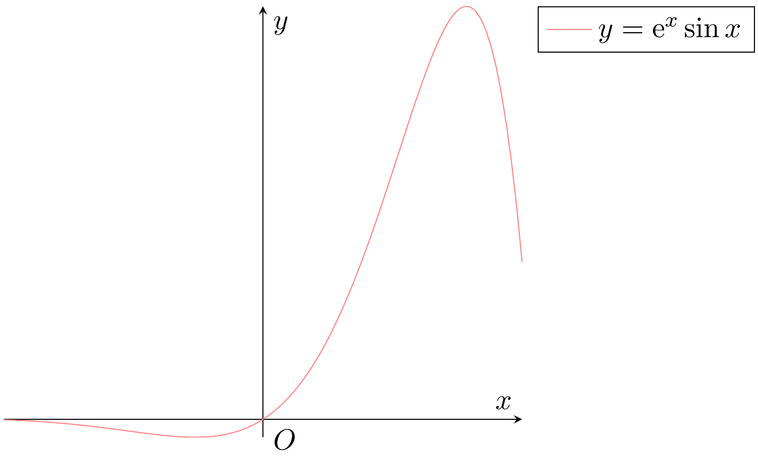
Denote the answer to part (b) by $g(x)$.

(c) On the same diagram, sketch the graph of $y=f(x)$ and $y=g(x)$. Label the two graphs clearly.

(d) Find, for $-3\le x\le 3$, the set of values of x for which the value of $g(x)$ is within ±0.5 of the value of $f(x)$.

Solution.

Part (a).



Part (b). Observe that

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots$$

and

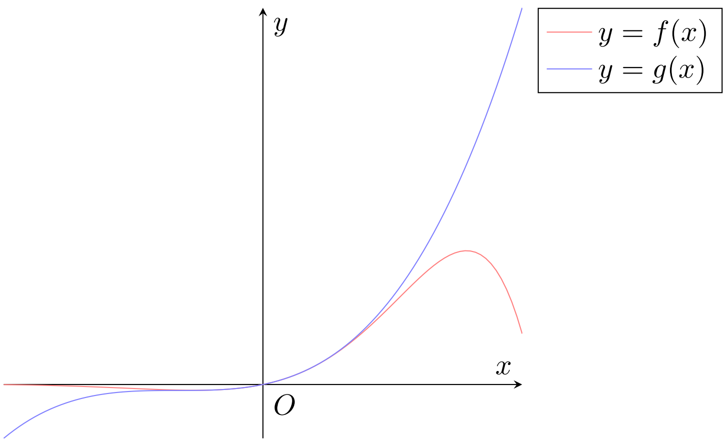
$$\sin x = x - \frac{x^3}{6} + \cdots .$$

Thus,

$$e^x \sin x = \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots\right)\left(x - \frac{x^3}{6} + \cdots\right)$$

$$= \left(x - \frac{x^3}{6}\right) + \left(x^2\right) + \left(\frac{x^3}{2}\right) + \cdots = x + x^2 + \frac{x^3}{3} + \cdots .$$

Part (c).



Part (d). Using G.C., $\{x \in \mathbb{R} : -1.96 \leq x \leq 1.56\}$.

* * * * *

Problem 7. It is given that $y = 1/(1 + \sin 2x)$. Show that, when $x = 0$, $\mathrm{d}^2y/\mathrm{d}x^2 = 8$. Find the first three terms of the MacLaurin series for y.

(a) Use the series to obtain an approximate value for $\int_{-0.1}^{0.1} y \, \mathrm{d}x$, leaving your answer as a fraction in its lowest terms.

(b) Find the first two terms of the MacLaurin series for $\mathrm{d}y/\mathrm{d}x$.

(c) Write down the equation of the tangent at the point where $x = 0$ on the curve $y = 1/(1 + \sin 2x)$.

Solution. Differentiating with respect to x, we get

$$y' = -\frac{2\cos 2x}{(1+\sin 2x)^2} = -2y^2 \cos 2x.$$

Differentiating once more, we get

$$y'' = -2\left(-2y^2 \sin 2x + 2y \cdot y' \cos 2x\right).$$

Evaluating the above at $x = 0$, we obtain

$$y(0) = 1, \quad y'(0) = -2, \quad y''(0) = 8.$$

Hence,

$$\frac{1}{1+\sin 2x} = \frac{1}{0!} + \frac{-2}{1!}x + \frac{8}{2!}x^2 + \cdots = 1 - 2x + 4x^2 + \cdots.$$

Part (a).

$$\int_{-0.1}^{0.1} y \,\mathrm{d}x \approx \int_{-0.1}^{0.1} \left(1 - 2x + 4x^2\right) \mathrm{d}x = \left[x - x^2 + \frac{4}{3}x^3\right]_{-0.1}^{0.1} = \frac{76}{275}.$$

Part (b).

$$y' = \frac{\mathrm{d}}{\mathrm{d}x}\left(1 - 2x + 4x^2 + \cdots\right) = -2 + 8x + \cdots.$$

Part (c). Using the point-slope formula,

$$y - 1 = -2(x - 0) \implies y = -2x + 1.$$

* * * * *

Problem 8. It is given that $y = \mathrm{e}^{\arcsin 2x}$.

(a) Show that $(1-4x^2)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 4x\frac{\mathrm{d}y}{\mathrm{d}x} = 4y$.

(b) By further differentiating this result, find the MacLaurin series for y in ascending powers of x, up to an including the term in x^3.

(c) Hence, find an approximation value of $\mathrm{e}^{\pi/2}$, by substituting a suitable value of x in the MacLaurin series for y.

(d) Suggest one way to improve the accuracy of the approximated value obtained.

Solution.

Part (a). Note that

$$y = \mathrm{e}^{\arcsin(2x)} \implies \ln y = \arcsin(2x).$$

Implicitly differentiating with respect to x,

$$\frac{1}{y} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2}{\sqrt{1-4x^2}} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2y}{\sqrt{1-4x^2}}.$$

Implicitly differentiating with respect to x once again,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\sqrt{1-4x^2}\left(2\cdot\frac{\mathrm{d}y}{\mathrm{d}x}\right) - 2y\left(\frac{-4x}{\sqrt{1-4x^2}}\right)}{1-4x^2}.$$

Now observe that

$$2\sqrt{1-4x^2}\cdot\frac{\mathrm{d}y}{\mathrm{d}x} + 4x\left(\frac{2y}{\sqrt{1-4x^2}}\right) = 4y + 4x\cdot\frac{\mathrm{d}y}{\mathrm{d}x}.$$

Hence,

$$\left(1-4x^2\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 4y + 4x\cdot\frac{\mathrm{d}y}{\mathrm{d}x} \implies \left(1-4x^2\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 4x\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = 4y.$$

Part (b). Implicitly differentiating with respect to x once again,

$$\left(1-4x^2\right)\frac{\mathrm{d}^3y}{\mathrm{d}x^3} - 8x\cdot\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 4\left(x\cdot\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + \frac{\mathrm{d}y}{\mathrm{d}x}\right) = 4\cdot\frac{\mathrm{d}y}{\mathrm{d}x}.$$

Rearranging,

$$\left(1-4x^2\right)\frac{\mathrm{d}^3y}{\mathrm{d}x^3} - 12x\cdot\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 8\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

Evaluating the above equations at $x = 0$, we get

$$y(0) = 1, \quad y'(0) = 2, \quad y''(0) = 4, \quad y'''(0) = 16.$$

Hence,

$$y = \frac{1}{0!} + \frac{2}{1!}x + \frac{4}{2!}x^2 + \frac{16}{3!}x^3 + \cdots = 1 + 2x + 2x^2 + \frac{8}{3}x^3 + \cdots.$$

Part (c). Consider $y = \mathrm{e}^{\pi/2}$.

$$y = \arcsin 2x = \mathrm{e}^{\pi/2} \implies x = \frac{1}{2}\sin\frac{\pi}{2} = \frac{1}{2}.$$

Substituting $x = 1/2$ into the MacLaurin series for y,

$$\mathrm{e}^{\pi/2} \approx 1 + 2\left(\frac{1}{2}\right) + 2\left(\frac{1}{2}\right)^2 + \frac{8}{3}\left(\frac{1}{2}\right)^3 = \frac{17}{6}.$$

Part (d). More terms of the MacLaurin series of y could be considered.

* * * * *

Problem 9. The curve $y = f(x)$ passes through the point $(0, 1)$ and satisfies the equation $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{6-2y}{\cos 2x}$.

(a) Find the MacLaurin series of $f(x)$, up to and including the term in x^3.

(b) Using standard results given in the List of Formulae (MF27), express $\frac{1-\sin x}{\cos x}$ as a power series of x, up to and including the term in x^3.

(c) Using the two power series you have found, show to this degree of approximation, that $f(x)$ can be expressed as $a(\tan 2x - \sec 2x) + b$, where a and b are constants to be determined.

Solution.

Part (a). Note that

$$y' = \frac{6-2y}{\cos 2x} \implies y' \cos 2x = 6 - 2y.$$

Implicitly differentiating with respect to x,

$$-2y' \sin 2x + y'' \cos 2x = -2y'.$$

Implicitly differentiating once more,

$$-2\left(y'' \sin 2x + 2y' \cos 2x\right) + \left(y''' \cos 2x - 2y'' \sin 2x\right) = -2y''$$

Hence,

$$y(0) = 1, \quad y'(0) = 4, \quad y''(0) = -8, \quad y'''(0) = 32,$$

whence

$$f(x) = \frac{1}{0!}x + \frac{4}{1!}x + \frac{-8}{2!}x^2 + \frac{32}{3!}x^3 + \cdots = 1 + 4x - 4x^2 + \frac{16}{3}x^3 + \cdots.$$

Part (b). Note that

$$\frac{1}{\cos x} \approx \left(1 - \frac{x^2}{2}\right)^{-1} \approx 1 + \frac{x^2}{2}.$$

Hence,

$$\frac{1-\sin x}{\cos x} \approx \left(1 - x + \frac{x^3}{6}\right)\left(1 + \frac{x^2}{2}\right) = 1 - x + \frac{x^2}{2} - \frac{x^3}{3} + \cdots.$$

Part (c). Note that

$$\frac{1-\sin x}{\cos x} = \sec x - \tan x.$$

Hence,

$$a(\tan 2x - \sec 2x) + b \approx -a\left[1 - 2x + \frac{(2x)^2}{2} - \frac{(2x)^3}{3}\right] + b$$

$$= a\left(-1 + 2x - 2x^2 + \frac{8}{3}x^3\right) + b = a\left(-\frac{3}{2} + \frac{f(x)}{2}\right) + b = -\frac{3}{2}a + b + \frac{a}{2}f(x).$$

Thus,

$$\frac{a}{2}f(x) - \frac{3}{2}a + b \approx a(\tan 2x - \sec 2x) + b.$$

In order to obtain an approximation for $f(x)$, we need $\frac{a}{2} = 1$ and $-\frac{3}{2}a + b = 0$, whence $a = 2$ and $b = 3$.

* * * * *

Problem 10. Given that x is sufficiently small for x^3 and higher powers of x to be neglected, and that $13 - 59\sin x = 10(2 - \cos 2x)$, find a quadratic equation for x and hence solve for x.

Solution. Note that

$$13 - 59\sin x = 10\left(2 - \cos 2x\right) = 10\left[2 - \left(1 - 2\sin^2 x\right)\right] = 10 + 20\sin^2 x.$$

Thus,

$$20\sin^2 x + 59\sin x - 3 = (20\sin x - 1)(\sin x + 3) = 0,$$

whence $\sin x = 1/20$. Note that we reject $\sin x = -3$ since $|\sin x| \leq 1$. Since x is sufficiently small for x^3 and higher powers of x to be neglected, $\sin x \approx x$. Thus, $x \approx 1/20$.

Problem 11. In triangle ABC, angle $A = \pi/3$ radians, angle $B = (\pi/3 + x)$ radians and angle $C = (\pi/3 - x)$ radians, where x is small. The lengths of the sides BC, CA and AB are denoted by a, b and c respectively. Show that $b - c \approx 2ax/\sqrt{3}$.

Solution. By the sine rule,

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}.$$

Hence,

$$b = a\left(\frac{\sin B}{\sin A}\right) = \frac{2a}{\sqrt{3}}\sin B, \quad c = a\left(\frac{\sin C}{\sin A}\right) = \frac{2a}{\sqrt{3}}\sin C.$$

Thus,

$$\begin{aligned} b - c = \frac{2a}{\sqrt{3}}(\sin B - \sin C) &= \frac{2a}{\sqrt{3}}\left[\sin\left(\frac{\pi}{3} + x\right) - \sin\left(\frac{\pi}{3} - x\right)\right] \\ &= \frac{2a}{\sqrt{3}}\left[2\sin x \cos\frac{\pi}{3}\right] = \frac{2a}{\sqrt{3}}\sin x. \end{aligned}$$

Since x is small, $\sin x \approx x$. Hence,

$$b - c \approx \frac{2ax}{\sqrt{3}}.$$

* * * * *

Problem 12. D'Alembert's ratio test states that a series of the form $\sum_{r=0}^{\infty} a_r$ converges when $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| < 1$, and diverges when $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| > 1$. When $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 1$, the test is inconclusive. Using the test, explain why the series $\sum_{r=0}^{\infty}\frac{x^r}{r!}$ converges for all real values of x and state the sum to infinity of this series, in terms of x.

Solution. Let $a_n = \frac{x^n}{n!}$ and consider $\lim_{n\to\infty}\left|\frac{a_{n+1}}{n}\right|$.

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{n}\right| = \lim_{n\to\infty}\left|\frac{x^{n+1}}{(n+1)!}\Big/\frac{x^n}{n!}\right| = \lim_{n\to\infty}\left|\frac{x^{n+1}}{x^n}\cdot\frac{n!}{(n+1)!}\right| = \lim_{n\to\infty}\left|\frac{x}{n+1}\right| = 0.$$

Since $\lim_{n\to\infty}\left|\frac{a_{n+1}}{n}\right| < 1$ for all $x \in \mathbb{R}$, it follows by D'Alembert's ratio test that $\sum_{r=0}^{\infty}\frac{x^r}{r!}$ converges for all real values of x. The sum to infinity of the series in question is e^x.

Self-Practice B6

Problem 1. Express $\frac{6x+4}{(1-2x)(1+3x^2)}$ in partial fractions. Hence, find the coefficients of x^5 and x^6 in the expansion, in ascending powers of x, of $\frac{6x+4}{(1-2x)(1+3x^2)}$.

Solution. Let

$$\frac{6x+4}{(1-2x)(1+3x^2)} = \frac{A}{1-2x} + \frac{Bx+C}{1+3x^2},$$

where A, B and C are constants to be determined. Using the cover-up rule, we immediately get

$$A = \frac{6(1/2)+4}{1+3(1/2)^2} = 4.$$

Clearing denominators, we get

$$6x+4 = 4\left(1+3x^2\right) + (Bx+C)(1-2x) = (12-2B)\,x^2\,(B-2C)\,x + (4+C)\,.$$

Comparing coefficients, we have $B = 6$ and $C = 0$. Hence,

$$\frac{6x+4}{(1-2x)(1+3x^2)} = \frac{4}{1-2x} + \frac{6x}{1+3x^2}.$$

Note that

$$\frac{4}{1-2x} = \dots + (2x)^5 + (2x)^6 + \dots = 4\left[\dots + 128x^5 + 256x^6 + \dots\right]$$

and

$$\frac{6x}{1+3x^2} = 6x\left[\dots + \left(-3x^2\right)^2 + \dots\right] = \dots + 54x^5 + \dots.$$

Hence,

$$\frac{6x+4}{(1-2x)(1+3x^2)} = \dots + 182x^5 + 256x^6 + \dots.$$

* * * * *

Problem 2. If x is so small that terms in x^n, $n \geq 3$, can be neglected and $\frac{3+ax}{3+bx} = (1-x)^{1/3}$, find the values of a and b. Hence, find an approximation for $\sqrt[3]{0.96}$ in the form $\frac{p}{q}$, where p and q are integers.

Solution. Rearranging,

$$3+ax = (3+bx)(1-x)^{1/3} = (3+bx)\left(1-\frac{x}{3}-\frac{x^2}{9}\right) = 3+(b-1)x - \frac{b+1}{3}x^2.$$

Comparing coefficients, we have $a = -2$ and $b = -1$. Thus,

$$\frac{3-2x}{3-x} = (1-x)^{1/3}.$$

Substituting $x = 0.04$, we get

$$\sqrt[3]{0.96} = \frac{3-2(0.04)}{3-0.04} = \frac{73}{74}.$$

* * * * *

Problem 3. Given that $y = \tan\left(\frac{1}{2}\arctan x\right)$, show that

$$\left(1+x^2\right)\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{2}\left(1+y^2\right).$$

By differentiating this result twice, show that, up to and including the term in x^3, the Maclaurin series for $\tan\left(\frac{1}{2}\arctan x\right)$ is $\frac{1}{2}x - \frac{1}{8}x^3$.

Solution. Note that $\arctan y = \frac{1}{2}\arctan x$. Differentiating with respect to x,

$$\frac{y'}{1+y^2} = \frac{1}{2}\cdot\frac{1}{1+x^2} \implies \left(1+x^2\right)y' = \frac{1}{2}\left(1+y^2\right).$$

Differentiating with respect to x,

$$\left(1+x^2\right)y'' + 2xy' = yy'.$$

Differentiating once more,

$$\left(1+x^2\right)y''' + 4xy'' + 2y' = y\cdot y'' + \left(y'\right)^2.$$

When $x=0$, we get

$$y(0)=0, \quad y'(0)=\frac{1}{2}, \quad y''(0)=0, \quad y'''(0)=-\frac{3}{4}.$$

Thus,

$$y = \tan\left(\frac{1}{2}\arctan x\right) = \frac{1}{2}x + \frac{-3/4}{3!}x^3 = \frac{1}{2}x - \frac{1}{8}x^3.$$

* * * * *

Problem 4. Given that $\cos y = \sqrt{1-\frac{1}{4}\mathrm{e}^x}$ and $0 < y < \frac{\pi}{2}$, show that $\sin(2y)\,\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{4}\mathrm{e}^x$. By further differentiation of this result, find the Maclaurin series for y, up to and including the term in x^2, leaving your answer in exact form. Deduce the equation of the tangent to the curve $y = \arccos\sqrt{1-\frac{1}{4}\mathrm{e}^x}$ at $x=0$.

Solution. Rearranging, we get

$$\cos^2 y = 1 - \frac{1}{4}\mathrm{e}^x.$$

Differentiating with respect to x,

$$-2\cos y\sin y\cdot y' = -\frac{1}{4}\mathrm{e}^x \implies \sin(2y)\,y' = \frac{1}{4}\mathrm{e}^x.$$

Differentiating once more,

$$\sin(2y)\,y'' + 2\cos(2y)\left(y'\right)^2 = \frac{1}{4}\mathrm{e}^x.$$

When $x=0$, we get

$$y(0)=\frac{\pi}{6}, \quad y'(0)=\frac{1}{2\sqrt{3}}, \quad y''(0)=\frac{1}{3\sqrt{3}}.$$

Thus,

$$y = \arccos\sqrt{1-\frac{1}{4}\mathrm{e}^x} = \frac{\pi}{6} + \left(\frac{1}{2\sqrt{3}}\right)x + \left(\frac{1}{3\sqrt{3}}\right)\left(\frac{x^2}{2}\right) + \cdots = \frac{\pi}{6} + \frac{x}{2\sqrt{3}} + \frac{x^2}{6\sqrt{3}} + \cdots.$$

The equation of the tangent at $x=0$ is simply

$$y = \frac{\pi}{6} + \frac{x}{2\sqrt{3}}.$$

* * * * *

Problem 5. By expressing $\sin\left(\frac{\pi}{3}+2x\right)$ in terms of $\sin 2x$ and $\cos 2x$, show that

$$\sin\left(\frac{\pi}{3}+2x\right) \approx \frac{\sqrt{3}}{2} + x - \sqrt{3}x^2$$

if x is sufficiently small. Hence, by using a suitable value of x, estimate the value of $\sin\frac{\pi}{9}$, giving your answer to 3 significant figures.

Solution. By the angle-sum formula,

$$\sin\left(\frac{\pi}{3}+2x\right)=\sin\frac{\pi}{3}\cos 2x+\cos\frac{\pi}{3}\sin 2x=\frac{\sqrt{3}}{2}\cos 2x+\frac{1}{2}\sin 2x.$$

For sufficiently small x, we have $\sin x\approx x$ and $\cos x=1-x^2/2$. Hence,

$$\sin\left(\frac{\pi}{3}+2x\right)\approx\frac{\sqrt{3}}{2}\left(1-\frac{(2x)^2}{2}\right)+\frac{1}{2}(2x)=\frac{\sqrt{3}}{2}+x-\sqrt{3}x^2.$$

Consider $\pi/3+2x=\pi/9$. Clearly $x=-\pi/9$. Substituting this into the above approximation, we get

$$\sin\frac{\pi}{9}\approx\frac{\sqrt{3}}{2}+\left(-\frac{\pi}{9}\right)-\sqrt{3}\left(-\frac{\pi}{9}\right)^2=0.306\text{ (3 s.f.)}.$$

* * * * *

Problem 6 (🌶). Consider the infinite series $\frac{1}{1!}+\frac{4}{2!}+\frac{7}{3!}+\frac{10}{4!}+\dots$.

(a) If the series continues with the same pattern, find an expression for the nth term.

(b) By rewriting the infinite series in terms of sigma notation and using the standard series for e^x, show that the series evaluates to $\mathrm{e}+2$.

Solution.
Part (a). The nth term is given by $\frac{3n-2}{n!}$, where $n\geq 1$.
Part (b). The infinite series is given by

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{3n-2}{n!}&=3\sum_{n=1}^{\infty}\frac{n}{n!}-2\sum_{n=1}^{\infty}\frac{1}{n!}\\&=3\sum_{n=1}^{\infty}\frac{1}{(n-1)!}-2\left(1+\sum_{n=1}^{\infty}\frac{1}{n!}\right)+2\\&=3\sum_{n=0}^{\infty}\frac{1}{n!}-2\sum_{n=0}^{\infty}\frac{1}{n!}+2\\&=3\mathrm{e}-2\mathrm{e}+2=\mathrm{e}+2.\end{aligned}$$

Problem 7 (🌶). Find the function represented by each of the following series by expressing it as a sum or difference of two standard series.

(a) $f(x) = 2 + x + \frac{x^3}{3!} + \frac{2x^4}{4!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \frac{2x^8}{8!} + \dots,\ x \in \mathbb{R}$.

(b) $g(x) = (a+b)x - \frac{a^2+b^2}{2}x^2 + \frac{a^3+b^3}{3}x^3 - \frac{a^4+b^4}{4}x^4 + \dots$, where a and b are positive constants such that $-\frac{1}{a} < x \leq \frac{1}{a}$ and $-\frac{1}{b} < x \leq \frac{1}{b}$.

Solution.

Part (a). Observe that $f(x)$ is defined for all $x \in \mathbb{R}$. This suggests that $f(x)$ is composed of e^x, $\cos x$ and $\sin x$. Also observe that the powers of 2, 6, 10, ... are missing. This suggests that we are adding $\cos x$ to e^x:

$$\begin{aligned}
f(x) &= 2 + x + \frac{x^3}{3!} + \frac{2x^4}{4!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \frac{2x^8}{8!} + \dots \\
&= \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \frac{x^7}{7!} + \frac{x^8}{8!} + \dots\right) \\
&+ \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} + \dots\right) \\
&= \mathrm{e}^x + \cos x.
\end{aligned}$$

Part (b). We can easily separate $g(x)$ as follows:

$$\begin{aligned}
g(x) &= (a+b)x - \frac{a^2+b^2}{2}x^2 + \frac{a^3+b^3}{3}x^3 - \frac{a^4+b^4}{4}x^4 + \dots \\
&= \left(ax - \frac{(ax)^2}{2} + \frac{(ax)^3}{3} - \frac{(ax)^4}{4} + \dots\right) + \left(bx - \frac{(bx)^2}{2} + \frac{(bx)^3}{3} - \frac{(bx)^4}{4} + \dots\right) \\
&= \ln(1+ax) + \ln(1+bx)\,.
\end{aligned}$$

Assignment B6

Problem 1. Expand $(1+2x)^{-\frac{1}{3}}$, where $|x|<\frac{1}{2}$, as a series of ascending powers of x, up to an including the term in x^2, simplifying the coefficients.

By choosing $x=\frac{1}{14}$, find an approximate value of $\sqrt[3]{7}$ in the form $\frac{p}{q}$, where p and q are to be determined.

Using your calculator, calculate the numerical value of $\sqrt[3]{7}$. Compare this value to the approximate value found, and with reference to the value of x chosen, comment on the accuracy of your approximation.

Solution.

$$(1+2x)^{-1/3}=1-\frac{1}{3}(2x)+\frac{\left(-\frac{1}{3}\right)\left(-\frac{1}{3}-1\right)}{2}(2x)^2+\cdots=1-\frac{2}{3}x+\frac{8}{9}x^2+\cdots.$$

Substituting $x=1/14$,

$$\left[1+2\left(\frac{1}{14}\right)\right]^{-1/3}=\frac{\sqrt[3]{7}}{2}\approx 1-\frac{2}{3}\left(\frac{1}{14}\right)+\frac{8}{9}\left(\frac{1}{14}\right)^2=\frac{422}{441}$$
$$\implies \sqrt[3]{7}\approx\frac{844}{441}=1.9138 \text{ (5 s.f.)}.$$

Since $\sqrt[3]{7}=1.9129$ (5 s.f.), the approximation is accurate.

* * * * *

Problem 2. In the triangle ABC, $AB=1$, $BC=3$ and angle $ABC=\theta$ radians. Given that θ is a sufficiently small angle, show that

$$AC\approx(4+3\theta^2)^{\frac{1}{2}}\approx a+b\theta^2$$

for constants a and b to be determined.

Solution. By the cosine rule,

$$AC^2=AB^2+BC^2-2(AB)(BC)\cos ABC=1^2+3^2-2(1)(3)\cos\theta=10-6\cos\theta.$$

Since θ is sufficiently small, $\cos\theta\approx 1-\theta^2/2$. Hence,

$$AC^2\approx 10-6\left(1-\frac{\theta^2}{2}\right)=4+3\theta^2$$
$$\implies AC=\left(4+3\theta^2\right)^{1/2}=2\left(1+\frac{3\theta^2}{4}\right)^{1/2}\approx 2\left[1+\frac{1}{2}\left(\frac{3\theta^2}{4}\right)\right]=2+\frac{3\theta^2}{4}.$$

Hence, $a=2$ and $b=\frac{3}{4}$.

* * * * *

Problem 3. Given that $y=\ln\sec x$, show that

(a) $\frac{\mathrm{d}^3y}{\mathrm{d}x^3}=2\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\frac{\mathrm{d}y}{\mathrm{d}x}$

(b) the value of $\frac{\mathrm{d}^4y}{\mathrm{d}x^4}$ when $x=0$ is 2.

Write down the MacLaurin series for $\ln\sec x$ up to and including the term in x^4. By substituting $x=\frac{\pi}{4}$, show that $\ln 2\approx\frac{\pi^2}{16}+\frac{\pi^4}{1536}$.

Solution.

Part (a). Note that

$$y = \ln \sec x = -\ln \cos x \implies \mathrm{e}^{-y} = \cos x.$$

Implicitly differentiating with respect to x,

$$-y'\mathrm{e}^{-y} = -\sin x \implies y' = \tan x.$$

Differentiating repeatedly,

$$y'' = \sec^2 x \implies y''' = 2\sec^2 x \tan x.$$

Thus,

$$y''' = 2\sec^2 x \tan x = 2y'' \cdot y'.$$

Part (b). Implicitly differentiating the above differential equation,

$$y^{(4)} = 2\left[y''' \cdot y' + \left(y''\right)^2\right].$$

Evaluating the above equations at $x = 0$, we see that

$$y(0) = 0, \quad y'(0) = 0, \quad y''(0) = 1, \quad y^{(3)}(0) = 0, \quad y^{(4)}(0) = 2.$$

We have

$$\ln \sec x = \frac{x^2}{2} + \frac{x^4}{12} + \cdots .$$

Substituting $x = \pi/4$,

$$\ln \sec \frac{\pi}{4} = \frac{1}{2}\ln 2 \approx \frac{1}{2}\left(\frac{\pi}{4}\right)^2 + \frac{1}{12}\left(\frac{\pi}{4}\right)^4 = \frac{\pi^2}{32} + \frac{\pi^4}{3072} \implies \ln 2 \approx \frac{\pi^2}{16} + \frac{\pi^4}{1536}.$$

B7 Integration Techniques

Tutorial B7

Problem 1. Find

(a) $\int \frac{1}{\sqrt{3-2x}}\,\mathrm{d}x$

(b) $\int \frac{1}{3-2x}\,\mathrm{d}x$

(c) $\int \frac{1}{3-2x^2}\,\mathrm{d}x$

(d) $\int \frac{1}{\sqrt{3-2x^2}}\,\mathrm{d}x$

(e) $\int \frac{x}{\sqrt{3-2x^2}}\,\mathrm{d}x$

(f) $\int \frac{1}{3+4x+2x^2}\,\mathrm{d}x$

Solution.

Part (a). Consider the substitution $u = 3 - 2x$.

$$\int \frac{1}{\sqrt{3-2x}}\,\mathrm{d}x = -\int \frac{1}{2\sqrt{u}}\,\mathrm{d}u = -\sqrt{u} + C = -\sqrt{3-2x} + C.$$

Part (b). Consider the substitution $u = 3 - 2x$.

$$\int \frac{1}{3-2x}\,\mathrm{d}x = -\frac{1}{2}\int \frac{1}{u}\,\mathrm{d}u = -\frac{1}{2}\ln|u| + C = -\frac{1}{2}\ln|3-2x| + C.$$

Part (c).

$$\begin{aligned}\int \frac{1}{3-2x^2}\,\mathrm{d}x = \frac{1}{2}\int \frac{1}{3/2-x^2}\,\mathrm{d}x &= \frac{1}{2}\left(\frac{1}{2\sqrt{3/2}}\right)\ln\left(\frac{\sqrt{3/2}+x}{\sqrt{3/2}-x}\right) + C \\ &= \frac{1}{2\sqrt{6}}\ln\left(\frac{\sqrt{3}+\sqrt{2}x}{\sqrt{3}-\sqrt{2}x}\right) + C.\end{aligned}$$

Part (d).

$$\begin{aligned}\int \frac{1}{\sqrt{3-2x^2}}\,\mathrm{d}x = \frac{1}{\sqrt{2}}\int \frac{1}{\sqrt{3/2-x^2}}\,\mathrm{d}x &= \frac{1}{\sqrt{2}}\arcsin\left(\frac{x}{\sqrt{3/2}}\right) + C \\ &= \frac{\sqrt{2}}{2}\arcsin\left(\frac{\sqrt{6}x}{3}\right) + C.\end{aligned}$$

Part (e). Consider the substitution $u = 3 - 2x^2$.

$$\int \frac{x}{\sqrt{3-2x^2}}\,\mathrm{d}x = -\frac{1}{2}\int \frac{1}{2\sqrt{u}}\,\mathrm{d}u = -\frac{\sqrt{u}}{2} + C = -\frac{\sqrt{3-2x^2}}{2} + C.$$

Part (f).

$$\int \frac{1}{3+4x+2x^2}\,\mathrm{d}x = \frac{1}{2}\int \frac{1}{(x+1)^2+1/2}\,\mathrm{d}x = \frac{1}{2}\left(\frac{1}{\sqrt{1/2}}\right)\arctan\left(\frac{x+1}{1/\sqrt{1/2}}\right)+C$$
$$= \frac{\arctan\big(\sqrt{2}(x+1)\big)}{\sqrt{2}}+C.$$

* * * * *

Problem 2. Find

(a) $\int \frac{\sec^2 3x}{\tan 3x}\,\mathrm{d}x$

(b) $\int \cos(3x+\alpha)\,\mathrm{d}x$, where α is a constant

(c) $\int \cos^2 3x\,\mathrm{d}x$

(d) $\int \mathrm{e}^{1-2x}\,\mathrm{d}x$

Solution.

Part (a).

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan 3x = 3\sec^2 3x \implies \int \frac{\sec^2 3x}{\tan 3x}\,\mathrm{d}x = \frac{1}{3}\int \frac{3\sec^2 3x}{\tan 3x}\,\mathrm{d}x = \frac{\ln\tan 3x}{3}+C.$$

Part (b).

$$\int \cos(3x+\alpha)\,\mathrm{d}x = \frac{\sin(3x+\alpha)}{3}+C$$

Part (c). Recall that

$$\cos^2\theta = \frac{1+\cos 2\theta}{2} \implies \cos^2 3x = \frac{1+\cos 6x}{2}.$$

Thus,

$$\int \cos^2 3x\,\mathrm{d}x = \frac{1}{2}\int (1+\cos 6x)\,\mathrm{d}x = \frac{1}{2}\left(x+\frac{\sin 6x}{6}\right)+C = \frac{x}{2}+\frac{\sin 6x}{12}+C.$$

Part (d).

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{1-2x} = -2\mathrm{e}^{1-2x} \implies \int \mathrm{e}^{1-2x}\,\mathrm{d}x = -\frac{1}{2}\mathrm{e}^{1-2x}+C.$$

* * * * *

Problem 3. Find

(a) $\int 2x\sqrt{3x^2-5}\,\mathrm{d}x$

(b) $\int \frac{x^2-1}{\sqrt{x^3-3x}}\,\mathrm{d}x$

(c) $\int \sin x\sqrt{\cos x}\,\mathrm{d}x$

(d) $\int \mathrm{e}^{2x}\left(1-\mathrm{e}^{2x}\right)^4\,\mathrm{d}x$

Solution.

Part (a). Consider the substitution $u = 3x^2 - 5$.

$$\int 2x\sqrt{3x^2-5}\,\mathrm{d}x = \frac{1}{3}\int \sqrt{u}\,\mathrm{d}u = \frac{1}{3}\left(\frac{2}{3}u^{3/2}\right) + C = \frac{2}{9}\left(3x^2-5\right)^{3/2} + C.$$

Part (b). Consider the substitution $u = x^3 - 3x$:

$$\int \frac{x^2-1}{\sqrt{x^3-3x}}\,\mathrm{d}x = \frac{2}{3}\int \frac{\mathrm{d}u}{2\sqrt{u}} = \frac{2}{3}\sqrt{u} + C = \frac{2}{3}\sqrt{x^3-3x} + C.$$

Part (c). Consider the substitution $u = \cos x$.

$$\int \sin x\sqrt{\cos x}\,\mathrm{d}x = -\int \sqrt{u}\,\mathrm{d}u = -\frac{2}{3}u^{3/2} + C = -\frac{2}{3}\cos^{3/2} x + C.$$

Part (d). Consider the substitution $u = 1 - \mathrm{e}^{2x}$.

$$\int \mathrm{e}^{2x}(1-\mathrm{e}^{2x})^4\,\mathrm{d}x = -\frac{1}{2}\int u^4\,\mathrm{d}u = -\frac{1}{2}\left(\frac{u^5}{5}\right) + C = -\frac{\left(1-\mathrm{e}^{2x}\right)^5}{10} + C.$$

* * * * *

Problem 4. Find

(a) $\int \frac{1}{\sqrt{x}(1-\sqrt{x})}\,\mathrm{d}x$

(b) $\int \frac{3x}{x+3}\,\mathrm{d}x$

(c) $\int \frac{\sin x + \cos x}{\sin x - \cos x}\,\mathrm{d}x$

Solution.

Part (a). Consider the substitution $u = 1 - \sqrt{x}$.

$$\int \frac{1}{\sqrt{x}(1-\sqrt{x})}\,\mathrm{d}x = -2\int \frac{1}{u}\,\mathrm{d}u = -2\ln|u| + C = -2\ln\left|1-\sqrt{x}\right| + C.$$

Part (b).

$$\int \frac{3x}{x+3}\,\mathrm{d}x = \int \left(3 - \frac{9}{x+3}\right)\mathrm{d}x = 3x - 9\ln|x+3| + C.$$

Part (c). Consider the substitution $u = \sin x - \cos x$.

$$\int \frac{\sin x + \cos x}{\sin x - \cos x}\,\mathrm{d}x = \int \frac{1}{u}\,\mathrm{d}u = \ln|u| + C = \ln|\sin x - \cos x| + C.$$

* * * * *

Problem 5. Find

(a) $\int \frac{\mathrm{e}^{-\sqrt{x}}}{\sqrt{x}}\,\mathrm{d}x$

(b) $\int (\sin x)(\cos x)(\mathrm{e}^{\cos 2x})\,\mathrm{d}x$

Solution.

Part (a). Consider the substitution $u = -\sqrt{x}$.

$$\int \frac{e^{-\sqrt{x}}}{\sqrt{x}}\,dx = -2\int e^u\,du = -2e^u + C = -2e^{-\sqrt{x}} + C.$$

Part (b). Consider the substitution $u = \cos 2x$.

$$\int (\sin x)(\cos x)(e^{\cos 2x})\,dx = \frac{1}{2}\int e^{\cos 2x}\sin 2x\,dx = -\frac{1}{4}\int e^u\,du$$
$$= -\frac{e^u}{4} + C = -\frac{e^{\cos 2x}}{4} + C.$$

* * * * *

Problem 6. Find

(a) $\int \tan^2 2x\,dx$

(b) $\int \frac{1}{1+\cos 2t}\,dt$

(c) $\int \sin\left(\frac{5}{2}\theta\right)\cos\left(\frac{1}{2}\theta\right)\,d\theta$

(d) $\int \tan^4 x\,dx$

Solution.

Part (a).

$$\int \tan^2 2x\,dx = \int \left(\sec^2 2x - 1\right)\,dx = \frac{\tan 2x}{2} - x + C.$$

Part (b). Note that

$$\frac{1}{1+\cos 2t} = \frac{1}{1+(2\cos^2 t - 1)} = \frac{\sec^2 t}{2}.$$

Hence,

$$\int \frac{1}{1+\cos 2t}\,dt = \frac{1}{2}\int \sec^2 t\,dt = \frac{\tan t}{2} + C.$$

Part (c). By the product-to-sum identity,

$$\sin\left(\frac{5\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) = \frac{\sin 3\theta + \sin 2\theta}{2}.$$

Hence,

$$\int \sin\left(\frac{5}{2}\theta\right)\cos\left(\frac{1}{2}\theta\right)\,d\theta = \frac{1}{2}\int (\sin 3\theta + \sin 2\theta)\,d\theta = -\frac{\cos 3\theta}{6} - \frac{\cos 2\theta}{4} + C$$

Part (d). Note that

$$\int \tan^2 x\,dx = \tan x - x + C$$

and

$$\int \tan^2 x \sec^2 x\,dx = \frac{\tan^3 x}{3} + C.$$

Hence,

$$\int \tan^4 x \, dx = \int \tan^2 x \left(\sec^2 x - 1\right) dx = \int \left(\tan^2 x \sec^2 x - \tan^2\right) dx$$
$$= \frac{\tan^3 x}{3} - \tan x + x + C$$

* * * * *

Problem 7. Find

(a) $\int \frac{1}{4x^2+2x+10} \, dx$

(b) $\int \frac{x^2}{1-x^2} \, dx$

(c) $\int \frac{1}{\sqrt{3+2x-x^2}} \, dx$

Solution.

Part (a).

$$\int \frac{1}{4x^2+2x+10} \, dx = 4\int \frac{1}{(4x+1)^2+39} \, dx = 4\left(\frac{1}{4}\right)\left(\frac{1}{\sqrt{39}}\right)\arctan\left(\frac{4x+1}{\sqrt{39}}\right) + C$$
$$= \frac{1}{\sqrt{39}} \arctan\left(\frac{4x+1}{\sqrt{39}}\right) + C.$$

Part (b).

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

Part (c).

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

* * * * *

Problem 8. Evaluate the following without the use of graphic calculator:

(a) $\int_{\pi/3}^{2\pi/3} 4\cot\frac{x}{2}\csc^2\frac{x}{2} \, dx$

(b) $\int_0^4 \frac{x+2}{\sqrt{2x+1}} \, dx$

(c) $\int_0^1 \frac{2}{(1+x)(1+x^2)} \, dx$

(d) $\int_{-4}^{-2} \frac{x^3+2}{x^2-1} \, dx$

Solution.

Part (a).

$$\int_{\frac{1}{3}\pi}^{\frac{2}{3}\pi} 4\cot\frac{x}{2}\csc^2\frac{x}{2} \, dx = -4\int_{\frac{1}{3}\pi}^{\frac{2}{3}\pi} \cot\frac{x}{2}\left(-\csc^2\frac{x}{2}\right) dx = -8\left[\frac{\tan^2(x/2)}{2}\right]_{\frac{1}{3}\pi}^{\frac{2}{3}\pi} = \frac{32}{3}.$$

Part (b). Consider the substitution $u = 2x + 1$.

$$\int_0^4 \frac{x+2}{\sqrt{2x+1}}\,\mathrm{d}x = \frac{1}{2}\int_0^4 \left(\sqrt{2x+1} + \frac{3}{\sqrt{2x+1}}\right)\mathrm{d}x$$
$$= \frac{1}{4}\int_1^9 \left(\sqrt{u} + \frac{3}{\sqrt{u}}\right)\mathrm{d}u = \frac{1}{4}\left[\frac{u^{3/2}}{3/2} + \frac{3u^{1/2}}{1/2}\right]_1^9 = \frac{22}{3}.$$

Part (c). Note that

$$\int_0^1 \frac{x}{1+x^2}\,\mathrm{d}x = \frac{1}{2}\int_0^1 \frac{2x}{1+x^2}\,\mathrm{d}x = \frac{1}{2}\left[\ln|1+x^2|\right]_0^1 = \frac{\ln 2}{2}.$$

Thus,

$$\int_0^1 \frac{2}{(1+x)(1+x^2)}\,\mathrm{d}x = \int_0^1 \left(\frac{1}{1+x} + \frac{1}{1+x^2} - \frac{x}{1+x^2}\right)\mathrm{d}x$$
$$= \left[\ln|1+x|\right]_0^1 + \left[\arctan x\right]_0^1 - \frac{1}{2}\ln 2 = \ln 2 + \frac{\pi}{4} - \frac{1}{2}\ln 2 = \frac{1}{2}\ln 2 + \frac{\pi}{4}.$$

Part (d).

$$\int_{-4}^{-2} \frac{x^3+2}{x^2-1}\,\mathrm{d}x = \int_{-4}^{-2}\left(x + \frac{3/2}{x-1} - \frac{1/2}{x+1}\right)\mathrm{d}x = \left[\frac{x^2}{2} + \frac{3}{2}\ln|x-1| - \frac{1}{2}\ln|x+1|\right]_{-4}^{-2}$$
$$= -6 + 2\ln 3 - \frac{3}{2}\ln 5.$$

* * * * *

Problem 9. Using the given substitution, find

(a) $\int \frac{x}{(2x+3)^3}\,\mathrm{d}x$ $[u = 2x+3]$

(b) $\int \frac{1}{\mathrm{e}^x + 4\mathrm{e}^{-x}}\,\mathrm{d}x$ $[u = \mathrm{e}^x]$

(c) $\int_0^{\sqrt{2}} \sqrt{4-y^2}\,\mathrm{d}y$ $[y = 2\sin\theta]$

(d) $\int_0^{\pi/2} \frac{1}{1+\sin\theta}\,\mathrm{d}\theta$ $\left[t = \tan\frac{\theta}{2}\right]$

Solution.

Part (a). Using the substitution $u = 2x + 3$,

$$\int \frac{x}{(2x+3)^3}\,\mathrm{d}x = \frac{1}{4}\int \frac{u-3}{u^3}\,\mathrm{d}x = \frac{1}{4}\int\left(\frac{1}{u^2} - \frac{3}{u^3}\right)\mathrm{d}u = \frac{1}{4}\left(\frac{u^{-1}}{-1} - \frac{3u^{-2}}{-2}\right) + C$$
$$= \frac{3}{8}(2x+3)^{-2} - \frac{1}{4}(2x+3)^{-1} + C.$$

Part (b). Using the substitution $u = \mathrm{e}^x$,

$$\int \frac{1}{\mathrm{e}^x + 4\mathrm{e}^{-x}}\,\mathrm{d}x = \int \frac{e^x}{e^{2x}+4}\,\mathrm{d}x = \int \frac{1}{u^2+4}\,\mathrm{d}u$$
$$= \frac{1}{2}\arctan\left(\frac{u}{2}\right) + C = \frac{1}{2}\arctan\left(\frac{\mathrm{e}^x}{2}\right) + C.$$

Part (c). Using the substitution $y = 2\sin\theta$,

$$\int_0^{\sqrt{2}} \sqrt{4-y^2}\,\mathrm{d}y = 2\int_0^{\pi/4} \cos\theta\sqrt{4-4\sin^2\theta}\,\mathrm{d}\theta = 4\int_0^{\pi/4} \cos\theta\sqrt{1-\sin^2\theta}\,\mathrm{d}\theta$$
$$= 4\int_0^{\pi/4} \cos^2\theta\,\mathrm{d}\theta = 4\int_0^{\pi/4} \frac{1+\cos 2\theta}{2}\,\mathrm{d}\theta = 2\left[\theta + \frac{\sin 2\theta}{2}\right]_0^{\pi/4} = 1 + \frac{\pi}{2}.$$

Part (d). Consider the substitution $t = \tan\frac{\theta}{2}$. Then

$$\theta = 2\arctan t \implies \mathrm{d}\theta = \frac{2}{1+t^2}\,\mathrm{d}t$$

and

$$\sin\theta = \sin(2\arctan t) = 2\sin(\arctan t)\cos(\arctan t) = 2\left(\frac{t}{\sqrt{1+t^2}}\right)\left(\frac{1}{\sqrt{1+t^2}}\right) = \frac{2t}{1+t^2}.$$

Hence,

$$\int_0^{\pi/2} \frac{1}{1+\sin\theta}\,\mathrm{d}\theta = \int_0^1 \frac{2/(1+t^2)}{1+2t/(1+t^2)}\,\mathrm{d}u = \int_0^1 \frac{2}{(t+1)^2}\,\mathrm{d}t = 2\left[-\frac{1}{t+1}\right]_0^1 = 1.$$

* * * * *

Problem 10. Find

(a) $\int \ln(2x+1)\ \mathrm{d}x$

(b) $\int x\arctan(x^2)\,\mathrm{d}x$

(c) $\int \mathrm{e}^{-2x}\cos 2x\,\mathrm{d}x$

(d) $\int_0^2 x^2e^{-x}\,\mathrm{d}x$

Solution.

Part (a). Consider the substitution $u = 2x+1$.

$$\int \ln(2x+1)\ \mathrm{d}x = \frac{1}{2}\int \ln u\,\mathrm{d}u.$$

Integrating by parts,

	D	I
$+$	$\ln u$	1
$-$	$1/u$	u

Thus,

$$\int \ln(2x+1)\ \mathrm{d}x = \frac{1}{2}\left(u\ln u - \int u\left(\frac{1}{u}\right)\,\mathrm{d}u\right) = \frac{u\ln u - u}{2} + C$$
$$= \frac{(2x+1)\ln(2x+1) - (2x+1)}{2} + C = x\ln(2x+1) + \frac{\ln(2x+1)}{2} - x + C.$$

Part (b). Consider the substitution $u = x^2$.

$$\int x\arctan(x^2)\,\mathrm{d}x = \frac{1}{2}\int \arctan u\,\mathrm{d}u.$$

Integrating by parts,

	D	I
$+$	$\arctan u$	1
$-$	$1/(1+u^2)$	u

Thus,

$$\int x \arctan(x^2)\,\mathrm{d}x = \frac{1}{2}\left(u \arctan u - \int \frac{u}{1+u^2}\,\mathrm{d}u\right) = \frac{1}{2}\left[u \arctan u - \frac{\ln\left(1+u^2\right)}{2}\right] + C$$
$$= \frac{x^2 \arctan x^2}{2} - \frac{\ln\left(1+x^4\right)}{4} + C.$$

Part (c). Let

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

Integrating by parts, we have

	D	I
$+$	e^{-2x}	$\cos 2x$
$-$	$-2\mathrm{e}^{-2x}$	$\sin(2x)/2$
$+$	$4\mathrm{e}^{-2x}$	$-\cos(2x)/4$

Thus,

$$I = \frac{\mathrm{e}^{-2x}\sin 2x}{2} - \frac{\mathrm{e}^{-2x}\cos 2x}{2} - I \implies I = \frac{\mathrm{e}^{-2x}\left(\sin 2x - \cos 2x\right)}{4} + C.$$

Part (d). Integrating by parts, we get

	D	I
$+$	x^2	e^{-x}
$-$	$2x$	$-\mathrm{e}^{-x}$
$+$	2	e^{-x}
$-$	0	$-\mathrm{e}^{-x}$

Thus,

$$\int_0^2 x^2\mathrm{e}^{-x}\,\mathrm{d}x = \left[-x^2\mathrm{e}^{-x} - 2x\mathrm{e}^{-x} - 2\mathrm{e}^{-x}\right]_0^2 = 2 - 10\mathrm{e}^{-2}.$$

* * * * *

Problem 11.

(a) Show that $\frac{\mathrm{d}}{\mathrm{d}x}\ln(\sec x + \tan x) = \sec x$.

(b) Find $\int x \sin x\,\mathrm{d}x$.

(c) Find the exact value of $\int_0^{\pi/4} (x \sin x) \ln(\sec x + \tan x)\ \mathrm{d}x$.

Solution.

Part (a).

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln(\sec x+\tan x)=\frac{\sec x\tan x+\sec^2 x}{\sec x+\tan x}=\sec x\left(\frac{\tan x+\sec x}{\sec x+\tan x}\right)=\sec x.$$

Part (b). Integrating by parts,

	D	I
$+$	x	$\sin x$
$-$	1	$-\cos x$
$+$	0	$-\sin x$

Hence,

$$\int x\sin x\,\mathrm{d}x=-x\cos x+\sin x+C.$$

Part (c). Integrating by parts,

	D	I
$+$	$\ln(\sec x+\tan x)$	$x\sin x$
$-$	$\sec x$	$-x\cos x+\sin x$

Thus,

$$\begin{aligned}
&\int_0^{\pi/4}(x\sin x)\ln(\sec x+\tan x)\ \mathrm{d}x\\
&=\left[\ln(\sec x+\tan x)\,(-x\cos x+\sin x)\right]_0^{\pi/4}-\int_0^{\pi/4}(-x+\tan x)\ \mathrm{d}x\\
&=\left[\ln(\sec x+\tan x)\,(-x\cos x+\sin x)-\frac{x^2}{2}-\ln|\cos x|\right]_0^{\pi/4}\\
&=\frac{\sqrt{2}}{2}\left(1-\frac{\pi}{4}\right)\ln\left(\sqrt{2}+1\right)+\frac{\pi^2}{32}-\frac{\ln 2}{2}
\end{aligned}$$

* * * * *

Problem 12.

(a) Use the fact that $7\cos x-4\sin x=\frac{3}{2}(\cos x+\sin x)+\frac{11}{2}(\cos x-\sin x)$ to find the exact value of $\int_0^{\pi/2}\frac{7\cos x-4\sin x}{\cos x+\sin x}\,\mathrm{d}x$.

(b) Use integration by parts to find the exact value of $\int_1^{\mathrm{e}}(\ln x)^2\,\mathrm{d}x$.

Solution.

Part (a). Note that

$$\frac{7\cos x-4\sin x}{\cos x+\sin x}=\frac{1}{2}\left(\frac{3(\cos x+\sin x)+11(\cos x-\sin x)}{\cos x+\sin x}\right)=\frac{3}{2}+\frac{11}{2}\left(\frac{\cos x-\sin x}{\cos x+\sin x}\right).$$

Thus,

$$\begin{aligned}
\int_0^{\pi/2}\frac{7\cos x-4\sin x}{\cos x+\sin x}\,\mathrm{d}x&=\frac{1}{2}\int_0^{\pi/2}\left(3+11\cdot\frac{\cos x-\sin x}{\cos x+\sin x}\right)\,\mathrm{d}x\\
&=\left[\frac{3x}{2}+\frac{11}{2}\ln|\cos x+\sin x|\right]_0^{\pi/2}=\frac{3\pi}{4}.
\end{aligned}$$

Part (b). Integrating by parts,

	D	I
$+$	$(\ln x)^2$	1
$-$	$\frac{2\ln x}{x}$	x

Thus,

$$\int_1^{\mathrm{e}} (\ln x)^2 \, \mathrm{d}x = \left[x(\ln x)^2\right]_1^{\mathrm{e}} - 2\int_1^{\mathrm{e}} \ln x \, \mathrm{d}x = \left[x(\ln x)^2 - 2\left(x \ln x - x\right)\right]_1^{\mathrm{e}} = \mathrm{e} - 2.$$

* * * * *

Problem 13.

(a) Solve the inequality $x^2 + 2x - 3 < 0$.

(b) Without using the graphing calculator, evaluate

  (i) $\int_{-4}^{4} \left|x^2 + 2x - 3\right| \, \mathrm{d}x$

  (ii) $\int_0^2 x\left|x^2 + 2x - 3\right| \, \mathrm{d}x$

Solution.

Part (a).

$$x^2 + 2x - 3 = (x+1)^2 - 4 < 0 \implies (x+1)^2 < 4 \implies -2 < x + 1 < 2 \implies -3 < x < 1.$$

Part (b).

Part (b)(i). Let $F(x) = \int \left(x^2 + 2x - 3\right) \, \mathrm{d}x = \frac{1}{3}x^3 + x^2 - 3x + C$. Then,

$$\begin{aligned}
&\int_{-4}^{4} \left|x^2 + 2x - 3\right| \, \mathrm{d}x \\
&\quad = \int_{-4}^{-3} \left|x^2 + 2x - 3\right| \, \mathrm{d}x + \int_{-3}^{1} \left|x^2 + 2x - 3\right| \, \mathrm{d}x + \int_{1}^{4} \left|x^2 + 2x - 3\right| \, \mathrm{d}x \\
&\quad = \int_{-4}^{-3} x^2 + 2x - 3 \, \mathrm{d}x - \int_{-3}^{1} x^2 + 2x - 3 \, \mathrm{d}x + \int_{1}^{4} x^2 + 2x - 3 \, \mathrm{d}x \\
&\quad = \Big[F(-3) - F(-4)\Big] - \Big[F(1) - F(-3)\Big] + \Big[F(4) - F(1)\Big] = 40.
\end{aligned}$$

Part (b)(ii). Let $F(x) = \int x\left(x^2 + 2x - 3\right) \, \mathrm{d}x = \frac{1}{4}x^4 + \frac{2}{3}x^3 - \frac{3}{2}x^2 + C$. Then,

$$\begin{aligned}
&\int_0^2 x\left|x^2 + 2x - 3\right| \, \mathrm{d}x \\
&\quad = \int_0^1 x\left|x^2 + 2x - 3\right| \, \mathrm{d}x + \int_1^2 x\left|x^2 + 2x - 3\right| \, \mathrm{d}x \\
&\quad = -\int_0^1 x(x^2 + 2x - 3) \, \mathrm{d}x + \int_1^2 x(x^2 + 2x - 3) \, \mathrm{d}x \\
&\quad = -\Big[F(1) - F(0)\Big] + \Big[F(2) - F(1)\Big] = \frac{9}{2}.
\end{aligned}$$

Problem 14. The indefinite integral $\int \frac{P(x)}{x^3+1}\,\mathrm{d}x$, where $P(x)$ is a polynomial in x, is denoted by I.

(a) Find I when $P(x) = x^2$.

(b) By writing $x^3 + 1 = (x+1)\left(x^2 + Ax + B\right)$, where A and B are constants, find I when

(i) $P(x) = x^2 - x + 1$

(ii) $P(x) = x + 1$

(c) Using the results of parts (a) and (b), or otherwise, find I when $P(x) = 1$.

Solution.

Part (a).

$$\int \frac{x^2}{x^3+1}\,\mathrm{d}x = \frac{1}{3}\frac{3x^2}{x^3+1}\,\mathrm{d}x = \frac{\ln\left|x^3+1\right|}{3} + C.$$

Part (b).

$$x^3 + 1 = (x+1)\left(x^2 - x + 1\right).$$

Part (b)(i).

$$\int \frac{x^2-x+1}{x^3+1}\,\mathrm{d}x = \int \frac{x^2-x+1}{(x+1)\left(x^2-x+1\right)}\,\mathrm{d}x = \int \frac{1}{x+1}\,\mathrm{d}x = \ln|x+1| + C.$$

Part (b)(ii).

$$\begin{aligned}\int \frac{x+1}{x^3+1}\,\mathrm{d}x &= \int \frac{x+1}{(x+1)\left(x^2-x+1\right)}\,\mathrm{d}x = \int \frac{1}{x^2-x+1}\,\mathrm{d}x = \int \frac{1}{(x-1/2)^2+3/4}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{3/4}}\arctan\left(\frac{x-1/2}{\sqrt{3/4}}\right) + C = \frac{2}{\sqrt{3}}\arctan\left(\frac{2x-1}{\sqrt{3}}\right) + C.\end{aligned}$$

Part (c). Observe that $1 = \frac{1}{2}\left[\left(x^2-x+1\right) - x^2 + (x+1)\right]$. Hence,

$$\begin{aligned}\int \frac{1}{x^3+1}\,\mathrm{d}x &= \frac{1}{2}\left(\int \frac{x^2-x+1}{x^3+1}\,\mathrm{d}x - \int \frac{x^2}{x^3+1}\,\mathrm{d}x + \int \frac{x+1}{x^3+1}\,\mathrm{d}x\right) \\ &= \frac{1}{2}\left[\ln|x+1| - \frac{\ln\left|x^3+1\right|}{3} + \frac{2}{\sqrt{3}}\arctan\left(\frac{2x-1}{\sqrt{3}}\right)\right] + C \\ &= \frac{1}{2}\ln|x+1| - \frac{\ln\left|x^3+1\right|}{6} + \frac{1}{\sqrt{3}}\arctan\left(\frac{2x-1}{\sqrt{3}}\right) + C.\end{aligned}$$

Self-Practice B7

Problem 1. Find

(a) $\int \tan\left(\frac{\pi}{6} - 3x\right) \mathrm{d}x$,

(b) $\int \tan x \sec^4 x \, \mathrm{d}x$,

(c) $\int \frac{1}{x^2+3x+2} \, \mathrm{d}x$,

(d) $\int \cos \frac{3x}{2} \cos \frac{5x}{2} \, \mathrm{d}x$,

(e) $\int \frac{2}{x \ln x^2} \, \mathrm{d}x$,

(f) $\int x\mathrm{e}^{-x^2} \, \mathrm{d}x$.

Solution.

Part (a).

$$\int \tan\left(\frac{\pi}{6} - 3x\right) \mathrm{d}x = -\frac{1}{3} \int -3\tan\left(\frac{\pi}{6} - 3x\right) \mathrm{d}x = -\frac{1}{3} \ln\left|\sec\left(\frac{\pi}{6} - 3x\right)\right| + C.$$

Part (b).

$$\int \tan x \sec^4 x \, \mathrm{d}x = \int (\sec x \tan x) \sec^3 x \, \mathrm{d}x = \frac{\sec^4 x}{4} + C.$$

Part (c).

$$\int \frac{1}{x^2 + 3x + 2} \, \mathrm{d}x = \int \left(\frac{1}{x+1} - \frac{1}{x+2}\right) \mathrm{d}x = \ln|x+1| - \ln|x+2| + C = \ln\left|\frac{x+1}{x+2}\right| + C.$$

Part (d).

$$\int \cos \frac{3x}{2} \cos \frac{5x}{2} \, \mathrm{d}x = \int (\cos 4x + \cos x) \, \mathrm{d}x = \frac{\sin 4x}{8} + \frac{\sin x}{2} + C.$$

Part (e).

$$\int \frac{2}{x \ln x^2} \, \mathrm{d}x = \int \frac{1/x}{\ln x} \, \mathrm{d}x = \ln|\ln x| + C.$$

Part (f).

$$\int x\mathrm{e}^{-x^2} \, \mathrm{d}x = -\frac{1}{2} \int -2x\mathrm{e}^{-x^2} \, \mathrm{d}x = -\frac{1}{2}\mathrm{e}^{-x^2} + C.$$

* * * * *

Problem 2. Using the substitution $x = \tan\theta$, find the exact value of $\int_0^1 \frac{1-x^2}{(1+x^2)^2} \, \mathrm{d}x$.

Solution. Note that $1 + x^2 = 1 + \tan^2\theta = \sec^2\theta$. Hence,

$$\int_0^1 \frac{1 - x^2}{(1 + x^2)^2} \, \mathrm{d}x = \int_0^{\pi/4} \frac{1 - \tan^2\theta}{\sec^4\theta} \left(\sec^2\theta \, \mathrm{d}\theta\right) = \int_0^{\pi/4} \frac{1 - \tan^2\theta}{\sec^2\theta} \, \mathrm{d}\theta.$$

Using trigonometric identities to simplify the integrand, we get

$$\int_0^{\pi/4} \frac{1 - \tan^2\theta}{\sec^2\theta} \, \mathrm{d}\theta = \int_0^{\pi/4} \cos 2\theta \, \mathrm{d}t = \left[\frac{1}{2} \sin 2\theta\right]_0^{\pi/4} = \frac{1}{2}.$$

Problem 3. State the derivative of $\sin x^2$. Hence, find $\int x^3 \cos x^2 \, \mathrm{d}x$.

Solution. We have

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

Consider the substitution $u = \sin x^2$. Using the above result, we have

$$\int x^2 \cos x^2 \, \mathrm{d}x = \frac{1}{2} \int \left(2x \cos x^2\right) x^2 \, \mathrm{d}x = \frac{1}{2} \int \arcsin u \, \mathrm{d}u.$$

Integrating by parts, we get

$$\frac{1}{2} \left(u \arcsin u - \int \frac{u}{\sqrt{1-u^2}} \, \mathrm{d}u \right).$$

The integral is fairly simple to evaluate:

$$\int \frac{u}{\sqrt{1-u^2}} = -\frac{1}{2} \int \frac{-2u}{\sqrt{1-u^2}} \, \mathrm{d}u = -\sqrt{1-u^2} + C.$$

Thus,

$$\int x^3 \cos x^2 \, \mathrm{d}x = \frac{1}{2} \left(x^2 \sin x^2 + \sqrt{1 - \sin^2 x^2} \right) + C = \frac{1}{2} \left(x^2 \sin x^2 + \cos x^2 \right) + C.$$

* * * * *

Problem 4. Find the exact value of p such that $\int_0^1 \frac{1}{4-x^2} \, \mathrm{d}x = \int_0^{1/2p} \frac{1}{\sqrt{1-p^2x^2}} \, \mathrm{d}x$.

Solution. Using standard integration results, the LHS evaluates to

$$\int_0^1 \frac{1}{4-x^2} \, \mathrm{d}x = \left[\frac{1}{4} \ln \frac{2+x}{2-x} \right]_0^1 = \frac{1}{4} \ln 3.$$

Meanwhile, under the substitution $u = px$, the RHS evaluates as

$$\int_0^{1/2p} \frac{1}{\sqrt{1-p^2x^2}} \, \mathrm{d}x = \frac{1}{p} \int_0^{1/2} \frac{1}{\sqrt{1-u^2}} \, \mathrm{d}u = \frac{1}{p} \left[\arcsin u \right]_0^{1/2} = \frac{\pi}{6p}.$$

Equating the two, we get

$$\frac{1}{4} \ln 3 = \frac{\pi}{6p} \implies p = \frac{2\pi}{3 \ln 3}.$$

Problem 5.

(a) Find $\int \frac{x+3}{\sqrt{4x-x^2}} \, \mathrm{d}x$.

(b) If $x = 4\cos^2\theta + 7\sin^2\theta$, show that $7 - x = 3\cos^2$, and find a similar expression for $x - 4$. By using the substitution $x = 4\cos^2\theta + 7\sin^2\theta$, evaluate $\int_4^7 \frac{1}{\sqrt{(x-4)(7-x)}} \, \mathrm{d}x$.

Solution.

Part (a). Note that

$$\int \frac{x+3}{\sqrt{4x-x^2}} \, \mathrm{d}x = -\frac{1}{2} \int \frac{-2x-6}{\sqrt{4x-x^2}} \, \mathrm{d}x = -\frac{1}{2} \int \frac{-2x-4}{\sqrt{4x-x^2}} \, \mathrm{d}x + 5 \int \frac{1}{\sqrt{4x-x^2}} \, \mathrm{d}x.$$

Also note that $4x - x^2 = 4 - (x-2)^2$. Hence,

$$\int \frac{x+3}{\sqrt{4x-x^2}} \, \mathrm{d}x = -\frac{1}{2} \int \frac{-2x-4}{\sqrt{4x-x^2}} \, \mathrm{d}x + 5 \int \frac{1}{\sqrt{4-(x-2)^2}} \, \mathrm{d}x,$$

which we can easily evaluate as

$$\int \frac{x+3}{\sqrt{4x-x^2}} \, \mathrm{d}x = -\sqrt{4x-x^2} + 5 \arcsin \frac{x-2}{2} + C.$$

Part (b). Clearly,

$$x = 4\cos^2\theta + 7\sin^2\theta = 7\left(\cos^2\theta + \sin^2\theta\right) - 3\cos^2\theta = 7 - 3\cos^2,$$

whence $7 - x = 3\cos^2\theta$ as desired. Similarly,

$$x = 4\left(\cos^2\theta + \sin^2\theta\right) + 3\sin^2\theta = 4 + 3\sin^2\theta,$$

whence $x - 4 = 3\sin^2\theta$.

Under the substitution $u = 4\cos^2\theta + 7\sin^2\theta$, the integral transforms as

$$\int_4^7 \frac{1}{\sqrt{(x-4)(7-x)}} \, \mathrm{d}x = \int_0^{\pi/2} \frac{6\cos\theta\sin\theta}{\sqrt{\left(3\cos^2\theta\right)\left(3\sin^2\theta\right)}} \, \mathrm{d}\theta = 2\int_0^{\pi/2} \mathrm{d}\theta = \pi.$$

* * * * *

Problem 6. Express $\frac{x^2+x+28}{(1-x)(x^2+9)}$ in partial fractions. Hence, show that $\int_0^3 \frac{x^2+x+28}{(1-x)(x^2+9)} \, \mathrm{d}x = \frac{\pi}{12} - 2\ln 2$.

Solution. Let

$$\frac{x^2+x+28}{(1-x)(x^2+9)} = \frac{A}{1-x} + \frac{Bx+C}{x^2+9},$$

where A, B and C are constants to be determined. By the cover-up rule, we immediately get

$$A = \frac{1+1+28}{1^2+9} = 3.$$

Clearing denominators, we get

$$x^2 + x + 28 = 3\left(x^2+9\right) + (Bx+C)(1-x) = (3-B)\,x^2 + (B-C)\,x + (27+C).$$

Comparing coefficients, we get $B = 2$ and $C = 1$, whence

$$\frac{x^2+x+28}{(1-x)(x^2+9)} = \frac{3}{1-x} + \frac{2x+1}{x^2+9}.$$

Using the above result on the integral, we have

$$\begin{aligned}\int_0^3 \frac{x^2+x+28}{(1-x)(x^2+9)}\,\mathrm{d}x &= \int_0^3 \left(\frac{3}{1-x} + \frac{2x}{x^2+9} + \frac{1}{x^2+9}\right)\mathrm{d}x \\ &= \left[-3\ln|1-x| + \ln\bigl(x^2+9\bigr) + \frac{1}{3}\arctan\frac{x}{3}\right]_0^3 \\ &= \frac{\pi}{12} - 2\ln 2.\end{aligned}$$

* * * * *

Problem 7. Find the derivative of $\arcsin x + x\sqrt{1-x^2}$, expressing your answer in its simplest form. Hence, evaluate the exact value of $\int_0^{1/2} \sqrt{1-x^2}\,\mathrm{d}x$.

Solution. We have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\arcsin x + x\sqrt{1-x^2}\right] = \frac{1}{\sqrt{1-x^2}} + \left(\sqrt{1-x^2} - \frac{x^2}{\sqrt{1-x^2}}\right) = 2\sqrt{1-x^2}.$$

Hence,

$$\int_0^{1/2} \sqrt{1-x^2}\,\mathrm{d}x = \frac{1}{2}\int_0^{1/2} 2\sqrt{1-x^2}\,\mathrm{d}x = \frac{1}{2}\left[\arcsin x + x\sqrt{1-x^2}\right]_0^{1/2} = \frac{\pi}{12} + \frac{\sqrt{3}}{8}.$$

Assignment B7

Problem 1.

(a) Find $\int \frac{6x^3+2}{x^2+1}\,\mathrm{d}x$.

(b) Evaluate $\int_2^4 x\ln x\,\mathrm{d}x$ exactly.

Solution.

Part (a). Note that

$$\frac{6x^3+2}{x^2+1} = 6x - \frac{6x}{x^2+1} + \frac{2}{x^2+1}.$$

Hence,

$$\int \frac{6x^3+2}{x^2+1}\,\mathrm{d}x = \int \left(6x - \frac{6x}{x^2+1} + \frac{2}{x^2+1}\right)\mathrm{d}x = 3x^2 - 3\ln\bigl(x^2+1\bigr) + 2\arctan x + C.$$

Part (b). Consider the substitution $u = x^2$.

$$\int_2^4 x\ln x\,\mathrm{d}x = \frac{1}{2}\int_4^{16} \ln\sqrt{u}\,\mathrm{d}u = \frac{1}{4}\int_4^{16} \ln u\,\mathrm{d}u = \frac{1}{4}\left[u\ln u - u\right]_4^{16} = 14\ln 2 - 3.$$

* * * * *

Problem 2.

(a) Use the derivative of $\cos\theta$ to show that $\frac{\mathrm{d}}{\mathrm{d}\theta}\sec\theta = \sec\theta\tan\theta$.

(b) Use the substitution $x = \sec\theta - 1$ to find the exact value of $\int_{\sqrt{2}-1}^{1} \frac{1}{(x+1)\sqrt{x^2+2x}}\,\mathrm{d}x$.

Solution.

Part (a).

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\sec\theta = \frac{\mathrm{d}}{\mathrm{d}\theta}\frac{1}{\cos\theta} = \frac{\sin\theta}{\cos^2\theta} = \frac{1}{\cos\theta}\cdot\frac{\sin\theta}{\cos\theta} = \sec\theta\tan\theta.$$

Part (b). Consider the substitution $x = \sec\theta - 1 \implies \mathrm{d}x = \sec\theta\tan\theta\,\mathrm{d}\theta$. When $x = 1$, we have $\theta = \pi/3$. When $x = \sqrt{2} - 1$, we have $\theta = \pi/4$. Also note that $x + 1 = \sec\theta$. Now observe that

$$x^2 + 2x = (\sec\theta - 1)^2 + 2(\sec\theta - 1) = \sec^2\theta - 1 = \tan^2\theta \implies \sqrt{x^2+2x} = \tan\theta.$$

Thus,

$$\int_{\sqrt{2}-1}^{1} \frac{1}{(x+1)\sqrt{x^2+2x}}\,\mathrm{d}x = \int_{\pi/4}^{\pi/3} \frac{\sec\theta\tan\theta}{\sec\theta\tan\theta}\,\mathrm{d}\theta = \frac{\pi}{3} - \frac{\pi}{4} = \frac{\pi}{12}.$$

Problem 3. The expression $\frac{x^2}{9-x^2}$ can be written in the form $A+\frac{B}{3-x}+\frac{C}{3+x}$.

(a) Find the values of constants A, B and C.

(b) Show that $\int_0^2 \frac{x^2}{9-x^2}\,\mathrm{d}x = \frac{3}{2}\ln 5 - 2$.

(c) Hence, find the value of $\int_0^2 \ln\bigl(9-x^2\bigr)\,\mathrm{d}x$, giving your answer in terms of $\ln 5$.

Solution.

Part (a).

$$\frac{x^2}{9-x^2} = -1+\frac{9}{9-x^2} = -1+\frac{9}{(3-x)(3+x)} = -1+\frac{3/2}{3-x}+\frac{3/2}{3+x}.$$

Thus, $A=-1$, $B=3/2$ and $C=3/2$.

Part (b).

$$\int_0^2 \frac{x^2}{9-x^2}\,\mathrm{d}x = \int_0^2 \left(-1+\frac{3/2}{3-x}+\frac{3/2}{3+x}\right)\mathrm{d}x$$
$$= \left[-x-\frac{3}{2}\ln(3-x)+\frac{3}{2}\ln(3+x)\right]_0^2 = \frac{3}{2}\ln 5 - 2.$$

Part (c). Integrating by parts,

	D	I
$+$	$\ln\left(9-x^2\right)$	1
$-$	$-2x/\left(9-x^2\right)$	x

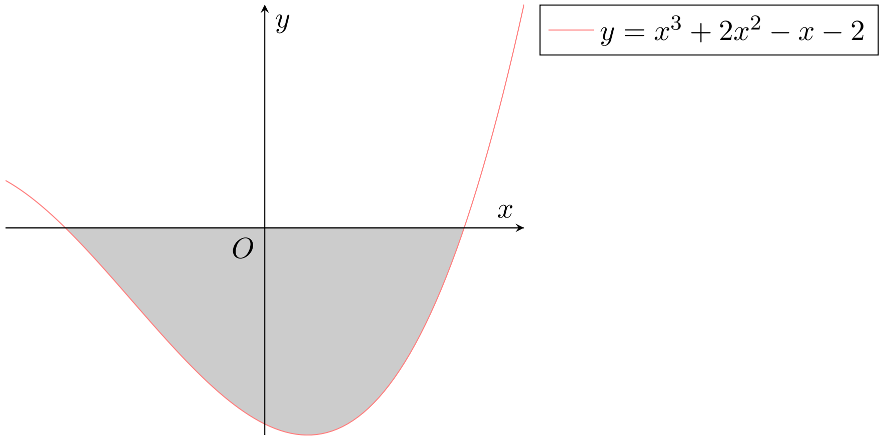
Thus,

$$\int_0^2 \ln\left(9-x^2\right)\,\mathrm{d}x = \left[x\ln\left(9-x^2\right)\right]_0^2 + 2\left(\frac{3}{2}\ln 5 - 2\right) = 5\ln 5 - 4.$$

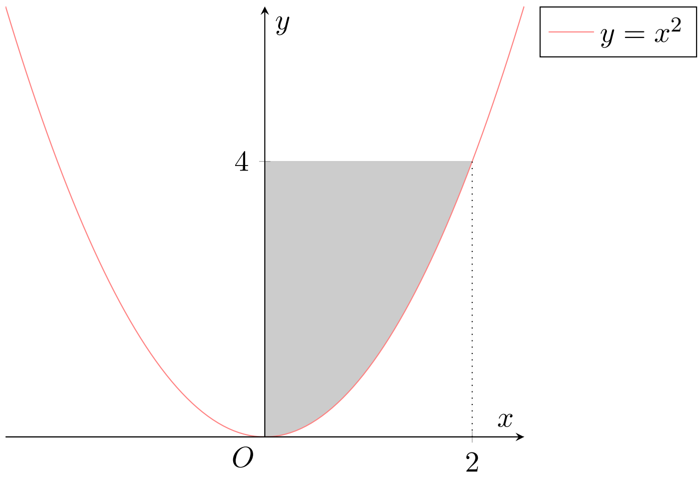
B8 Applications of Integration I - Area and Volume

Tutorial B8

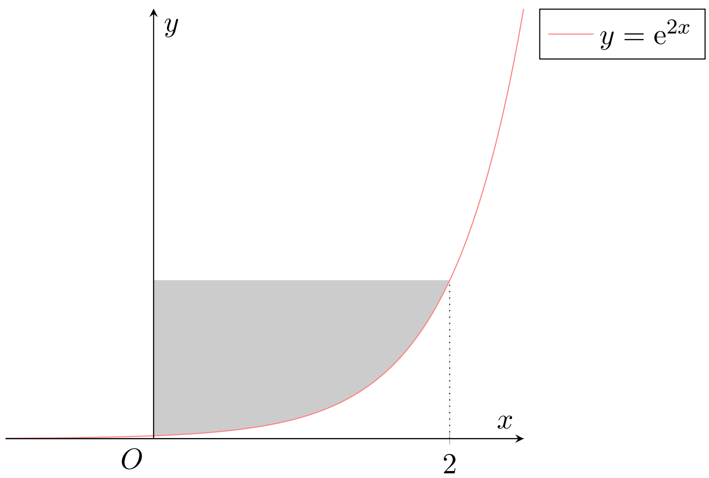
Problem 1. Write down the integral for the area of the shaded region for each of the figure below and use the GC to evaluate it, to 3 significant figures.



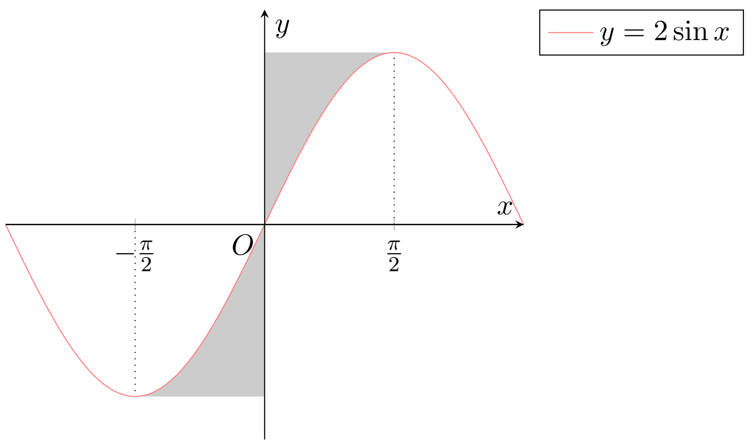
(a)



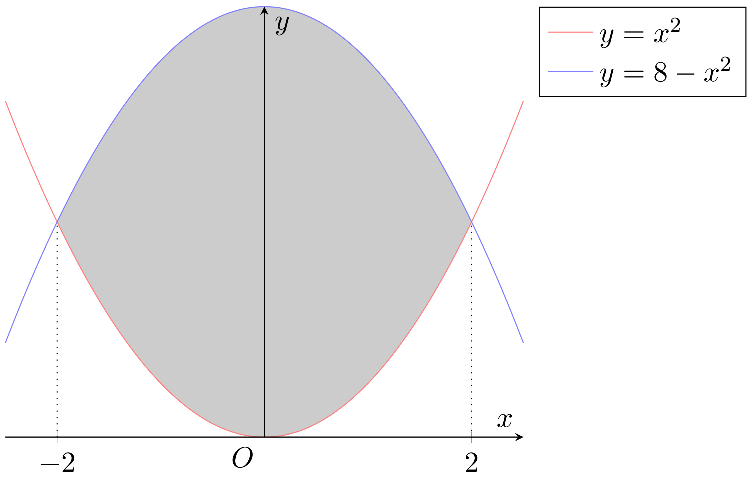
(b)



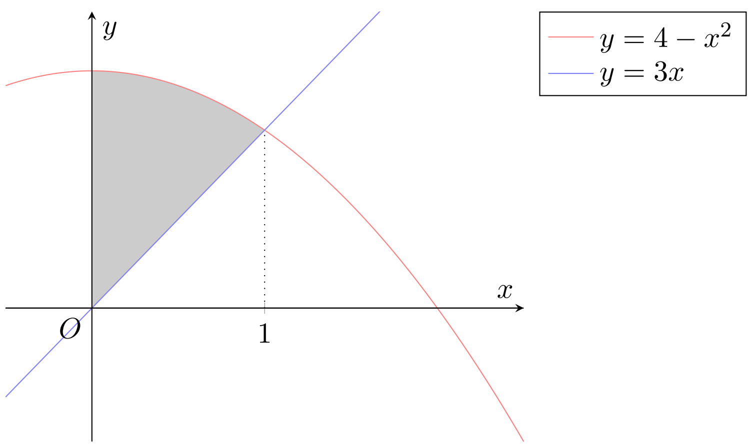
(c)



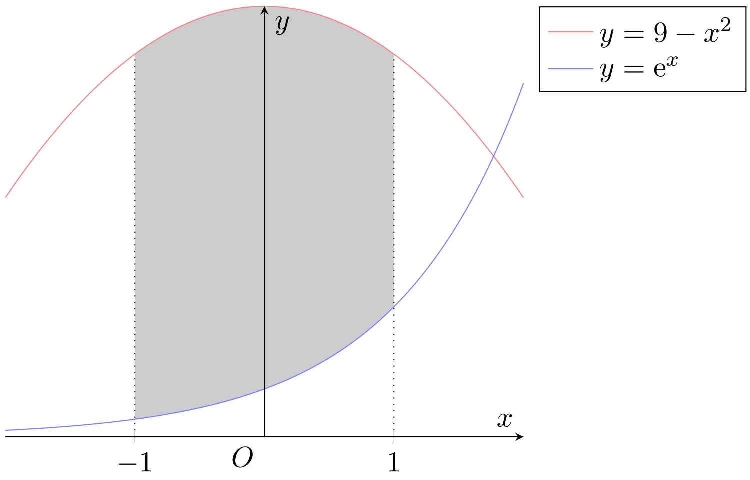
(d)



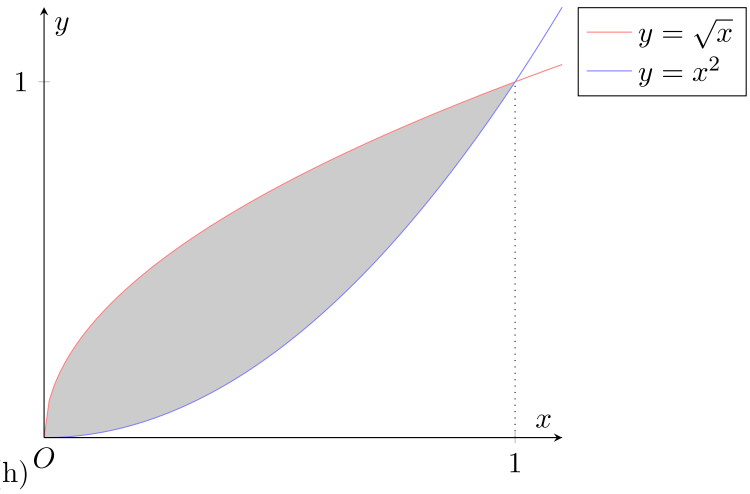
(e)



(f)



(g)



(h)

Solution.

Part (a).

$$\text{Area} = -\int_{-1}^{1} \left(x^3 + 2x^2 - x - 2\right)\,\mathrm{d}x = 2.67 \text{ units}^2 \text{ (3 s.f.)}.$$

Part (b). Note that $y = x^2 \implies x = \sqrt{y}$.

$$\text{Area} = \int_0^4 \sqrt{y}\,\mathrm{d}y = 5.33 \text{ units}^2 \text{ (3 s.f.)}.$$

Part (c). Note that $y = \mathrm{e}^{2x} \implies x = \frac{1}{2}\ln y$. Also, when $x = 0$, we have $y = 1$. Further, when $x = 2$, we have $y = \mathrm{e}^4$. Thus,

$$\text{Area} = \int_0^{\mathrm{e}^4} \frac{1}{2}\ln y\,\mathrm{d}y = 82.4 \text{ units}^2 \text{ (3 s.f.)}.$$

Part (d). Note that when $x = \pi/2$, we have $y = 2$. Thus,

$$\text{Area} = 2\int_0^2 \arcsin\frac{y}{2}\,\mathrm{d}y = 2.28 \text{ units}^2 \text{ (3 s.f.)}.$$

Part (e).

$$\text{Area} = \int_{-2}^{2} \left[\left(8 - x^2\right) - x^2\right]\,\mathrm{d}x = 21.3 \text{ units}^2 \text{ (3 s.f.)}.$$

Part (f).

$$\text{Area} = \int_0^1 \left[\left(4 - x^2\right) - 3x\right] \, \mathrm{d}x = 2.17 \text{ units}^2 \text{ (3 s.f.)}.$$

Part (g).

$$\text{Area} = \int_{-1}^1 \left[\left(9 - x^2\right) - \mathrm{e}^x\right] \, \mathrm{d}x = 15.0 \text{ units}^2 \text{ (3 s.f.)}.$$

Part (h).

$$\text{Area} = \int_0^1 \left(\sqrt{x} - x^2\right) \, \mathrm{d}x = 0.333 \text{ units}^2 \text{ (3 s.f.)}.$$

* * * * *

Problem 2.

(a) Write down the integral for the volume of the solid generated when the shaded region is rotated about the x-axis through 2π for questions 1(a), (e), (f) and (h) using the disc method and use the GC to evaluate it.

(b) Write down the integral for the volume of the solid generated when the shaded region is rotated about the y-axis through 2π for questions 1(b), (d) and (f) using the disc method and use the GC to evaluate it.

Solution.

Part (a).

Part (a)(i).

$$\text{Volume} = \pi \int_{-1}^1 \left(x^3 + 2x^2 - x - 2\right)^2 \, \mathrm{d}x = 13.9 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (a)(ii).

$$\text{Volume} = \pi \int_{-2}^2 \left[\left(8 - x^2\right)^2 - x^2\right] \, \mathrm{d}x = 536 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (a)(iii).

$$\text{Volume} = \pi \int_0^1 \left[\left(4 - x^2\right)^2 - (3x)^2\right] \, \mathrm{d}x = 33.1 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (a)(iv).

$$\text{Volume} = \pi \int_0^1 \left[\left(\sqrt{x}\right)^2 - \left(x^2\right)^2\right] \, \mathrm{d}x = 0.942 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (b).

Part (b)(i).

$$\text{Volume} = \pi \int_0^4 \left(\sqrt{y}\right)^2 \, \mathrm{d}y = 25.1 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (b)(ii).

$$\text{Volume} = 2\pi \int_0^2 \arcsin^2 \frac{y}{2} \, \mathrm{d}y = 5.87 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (b)(iii).

$$\text{Volume} = \pi \int_3^4 (4 - y) \, \mathrm{d}y + \frac{\pi \left(1^2\right)(3)}{3} = 4.71 \text{ units}^3 \text{ (3 s.f.)}.$$

* * * * *

Problem 3.

(a) Write down the integral for the volume of the solid generated when the shaded region is rotated about the x-axis through 2π for questions 1(e), (f) and (h) using the shell method and use the GC to evaluate it.

(b) Write down the integral for the volume of the solid generated when the shaded region is rotated about the y-axis through 2π for questions 1(b), (d) and (f) using the shell method and use the GC to evaluate it.

Solution.

Part (a).

Part (a)(i). Note that $y = x^2 \implies x = \sqrt{y}$ and $y = 8 - x^2 \implies x = \sqrt{8-y}$ for $x > 0$. Thus,

$$\text{Volume} = 2\left(2\pi \int_0^4 \sqrt{y} \cdot y \,\mathrm{d}y + 2\pi \int_4^8 \sqrt{8-y} \cdot y \,\mathrm{d}y\right) = 536 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (a)(ii). Note that $y = 3x \implies x = y/3$ and $y = 4 - x^2 \implies x = \sqrt{4-y}$ for $x > 0$. Thus,

$$\text{Volume} = 2\pi \int_0^3 \frac{1}{3} y \cdot y \,\mathrm{d}y + 2\pi \int_3^4 \sqrt{4-y} \cdot y \,\mathrm{d}y = 33.1 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (a)(iii). Note that $y = \sqrt{x} \implies x = y^2$ and $y = x^2 \implies x = \sqrt{y}$ for $x > 0$. Thus,

$$\text{Volume} = 2\pi \int_0^1 \left(\sqrt{y} - y^2\right) y \,\mathrm{d}y = 0.942 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (b).

Part (b)(i).

$$\text{Volume} = 2\pi \int_0^2 x \cdot x^2 \,\mathrm{d}x = 25.1 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (b)(ii).

$$\text{Volume} = 2 \cdot 2\pi \int_0^{\pi/2} x\,(2 - 2\sin x)\,\mathrm{d}x = 5.87 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (b)(iii).

$$\text{Volume} = 2\pi \int_0^1 x\left[\left(4 - x^2\right) - 3x\right] \mathrm{d}x = 4.71 \text{ units}^3 \text{ (3 s.f.)}.$$

* * * * *

Problem 4. Calculate the area enclosed by the petals of the curve $r = \sin 2\theta$ where $r \geq 0$.

Solution. Note that $r \geq 0 \implies \sin 2\theta \geq 0 \implies r \in \left[0, \frac{\pi}{2}\right] \cup \left[\pi, \frac{3\pi}{2}\right]$. Thus,

$$\text{Area} = 2 \cdot \frac{1}{2} \int_0^{\pi/2} \sin^2 2\theta \, \mathrm{d}\theta = \int_0^{\pi/2} \frac{1 - \cos 4\theta}{2} \, \mathrm{d}\theta = \frac{1}{2} \left[\theta - \frac{\sin 4\theta}{4}\right]_0^{\pi/2} = \frac{\pi}{4} \text{ units}^2.$$

* * * * *

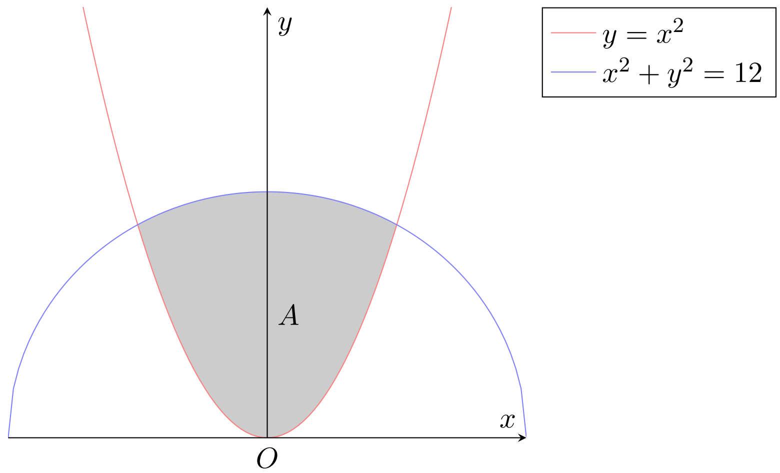
Problem 5. The finite region A is bounded by the curve $y = x^2$ and a minor arc of the circle $x^2 + y^2 = 12$.

(a) Find the numerical value of the area of A, correct to 2 decimal places.

(b) Find the exact volume of the solid obtained when A is rotated about the x-axis through 2π radians.

(c) Find the exact volume of the solid obtained when A is rotated about the y-axis through π radians.

Solution.



Part (a). Consider the intersections between $y = x^2$ and $x^2 + y^2 = 12$.

$$\begin{aligned} x^2 + y^2 = x^2 + \left(x^2\right)^2 = 12 &\implies x^4 + x^2 - 12 = (x^2 - 3)(x^2 + 4) = 0 \\ &\implies \left(x - \sqrt{3}\right)\left(x + \sqrt{3}\right)\left(x^2 + 4\right) = 0. \end{aligned}$$

Hence, the two curves intersect at $x = -\sqrt{3}$ and $x = \sqrt{3}$. Note that $x^2 + 4 = 0$ has no solution since $x^2 + 4 > 0$. Also note that $x^2 + y^2 = 12 \implies y = \sqrt{12 - x^2}$ for $y > 0$. Thus,

$$\text{Area} = 2 \int_0^{\sqrt{3}} \left(\sqrt{12 - x^2} - x^2\right) \mathrm{d}x = 8.02 \text{ units}^2 \text{ (3 s.f.)}.$$

Part (b). Note that $x^2 + y^2 = 12 \implies y^2 = 12 - x^2$.

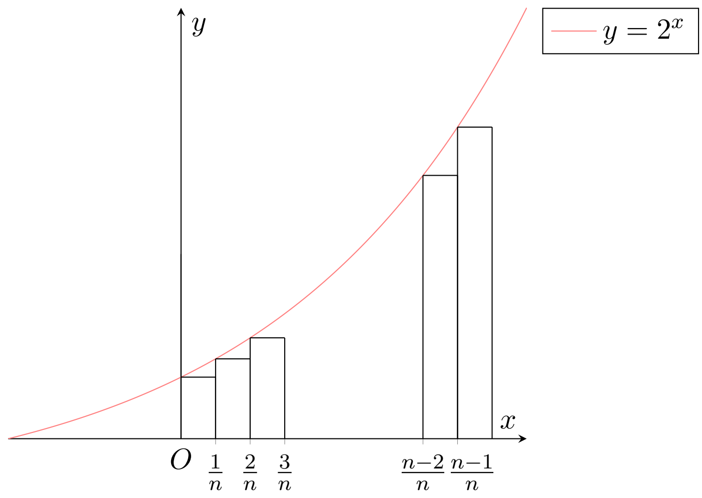
$$\text{Volume} = 2\pi \int_0^{\sqrt{3}} \left[\left(12 - x^2\right) - \left(x^2\right)^2\right] \mathrm{d}x = 2\pi \left[12x - \frac{x^3}{3} - \frac{x^5}{5}\right]_0^{\sqrt{3}} = \frac{92\sqrt{3}\pi}{5} \text{ units}^3.$$

Part (c). Note that when the curves intersect at $x = \sqrt{3}$, we have $y = 3$. Furthermore, when $x = 0$, we have $y = \sqrt{12}$. Also note that $x^2 + y^2 = 12 \implies x^2 = 12 - y^2$.

$$\text{Volume} = \pi \int_0^3 y \, \mathrm{d}y + \pi \int_3^{\sqrt{12}} \left(12 - y^2\right) \mathrm{d}y = \pi \left(16\sqrt{3} - \frac{45}{2}\right) \text{ units}^3.$$

* * * * *

Problem 6.



(a) The graph of $y = 2^x$, for $0 \le x \le 1$ is shown in the diagram. Rectangles, each of width $\frac{1}{n}$, are drawn under the curve. Given that $\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$, show that the total area A of all n rectangles is given by $\frac{1}{n}\left(\frac{1}{2^{\frac{1}{n}}-1}\right)$.

(b) Find the limit of A in exact form as $n \to \infty$.

Let V be the volume of all n rectangles rotated about the x-axis.

(c) Find V in terms of n.

(d) State the limit of V in exact form as $n \to \infty$.

Solution.

Part (a).

$$A = \sum_{k=0}^{n-1} \frac{2^{k/n}}{n} = \frac{1}{n}\sum_{k=0}^{n-1}\left(2^{1/n}\right)^k = \frac{1}{n}\cdot\frac{1-\left(2^{1/n}\right)^n}{1-2^{1/n}} = \frac{1}{n}\left(\frac{1-2}{1-2^{1/n}}\right) = \frac{1}{n}\left(\frac{1}{2^{1/n}-1}\right).$$

Part (b).

$$\lim_{n\to\infty} A = \lim_{n\to\infty}\frac{1/n}{2^{1/n}-1} = \lim_{m\to 0}\frac{m}{2^m-1} = \lim_{m\to 0}\frac{1}{\ln 2\cdot 2^m} = \frac{1}{\ln 2}.$$

Part (c).

$$V = \pi\sum_{k=0}^{n-1}\frac{1}{n}\left(2^{k/n}\right)^2 = \frac{\pi}{n}\sum_{k=0}^{n-1}\left(2^{2/n}\right)^k = \frac{\pi}{n}\left(\frac{1-\left(2^{2/n}\right)^n}{1-2^{2/n}}\right)$$
$$= \frac{\pi}{n}\left(\frac{1-4}{1-2^{2/n}}\right) = \frac{3\pi}{n\left(4^{1/n}-1\right)}.$$

Part (d).

$$\lim_{n\to\infty} V = \lim_{n\to\infty} \frac{3\pi}{n\left(4^{1/n}-1\right)} = 3\pi \lim_{n\to\infty} \frac{1/n}{4^{1/n}-1} = 3\pi \lim_{m\to 0} \frac{m}{4^m-1}$$
$$= 3\pi \lim_{m\to 0} \frac{1}{4^m \ln 4} = 3\pi\left(\frac{1}{\ln 4}\right) = \frac{3\pi}{2\ln 2}.$$

* * * * *

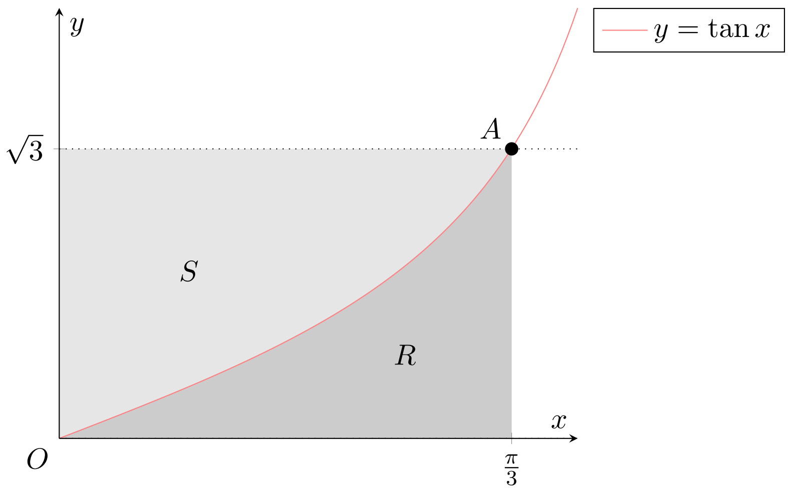
Problem 7. O is the origin and A is the point on the curve $y = \tan x$ where $x = \pi/3$.

(a) Calculate the area of the region R enclosed by the arc OA, the x-axis and the line $x = \pi/3$, giving your answer in an exact form.

(b) The region S is enclosed by the arc OA, the y-axis and the line $y = \sqrt{3}$. Find the volume of the solid of revolution formed when S is rotated through 360° about the x-axis, giving your answer in an exact form.

(c) Find $\int_0^{\sqrt{3}} \arctan y \, dy$ in exact form.

Solution.



Part (a).

$$[R] = \int_0^{\pi/3} \tan x \, dx = [\ln \sec x]_0^{\pi/3} = \ln 2 \text{ units}^2.$$

Part (b).

$$\text{Volume} = \pi \int_0^{\pi/3} \left[\left(\sqrt{3}\right)^2 - \tan^2 x\right] dx = \pi \int_0^{\pi/3} \left(3 - \sec^2 x + 1\right) dx$$
$$= \pi \left[4x - \tan x\right]_0^{\pi/3} = \left(\frac{4\pi^2}{3} - \sqrt{3}\pi\right) \text{ units}^3.$$

Part (c). Observe that $\int_0^{\sqrt{3}} \arctan y \, dy = [S] = [R \cup S] - [R] = (\pi/3) \cdot \sqrt{3} - \ln 2$.

* * * * *

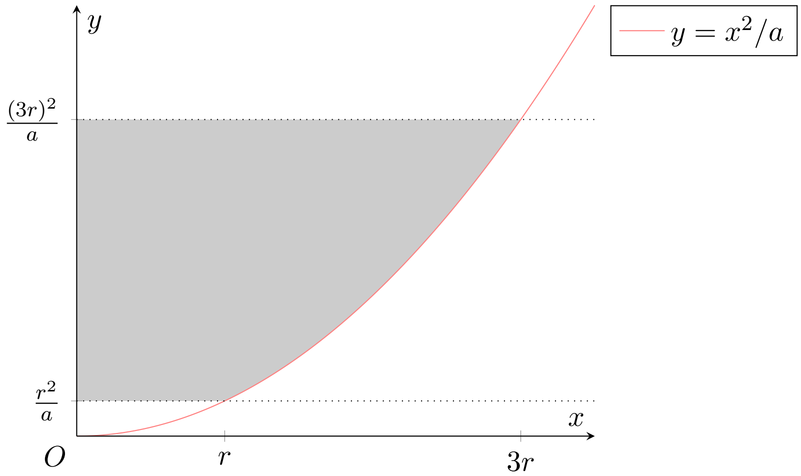
Problem 8. A portion of the curve $ay = x^2$, where a is a positive constant, is rotated about the vertical axis Oy to form the curved surface of an open bowl. The bowl has a horizontal circular base of radius r and a horizontal circular rim of radius $3r$.

(a) Prove that the depth of the bowl is $\frac{8r^2}{a}$.

(b) Find the volume of the bowl in terms of r and a.

(c) Given that the volume of the bowl is $\frac{\pi a^3}{10}$, find the depth of the bowl in terms of a only.

Solution. Note that $ay = x^2 \implies y = \frac{x^2}{a}$.



Part (a).

$$\text{Depth of bowl} = \frac{(3r)^2}{a} - \frac{r^2}{a} = \frac{8r^2}{a} \text{ units.}$$

Part (b).

$$\text{Volume} = \pi \int_{r^2/a}^{9r^2/a} ay \, \mathrm{d}y = \pi \left[\frac{a}{2}y^2\right]_{r^2/a}^{9r^2/a} = \frac{a\pi}{2} \cdot \frac{80r^4}{a^2} = \frac{40\pi r^4}{a} \text{ units}^3.$$

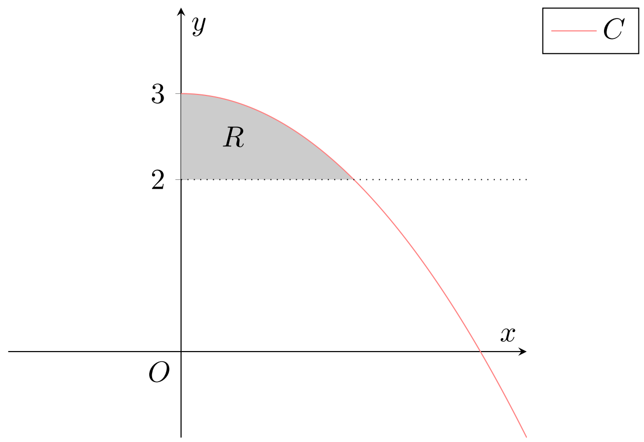
Part (c).

$$\frac{40\pi r^4}{a} = \frac{\pi a^3}{10} \implies 400r^4 = a^4 \implies 20r^2 = a^2 \implies r^2 = \frac{1}{20}a^2.$$

Hence, the depth of the bowl is

$$\frac{8}{a}\left(\frac{1}{20}a^2\right) = \frac{2}{5}a \text{ units.}$$

Problem 9. The diagram shows the region R bounded by part of the curve C with equation $y = 3 - x^2$, the y-axis and the line $y = 2$, lying in the first quadrant.



Write down the equation of the curve obtained when C is translated by 2 units in the negative y-direction.

Hence, or otherwise, show that the volume of the solid formed when R is rotated completely about the line $y = 2$ is given by $\pi \int_0^1 \left(1 - 2x^2 + x^4\right) \mathrm{d}x$ and evaluate this integral exactly.

Solution. Clearly, $C : y = 1 - x^2$.

Note that $3 - x^2 = 2 \implies x = \pm 1$, whence $x = 1$ since $x > 0$.

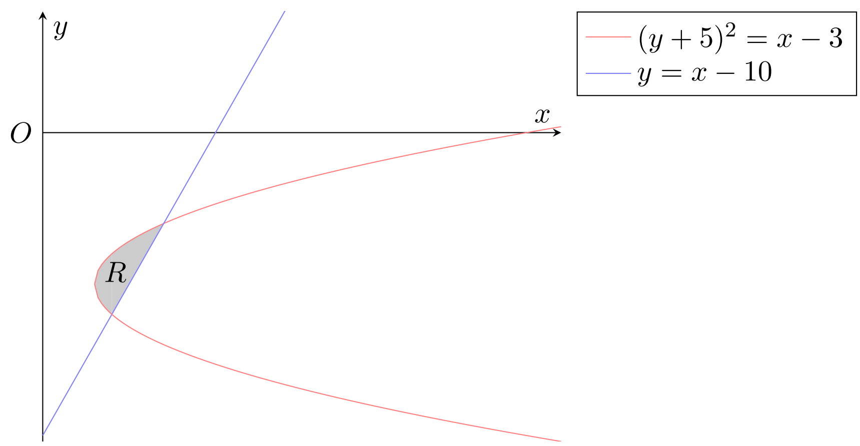
$$\begin{aligned} \text{Volume} &= \pi \int_0^1 \left(1 - x^2\right)^2 \mathrm{d}x = \pi \int_0^1 \left(1 - 2x^2 + x^4\right) \mathrm{d}x \\ &= \pi \left[x - \frac{2}{3}x^3 + \frac{1}{5}x^5\right]_0^1 = \frac{8}{15}\pi \text{ units}^3. \end{aligned}$$

* * * * *

Problem 10. The diagram below shows a region R bounded by the curve $(y+5)^2 = x-3$ and the line $y = x - 10$. Find the volume of solid formed when R is rotated four right angles about

(a) the y-axis, and

(b) the x-axis.



Solution.

Part (a). Consider the intersections between $(y+5)^2 = x-3$ and $y = x-10$.

$$(y+5)^2 = (x-5)^2 = x-3 \implies x^2 - 11x + 28 = (x-4)(x-7) = 0.$$

Hence, $x = 4$ and $x = 7$, whence $y = -6$ and $y = -3$. Thus, the two curves intersect at $(4, -6)$ and $(7, -3)$.

Note that $(y+5)^2 = x-3 \implies x = 3 + (y+5)^2$ and $y = x - 10 \implies x = y + 10$.

$$\text{Volume} = \pi \int_{-6}^{-3} \left[(y+10)^2 - \left(3 + (y+5)^2\right)^2\right] \mathrm{d}y = 130 \text{ units}^3 \text{ (3 s.f.)}.$$

Part (b). Note that

$$(y+5)^2 = x-3 \implies \begin{cases} y = -5 + \sqrt{x-3}, & y \geq -5 \\ y = -5 - \sqrt{x-3}, & y < -5 \end{cases}$$

Thus,

$$\begin{aligned} \text{Volume} &= \pi \int_3^4 \left[\left(-5 - \sqrt{x-3}\right)^2 - \left(-5 + \sqrt{x+3}\right)^2\right] \mathrm{d}x \\ &\quad + \pi \int_4^7 \left[(x-10)^2 - \left(-5 + \sqrt{x-3}\right)^2\right] \mathrm{d}x = 127 \text{ units}^3 \text{ (3 s.f.)}. \end{aligned}$$

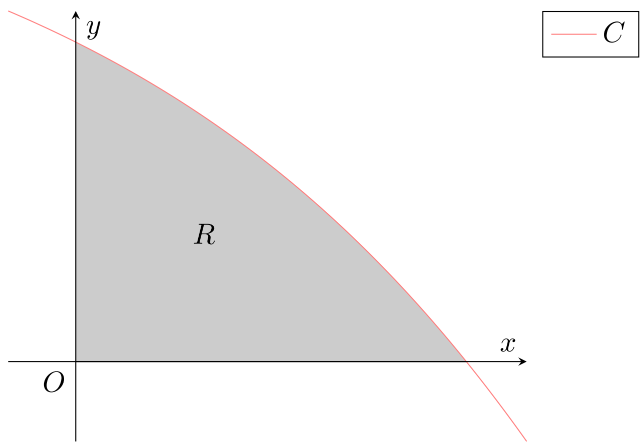
* * * * *

Problem 11. The curve C is defined by the following pair of parametric equations.

$$x = t - \frac{1}{t^2},\ y = 2 - t^2, \qquad t > 0.$$

Find the area of the finite region R enclosed by the curve C and the axes as well as the volume of solid obtained when R is rotated about the x-axis through 4 right-angles.

Solution.



Note that when $x = 0$, we have $t = 1$. Also note that when $y = 0$, we have $t = \sqrt{2}$, whence $x = \sqrt{2} - 1/2$. Thus,

$$\begin{aligned} [R] &= \int_0^{\sqrt{2}-1/2} y \,\mathrm{d}x = \int_1^{\sqrt{2}} \left(2 - t^2\right) \frac{\mathrm{d}x}{\mathrm{d}t} \,\mathrm{d}t \\ &= \int_1^{\sqrt{2}} \left(2 - t^2\right) \left(1 + \frac{2}{t^3}\right) \mathrm{d}t = 0.526 \text{ units}^2 \text{ (3 s.f.)}. \end{aligned}$$

Also,

$$\text{Volume} = \pi \int_0^{\sqrt{2}-\frac{1}{2}} y^2 \, \mathrm{d}x = \pi \int_1^{\sqrt{2}} \left(2 - t^2\right) \frac{\mathrm{d}x}{\mathrm{d}t} \, \mathrm{d}t$$
$$= \pi \int_1^{\sqrt{2}} \left(2 - t^2\right) \left(1 + \frac{2}{t^3}\right) \mathrm{d}t = 1.19 \text{ units}^3 \text{ (3 s.f.)}.$$

* * * * *

Problem 12. Find the area enclosed by the ellipse $x = a \cos t$, $y = b \sin t$, where a and b are positive constants. Find also the volume of solid obtained when the region enclosed by the ellipse is rotated through π radians about the x-axis.

Solution. By symmetry, we only need to consider the area of the ellipse in the first quadrant. Note that $x = 0 \implies t = \pi/2$ and $x = a \implies t = 0$. Hence,

$$\text{Area} = 4 \int_0^a y \, \mathrm{d}x = 4 \int_{\pi/2}^0 y \cdot \frac{\mathrm{d}x}{\mathrm{d}t} \, \mathrm{d}t = 4 \int_{\pi/2}^0 (b \sin t)(-a \sin t) \, \mathrm{d}t = 4ab \int_0^{\pi/2} \sin^2 t \, \mathrm{d}t$$
$$= 4ab \int_0^{\pi/2} \frac{1 - \cos 2t}{2} \, \mathrm{d}t = 2ab \left[t - \frac{\sin 2t}{2}\right]_0^{\pi/2} = \pi ab \text{ units}^2.$$

Also,

$$\text{Volume} = 2\pi \int_0^a y^2 \, \mathrm{d}x = 2\pi \int_{\pi/2}^0 y^2 \cdot \frac{\mathrm{d}x}{\mathrm{d}t} \, \mathrm{d}t = 2\pi \int_{\pi/2}^0 (b \sin t)^2 (-a \sin t) \, \mathrm{d}t$$
$$= 2\pi ab^2 \int_0^{\pi/2} \sin^3 t \, \mathrm{d}t = 2\pi ab^2 \int_0^{\pi/2} \frac{3 \sin t - \sin 3t}{4} \, \mathrm{d}t$$
$$= \frac{1}{2} \pi ab^2 \left[-3 \cos t + \frac{1}{3} \cos 3t\right]_0^{\pi/2} = \frac{4\pi}{3} ab^2 \text{ units}^3.$$

* * * * *

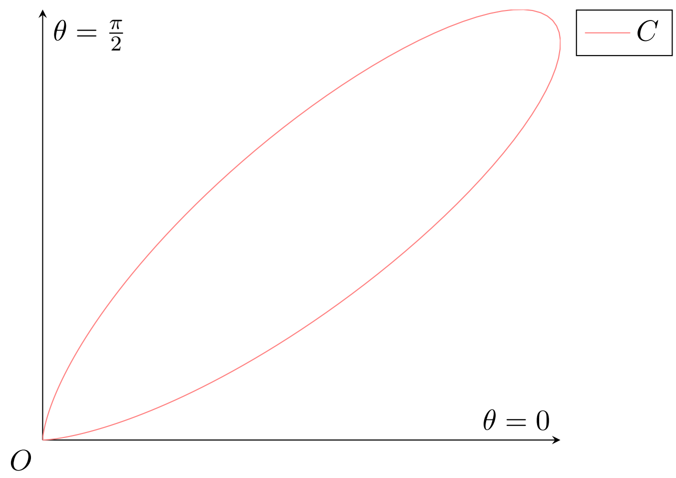
Problem 13. Find the polar equation of the curve C with equation $x^5 + y^5 = 5bx^2y^2$, where b is a positive constant. Sketch the part of the curve C where $0 \leq \theta \leq \frac{\pi}{2}$. Show, using polar coordinates, that the area A of the region enclosed by this part of the curve is given by

$$A = \frac{25b^2}{2} \int_0^{\frac{\pi}{2}} \frac{\sin^4 \theta \cos^4 \theta}{(\cos^5 \theta + \sin^5 \theta)^2} \, \mathrm{d}\theta$$

By differentiating $\frac{1}{1+\tan^5 \theta}$ with respect to θ, or otherwise, find the exact value of A in terms of b.

Solution.

$$x^5 + y^5 = 5bx^2y^2 \implies (r \cos \theta)^5 + (r \sin \theta)^5 = 5b(r \cos \theta)^2 (r \sin \theta)^2$$
$$\implies r \left(\cos^5 \theta + \sin^5 \theta\right) = 5b \cos^2 \theta \sin^2 \theta \implies r = \frac{5b \cos^2 \theta \sin^2 \theta}{\cos^5 \theta + \sin^5 \theta}.$$



We have

$$A = \frac{1}{2}\int_0^{\pi/2}\left(\frac{5b\cos^2\theta\sin^2\theta}{\cos^5\theta+\sin^5\theta}\right)^2 \mathrm{d}\theta = \frac{1}{2}\int_0^{\pi/2}\frac{25b^2\cos^4\theta\sin^4\theta}{\left(\cos^5\theta+\sin^5\theta\right)^2}\mathrm{d}\theta$$
$$= \frac{25b^2}{2}\int_0^{\pi/2}\frac{\cos^4\theta\sin^4\theta}{\left(\cos^5\theta+\sin^5\theta\right)^2}\mathrm{d}\theta.$$

Note that

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\frac{1}{1+\tan^5\theta} = -\frac{5\tan^4\theta\sec^2\theta}{\left(1+\tan^5\theta\right)^2} = -5\left(\frac{\cos^{10}\theta}{\cos^5\theta+\sin^5\theta}\right)\left(\frac{\sin^4\theta}{\cos^6\theta}\right) = -\frac{5\cos^4\theta\sin^4\theta}{\cos^5\theta+\sin^5\theta}.$$

Hence,

$$A = \frac{-5b^2}{2}\int_0^{\pi/2} -\frac{5\cos^4\theta\sin^4\theta}{\left(\cos^5\theta+\sin^5\theta\right)^2}\mathrm{d}\theta = -\frac{5b^2}{2}\left[\frac{1}{1+\tan^5\theta}\right]_0^{\pi/2} = \frac{5b^2}{2}.$$

* * * * *

Problem 14. The polar equation of a curve is given by $r = \mathrm{e}^\theta$ where $0 \le \theta \le \pi/2$. Cartesian axes are taken at the pole O. Express x and y in terms of θ and hence find the Cartesian equation of the tangent at $\left(\mathrm{e}^{\pi/2}, \pi/2\right)$. The region R is bounded by the polar curve, tangent and the x-axis. Find the exact area of the region R.

Solution. We have $x = \mathrm{e}^\theta\cos\theta$ and $y = \mathrm{e}^\theta\sin\theta$. Thus,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = \frac{e^\theta\cos\theta+e^\theta\sin\theta}{-e^\theta\sin\theta+e^\theta\cos\theta} = \frac{\cos\theta+\sin\theta}{\cos\theta-\sin\theta}.$$

Hence,

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{\theta=\pi/2} = \frac{\cos(\pi/2)+\sin(\pi/2)}{\cos(\pi/2)-\sin(\pi/2)} = -1.$$

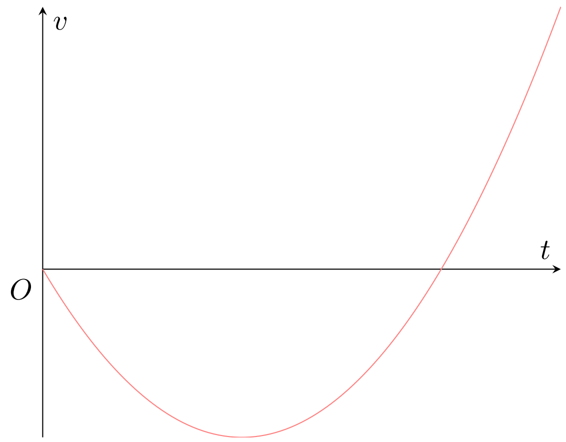
When $\theta = \pi/2$, we have $x = 0$ and $y = \mathrm{e}^{\pi/2}$. Hence, the tangent is given by

$$y - e^{\pi/2} = -(x-0) \implies y = -x + e^{\pi/2}.$$

Thus,

$$[R] = \frac{1}{2}\left(\mathrm{e}^{\pi/2}\right)\left(\mathrm{e}^{\pi/2}\right) - \frac{1}{2}\int_0^{\pi/2}\left(e^\theta\right)^2 \mathrm{d}\theta = \frac{\mathrm{e}^\pi}{2} - \frac{1}{2}\left[\frac{\mathrm{e}^{2\theta}}{2}\right]_0^{\pi/2} = \frac{e^\pi+1}{4} \text{ units}^2.$$

Problem 15.



The diagram shows the velocity-time graph of a particle moving in a straight line. The equation of the curve shown is $v = t(t-10)$ where t seconds is the time and v ms^{-1} is the velocity. The particle starts at a point A on the line when $t = 0$.

Calculate

(a) the distance travelled by the particle before coming to instantaneous rest, and

(b) the time at which the particle returns to A.

Solution.

Part (a). For instantaneous rest, $v = 0$. Hence, $t(t-10) = 0$, whence $t = 10$. Note that we reject $t = 0$ since $t > 0$. The distance travelled by the particle before coming to instantaneous rest is hence

$$-\int_0^{10} v \, \mathrm{d}t = -\int_0^{10} t(t-10) \, \mathrm{d}t = -\int_0^{10} \left(t^2 - 10t\right) \, \mathrm{d}t = -\left[\frac{t^3}{3} - \frac{10t^2}{2}\right]_0^{10} = \frac{500}{3} \text{ m}.$$

Part (b). When the particle returns to A, $s = 0$. Let the time at which the particle returns to A be t_0.

$$\int_0^{t_0} v \, \mathrm{d}t = \int_0^{t_0} t(t-10) \, \mathrm{d}t = \left[\frac{t_0^3}{3} - \frac{10t_0^2}{2}\right]_0^{t_0} = \frac{1}{3}t_0^3 - 5t_0^2 = \frac{1}{3}t_0^2\left(t_0 - 15\right) = 0.$$

Thus, $t_0 = 15$. Note that we reject $t_0 = 0$ since $t_0 > 0$. It hence takes the particle 15 seconds to return to A.

Self-Practice B8

Problem 1.

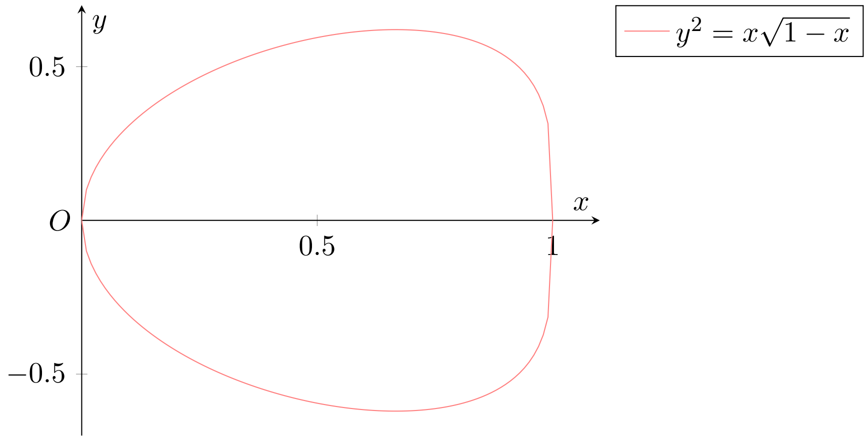
(a) Find $\int x \sin^2 x \,\mathrm{d}x$.

(b) The region R is bounded by the curve $y = \sqrt{x}\sin x$, the lines $x = 0$ and $x = \pi$, and the x-axis. Find the volume of the solid of revolution formed when R is rotated through 4 right angles about the x-axis.

(c) Hence, calculate the volume of the solid of revolution formed when S is rotated through 4 right angles about the x-axis, where S is the region bounded by the curve $y = \sqrt{x}\sin x$, the lines $x = \pi$ and $y = \sqrt{\pi}$, and the y-axis.

* * * * *

Problem 2. The diagram shows the curve C with the equation $y^2 = x\sqrt{1-x}$. The region enclosed by C is denoted by R.



(a) Write down an integral that gives the area of R, and evaluate this integral numerically.

(b) The part of R above the x-axis is rotated through 2π radians about the x-axis. By using the substitution $u = 1 - x$, or otherwise, find the exact value of the volume obtained.

(c) Find the exact x-coordinate of the maximum point of C.

* * * * *

Problem 3.

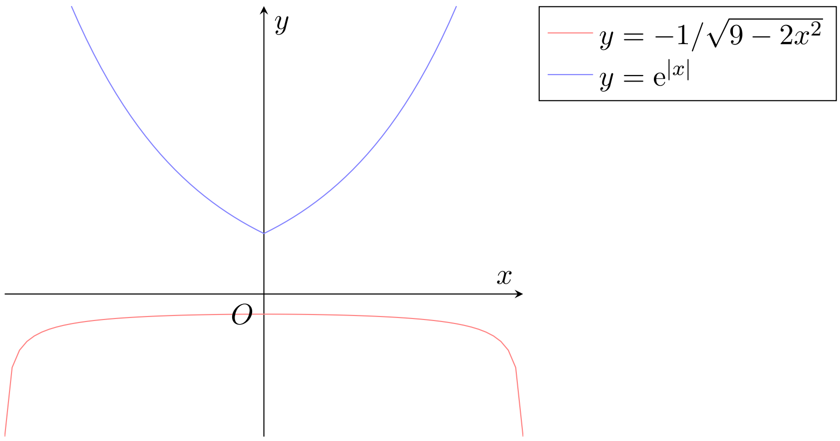
(a) Find the exact value of $\int_0^{5\pi/3} \sin^2 x \,\mathrm{d}x$. Hence, find the exact value of $\int_0^{5\pi/3} \cos^2 x \,\mathrm{d}x$.

(b) The region R is bounded by the curve $y = x^2 \sin x$, the line $x = \frac{1}{2}\pi$ and the part of the x-axis between 0 and $\frac{1}{2}\pi$. Find

(i) the exact area of R,

(ii) the numerical value of the volume of revolution formed when R is rotated completely about the x-axis, giving your answer correct to 3 decimal places.

Problem 4. The diagram below shows the graphs of $y = -\frac{1}{\sqrt{9-2x^2}}$ and $y = \mathrm{e}^{|x|}$.



(a) The region A is bounded by the curves $y = -\frac{1}{\sqrt{9-2x^2}}$ and $y = \mathrm{e}^{|x|}$, and the lines $x = -1$ and $x = 2$. Find the area of A, giving your answer to 3 significant figures.

(b) The region bounded by the curves $y = -\frac{1}{\sqrt{9-2x^2}}$, $y = \mathrm{e}^{|x|}$, the y-axis and the line $x = 2$ is rotated through 2π radians about the y-axis. Prove that the volume generated is $2\pi\left(\mathrm{e}^2 + 2\right)$.

* * * * *

Problem 5.

(a) (i) The region S, is enclosed by the x-axis, the line $x = 1$ and the curve given by the parametric equations

$$x = (1+t)^{3/2}, \quad y = (1-t)^{1/2}, \quad t \in [0, 1].$$

Find the exact area of S.

(ii) Find also the volume of the solid obtained when the region S is rotated about the y-axis.

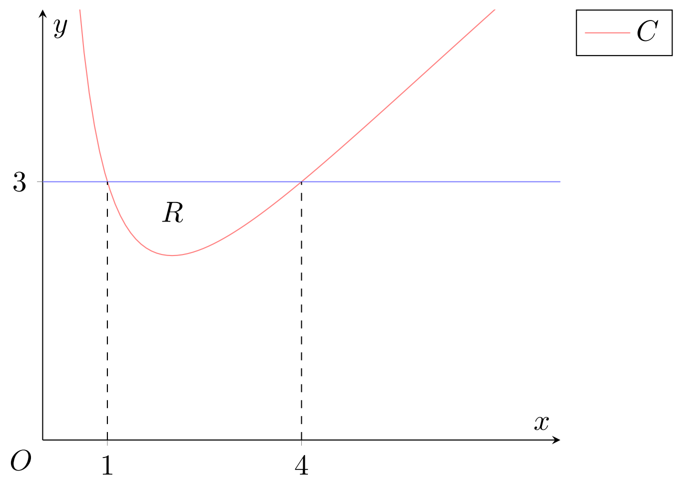
(b) The region R is bounded by the curve $y = \left(\frac{x-2}{4-x}\right)^{1/4}$, the line $x = 2$ and the line $y = 1$. By using the substitution $x = 2\left(1 + \cos^2\theta\right)$, or otherwise, find the exact volume of the solid generated when R is rotated through four right angles about the x-axis.

* * * * *

Problem 6. The diagram shows the region R in the first quadrant bounded by the curve C with equation $y = \sqrt{x} + \frac{2}{\sqrt{x}}$ and the line $y = 3$. The line and the curve intersect at the points $(1, 3)$ and $(4, 3)$. Calculate the exact area of R. Write down the equation of the curve obtained when C is translated by 3 units in the negative y-direction. Hence, or otherwise, show that the volume of the solid formed when R is rotated completely about the line $y = 3$ is given by

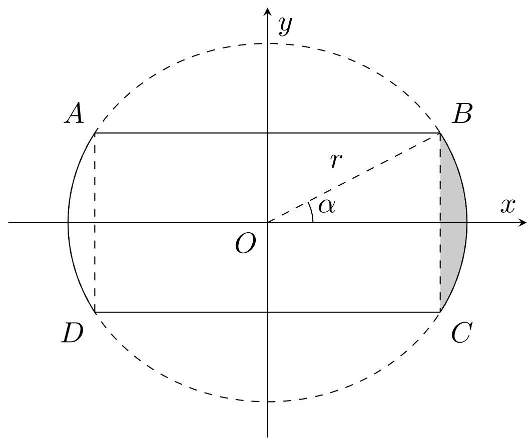
$$\pi \int_1^4 \left(x - 6\sqrt{x} + 13 - \frac{12}{\sqrt{x}} + \frac{4}{x}\right) \mathrm{d}x,$$

and evaluate this integral exactly.



* * * * *

Problem 7. The diagram shows the circle, centre O and radius r, with equation $x^2+y^2=r^2$. The points A, B, C, D on the circle form a rectangle with sides parallel to the axes. $\angle AOD = \angle BOC = 2\alpha$. The region bounded by the line AB, the line DC and the circular arc BC and AD is rotated about the x-axis to form a solid of rotation S.



(a) Show that the volume obtained by rotating the shaded part of the region about the x-axis is $\frac{1}{3}\pi r^3\left(\cos^3\alpha - 3\cos\alpha + 2\right)$.

(b) Show that the total volume of S is $\frac{4}{3}\pi r^3\left(1-\cos^3\alpha\right)$.

(c) Given that the volume of S is half the volume of a sphere of radius r, find the value of α.

* * * * *

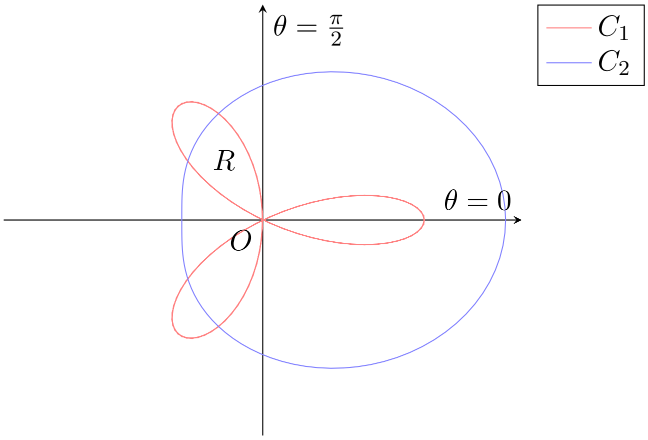
Problem 8. An ellipse E has equation $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, where a and b are positive constants. Show that the area A of the region enclosed by E is given by $A=\frac{4b}{a}\int_0^a \sqrt{(a^2-x^2)}\,\mathrm{d}x$. By using the substitution $x=a\sin\theta$, or otherwise, find the value of A in terms of a, b, and π. Show on a sketch the region R of points inside the ellipse E such that $x>0$ and $y<x$. Given that $a^2=3b^2$, find the area of R in terms of a and π.

* * * * *

Problem 9. Sketch the polar curve $r=a(1-\sin 2\theta)$, where $a>0$ and $0\le\theta<2\pi$. Prove that the area enclosed by each loop of the curve is $\frac{3}{4}\pi a^2$.

Problem 10. The diagram shows the curves C_1 and C_2 whose respective polar equations are

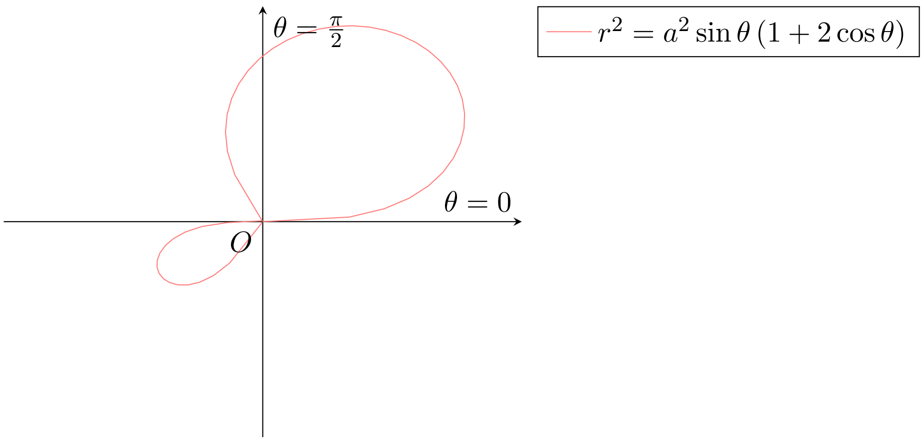
$$r = \cos 3\theta, \qquad (0 \le \theta \le 2\pi)$$
$$r = 1 + \frac{1}{2}\cos\theta. \qquad (0 \le \theta \le 2\pi)$$



R is the region bounded by the curve C_2 and one loop of the curve C_1. Find the area of the region R.

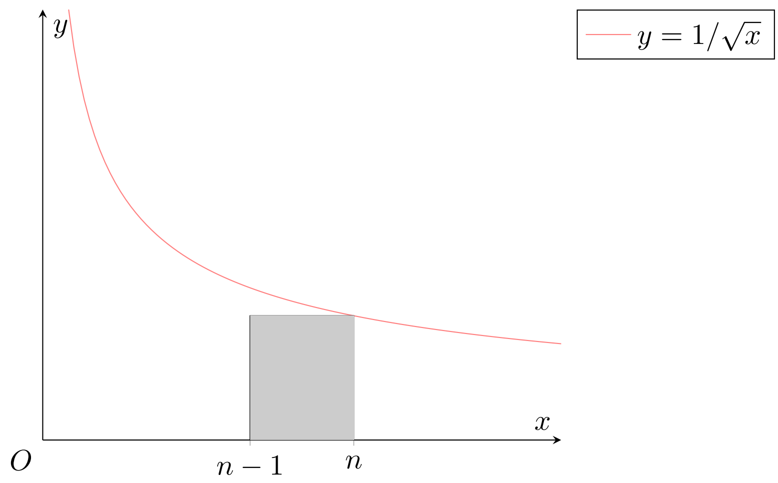
* * * * *

Problem 11. The curve with polar equation $r^2 = a^2 \sin\theta(1 + 2\cos\theta)$, where $r \ge 0$ and a is a positive constant, is shown. Show that the area of the larger loop is nine times that of the smaller loop.



* * * * *

Problem 12. The diagram shows a sketch of the graph of $y = 1/\sqrt{x}$.



By considering the shaded rectangle, and the area of the region between the graph and the x-axis for $n-1 \le x \le n$, where $n \ge 1$, show that

$$\frac{1}{\sqrt{n}} < 2\left(\sqrt{n} - \sqrt{n-1}\right).$$

Deduce that

$$1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} < 2\sqrt{n}.$$

Show also that

$$\frac{1}{\sqrt{n}} > 2\left(\sqrt{n+1} - \sqrt{n}\right).$$

Deduce that

$$1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} > 2\sqrt{n+1} - 2.$$

Hence, find a value of N for which

$$1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{N}} > 1000.$$

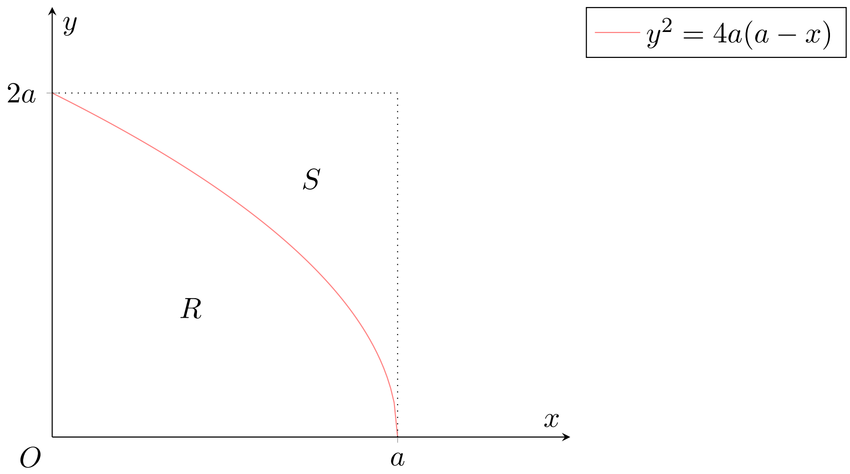
Assignment B8

Problem 1. The diagram shows the region R, which is bounded by the axes and the part of the curve $y^2 = 4a(a - x)$ lying in the first quadrant.

Find, in terms of a, the volume, V_x, of the solid formed when R is rotated completely about the x-axis.

The volume of the solid formed when R is rotated completely about the y-axis is V_y. Show that $V_y = \frac{8}{15} V_x$.

The region S, lying in the first quadrant, is bounded by the curve $y^2 = 4a(a - x)$ and the lines $x = a$ and $y = 2a$. Find, in terms of a, the volume of the solid formed when S is rotated completely about the y-axis.



Solution.

$$V_x = \pi \int_0^a y^2 \, \mathrm{d}x = \pi \int_0^a 4a(a - x) \, \mathrm{d}x = 4\pi a \left[ax - \frac{1}{2} x^2 \right]_0^a = 2\pi a^3 \text{ units}^3.$$

Note that

$$x = a - \frac{y^2}{4a} \implies x^2 = \left(a - \frac{y^2}{4a} \right)^2 = a^2 - \frac{1}{2} y^2 + \frac{1}{16a^2} y^4.$$

Hence,

$$\begin{aligned} V_y &= \pi \int_0^{2a} x^2 \, \mathrm{d}y = \pi \int_0^{2a} \left(a^2 - \frac{1}{2} y^2 + \frac{1}{16a^2} y^4 \right) \mathrm{d}y \\ &= \pi \left[a^2 y - \frac{1}{2} \left(\frac{y^3}{3} \right) + \frac{1}{16a^2} \left(\frac{y^5}{5} \right) \right]_0^{2a} = \frac{16}{15} \pi a^3 = \frac{8}{15} V_x. \end{aligned}$$

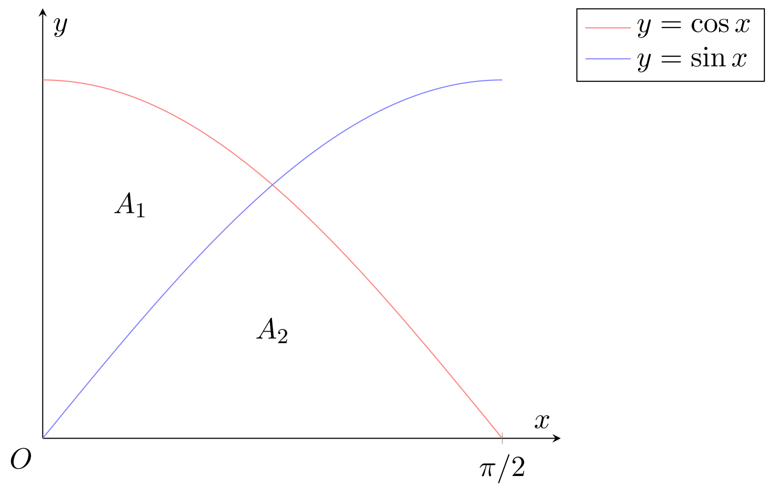
We have

$$\text{Volume} = \text{Volume of cylinder} - V_y = \pi \left(a^2 \right) (2a) - \frac{16}{15} \pi a^3 = \frac{14}{15} \pi a^3 \text{ units}^3.$$

Problem 2. The region bounded by the axes and the curve $y = \cos x$ from $x = 0$ to $x = \frac{1}{2}\pi$ is divided into two parts, of areas A_1 and A_2, by the curve $y = \sin x$.

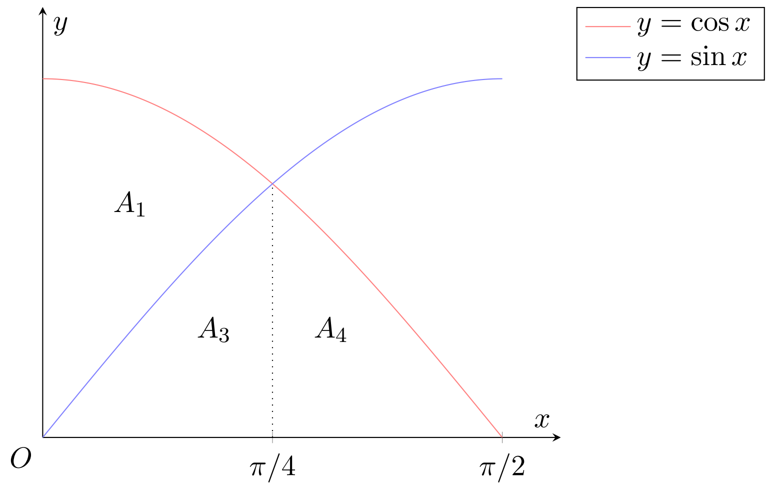
(a) Prove that $A_2 = \sqrt{2}A_1$.

(b) Find the volume of the solid obtained when the region with area A_2 is rotated about the y-axis through 2π radians. Give your answer in exact form.



Solution.

Part (a).



Let A_3 and A_4 be the areas as defined on the diagram above. By the symmetry of $y = \sin x$ and $y = \cos x$ about $x = \pi/4$, we have $A_3 = A_4$.

$$A_3 = \int_0^{\pi/4} \sin x \, \mathrm{d}x = [-\cos x]_0^{\pi/4} = 1 - \frac{\sqrt{2}}{2}.$$

Hence,

$$A_1 = \int_0^{\pi/4} \cos x \, \mathrm{d}x - A_3 = [\sin x]_0^{\pi/4} - \left(1 - \frac{\sqrt{2}}{2}\right) = \frac{\sqrt{2}}{2} - 1 + \frac{\sqrt{2}}{2} = \sqrt{2} - 1.$$

Thus,

$$A_2 = 2A_3 = 2\left(1 - \frac{\sqrt{2}}{2}\right) = \sqrt{2}\left(\sqrt{2} - 1\right) = \sqrt{2}A_1.$$

Part (b). Let V_3 and V_4 be the volumes of the solids obtained when A_3 and A_4 are rotated about the y-axis through 2π radians, respectively.

$$V_3 = 2\pi \int_0^{\pi/4} xy \, \mathrm{d}x = 2\pi \int_0^{\pi/4} x \sin x \, \mathrm{d}x.$$

Integrating by parts,

	D	I
$+$	x	$\sin x$
$-$	1	$-\cos x$
$+$	0	$-\sin x$

Thus,

$$V_3 = 2\pi \left[-x \cos x + \sin x\right]_0^{\pi/4} = \sqrt{2}\pi \left(1 - \frac{\pi}{4}\right).$$

Also,

$$V_4 = 2\pi \int_{\pi/4}^{\pi/2} xy \, \mathrm{d}x = 2\pi \int_{\pi/4}^{\pi/2} x \cos x \, \mathrm{d}x.$$

Integrating by parts,

	D	I
$+$	x	$\cos x$
$-$	1	$\sin x$
$+$	0	$-\cos x$

Thus,

$$V_4 = 2\pi \left[x \sin x + \cos x\right]_{\pi/4}^{\pi/2} = \pi^2 - \sqrt{2}\pi \left(1 + \frac{\pi}{4}\right).$$

Hence, the required volume is

$$V_3 + V_4 = \sqrt{2}\pi \left(1 - \frac{\pi}{4}\right) + \pi^2 - \sqrt{2}\pi \left(1 + \frac{\pi}{4}\right) = \pi^2 \left(1 - \frac{\sqrt{2}}{2}\right) \text{ units}^3.$$

Problem 3. A curve has parametric equations

$$x = \cos^2 t,\ y = \sin^3 t,\ 0 \le t \le \frac{1}{2}\pi.$$

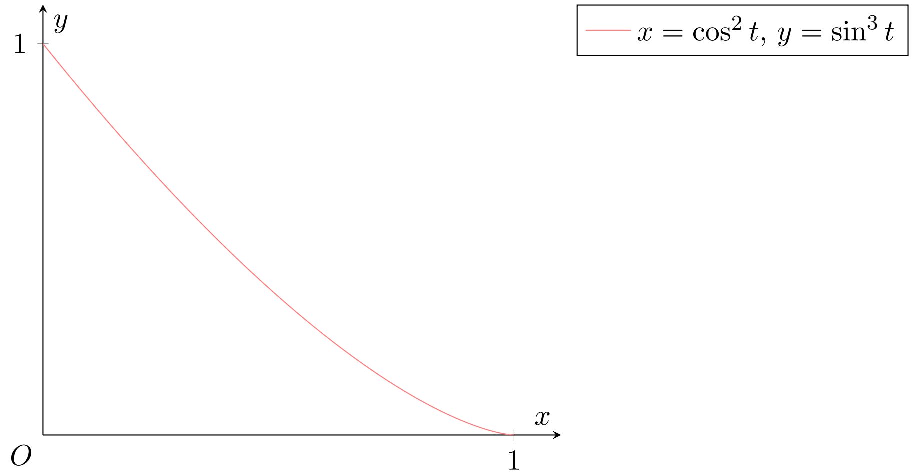
(a) Sketch the curve.

(b) Show that the area under the curve for $0 \le t \le \frac{1}{2}\pi$ is $2\int_0^{\pi/2} \cos t \sin^4 t \,\mathrm{d}t$, and find the exact value of the area.

(c) Find the volume of the solid obtained when the region in (b) is rotated about the y-axis through 2π radians.

Solution.

Part (a).



Part (b). Note that $x = 0 \implies t = \frac{\pi}{2}$ and $x = 1 \implies t = 0$. Hence,

$$\begin{aligned}\text{Area} &= \int_0^1 y \,\mathrm{d}x = \int_{\pi/2}^0 y \frac{\mathrm{d}x}{\mathrm{d}t} \,\mathrm{d}t = \int_{\pi/2}^0 \sin^3 t(-2\cos t \sin t) \,\mathrm{d}t = 2\int_0^{\pi/2} \cos t \sin^4 t \,\mathrm{d}t \\ &= 2\left[\frac{\sin^5 t}{5}\right]_0^{\pi/2} = \frac{2}{5} \text{ units}^2.\end{aligned}$$

Part (c).

$$\begin{aligned}\text{Volume} &= 2\pi \int_0^1 xy \,\mathrm{d}x = 2\pi \int_{\pi/2}^0 \cos^2 t \sin^3 t(-2\cos t \sin t) \,\mathrm{d}t = 4\pi \int_0^{\pi/2} \cos^3 t \sin^4 t \,\mathrm{d}t \\ &= 4\pi \int_0^{\pi/2} \sin^4 t \left(1 - \sin^2 t\right) \cos t \,\mathrm{d}t = 4\pi \int_0^{\pi/2} \left(\sin^4 t - \sin^6 t\right) \cos t \,\mathrm{d}t \\ &= 4\pi \left[\frac{\sin^5 t}{5} - \frac{\sin^7 t}{7}\right]_0^{\pi/2} = \frac{8\pi}{35} \text{ units}^3.\end{aligned}$$

* * * * *

Problem 4.

(a) Given that f is a continuous function, explain, with the aid of a sketch, why the value of

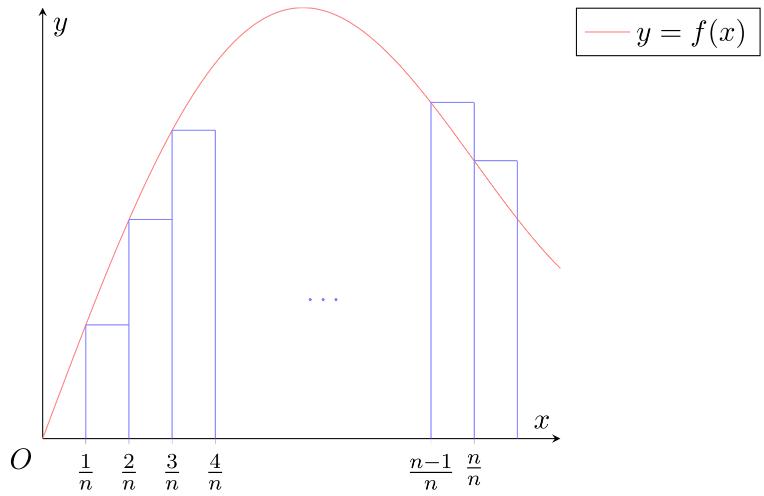
$$\lim_{n\to\infty} \frac{1}{n}\left[f\left(\frac{1}{n}\right) + f\left(\frac{2}{n}\right) + \ldots + f\left(\frac{n}{n}\right)\right]$$

is $\int_0^1 f(x) \,\mathrm{d}x$.

(b) Hence, evaluate $\lim_{n\to\infty} \frac{1}{n}\left(\frac{\sqrt[3]{1}+\sqrt[3]{2}+\ldots+\sqrt[3]{n}}{\sqrt[3]{n}}\right)$.

Solution.

Part (a).



The area of the rectangles in the above figure is given by

$$\frac{1}{n}\left[f\left(\frac{1}{n}\right)+f\left(\frac{2}{n}\right)+\ldots+f\left(\frac{n}{n}\right)\right].$$

This gives an approximation of the signed area under the curve from $x=\frac{1}{n}$ to $x=\frac{n}{n}=1$. As $n\to\infty$, the widths of the rectangles become smaller and the approximation becomes exact. Hence,

$$\lim_{n\to\infty}\frac{1}{n}\left[f\left(\frac{1}{n}\right)+f\left(\frac{2}{n}\right)+\ldots+f\left(\frac{n}{n}\right)\right]=\int_0^1 f(x)\,\mathrm{d}x.$$

Part (b).

$$\lim_{n\to\infty}\frac{1}{n}\left(\frac{\sqrt[3]{1}+\sqrt[3]{2}+\ldots+\sqrt[3]{n}}{\sqrt[3]{n}}\right)=\lim_{n\to\infty}\frac{1}{n}\left[\sqrt[3]{\frac{1}{n}}+\sqrt[3]{\frac{2}{n}}+\ldots+\sqrt[3]{\frac{n}{n}}\right]$$
$$=\int_0^1\sqrt[3]{x}\,\mathrm{d}x=\left[\frac{x^{4/3}}{4/3}\right]_0^1=\frac{3}{4}.$$

* * * * *

Problem 5. The function f satisfies $f'(x)>0$ for $a\le x\le b$, and g is the inverse of f. By making a suitable change of variable, prove that

$$\int_a^b f(x)\,\mathrm{d}x=b\beta-a\alpha-\int_\alpha^\beta g(y)\,\mathrm{d}y$$

where $\alpha=f(a)$ and $\beta=f(b)$. Interpret this formula geometrically by means of a sketch where α and a are positive. Verify this result for the case where $f(x)=\mathrm{e}^{2x}$, $a=0$, $b=1$.

Prove similarly and interpret geometrically the formula

$$2\pi\int_a^b xf(x)\,\mathrm{d}x=\pi(b^2\beta-a^2\alpha)-\pi\int_\alpha^\beta[g(y)]^2\,\mathrm{d}y.$$

Solution. Observe that $y = f(x) \implies \mathrm{d}y = f'(x)\,\mathrm{d}x$. Hence,

$$\int_\alpha^\beta g(y)\,\mathrm{d}y = \int_a^b f^{-1}(f(x))f'(x)\,\mathrm{d}x = \int_a^b xf'(x)\,\mathrm{d}x.$$

Integrating by parts,

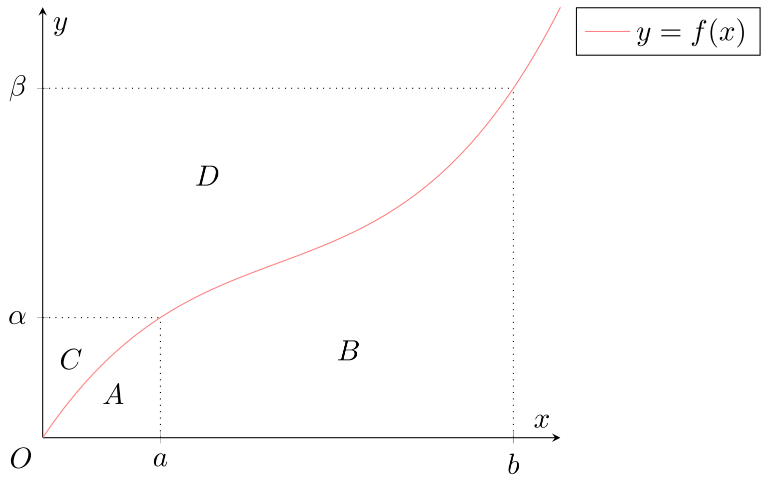
	D	I
$+$	x	$f'(x)$
$-$	1	$f(x)$

Hence,

$$\int_\alpha^\beta g(y)\,\mathrm{d}y = [xf(x)]_a^b - \int_a^b f(x)\,\mathrm{d}x = b\beta - a\alpha - \int_a^b f(x)\,\mathrm{d}x.$$

Thus,

$$\int_a^b f(x)\,\mathrm{d}x = b\beta - a\alpha - \int_\alpha^\beta g(y)\,\mathrm{d}y.$$



Consider the above diagram. We clearly have $[A \cup C] = a\alpha$, $[A \cup B \cup C \cup D] = b\beta$, $[B] = \int_a^b f(x)\,\mathrm{d}x$ and $[D] = \int_\alpha^\beta g(y)\,\mathrm{d}y$. Thus,

$$\int_a^b f(x)\,\mathrm{d}x = [B] = [A \cup B \cup C \cup D] - [A \cup C] - [D] = b\beta - a\alpha - \int_\alpha^\beta g(y)\,\mathrm{d}y.$$

Using the standard way, we get

$$\int_0^1 \mathrm{e}^{2x}\,\mathrm{d}x = \left[\frac{1}{2}\mathrm{e}^{2x}\right]_0^1 = \frac{\mathrm{e}^2 - 1}{2}.$$

We now use the formula. Let $f(x) = e^{2x}$. Then $g(x) = \frac{1}{2}\ln x$. Hence, $\alpha = g(0) = 1$ and $\beta = g(1) = e^2$. Invoking the above formula,

$$\int_0^1 \mathrm{e}^{2x}\,\mathrm{d}x = 1\left(\mathrm{e}^2\right) - 0(1) - \int_1^{\mathrm{e}^2} \frac{1}{2}\ln x\,\mathrm{d}x = \mathrm{e}^2 - \frac{1}{2}\left[x\ln x - x\right]_1^{\mathrm{e}^2} = \frac{\mathrm{e}^2 - 1}{2}.$$

Hence, the formula holds for the above case.

Similar to the above part, we have

$$\int_\alpha^\beta [g(y)]^2 \, \mathrm{d}y = \int_\alpha^\beta \left[f^{-1}(f(x))\right]^2 f'(x) \, \mathrm{d}x = \int_a^b x^2 f'(x) \, \mathrm{d}x.$$

Integrating by parts,

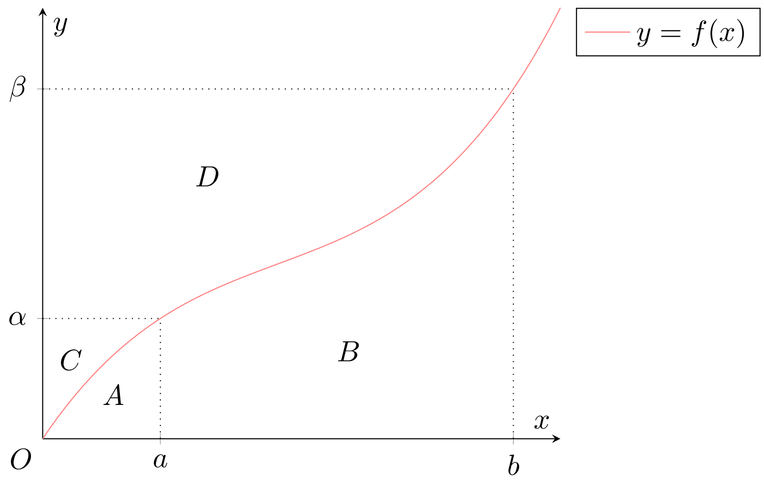
	D	I
$+$	x^2	$f'(x)$
$-$	$2x$	$f(x)$

Thus,

$$\int_\alpha^\beta [g(y)]^2 \, \mathrm{d}y = \left[x^2 f(x)\right]_a^b - 2\int_a^b x f(x) \, \mathrm{d}x = b^2\beta - a^2\alpha - 2\int_a^b x f(x) \, \mathrm{d}x.$$

Rearranging,

$$2\pi \int_a^b x f(x) \, \mathrm{d}x = \pi\left(b^2\beta - a^2\alpha\right) - \pi \int_\alpha^\beta [g(y)]^2 \, \mathrm{d}y.$$



Let $V(R)$ represent the volume of the solid obtained when a region R is rotated completely about the y-axis.

We clearly have $V(A \cup B \cup C \cup D) = \pi b^2 \beta$, $V(A \cup C) = \pi a^2 \alpha$, $V(B) = 2\pi \int_a^b x f(x) \, \mathrm{d}x$ (using the shell method), and $V(D) = \pi \int_\alpha^\beta [g(y)]^2 \, \mathrm{d}y$ (using the disc method). Thus,

$$\begin{aligned} 2\pi \int_a^b x f(x) \, \mathrm{d}x &= V(B) = V(A \cup B \cup C \cup D) - V(A \cup C) - V(D) \\ &= \pi b^2 \beta - \pi a^2 \alpha - \pi \int_\alpha^\beta [g(y)]^2 \, \mathrm{d}y = \pi\left(b^2\beta - a^2\alpha\right) - \pi \int_\alpha^\beta [g(y)]^2 \, \mathrm{d}y. \end{aligned}$$

B9 Applications of Integration II - Arc Length and Surface Area

Tutorial B9

Problem 1. Calculate the exact length of each of the arcs of the following curves.

(a) $y^3 = x^2$ for $-1 \le x \le 1$.

(b) $x = t^2 - 1$, $y = t^3 + 1$ from $t = 0$ to $t = 1$.

(c) $r = a\cos\theta$ from $\theta = 0$ to $\theta = \pi/2$.

Solution.
Part (a). Note that

$$y^3 = x^2 \implies y = x^{2/3} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2}{3}x^{-1/3}.$$

Hence,

$$\sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} = \sqrt{1 + \left(\frac{2}{3}x^{-1/3}\right)^2} = \sqrt{1 + \frac{4}{9}x^{-2/3}}.$$

Thus,

$$\begin{aligned}\text{Length} &= \int_{-1}^{1} \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = \int_{-1}^{1} \sqrt{1 + \frac{4}{9}x^{-2/3}}\,\mathrm{d}x = 2\int_0^1 \sqrt{1 + \frac{4}{9}x^{-2/3}}\,\mathrm{d}x \\ &= 3\int_0^1 \frac{2}{3}x^{-1/3}\sqrt{x^{2/3} + \frac{4}{9}}\,\mathrm{d}x = 3\left[\frac{2}{3}\left(x^{2/3} + \frac{4}{9}\right)^{3/2}\right]_0^1 = \frac{2}{27}\left(13\sqrt{13} - 8\right) \text{ units.}\end{aligned}$$

Part (b). Since the arc length of a curve is invariant under translation, it suffices to find the arc length of the curve with parametric equations $x = t^2, y = t^3$, $0 \le t \le 1$. The Cartesian equation of this curve is $y = x^{3/2}$, $0 \le x \le 1$, which is the inverse of $y = x^{2/3}$, $0 \le x \le 1$. From part (a), the required arc length is

$$\frac{1}{2} \cdot \frac{2}{27}\left(13\sqrt{13} - 8\right) = \frac{1}{27}\left(13\sqrt{13} - 8\right) \text{ units.}$$

Part (c). Since $r = a\cos\theta$, $0 \le \theta \le \pi/2$ describes the top half of a circle with centre $(a/2, 0)$ and diameter a, the arc length of the curve is $\pi a/2$ units.

* * * * *

Problem 2. Find the exact areas of the surfaces generated by completely rotating the following arcs about the (i) x-axis and (ii) y-axis.

(a) The line $2y = x$ between the origin and the point $(4, 2)$.

(b) The curve $x = t^3 - 3t + 2$, $y = 3\left(t^2 - 1\right)$, $t \in \mathbb{R}$ from $t = 1$ to $t = 2$.

Solution.

Part (a).

Part (a)(i). When rotated about the x-axis, the curve forms a cone with slant height $\sqrt{4^2+2^2} = 2\sqrt{5}$ and radius 2. Hence, the required surface area is $\pi(2)(2\sqrt{5}) = 4\sqrt{5}\pi$ units^2.

Part (a)(ii). When rotated about the y-axis, the curve forms a cone with slant height $\sqrt{4^2+2^2} = 2\sqrt{5}$ and radius 4. Hence, the required surface area is $\pi(4)(2\sqrt{5}) = 8\sqrt{5}\pi$ units^2.

Part (b). Note that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = 3t^2 - 3, \quad \frac{\mathrm{d}y}{\mathrm{d}t} = 6t.$$

Hence,

$$\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2} = \sqrt{(3t^2-3)^2 + (6t)^2} = \sqrt{(3t^2+3)^2} = 3t^2 + 3.$$

Part (b)(i).

$$\begin{aligned}\text{Area} &= 2\pi \int_1^2 y \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t = 2\pi \int_1^2 3\left(t^2-1\right)\left(3t^2+3\right)\,\mathrm{d}t \\ &= 18\pi \int_1^2 \left(t^4 - 1\right)\,\mathrm{d}t = 18\pi \left[\frac{1}{5}t^5 - t\right]_1^2 = \frac{468}{5}\pi \text{ units}^2.\end{aligned}$$

Part (b)(ii).

$$\text{Area} = 2\pi \int_1^2 x \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t = 2\pi \int_1^2 \left(t^3 - 3t + 2\right)\left(3t^2+3\right)\,\mathrm{d}t$$

$$= 6\pi \int_1^2 \left(t^5 - 2t^3 + 2t^2 - 3t + 2\right)\,\mathrm{d}t = 6\pi \left[\frac{1}{6}t^6 - \frac{2}{4}t^4 - \frac{2}{3}t^3 - \frac{3}{2}t^2 + 2t\right]_1^2 = 31\pi \text{ units}^2.$$

* * * * *

Problem 3. The section of the curve $y = \mathrm{e}^x$ between $x = 0$ and $x = 1$ is rotated through one revolution about

(a) the x-axis.

(b) the y-axis.

Find the numerical values of the areas of the surfaces obtained.

Solution.

Part (a).

$$\text{Area} = 2\pi \int_0^1 y \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = 2\pi \int_0^1 \mathrm{e}^x \sqrt{1 + \mathrm{e}^{2x}}\,\mathrm{d}x = 22.9 \text{ units}^2 \text{ (3 s.f.)}.$$

Part (b). Note that $y = \mathrm{e}^x \implies x = \ln y$ and $\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^x \implies \frac{\mathrm{d}x}{\mathrm{d}y} = \mathrm{e}^{-x}$.

$$\text{Area} = 2\pi \int_1^{\mathrm{e}} x \sqrt{1 + \left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2}\,\mathrm{d}y = 2\pi \int_0^1 \ln y \sqrt{1 + \mathrm{e}^{-2x}}\,\mathrm{d}x = 7.05 \text{ units}^2 \text{ (3 s.f.)}.$$

Problem 4. The curve $y^2 = \frac{1}{3}x(1-x)^2$ has a loop between $x=0$ and $x=1$. Prove that the total length of the loop is $\frac{4\sqrt{3}}{3}$.

Solution. Since the curve is even with respect to y, it is symmetric about the x-axis. We thus only consider the part of the curve above the x-axis, i.e. $y \geq 0$, where $y = (1-x)\sqrt{x/3}$. Differentiating,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\sqrt{3}}\left(-\sqrt{x} + \frac{1-x}{2\sqrt{x}}\right) = \frac{1-3x}{2\sqrt{3x}} \implies 1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 = 1 + \frac{(1-3x)^2}{12x} = \frac{(1+3x)^2}{12x}.$$

Thus,

$$\text{Length} = 2\int_0^1 \sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = 2\int_0^1 \frac{1+3x}{\sqrt{12x}}\,\mathrm{d}x = \frac{1}{\sqrt{12}}\left[\frac{x^{1/2}}{1/2} + \frac{3x^{3/2}}{3/2}\right]_0^1 = \frac{4\sqrt{3}}{3} \text{ units.}$$

* * * * *

Problem 5. The tangent at a point P on the curve $x = a\left(t - \frac{1}{3}t^3\right)$, $y = at^2$ cuts the x-axis at T. Prove that the distance of the point T from the origin O is half the length of the arc OP.

Solution. Let P be the point on the curve with parameter $t = t_P$. Note that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = a\left(1-t^2\right), \quad \frac{\mathrm{d}y}{\mathrm{d}t} = 2at.$$

Thus,

$$\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 = \left[a\left(1-t^2\right)\right]^2 + (2at)^2 = a^2\left(t^2+1\right)^2.$$

Thus,

$$\begin{aligned}\text{Length of arc } OP &= \int_0^{t_P} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t = a\int_0^{t_P} \left(t^2+1\right)\,\mathrm{d}t \\ &= a\left[\frac{t^3}{3} + t\right]_0^{t_P} = a\left(\frac{t_P^3}{3} + t_P\right) \text{ units.}\end{aligned}$$

Note that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{2at}{a\left(1-t^2\right)} = \frac{2t}{1-t^2}.$$

Hence, the equation of the tangent at P is given by

$$y - at_P^2 = \frac{2t_P}{1-t_P^2}\left[x - a\left(t_P - \frac{t_P^3}{3}\right)\right].$$

At T, $x = OT$ and $y = 0$. Hence,

$$0 - at_P^2 = \frac{2t_P}{1-t_P^2}\left[OT - a\left(t_P - \frac{t_P^3}{3}\right)\right],$$

whence

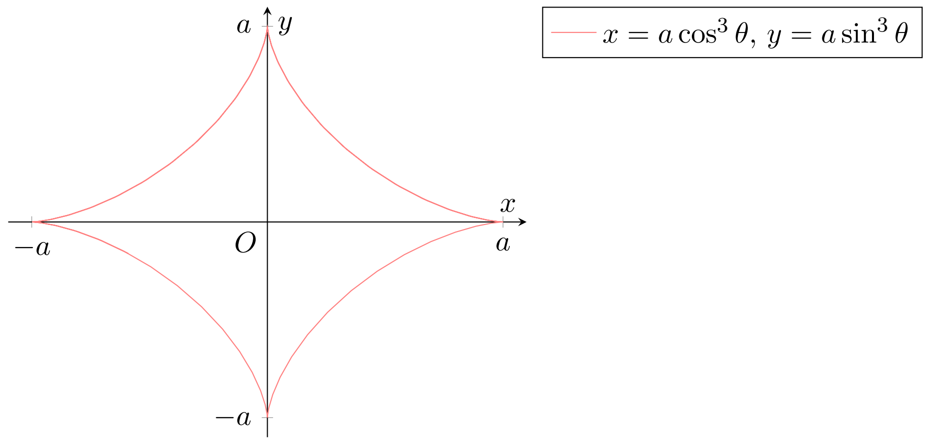
$$\begin{aligned}OT &= \frac{-at_P^2\left(1-t_P^2\right)}{2t_P} + a\left(t_P - \frac{t_P^3}{3}\right) = \frac{a}{2}\left[\left(-t_P + t_P^3\right) + \left(2t_P - \frac{2t_P^3}{3}\right)\right] \\ &= \frac{a}{2}\left(\frac{t_P^3}{3} + t_P\right) = \frac{OP}{2}.\end{aligned}$$

Problem 6. Sketch the curve whose parametric equations are $x = a\cos^3\theta$, $y = a\sin^3\theta$, $a > 0$.

(a) Find the total length of the curve.

(b) The portion of the curve in the first quadrant is revolved through four right angles about the x-axis. Prove that the area of the surface thus formed is $\frac{6}{5}\pi a^2$.

Solution.



Part (a). By symmetry, we only consider the length of the curve in the first quadrant. Note that $x = 0 \implies \theta = \pi/2$ and $x = a \implies \theta = 0$. Also,

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = -3a\cos^2\theta\sin\theta, \quad \frac{\mathrm{d}y}{\mathrm{d}\theta} = 3a\sin^2\theta\cos\theta.$$

Hence,

$$\begin{aligned}\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2 &= (-3a\cos^2\theta\sin\theta)^2 + (3a\sin^2\theta\cos\theta)^2 \\ &= 9a^2\left(\cos^4\theta\sin^2\theta + \sin^4\theta\cos^2\theta\right) = (3a\cos\theta\sin\theta)^2.\end{aligned}$$

Thus,

$$\begin{aligned}\text{Length} = 4\int_0^{\pi/2}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta &= 12a\int_0^{\pi/2}\cos\theta\sin\theta\,\mathrm{d}\theta \\ &= 12a\left[\frac{\sin^2\theta}{2}\right]_0^{\pi/2} = 6a \text{ units.}\end{aligned}$$

Part (b).

$$\begin{aligned}\text{Area} = 2\pi\int_0^{\pi/2} x\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta &= 2\pi\int_0^{\pi/2} a\cos^3\theta\,(3a\cos\theta\sin\theta)\,\mathrm{d}t \\ = 6\pi a^2\int_0^{\pi/2}\sin\theta\cos^4\theta\,\mathrm{d}\theta &= 6\pi a^2\left[-\frac{\cos^5\theta}{5}\right]_0^{\pi/2} = \frac{6}{5}\pi a^2 \text{ units}^2.\end{aligned}$$

Problem 7. The parametric equations of a curve are given by

$$x = \tan t - t,\ y = \ln \sec t,\ t \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

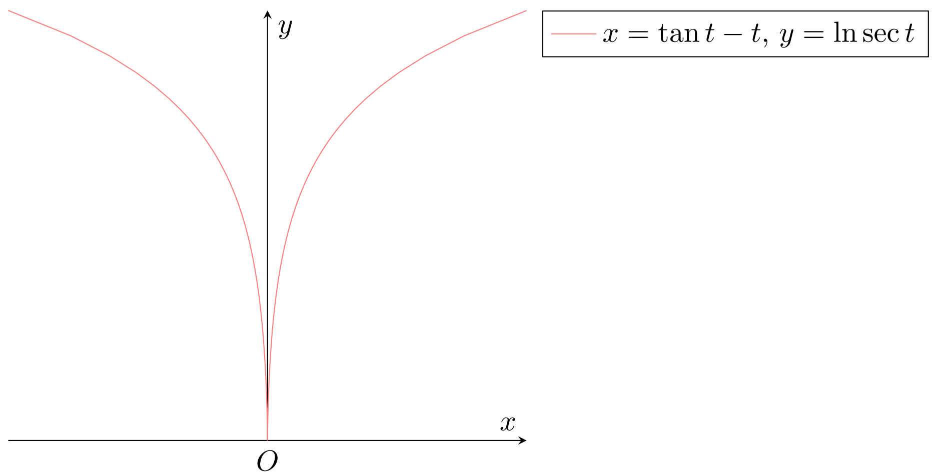
(a) Sketch the curve.

(b) Prove that the arc length of the curve measured from the origin to the point $\left(1 - \frac{\pi}{4}, \frac{1}{2}\ln 2\right)$ is $\sqrt{2} - 1$.

(c) The arc in (b) is rotated about the x-axis through an angle of 360°. Find the exact surface area formed.

Solution.

Part (a).



Part (b). Note that $x = 0 \implies t = 0$ and $x = 1 - \pi/4 \implies = t = \pi/4$. Further,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sec^2 t - 1 = \tan^2 t, \quad \frac{\mathrm{d}y}{\mathrm{d}t} = \tan t.$$

Thus,

$$\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 = \left(\tan^2 t\right)^2 + (\tan t)^2 = \tan^2 t\left(\tan^2 t + 1\right) = \tan^2 t \sec^2 t.$$

Hence,

$$\text{Length} = \int_0^{\pi/4} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t = \int_0^{\pi/4} \tan t \sec t\,\mathrm{d}t = [\sec t]_0^{\pi/4} = \sqrt{2} - 1 \text{ units.}$$

Part (c). We have

$$\text{Area} = 2\pi \int_0^{\pi/4} y\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t = 2\pi \int_0^{\pi/4} \ln \sec t \cdot \tan t \sec t\,\mathrm{d}t.$$

Integrating by parts,

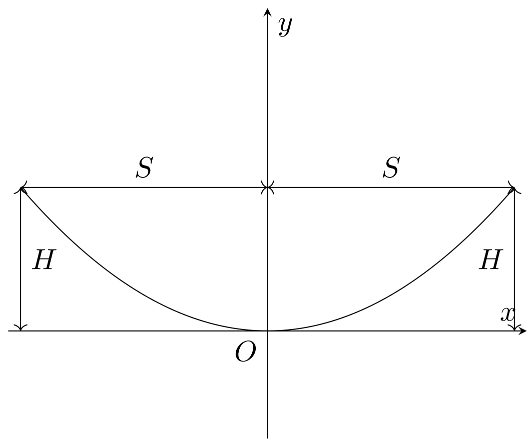
	D	I
$+$	$\ln \sec t$	$\tan t \sec t$
$-$	$\tan t$	$\sec t$

Thus,

$$\text{Area} = 2\pi\left[[\sec t \ln \sec t]_0^{\pi/4} - \int_0^{\pi/4} \tan t \sec t \,\mathrm{d}t\right] = 2\pi\left[\sqrt{2}\ln\sqrt{2} - \left(\sqrt{2}-1\right)\right]$$
$$= \sqrt{2}\pi\left(\ln 2 - 2 + \sqrt{2}\right) \text{ units}^2.$$

* * * * *

Problem 8.



The diagram shows a cable for a suspension bridge, which has the shape of a parabola with equation $y = kx^2$. The suspension bridge has a total span $2S$ and the height of the cable relative to the lowest point is H at each end. Show that the total length of the cable is $L = 2\int_0^S \sqrt{1 + \frac{4H^2}{S^4}x^2}\,\mathrm{d}x$.

(a) Engineers from country A proposed a suspension bridge across a strait of 8 km wide to country B. The plan included suspension towers 380 m high at each end. Find the length of the parabolic cable for this proposed bridge to the nearest metre.

(b) By using the result $\frac{\mathrm{d}}{\mathrm{d}x}\ln\left(x + \sqrt{a^2+x^2}\right) = \frac{1}{\sqrt{a^2+x^2}}$ or otherwise, find L in terms of S and H.

Solution. By symmetry, we only need to consider the length of the curve where $x \geq 0$. Since (S, H) is on the curve, $H = kS^2 \implies k = \frac{H}{S^2}$. Note that

$$y = kx^2 \implies \frac{\mathrm{d}y}{\mathrm{d}x} = 2kx = \frac{2H}{S^2}x.$$

Hence,

$$L = 2\int_0^S \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = 2\int_0^S \sqrt{1 + \frac{4H^2}{S^4}x^2}\,\mathrm{d}x.$$

Part (a). Note that $2S = 8000 \implies S = 4000$ and $H = 380$. Hence,

$$L = 2\int_0^{4000} \sqrt{1 + \frac{4(380)^2}{(4000)^4}x^2}\,\mathrm{d}x = 8048 \text{ (to the nearest integer)}.$$

The bridge is thus 8048 m long.

Part (b). Consider the integral $I = \int \sqrt{1+(kx)^2}\,\mathrm{d}x$. Under the substitution $kx = \tan\theta$, we get

$$I = \int \sqrt{1+(kx)^2}\,\mathrm{d}x = \frac{1}{k}\int \sqrt{1+\tan^2\theta}\sec^2\theta\,\mathrm{d}\theta = \frac{1}{k}\int \sec^3\theta\,\mathrm{d}\theta.$$

Integrating by parts,

	D	I
$+$	$\sec\theta$	$\sec^2\theta$
$-$	$\sec\theta\tan\theta$	$\tan t$

Hence,

$$kI = \sec\theta\tan\theta - \int \sec\theta\tan^2\theta\,\mathrm{d}\theta = \sec\theta\tan\theta - \int \sec\theta\left(\sec^2\theta - 1\right)\,\mathrm{d}\theta$$

$$= \sec\theta\tan\theta - \int \sec^3\theta\,\mathrm{d}\theta + \int \sec\theta\,\mathrm{d}\theta = \sec\theta\tan\theta - kI + \ln\left|\sec\theta + \tan\theta\right|.$$

Thus,

$$I = \frac{\sec\theta\tan\theta + \ln\left|\sec\theta+\tan\theta\right|}{2k} + C = \frac{1}{2k}\left[kx\sqrt{(kx)^2+1} + \ln\left|\sqrt{(kx)^2+1}+kx\right|\right] + C.$$

In our case, $k = \frac{2H}{S^2} > 0$. Hence,

$$L = 2\left[\frac{1}{2}\left(\frac{S^2}{2H}\right)\left[\left(\frac{2H}{S^2}x\right)\sqrt{\left(\frac{2H}{S^2}x\right)^2+1} + \ln\left(\sqrt{\left(\frac{2H}{S^2}x\right)^2+1}+\frac{2H}{S^2}x\right)\right]\right]_0^S$$

$$= \frac{S^2}{2H}\left[\left(\frac{2H}{S}\right)\sqrt{\left(\frac{2H}{S}\right)^2+1} + \ln\left(\sqrt{\left(\frac{2H}{S}\right)^2+1}+\frac{2H}{S}\right)\right]$$

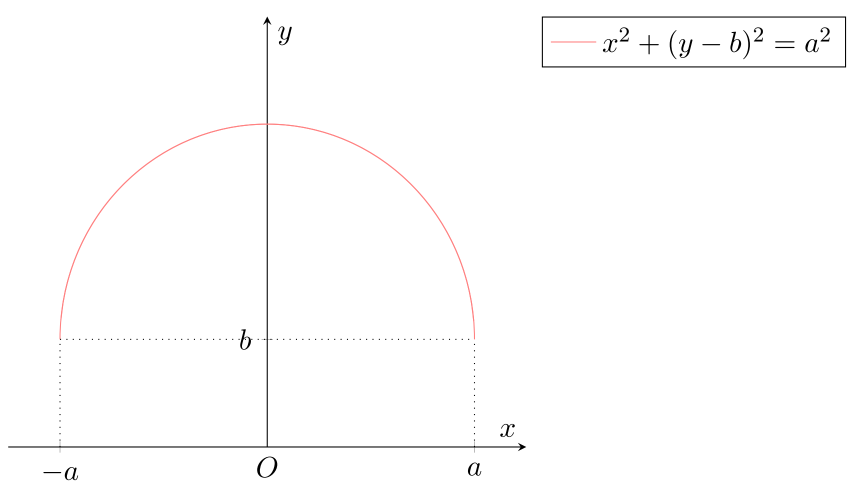
$$= \sqrt{4H^2+S^2} + \frac{S^2}{2H}\ln\left(\frac{\sqrt{4H^2+S^2}+2H}{S}\right).$$

* * * * *

Problem 9. Sketch the semicircle with equation $x^2+(y-b)^2 = a^2$, $y \geq b$ where a and b are positive constants.

A solid is formed by rotating the region bounded by the semicircle and its diameter on the line $y = b$ about the x-axis through 4 right angles. Find the total surface area of the solid.

Solution.



Observe that
$$x^2 + (y-b)^2 = a^2 \implies y = b + \sqrt{a^2 - x^2},$$
whence
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{-2x}{2\sqrt{a^2-x^2}} = -\frac{x}{\sqrt{a^2-x^2}}.$$
Thus,
$$1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 = 1 + \frac{x^2}{a^2-x^2} = \frac{a^2}{a^2-x^2}.$$
Hence,
$$\text{Area} = 2\pi \int_{-a}^{a} y\sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x + 2\pi(b)(2a) = 4\pi \int_0^a y\sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x + 4\pi ab$$
$$= 4\pi \int_0^a \left(b + \sqrt{a^2-x^2}\right)\left(\frac{a}{\sqrt{a^2-x^2}}\right)\mathrm{d}x + 4\pi ab = 4\pi a \int_0^a \left(\frac{b}{\sqrt{a^2-x^2}} + 1\right)\mathrm{d}x + 4\pi ab$$
$$= 4\pi a \left[b \arcsin\frac{x}{a} + x\right]_0^a + 4\pi ab = \left(2\pi^2 ab + 4\pi a^2 + 4\pi ab\right)\ \text{units}^2$$

* * * * *

Problem 10. Using polar coordinates with pole O, the curve C has the equation $r = a\mathrm{e}^{\theta/k}$, where a and k are positive constants and $0 \le \theta \le 2\pi$. The points A and B on the curve corresponds to $\theta = 0$ and $\theta = \beta$ respectively where $0 < \beta < \pi$. The length of the arc AB is denoted by q and the area of the sector OAB is denoted by Q.

(a) Show that $Q = \frac{1}{4}ka^2\left(\mathrm{e}^{2\beta/k} - 1\right)$.

(b) Show that $q = a(1+k^2)^{1/2}\left(\mathrm{e}^{\beta/k} - 1\right)$.

(c) Deduce from the results of parts (a) and (b) that, for large values of k, $\frac{Q}{q} \approx \frac{1}{2}a$.

(d) Draw a sketch of C for the case where k is large and explain how the result in part (c) can be deduced from the sketch.

Solution.

Part (a).
$$Q = \frac{1}{2}\int_0^\beta r^2\,\mathrm{d}\theta = \frac{a^2}{2}\int_0^\beta \mathrm{e}^{2\theta/k}\,\mathrm{d}\theta = \frac{a^2}{2}\left[\frac{\mathrm{e}^{2\theta/k}}{2/k}\right]_0^\beta = \frac{a^2k}{4}\left(\mathrm{e}^{2\beta/k} - 1\right).$$

Part (b). Note that
$$r = a\mathrm{e}^{\theta/k} \implies \frac{\mathrm{d}r}{\mathrm{d}\theta} = \frac{a\mathrm{e}^{\theta/k}}{k} = \frac{r}{k}.$$
Hence,
$$q = \int_0^\beta \sqrt{r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta = \int_0^\beta \sqrt{r^2 + \frac{r^2}{k^2}}\,\mathrm{d}\theta = \sqrt{1+k^{-2}}\int_0^\beta r\,\mathrm{d}\theta$$
$$= \sqrt{1+k^{-2}}\int_0^\beta a\mathrm{e}^{\theta/k}\,\mathrm{d}\theta = a\sqrt{1+k^{-2}}\left[\frac{\mathrm{e}^{\theta/k}}{1/k}\right]_0^\beta = a\sqrt{k^2+1}\left(\mathrm{e}^{\beta/k} - 1\right).$$

Part (c).

$$\lim_{k\to\infty} \frac{Q}{q} = \lim_{k\to\infty} \frac{\frac{1}{4}a^2 k \left(\mathrm{e}^{2\beta/k} - 1\right)}{a\sqrt{k^2+1}\left(\mathrm{e}^{\beta/k} - 1\right)} = \frac{a}{4} \lim_{k\to\infty} \left(\frac{k}{\sqrt{k^2+1}}\right) \lim_{k\to\infty} \left(\frac{\mathrm{e}^{2\beta/k} - 1}{\mathrm{e}^{\beta/k} - 1}\right).$$

Now observe that

$$\lim_{k\to\infty} \left(\frac{k}{\sqrt{k^2+1}}\right) = \lim_{k\to\infty} \frac{1}{1+k^{-2}} = 1,$$

and by the difference of squares identity,

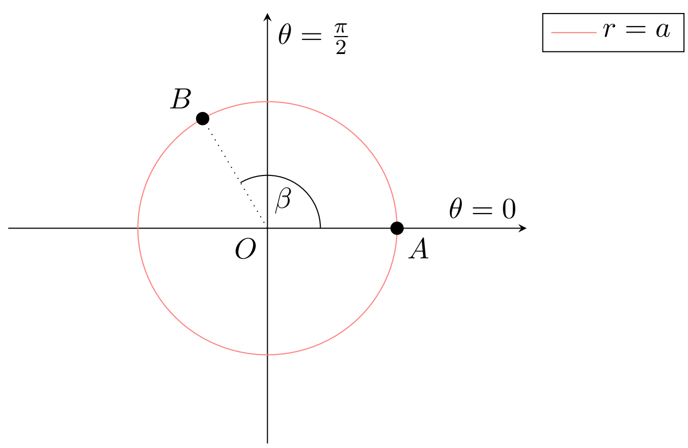
$$\lim_{k\to\infty} \left(\frac{e^{2\beta/k} - 1}{e^{\beta/k} - 1}\right) = \lim_{k\to\infty} \left(\mathrm{e}^{\beta/k} + 1\right) = 2.$$

Hence,

$$\lim_{k\to\infty} \frac{Q}{q} = \frac{a}{2}.$$

Part (d). Note that

$$\lim_{k\to\infty} r = \lim_{k\to\infty} a\mathrm{e}^{\theta/k} = a.$$



As $k \to \infty$, the curve becomes a circle. Hence, Q is the area of a sector with angle β, and q is the arc length of a sector with angle β. Thus,

$$\frac{Q}{q} = \left(\frac{\beta}{2\pi} \cdot \pi a^2\right) \Big/ \left(\frac{\beta}{2\pi} \cdot 2\pi a\right) = \frac{a}{2}.$$

Self-Practice B9

Problem 1. The arc of the curve $y^2 = 4ax$, for which $y \geq 0$ and $0 \leq x \leq a$, is rotated through 2π radians about the x-axis. Prove that the area of the surface so generated is $\frac{8}{3}(2\sqrt{2}-1)\pi a^2$.

Solution. Note that $y = \sqrt{4ax}$, so $\mathrm{d}y/\mathrm{d}x = \sqrt{a/x}$, hence the surface area of the solid generated is given by

$$\begin{aligned}\text{Area} &= 2\pi \int_0^a y\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = 2\pi \int_0^a \sqrt{4ax}\sqrt{1+\frac{a}{x}}\,\mathrm{d}x = 4\pi\sqrt{a}\int_0^a \sqrt{x+a}\,\mathrm{d}x \\ &= 4\pi\sqrt{a}\left[\frac{2}{3}(x+a)^{3/2}\right]_0^a = \frac{8}{3}\pi\sqrt{a}\left[(2a)^{3/2}-a^{3/2}\right] = \frac{8}{3}\left(2\sqrt{2}-1\right)\pi a^2 \text{ units}^2.\end{aligned}$$

* * * * *

Problem 2. The area bounded by the ellipse with parametric equations $x = 3\cos\theta, y = 2\sqrt{2}\sin\theta$ and the positive x- and y-axis is rotated completely about the y-axis. Find the curved surface area of the solid.

Solution. Note that $\mathrm{d}x/\mathrm{d}\theta = -3\sin\theta$ and $\mathrm{d}y/\mathrm{d}\theta = 2\sqrt{2}\cos\theta$, so the surface area of the solid is

$$2\pi\int_0^{\pi/2} x\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta = 2\pi\int_0^{\pi/2}(3\cos\theta)\sqrt{9\sin^2\theta+8\cos^2\theta}\,\mathrm{d}\theta = 54.4 \text{ units}^2.$$

* * * * *

Problem 3. A curve is defined parametrically by $x = 2\sqrt{2}a\sin\theta$, $y = \frac{1}{2}a\sin 2\theta$.

Show that

$$\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2 = a^2\left(2+\cos 2\theta\right)^2.$$

The portion of the curve from $\theta = 0$ to $\theta = \pi/3$ is rotated completely about the x-axis. Find the exact surface area generated.

Solution. Note that

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = 2\sqrt{2}a\cos\theta \quad \text{and} \quad \frac{\mathrm{d}y}{\mathrm{d}\theta} = a\cos 2\theta,$$

so

$$\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2 = 8a^2\cos^2\theta+\alpha^2\cos^2 2\theta = a^2\left(4+4\cos 2\theta+\cos^2 2\theta\right) = a^2\left(\cos 2\theta+2\right)^2.$$

The surface area generated is

$$\begin{aligned}\text{Area} &= 2\pi\int_0^{\pi/3} y\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta = 2\pi\int_0^{\pi/3}\frac{1}{2}a^2\sin 2\theta\left(\cos 2\theta+2\right)\,\mathrm{d}\theta \\ &= \pi a^2\int_0^{\pi/3}\left(\frac{1}{2}\sin 4\theta+2\sin 2\theta\right)\,\mathrm{d}\theta = \pi a^2\left[-\frac{1}{8}\cos 4\theta-\cos 2\theta\right]_0^{\pi/3} \\ &= \frac{27}{16}\pi a^2 \text{ units}^2.\end{aligned}$$

* * * * *

Problem 4. A curve is defined parametrically by $x = t^2 - 2\ln t$, $y = 4(t-1)$, where $t \in \mathbb{R}$, $t \geq 1$.

(a) The points A and B on the curve are given by $t = 1$ and $t = 2$ respectively. Show that the length of the arc AB of the curve is $3 + 2\ln 2$.

(b) The arc AB is rotated through one revolution about the x-axis. Show that the area of the curved surface generated is $\frac{8}{3}\pi(11 - 6\ln 2)$.

Solution.

Part (a). Note that $\mathrm{d}x/\mathrm{d}t = 2t - 2/t$ and $\mathrm{d}y/\mathrm{d}t = 4$, so

$$\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2 = \left(2t - \frac{2}{t}\right)^2 + 4^2 = \left(2t + \frac{2}{t}\right)^2.$$

Thus, the arc length AB is given by

$$\widehat{AB} = \int_1^2 \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}t = \int_1^2 \left(2t + \frac{2}{t}\right)\mathrm{d}t = \left[t^2 + 2\ln t\right]_1^2 = 3 + 2\ln 2 \text{ units.}$$

Part (b). The surface area of the solid generated is given by

$$\begin{aligned}\text{Area} &= 2\pi \int_1^2 y\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}t = 8\pi \int_1^2 (t-1)\left(2t + \frac{2}{t}\right)\mathrm{d}t \\ &= 16\pi \int_1^2 bpt^2 - t + 1 - \frac{1}{t}\,\mathrm{d}t = 16\pi\left[\frac{1}{3}t^3 - \frac{1}{2}t^2 + t - \ln t\right]_1^2 \\ &= \frac{8\pi}{3}(11 - 6\ln 2) \text{ units}^2.\end{aligned}$$

* * * * *

Problem 5. The curve Γ has polar equation $r = ke^\theta$, where k is a positive constant and $0 \le \theta \le \pi$. The points P and Q on Γ correspond to $\theta = \alpha$ and $\theta = \beta$ respectively ($\beta > \alpha$). The area of the region bounded by the lines $\theta = \alpha$, $\theta = \beta$ and the arc PQ is denoted by A. The length of the arc PQ is denoted by s.

(a) Find expressions for A and s in terms of α, β and k.

(b) Deduce that

$$\frac{A}{s^2} = \frac{1}{8}\left(\frac{\mathrm{e}^\beta + \mathrm{e}^a}{\mathrm{e}^b - \mathrm{e}^a}\right).$$

Solution.

Part (a). We have

$$A = \frac{1}{2}\int_\alpha^\beta r^2\,\mathrm{d}\theta = \frac{1}{2}\int_\alpha^\beta k^2\mathrm{e}^{2\theta}\,\mathrm{d}\theta = \frac{1}{2}k^2\left[\frac{1}{2}\mathrm{e}^{2\theta}\right]_\alpha^\beta = \frac{1}{4}k^2\left(\mathrm{e}^{2\beta} - \mathrm{e}^{2\alpha}\right) \text{ units}^2.$$

Note that $\mathrm{d}r/\mathrm{d}\theta = k\mathrm{e}^\theta = r$, so

$$\begin{aligned}s &= \int_\alpha^\beta \sqrt{r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta = \int_\alpha^\beta \sqrt{r^2 + r^2}\,\mathrm{d}\theta = \sqrt{2}\int_\alpha^\beta r\,\mathrm{d}\theta \\ &= \sqrt{2}\int_\alpha^\beta k\mathrm{e}^\theta\,\mathrm{d}t = \sqrt{2}k\left[\mathrm{e}^\theta\right]_\alpha^\beta = \sqrt{2}k\left(\mathrm{e}^\beta - \mathrm{e}^\alpha\right) \text{ units.}\end{aligned}$$

Part (b). We have

$$\frac{A}{s^2} = \frac{\frac{1}{4}k^2\left(\mathrm{e}^{2\beta} - \mathrm{e}^{2\alpha}\right)}{2k^2\left(\mathrm{e}^\beta - \mathrm{e}^\alpha\right)^2} = \frac{1}{8}\frac{\left(\mathrm{e}^\beta - \mathrm{e}^\alpha\right)\left(\mathrm{e}^\beta + \mathrm{e}^\alpha\right)}{\left(\mathrm{e}^\beta - \mathrm{e}^\alpha\right)^2} = \frac{1}{8}\left(\frac{\mathrm{e}^\beta + \mathrm{e}^a}{\mathrm{e}^b - \mathrm{e}^a}\right).$$

Assignment B9

Problem 1. The curve C is defined parametrically by $x = a(2\cos\theta + \cos 2\theta)$, $y = a(2\sin\theta + \sin 2\theta)$ where $0 \le \theta \le \pi$ and a is a positive constant.

(a) Find the coordinates of the points at which C meets the x-axis.

(b) Sketch C.

(c) Find the exact total length of C.

(d) Find the exact area of the curve surface generated when C is rotated through 2π radians about the x-axis.

Solution.

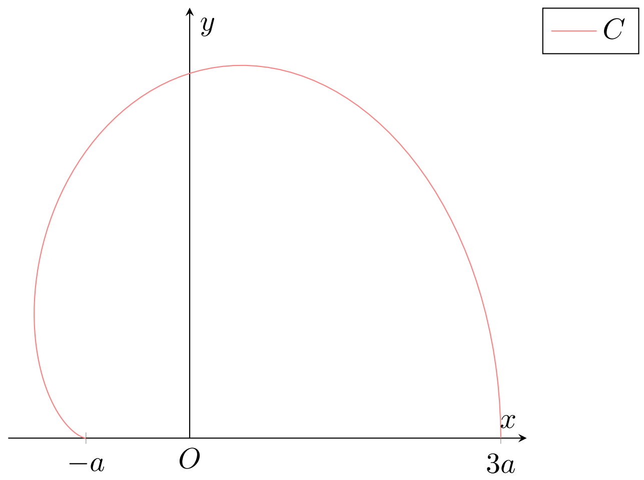
Part (a). When C meets the x-axis, $y = 0$.

$$y = a(2\sin\theta + \sin 2\theta) = a\left(2\sin\theta + 2\sin\theta\cos\right) = 2a\sin\theta\left(1 + \cos\theta\right) = 0.$$

Thus, $\theta = 0$ or $\theta = \pi$.

At $\theta = 0$, $x = 3a$. At $\theta = \pi$, $x = -a$. Hence, C meets the x-axis at $(3a, 0)$ and $(-a, 0)$.

Part (b).



Part (c). Note that

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = -2a(\sin\theta + \sin 2\theta), \quad \frac{\mathrm{d}y}{\mathrm{d}\theta} = 2a(\cos\theta + \cos 2\theta).$$

Hence,

$$\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2 = (2a)^2\left(2 + 2\sin\theta\sin 2\theta + 2\cos\theta\cos 2\theta\right)$$
$$= (2a)^2\left(2 + 2\cos\theta\right) = (2a)^2\left[2 + \left(4\cos^2\frac{\theta}{2} - 2\right)\right] = \left(4a\cos\frac{\theta}{2}\right)^2.$$

Thus,

$$\text{Length} = \int_0^\pi \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta = 4a\int_0^\pi \cos\frac{\theta}{2}\,\mathrm{d}t = 4a\left[2\sin\frac{\theta}{2}\right]_0^\pi = 8a \text{ units}.$$

Part (d).

$$\text{Area} = 2\pi \int_0^\pi y\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta = 16\pi a^2 \int_0^\pi \sin\theta\,(1+\cos\theta)\cos\frac{\theta}{2}\,\mathrm{d}\theta$$

$$= 16\pi a^2 \int_0^\pi \left(2\sin\frac{\theta}{2}\cos\frac{\theta}{2}\right)\left(2\cos^2\frac{\theta}{2}\right)\cos\frac{\theta}{2}\,\mathrm{d}\theta = 64\pi a^2 \int_0^\pi \cos^4\frac{\theta}{2}\sin\frac{\theta}{2}\,\mathrm{d}\theta$$

$$= -128\pi a^2 \left[-\frac{\cos^5(\theta/2)}{5}\right]_0^\pi = \frac{128}{5}\pi a^2 \text{ units}^2.$$

* * * * *

Problem 2. The curve C is given by the equation $y = \frac{1}{2}(\mathrm{e}^x + \mathrm{e}^{-x})$.

(a) Sketch the curve C.

(b) Find the exact area bounded by C, the lines $x = 2$ and $x = -2$ and the x-axis.

(c) Points A and B are on C where $x = 2$ and $x = -2$ respectively. Find the exact length of the arc AB.

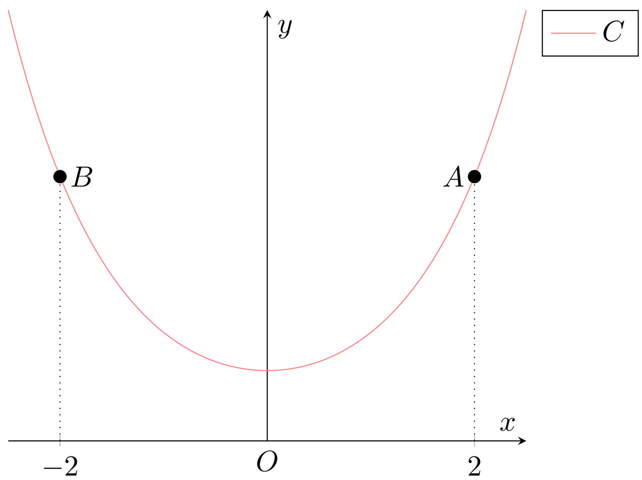
A solid, made of a certain material, is of the shape obtained by rotating the region bounded by C, the lines $x = 2$ and $x = -2$ and the x-axis about the y-axis through π radians.

(d) Find the exact amount of material required to make this solid if x is measured in cm.

(e) The solid is painted with a brush that uses 2 cm^3 of paint for every cm^2 of surface painted. Find the exact amount of paint required.

Solution.

Part (a).



Part (b). Note that $y = \frac{1}{2}(\mathrm{e}^x + \mathrm{e}^{-x}) = \cosh x$ is an even function. Hence,

$$\text{Area} = \int_{-2}^{2} y\,\mathrm{d}x = 2\int_0^2 \cosh x\,\mathrm{d}x = 2\left[\sinh x\right]_0^2$$

$$= 2(\sinh 2 - \sinh 0) = 2\left(\frac{e^2 - e^{-2}}{2} - 0\right) = \mathrm{e}^2 - \mathrm{e}^{-2} \text{ units}^2.$$

Part (c). Note that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\cosh x = \sinh x,$$

whence

$$\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} = \sqrt{1+\sinh^2 x} = \sqrt{\cosh^2 x} = \cosh x$$

Hence,

$$\text{Length} = \int_{-2}^{2}\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = \int_{-2}^{2}\cosh x\,\mathrm{d}x = 2\int_{0}^{2}\cosh x\,\mathrm{d}x = \mathrm{e}^2 - \mathrm{e}^{-2} \text{ units.}$$

Part (d). We have

$$\text{Volume} = 2\pi\int_0^2 xy\,\mathrm{d}x = 2\pi\int_0^2 x\cosh x\,\mathrm{d}x.$$

Integrating by parts,

	D	I
$+$	x	$\cosh x$
$-$	1	$\sinh x$
$+$	0	$\cosh x$

Thus,

$$\begin{aligned}\text{Volume} &= 2\pi\left[x\sinh x - \cosh x\right]_0^2 = 2\pi\left[(2\sinh 2 - \cosh 2) - (0\sinh 0 - \cosh 0)\right]\\ &= 2\pi\left[2\left(\frac{e^2-e^2}{2}\right) - \frac{e^2+e^{-2}}{2} + 1\right] = \pi\left(e^2 - 3e^{-2} + 2\right).\end{aligned}$$

Thus, $\pi\left(e^2 - 3e^{-2} + 2\right)$ cm^3 of material is required.

Part (e).

$$\begin{aligned}&\text{Area} = \text{Area of curved surface} + \text{Area of side} + \text{Area of bottom}\\ &= 2\pi\int_0^2 x\cosh x\,\mathrm{d}x + 2^2\pi + 2^2\pi\cosh 2 = \pi\left(e^2 - 3e^{-2} + 2\right) + 4\pi + 4\pi\left(\frac{e^2+e^{-2}}{2}\right)\\ &= \pi\left[3e^2 - e^{-2} + 6\right].\end{aligned}$$

Thus, $2\pi\left(3e^2 - e^{-2} + 6\right)$ cm^3 of paint is required.

B10 Applications of Integration III - Trapezium and Simpson's Rule

Tutorial B10

Problem 1. Estimate, using the trapezium rule, the values of the following definite integrals, taking the number ordinates given in each case.

(a) $\int_{-\pi/2}^{0} \frac{1}{1+\cos\theta}\,\mathrm{d}\theta$, 3 ordinates

(b) $\int_{-0.4}^{0.2} \frac{x^2-4x+1}{4x-4}$, 4 ordinates

Solution.
Part (a). Let $f(\theta) = \frac{1}{1+\cos\theta}$.

$$\int_{-\pi/2}^{0} \frac{1}{1+\cos\theta}\,\mathrm{d}\theta \approx \frac{1}{2}\cdot\frac{0-(-\pi/2)}{3-1}\cdot\left[f\left(-\frac{\pi}{2}\right)+2f\left(-\frac{\pi}{4}\right)+f(0)\right] = 1.05.$$

Part (b). Let $f(x) = \frac{x^2-4x+1}{4x-4}$.

$$\int_{-0.4}^{0.2} \frac{x^2-4x+1}{4x-4}\,\mathrm{d}x \approx \frac{1}{2}\cdot\frac{0.2-(-0.4)}{4-1}\cdot\left[f(-0.4)+2\left[f(-0.2)+f(0)\right]+f(0.2)\right] = -0.183.$$

* * * * *

Problem 2. Use the trapezium rule with intervals of width 0.5 to obtain an approximation to $\int_{2}^{3.5} \ln\frac{1}{x}\,\mathrm{d}x$, giving your answer to 2 decimal places.

Solution.

$$\int_{2}^{3.5} \ln\frac{1}{x}\,\mathrm{d}x \approx \frac{1}{2}\cdot\frac{3.5-2}{4-1}\cdot\left[\ln\frac{1}{2}+2\left(\ln\frac{1}{2.5}+\ln\frac{1}{3}\right)+\ln\frac{1}{3.5}\right] = -1.49 \text{ (2 d.p.)}.$$

* * * * *

Problem 3. Estimate, using Simpson's rule, the values of the following definite integrals, taking the number of ordinates given in each case.

(a) $\int_{-\pi/2}^{0} \frac{1}{1+\cos\theta}\,\mathrm{d}\theta$, 3 ordinates

(b) $\int_{0}^{0.4} \sqrt{1-x^2}\,\mathrm{d}x$, 5 ordinates

Solution.
Part (a). Let $f(\theta) = \frac{1}{1+\cos\theta}$.

$$\int_{-\pi/2}^{0} \frac{1}{1+\cos\theta}\,\mathrm{d}\theta \approx \frac{1}{3}\cdot\frac{0-(-\pi/2)}{3-1}\cdot\left[f\left(-\frac{\pi}{2}\right)+4f\left(-\frac{\pi}{4}\right)+f(0)\right] = 1.01.$$

Part (b). Let $f(x) = \sqrt{1-x^2}$.

$$\int_0^{0.4} \sqrt{1-x^2}\,\mathrm{d}x \approx \frac{1}{3}\cdot\frac{0.4-0}{5-1}\cdot\Big[f(0)+4f(0.1)+2f(0.2)+4f(0.3)+f(0.4)\Big] = 0.389.$$

* * * * *

Problem 4. Show, by means of substitution $u = \sqrt{x}$, that

$$\int_0^{0.25} \frac{1}{\sqrt{x}}\mathrm{e}^{-x}\,\mathrm{d}x = \int_0^{0.5} 2\mathrm{e}^{-u^2}\,\mathrm{d}u$$

Use the trapezium rule, with ordinates at $u=0$, $u=0.1$, $u=0.2$, $u=0.3$, $u=0.4$ and $u=0.5$, to estimate the value of $I = \int_0^{0.5} 2\mathrm{e}^{-u^2}\,\mathrm{d}u$, giving three decimal places in your answer.

Explain briefly why the trapezium rule cannot be used directly to estimate the value of $\int_0^{0.25} \frac{1}{\sqrt{x}}\mathrm{e}^{-x}\,\mathrm{d}x$.

By using the first four terms of the expansion of e^{-x}, obtain an estimate for the integral $\int_0^{0.25} \frac{1}{\sqrt{x}}\mathrm{e}^{-x}\,\mathrm{d}x$, giving three decimal places in your answer.

Solution. Note that

$$u = \sqrt{x} \implies u^2 = x \implies 2u\,\mathrm{d}u = \mathrm{d}x.$$

Furthermore,

$$x = 0 \implies u = 0, \quad x = 0.25 \implies u = 0.5.$$

Hence,

$$\int_0^{0.25} \frac{1}{\sqrt{x}}\mathrm{e}^{-x}\,\mathrm{d}x = \int_0^{0.5} \frac{1}{u}\mathrm{e}^{-u^2}\cdot 2u\,\mathrm{d}u = \int_0^{0.5} 2\mathrm{e}^{-u^2}\,\mathrm{d}u.$$

Let $f(u) = 2\mathrm{e}^{-u^2}$. Using the trapezium rule,

$$I \approx \frac{1}{2}\cdot\frac{0.5-0}{5}\Big[f(0)+2\big[f(0.1)+f(0.2)+f(0.3)+f(0.4)\big]+f(0.5)\Big] = 0.921 \text{ (3 d.p.)}.$$

At $x = 0$, $\frac{1}{\sqrt{x}}\mathrm{e}^{-x}$ is undefined. Hence, the trapezium rule cannot be used.

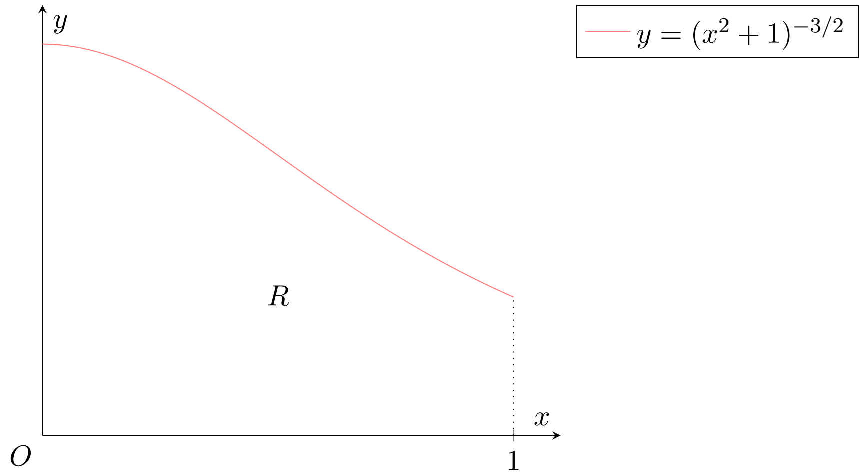
Recall that

$$\mathrm{e}^{-x} = 1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 + \cdots.$$

Hence,

$$\int_0^{0.25} \frac{1}{\sqrt{x}}\mathrm{e}^{-x}\,\mathrm{d}x \approx \int_0^{0.25} \frac{1}{\sqrt{x}}\left(1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3\right)\mathrm{d}x = 0.923 \text{ (3 d.p.)}.$$

Problem 5.



The diagram (not to scale) show the region R bounded by the axes, the curve $y = (x^2+1)^{-3/2}$ and the line $x = 1$. The integral $\int_0^1 (x^2+1)^{-3/2}$ is denoted by I.

(a) Use the trapezium rule and Simpson's rule, with ordinates at $x = 0$, $x = 0.5$ and $x = 1$, to estimate the value of I correct to 4 significant figures.

(b) Use the substitution $x = \tan\theta$ to show that $I = \frac{1}{2}\sqrt{2}$. Comment on the approximations using the 2 rules and give a reason why one gives a better approximation than the other.

(c) By using the trapezium rule, with the same ordinates as in part (a), or otherwise, estimate the volume of the solid formed when R is rotated completely about the x-axis, giving your answer to 2 significant figures.

Solution.

Part (a). Let $f(x) = (x^2+1)^{-3/2}$. Using the trapezium rule,

$$I \approx \frac{1}{2} \cdot \frac{1-0}{3-1} \cdot \Big[f(0) + 2f(0.5) + f(1)\Big] = 0.6962 \text{ (4 s.f.)}.$$

Using Simpson's rule,

$$I \approx \frac{1}{3} \cdot \frac{1-0}{3-1} \cdot \Big[f(0) + 4f(0.5) + f(1)\Big] = 0.7026 \text{ (4 s.f.)}.$$

Part (b). Using the substitution $x = \tan\theta$, we get

$$\int_0^1 (x^2+1)^{-3/2} = \int_0^{\pi/4} \left(\tan^2\theta + 1\right)^{-3/2} \sec^2\theta \, \mathrm{d}\theta = \int_0^{\pi/4} (\sec^2\theta)^{-3/2} \sec^2\theta \, \mathrm{d}\theta$$
$$= \int_0^{\pi/4} (\sec\theta)^{-1} \, \mathrm{d}\theta = \int_0^{\pi/4} \cos\theta \, \mathrm{d}\theta = [\sin\theta]_0^{\pi/4} = \frac{1}{2}\sqrt{2}.$$

The approximation given by Simpson's rule is closer to the actual value than the approximation given by the trapezium rule. This is because Simpson's rule accounts for the concavity of the curve, which produces a better estimate.

Part (c). Let $g(x) = (x^2+1)^{-3}$.

$$\text{Volume} = \pi \int_0^1 y^2 \, \mathrm{d}x = \pi \int_0^1 (x^2+1)^{-3} \, \mathrm{d}x$$
$$\approx \pi \left(\frac{1}{2} \cdot \frac{1-0}{3-1} \Big[g(0) + 2g(0.5) + g(1)\Big]\right) = 1.7 \text{ units}^3 \text{ (2 s.f.)}.$$

Problem 6. It is given that $f(x) = \frac{1}{\sqrt{1+\sqrt{x}}}$, and the integral $\int_0^1 f(x)\,\mathrm{d}x$ is denoted by I.

(a) Using the trapezium rule, with four trapezia of equal width, obtain an approximation I_1 to the value of I, giving 3 decimal places in your answer.

(b) Explain, with the aid of a sketch, why $I < I_1$.

(c) Evaluate I_2, where $I_2 = \frac{1}{3}\sum_{r=1}^{3} f\left(\frac{1}{3}r\right)$, giving 3 decimal places in your answer, and use the sketch in (b) to justify the inequality $I > I_2$.

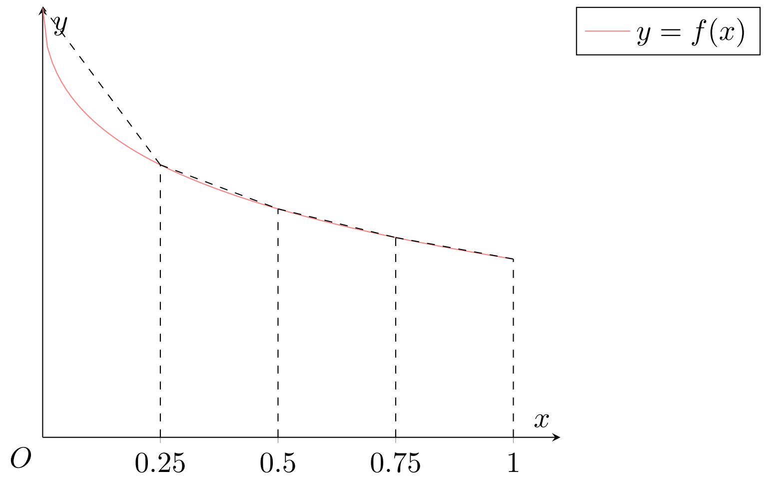
(d) By means of a substitution $\sqrt{x} = u - 1$, show that the value of I is $\frac{4}{3}(2-\sqrt{2})$.

Solution.

Part (a).

$$I_1 = \frac{1}{2}\cdot\frac{1-0}{4}\Big[f(0) + 2\big[f(0.25) + f(0.5) + f(0.75)\big] + f(1)\Big] = 0.792 \text{ (3 d.p.)}.$$

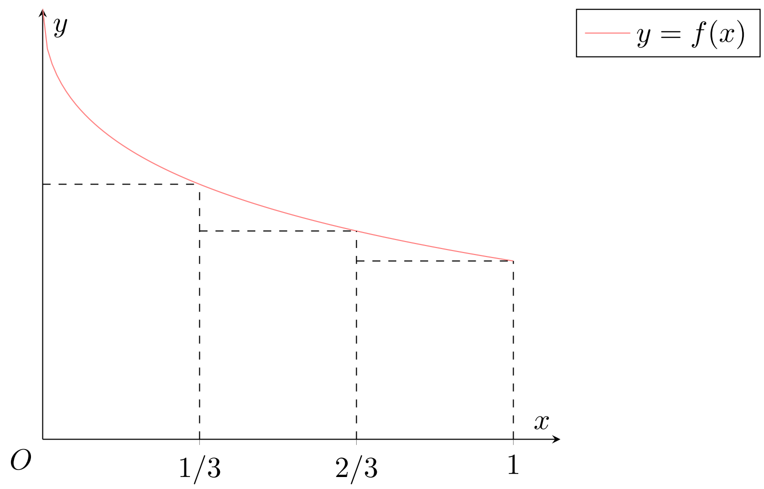
Part (b).



I is the area under the curve $y = f(x)$, while I_1 is the sum of the areas of the trapeziums. Hence, from the sketch, $I_1 > I$.

Part (c).

$$I_2 = \frac{1}{3}\sum_{r=1}^{3} f\left(\frac{1}{3}r\right) = \frac{1}{3}\left[f\left(\frac{1}{3}\right) + f\left(\frac{2}{3}\right) + f(1)\right] = 0.748 \text{ (3 d.p.)}.$$



I is the area under the curve $y = f(x)$, while I_2 is the sum of the areas of the rectangles. Hence, from the sketch, $I_2 < I$.

Part (d). Note

$$\sqrt{x} = u - 1 \implies x = u^2 - 2u + 1 \implies \mathrm{d}x = (2u - 2)\,\mathrm{d}u.$$

Furthermore,

$$x = 0 \implies u = 1, \quad x = 1 \implies u = 2.$$

Thus,

$$\int_0^1 \frac{1}{\sqrt{1+\sqrt{x}}}\,\mathrm{d}x = 2\int_1^2 \frac{u-1}{\sqrt{u}}\,\mathrm{d}u = 2\left[\frac{u^{3/2}}{3/2} - \frac{u^{1/2}}{1/2}\right]_1^2 = \frac{4}{3}(2 - \sqrt{2}).$$

* * * * *

Problem 7. For $0 < x < \pi$, the curve C has the equation $y = \ln\sin x$. The region of the plane bounded by C, the x-axis and the lines $x = \frac{\pi}{4}$ and $x = \frac{\pi}{2}$ is rotated through 2π radians about the x-axis.

Show that the surface area of the solid generated in this way is given by S, where

$$S = 2\pi \int_{\pi/4}^{\pi/2} \left|\frac{\ln\sin x}{\sin x}\right| \mathrm{d}x$$

Use Simpson's rule with 5 ordinates to find an approximate value of S, giving your answer to 3 decimal places.

Solution. Note that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\cos x}{\sin x} = \cot x \implies 1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 = 1 + \cot^2 x = \csc^2 x.$$

Thus,

$$\begin{aligned} S &= 2\pi \int_{\pi/4}^{\pi/2} |y| \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = 2\pi \int_{\pi/4}^{\pi/2} |\ln\sin x|\,|\csc x|\,\mathrm{d}x \\ &= 2\pi \int_{\pi/4}^{\pi/2} |\ln\sin x| \left|\frac{1}{\sin x}\right| \mathrm{d}x = 2\pi \int_{\pi/4}^{\pi/2} \left|\frac{\ln\sin x}{\sin x}\right| \mathrm{d}x. \end{aligned}$$

Let $f(x) = \left|\frac{\ln\sin x}{\sin x}\right|$.

$$S \approx \frac{2\pi}{3} \cdot \frac{\pi/4}{4}\left[f\left(\frac{4\pi}{16}\right) + 4f\left(\frac{5\pi}{16}\right) + 2f\left(\frac{6\pi}{16}\right) + 4f\left(\frac{7\pi}{16}\right) + f\left(\frac{8\pi}{16}\right)\right] = 0.670 \text{ (3 d.p.)}.$$

* * * * *

Problem 8. The value of the integral $\int_{0.2}^{0.4} f(x)\,\mathrm{d}x$ is to be estimated from information in the table below.

x	0.2	0.3	0.4
$f(x)$	1.2030	1.2441	1.2777

(a) Find the best possible estimate for the integral using the trapezium rule.

(b) Using the table of values above, find an approximate value for $f''(0.3)$ and use your answer to explain why the estimate found in part (a) is likely to be smaller than the actual value.

(c) Estimate the integral using Simpson's rule and determine the equation of the curve used in this method.

Solution.

Part (a).

$$\int_{0.2}^{0.4} f(x)\,\mathrm{d}x \approx \frac{1}{2}\cdot\frac{0.4-0.2}{3-1}\Big[f(0.2)+2f(0.3)+f(0.4)\Big] = 0.248.$$

Part (b). Note that $f'(0.25) \approx \frac{f(0.3)-f(0.2)}{0.3-0.2} = 0.411$ and $f'(0.35) \approx \frac{f(0.4)-f(0.3)}{0.4-0.3} = 0.336$. Hence,

$$f''(0.30) \approx \frac{f'(0.35)-f'(0.25)}{0.35-0.25} = -0.75.$$

Since $f''(0.3) < 0$, $f(x)$ is concave downwards around $x = 0.3$. Hence, the estimate is likely to be smaller than the actual value.

Part (c).

$$\int_{0.2}^{0.4} f(x)\,\mathrm{d}x \approx \frac{1}{3}\cdot\frac{0.4-0.2}{3-1}\cdot\Big[f(0.2)+4f(0.3)+f(0.4)\Big] = 0.249.$$

Let the equation of the quadratic used be $P(x) = ax^2+bx+c$, where $a, b, c \in \mathbb{R}$. Since $P(0.2) = f(0.2)$, $P(0.3) = f(0.3)$ and $P(0.4) = f(0.4)$, we obtain the system

$$\begin{cases} (0.2)^2a + 0.2b + c = 1.2030 \\ (0.3)^2a + 0.3b + c = 1.2441 \\ (0.4)^2a + 0.4b + c = 1.2777 \end{cases}$$

which has the unique solution $a = -0.375$, $b = 0.5985$, $c = 1.0983$. Thus, the required equation is

$$y = -0.375x^2 + 0.5985x + 1.0983.$$

* * * * *

Problem 9. The curve C is given by $y = \frac{1}{x}$, where $x > 0$.

(a) Apply the trapezium rule with ordinates at unit intervals to the function $f : x \mapsto \frac{1}{x}$, $x \in \mathbb{R}^+$, to show that $\ln n < \frac{1}{2} + \frac{1}{2n} + \sum_{r=2}^{n-1} \frac{1}{r}$ where $n \geq 3$.

(b) Obtain the area of the trapezium bounded by the axis, the lines $x = r \pm \frac{1}{2}$, and the tangent to the curve $y = \frac{1}{x}$ at the point $\left(r, \frac{1}{r}\right)$.

Hence, show that $\sum_{r=2}^{n-1} \frac{1}{r} < \ln\left(\frac{2n-1}{3}\right)$, where $n \geq 3$.

(c) From these results, obtain numerical values between which the value of $\sum_{r=2}^{99} \frac{1}{r}$ lies, and show that $4.110 < \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{100} < 4.205$.

Solution.

Part (a). Applying the trapezium rule,

$$\int_1^n \frac{1}{x}\,\mathrm{d}x \approx \frac{1}{2}\left[\frac{1}{1} + 2\left(\frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1}\right) + \frac{1}{n}\right] = \frac{1}{2} + \frac{1}{2n} + \sum_{r=2}^{n-1} \frac{1}{r}.$$

Note that $\mathrm{d}^2y/\mathrm{d}x^2 = 2x^{-3} > 0$ for $x > 0$. Hence, $y = 1/x$ is concave upwards. Thus,

$$\frac{1}{2} + \frac{1}{2n} + \sum_{r=2}^{n-1} \frac{1}{r} > \int_1^n \frac{1}{x}\,\mathrm{d}x = \ln n.$$

Part (b). Since $\mathrm{d}y/\mathrm{d}x = -x^{-2}$, the equation of the tangent at $x = r$ is given by

$$y - \frac{1}{r} = -\frac{1}{r^2}(x - r) \implies y = -\frac{x}{r^2} + \frac{2}{r}.$$

The area of the trapezium centred at r is hence given by

$$\int_{r-1/2}^{r+1/2} \left(-\frac{x}{r^2} + \frac{2}{r}\right) \mathrm{d}x = \left[-\frac{1}{r^2}\left(\frac{x^2}{2}\right) + \frac{2x}{r}\right]_{r-1/2}^{r+1/2} = \frac{1}{r} \text{ units}^2.$$

Observe that the area of the trapezium centred at r is less than the area under the curve $y = \frac{1}{x}$ from $r - \frac{1}{2}$ to $r + \frac{1}{2}$. That is,

$$\frac{1}{r} < \int_{r-1/2}^{r+1/2} \frac{1}{x}\,\mathrm{d}x = \ln\left(r + \frac{1}{2}\right) - \ln\left(r - \frac{1}{2}\right).$$

Summing from $r = 2$ to $n - 1$,

$$\sum_{r=2}^{n-1} \frac{1}{r} < \sum_{r=2}^{n-1} \left[\ln\left(r + \frac{1}{2}\right) - \ln\left(r - \frac{1}{2}\right)\right] = \ln\left(n - \frac{1}{2}\right) - \ln\left(2 - \frac{1}{2}\right) = \ln\left(\frac{2n-1}{3}\right).$$

Part (c). Taking $n = 100$, we have

$$\frac{1}{2} + \frac{1}{2(100)} + \sum_{r=2}^{100-1} \frac{1}{r} > \ln 100 \implies \sum_{r=2}^{99} \frac{1}{r} > \ln 100 - \frac{1}{2} - \frac{1}{200} = 4.100$$

We also have

$$\sum_{r=2}^{100-1} \frac{1}{r} < \ln\left(\frac{2(100) - 1}{3}\right) \implies \sum_{r=2}^{100-1} \frac{1}{r} < \ln\left(\frac{199}{3}\right) = 4.195$$

Putting both inequalities together, we obtain

$$4.100 < \sum_{r=2}^{99} \frac{1}{r} < 4.195.$$

Adding $\frac{1}{100} = 0.01$ to all sides of the inequality, we see that

$$4.110 < \sum_{r=2}^{100} \frac{1}{r} < 4.205.$$

Self-Practice B10

Problem 1. Use the trapezium rule, with 6 intervals to estimate the value of $\int_0^3 \ln(1+x)\,\mathrm{d}x$, showing your working. Give your answer to 3 significant figures. Hence, write down an approximate value for $\int_0^3 \ln\sqrt{1+x}\,\mathrm{d}x$.

Solution. Let $f(x) = \ln(1 = x)$. Then

$$\begin{aligned}\int_0^3 \ln(1+x)\ \mathrm{d}x &\approx \frac{0.5}{2}\left[f(0) + 2\left(f(0.5) + f(1) + f(1.5) + f(2) + f(2.5)\right) + f(3)\right] \\ &= 2.5297 = 2.53\ (3\text{ s.f.}).\end{aligned}$$

Thus,

$$\int_0^3 \ln\sqrt{1+x}\,\mathrm{d}x = \frac{1}{2}\int_0^3 \ln(1+x)\ \mathrm{d}x \approx \frac{2.5297}{2} = 1.26\ (3\text{ s.f.}).$$

* * * * *

Problem 2. Use the trapezium rule with 5 intervals to estimate the value of

$$\int_0^{0.5} \sqrt{1+x^2}\,\mathrm{d}x,$$

showing your working. Give your answer to 2 decimal places.

By expanding $(1+x^2)^{1/2}$ in powers of x as far as the term in x^4, obtain a second estimate for the value of $\int_0^{0.5} \sqrt{1+x^2}\,\mathrm{d}x$ giving this answer also correct to 2 decimal places.

Solution. Let $f(x) = \sqrt{1+x^2}$. Then

$$\int_0^{0.5} \sqrt{1+x^2}\,\mathrm{d}x \approx \frac{0.1}{2}\sum_{i=0}^{4}\left[f(0.1i) + f(0.1(i+1))\right] = 0.52\ (2\text{ d.p.}).$$

We have

$$\left(1+x^2\right)^{1/2} = 1 + \frac{1}{2}x^2 + \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)}{2}\left(x^2\right)^2 + \cdots \approx 1 + \frac{1}{2}x^2 - \frac{1}{8}x^4,$$

hence

$$\int_0^{0.5} \sqrt{1+x^2}\,\mathrm{d}x \approx \int_0^{0.5}\left(1 + \frac{1}{2}x^2 - \frac{1}{8}x^4\right)\,\mathrm{d}x = 0.52\ (2\text{ d.p.}).$$

* * * * *

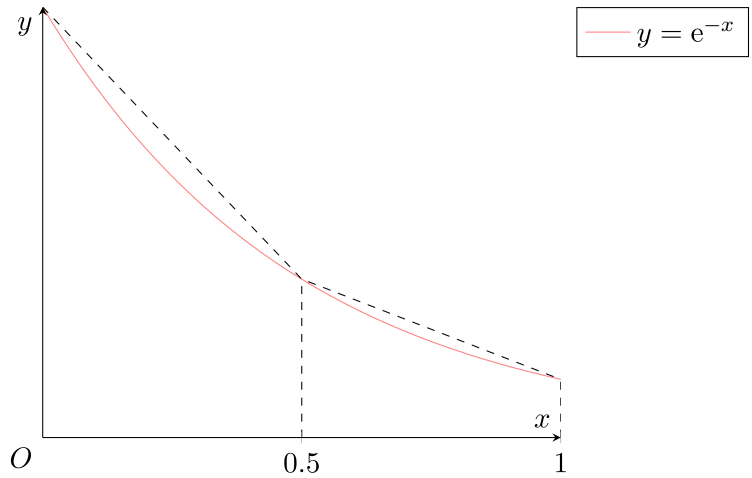
Problem 3. The trapezium rule, with 2 intervals of equal width, is to be used to find an approximate value for $\int_0^1 \mathrm{e}^{-x}\,\mathrm{d}x$. Explain, with the aid of a sketch, why the approximation will be greater than the exact value of the integral. Calculate the approximate value and the exact value, giving each answer correct to 3 decimal places.

Another approximation to $\int_0^1 \mathrm{e}^{-x}\,\mathrm{d}x$ is to be calculated by using two trapezia of unequal width. The first trapezium has width h and the second has width $1-h$, so that the three ordinates are at $x = 0$, $x = h$ and $x = 1$. Show that the total area T of these two trapezia is given by

$$T = \frac{1}{2}\left[\mathrm{e}^{-1} + h\left(1 - \mathrm{e}^{-1}\right) + \mathrm{e}^{-h}\right].$$

Show that the value of h for which T is a minimum is given by $h = \ln\frac{\mathrm{e}}{\mathrm{e}-1}$.

Solution. Let $f(x) = \mathrm{e}^{-x}$. Note that $\mathrm{d}^2 f/\mathrm{d}x^2 = \mathrm{e}^{-x}$, which is positive for $0 \leq x \leq 1$. Thus, the graph of $y = f(x)$ is convex, as shown below.



The secant lines lie above the curve $y = f(x)$, thus, the area of the trapeziums is larger than the area under the curve. Thus, the trapezium rule gives an overestimate.

The exact value of the integral is

$$\int_0^1 \mathrm{e}^{-x}\,\mathrm{d}x = \left[-\mathrm{e}^{-x}\right]_0^1 = 1 - \mathrm{e}^{-1} = 0.632 \text{ (3 d.p.)}.$$

Using the trapezium rule, we have

$$\int_0^1 \mathrm{e}^{-x}\,\mathrm{d}x \approx \frac{0.5}{2}\left[f(0) + 2f(0.5) + f(1)\right] = 0.645 \text{ (3 d.p.)}.$$

The total area is given by

$$\begin{aligned} T &= \frac{h}{2}\left(f(0) + f(h)\right) + \frac{1-h}{2}\left(f(h) + f(1)\right) \\ &= \frac{h}{2}\left(1 + \mathrm{e}^{-h}\right) + \frac{1-h}{2}\left(\mathrm{e}^{-h} + \mathrm{e}^{-1}\right) \\ &= \frac{1}{2}\left[\mathrm{e}^{-1} + h\left(1 - \mathrm{e}^{-1}\right) + \mathrm{e}^{-h}\right]. \end{aligned}$$

Differentiating with respect to h, we obtain

$$\frac{\mathrm{d}T}{\mathrm{d}h} = \frac{1}{2}\left(1 - \mathrm{e}^{-1} - \mathrm{e}^{-h}\right).$$

For stationary points, $\mathrm{d}T/\mathrm{d}h = 0$, so

$$\mathrm{e}^{-h} = 1 - \mathrm{e}^{-1} \implies -h = \ln\left(1 - \mathrm{e}^{-1}\right) \implies h = \ln\frac{1}{1 - \mathrm{e}^{-1}} = \ln\frac{\mathrm{e}}{\mathrm{e} - 1}.$$

Note further that

$$\frac{\mathrm{d}^2 T}{\mathrm{d}h^2} = \frac{1}{2}\mathrm{e}^{-h},$$

which is positive at $h = \ln(e/(e-1))$, thus it is a minimum.

* * * * *

Problem 4. Derive Simpson's rule with 2 strips for evaluating $\int_a^b f(x)\,\mathrm{d}x$.

Use Simpson's composite rule with 4 strips to obtain an estimate of $\int_2^3 \cos(x-2)\ln x\,\mathrm{d}x$, giving your answer to 5 decimal places.

Solution. Let $c = (b-a)/2$. Define

$$g(x) = f\left(x + \frac{a+b}{2}\right).$$

Then

$$g(-c) = f(a), \quad g(0) = f\left(\frac{a+b}{2}\right), \quad g(c) = f(b).$$

Let $h(x) = \alpha x^2 + \beta x + \gamma$ be the quadratic such that $h(x) = g(x)$ at $x = -c, 0, c$. Then

$$\begin{aligned} g(0) &= \gamma, \\ g(-c) &= \alpha c^2 - \beta c + \gamma, \\ g(c) &= \alpha c^2 + \beta c + \gamma. \end{aligned}$$

From the last two equations, it follows that

$$\alpha c^2 = \frac{g(c) + g(-c)}{2} - g(0).$$

Our estimate is thus

$$\begin{aligned} \int_a^b f(x)\,\mathrm{d}x &= \int_{-c}^{c} g(x)\,\mathrm{d}x \\ &\approx \int_{-c}^{c} h(x)\,\mathrm{d}x = \int_{-c}^{c} \left(\alpha x^2 + \beta x + \gamma\right)\,\mathrm{d}x \\ &= \left[\frac{\alpha}{3}x^3 + \frac{\beta}{2}x^2 + \gamma c\right]_{-c}^{c} \\ &= \frac{2c}{3}\left(\alpha c^2 + 3\gamma\right) \\ &= \frac{2\left(\frac{b-a}{2}\right)}{3}\left[\frac{g(c) + g(-c)}{2} - g(0) + 3g(0)\right] \\ &= \frac{b-a}{6}\left[g(-c) + 4g(0) + g(c)\right] \\ &= \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right], \end{aligned}$$

which is precisely Simpson's rule for two strips.

Let $f(x) = \cos(x-2)\ln x$. Using Simpson's rule, we get

$$\int_2^3 \cos(x-2)\ln x\,\mathrm{d}x = \frac{0.5}{6}\left[f(2) + 4f(2.25) + 2f(2.5) + 4f(2.75) + f(3)\right] = 0.74988 \text{ (5 d.p.)}.$$

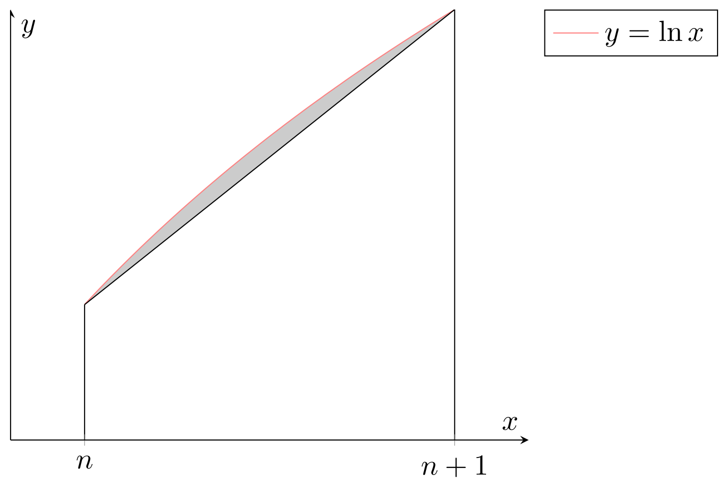
* * * * *

Problem 5.

(a) Show that $\int_n^{n+1} \ln x\,\mathrm{d}x = (n+1)\ln(n+1) - n\ln n - 1$.

(b) The diagram below shows the graph of $y = \ln x$ between $x = n$ and $x = n+1$. The area of the shaded region represents the error when the value of the integral in part (a) is approximated by using a single trapezium. Show that the area of the shaded region is

$$\left(n + \frac{1}{2}\right)\ln\left(1 + \frac{1}{n}\right) - 1.$$



(c) Use a series expansion to show that if n is large enough for $\frac{1}{n^3}$ and higher powers of $\frac{1}{n}$ to be neglected, then the area in part (b) is approximately equal to $\frac{k}{n^2}$, where k is a constant to be determined.

Solution.

Part (a). We have

$$\begin{aligned}\int_n^{n+1} \ln x \,\mathrm{d}x &= [x(\ln x - 1)]_n^{n+1} \\ &= (n+1)(\ln(n+1) - 1) - n(\ln n - 1) \\ &= (n+1)\ln(n+1) - n\ln n - 1.\end{aligned}$$

Part (b). The estimate of the area under the curve is

$$\frac{\ln n + \ln(n+1)}{2}.$$

The error is thus

$$\begin{aligned}&(n+1)\ln(n+1) - n\ln n - 1 - \frac{\ln n + \ln(n+1)}{2} \\ &= \left(n + \frac{1}{2}\right)\ln(n+1) - \left(n + \frac{1}{2}\right)\ln n - 1 \\ &= \left(n + \frac{1}{2}\right)\ln\left(\frac{n+1}{n}\right) - 1 = \left(n + \frac{1}{2}\right)\ln\left(1 + \frac{1}{n}\right) - 1.\end{aligned}$$

Part (c). Note that

$$\ln\left(1 + \frac{1}{n}\right) \approx \frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3}.$$

Thus,

$$\left(n + \frac{1}{2}\right)\ln\left(1 + \frac{1}{n}\right) - 1 \approx \left(n + \frac{1}{2}\right)\left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3}\right) - 1 = \frac{1}{12n^2},$$

so $k = 1/12$.

Assignment B10

Problem 1. Given that $y = e^{-x}\cos x$, show that $\frac{d^2y}{dx^2} = -2\left(y + \frac{dy}{dx}\right)$. By further differentiation, find the series expansion of y, in ascending powers of x, up to and including the term in x^3. Use the series to obtain an approximate value for $\int_0^{0.2} \frac{\cos x^2}{e^{x^2}}\,dx$, giving your answer correct to 4 decimal places.

Using the trapezium rule with 4 trapezia of equal width, find another approximation for $\int_0^{0.2} \frac{\cos x^2}{e^{x^2}}\,dx$, giving your answer correct to 4 decimal places.

Solution. Differentiating with respect to x, we get

$$y' = -e^{-x}\sin x - e^{-x}\cos x \implies y' = -e^{-x}\sin x - y.$$

Differentiating once more,

$$y'' = -e^{-x}\cos x + e^{-x}\sin x - y' = -y + \left(-y' - y\right) - y' = -2\left(y + y'\right).$$

Further differentiating, we obtain $y''' = -2(y' + y'')$. Evaluating y and its derivatives at $x = 0$, we get

$$y(0) = 1, \quad y'(0) = -1, \quad y''(0) = 0, \quad y'''(0) = 2.$$

Thus,

$$e^{-x}\cos x = 1 - x + \frac{x^3}{3} + \cdots.$$

Hence,

$$\int_0^{0.2} \frac{\cos x^2}{e^{x^2}}\,dx = \int_0^{0.2} e^{-x^2}\cos x^2\,dx \approx \int_0^{0.2} \left(1 - \left(x^2\right) + \frac{\left(x^2\right)^3}{3}\right)dx = 0.1973 \text{ (4 d.p.)}.$$

Let $g(x) = \frac{\cos x^2}{e^{x^2}}$. By the trapezium rule, we have

$$\int_0^{0.2} \frac{\cos x^2}{e^{x^2}}\,dx \approx \frac{1}{2}\cdot\frac{0.2}{4}\Big[g(0) + 2\big[g(0.05) + g(0.1) + g(0.15)\big] + g(0.2)\Big] = 0.1973 \text{ (4 d.p.)}.$$

* * * * *

Problem 2. The curve C has equation $y^2 = \frac{x}{\sqrt{1+x^2}}$, $y \geq 0$.

The finite region R is bounded by C, the x-axis and the lines $x = 0$ and $x = 2$. R is rotated through 2π radians about the x-axis.

(a) Find the exact volume of the solid formed.

An estimate for the volume in (a) is found using the trapezium rule with 7 ordinates.

(b) Find the percentage error resulting from using this estimate, giving your answer to 3 decimal places.

Explain, with the help of a sketch, why the estimate given by the trapezium rule is less than the actual value.

Solution.

Part (a).

$$\text{Volume} = \pi\int_0^2 y^2\,\mathrm{d}x = \pi\int_0^2 \frac{2x}{2\sqrt{1+x^2}}\,\mathrm{d}x = \pi\left[\sqrt{1+x^2}\right]_0^2 = \pi(\sqrt{5}-1)\text{ units}^3.$$

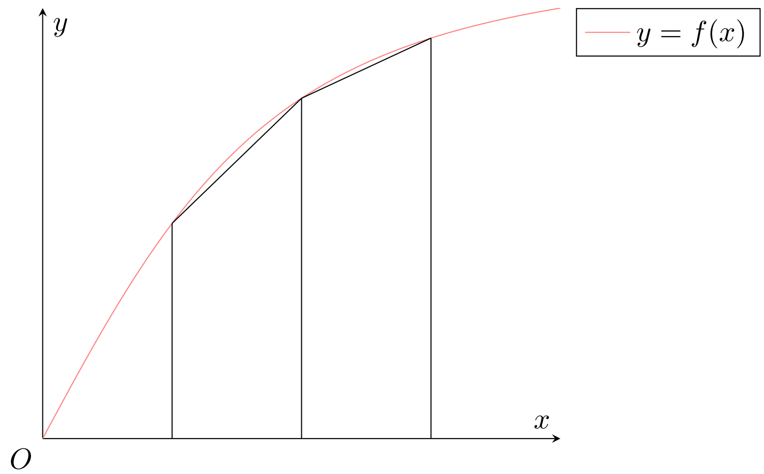
Part (b). Let $f(x) = \frac{x}{\sqrt{1+x^2}}$. By the trapezium rule,

$$\text{Volume} = \pi\int_0^2 f(x)\,\mathrm{d}x \approx \pi\cdot\frac{1}{2}\cdot\frac{2-0}{6}\sum_{n=0}^{5}\left[f\left(\frac{n}{3}\right)+f\left(\frac{n+1}{3}\right)\right] = 3.8566\text{ (5 s.f.)}.$$

Hence, the percentage error is

$$\left|\frac{\pi(\sqrt{5}-1)-3.8566}{\pi(\sqrt{5}-1)}\right| = 0.686\%\text{ (3 d.p.)}.$$

Consider the following graph of $y = f(x)$.



From the graph, the curve $y = f(x)$ is clearly concave downwards. Hence, the approximation given by the trapezium rule is an underestimate and is thus less than the actual value.

* * * * *

Problem 3. Prove that $\int_{-h}^{h} f(x)\,\mathrm{d}x = \frac{1}{3}h(y_{-1}+4y_0+y_1)$, where $y = f(x)$ is the quadratic curve passing through the points $(-h, y_{-1})$, $(0, y_0)$ and (h, y_1).

Use Simpson's rule with 5 ordinates to find an approximation to

$$\int_{-3}^{1}\left(x^4-7x^3+3x^2+6x+4\right)^{1/3}\mathrm{d}x$$

Find another approximation to the same integral using the trapezium rule with 5 ordinates.

Which of these approximations would you expect to be more accurate? Justify your answer.

Solution. Let $f(x) = ax^2+bx+c$ be the quadratic such that the graph $y = f(x)$ passes through the points $(-h, y_{-1})$, $(0, y_0)$ and (h, y_1).

Note that we have $y_0 = f(0) = c$. We also have

$$y_{-1}+y_1 = f(-h)+f(h) = \left[a(-h)^2+b(-h)+c\right]+\left[ah^2+bh+c\right] = 2ah^2+2c.$$

Hence,

$$\int_{-h}^{h} f(x)\,\mathrm{d}x = \int_{-h}^{h} (ax^2+bx+c)\,\mathrm{d}x = \left[\frac{1}{3}x^3 + \frac{1}{2}bx^2 + cx\right]_{-h}^{h} = \frac{1}{3}h\left(2h^2+6c\right)$$
$$= \frac{1}{3}h\left(2h^2+2c+4c\right) = \frac{1}{3}h\left(y_{-1}+y_1+4y_0\right) = \frac{1}{3}h\left(y_{-1}+4y_0+y_1\right).$$

Let $f(x) = \left(x^4 - 7x^3 + 3x^2 + 6x + 4\right)^{1/3}$. By Simpson's rule,

$$\int_{-3}^{1} \left(x^4 - 7x^3 + 3x^2 + 6x + 4\right)^{1/3}\,\mathrm{d}x$$
$$\approx \frac{1}{3}\cdot\frac{1-(-3)}{4}\Big[f(-3)+4f(-2)+2f(-1)+4f(0)+f(1)\Big] = 11.977 \text{ (5 s.f.)}$$

By the trapezium rule,

$$\int_{-3}^{1} \left(x^4 - 7x^3 + 3x^2 + 6x + 4\right)^{1/3}\,\mathrm{d}x$$
$$\approx \frac{1}{2}\cdot\frac{1-(-3)}{4}\Big[f(-3)+2f(-2)+2f(-1)+2f(0)+f(1)\Big] = 12.142 \text{ (5 s.f.)}$$

The approximation given by Simpson's rule should be more accurate as Simpson's rule accounts for the concavity of the curve $y = f(x)$.

* * * * *

Problem 4.

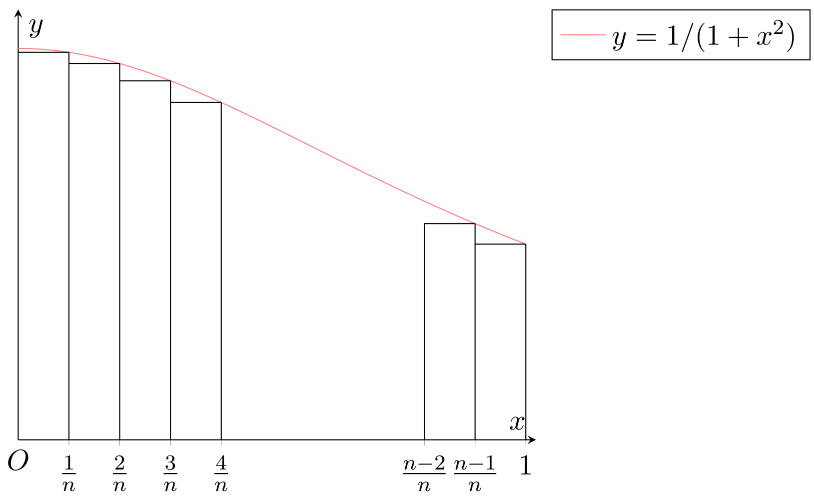
(a) Find the exact value of $\int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x$.

(b) The graph of $y = \frac{1}{1+x^2}$ is shown in the diagram below. Rectangles, each of width $\frac{1}{n}$, are drawn under the curve.

Show that the total area A of all n rectangles is given by

$$A = \frac{1}{n}\left[\frac{1}{1+\left(\frac{1}{n}\right)^2} + \frac{1}{1+\left(\frac{2}{n}\right)^2} + \frac{1}{1+\left(\frac{3}{n}\right)^2} + \ldots + \frac{1}{2}\right]$$

State the limit of A as $n \to \infty$.



(c) It is given that

$$B = \frac{1}{n}\left[\frac{1}{1+\left(\frac{1}{n}\right)^4} + \frac{1}{1+\left(\frac{2}{n}\right)^4} + \frac{1}{1+\left(\frac{3}{n}\right)^4} + \ldots + \frac{1}{2}\right]$$

Find an approximation for the limit of B as $n \to \infty$ by considering an appropriate graph and using the trapezium rule with 5 intervals. Given your answer correct to 2 decimal places.

Solution.

Part (a).

$$\int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x = [\arctan x]_0^1 = \frac{\pi}{4}.$$

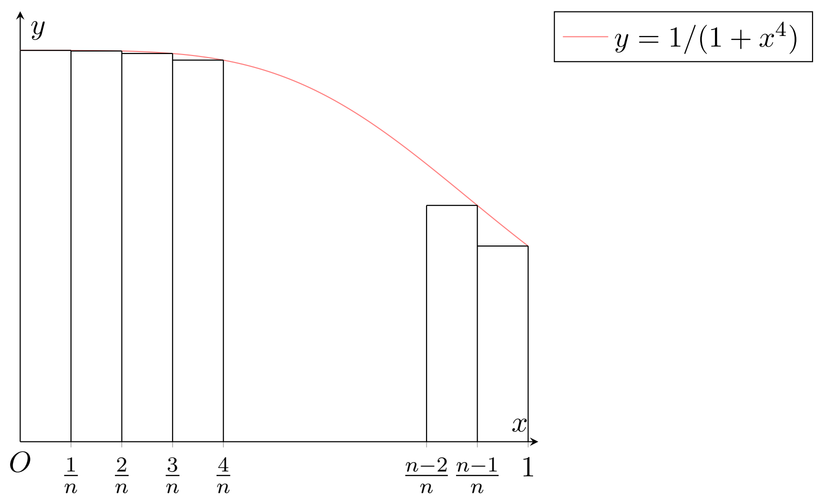
Part (b). Observe that the kth rectangle has height $\frac{1}{1+(k/n)^2}$ and width $1/n$. Hence,

$$\begin{aligned} A &= \sum_{k=1}^{n} \frac{1}{n} \cdot \frac{1}{1+(k/n)^2} = \frac{1}{n}\sum_{k=1}^{n} \frac{1}{1+(k/n)^2} \\ &= \frac{1}{n}\left[\frac{1}{1+\left(\frac{1}{n}\right)^2} + \frac{1}{1+\left(\frac{2}{n}\right)^2} + \frac{1}{1+\left(\frac{3}{n}\right)^2} + \ldots + \frac{1}{1+\left(\frac{n}{n}\right)^2}\right] \\ &= \frac{1}{n}\left[\frac{1}{1+\left(\frac{1}{n}\right)^2} + \frac{1}{1+\left(\frac{2}{n}\right)^2} + \frac{1}{1+\left(\frac{3}{n}\right)^2} + \ldots + \frac{1}{2}\right] \end{aligned}$$

Thus,

$$\lim_{n\to\infty} A = \int_0^1 \frac{1}{1+x^2} = \frac{\pi}{4}.$$

Part (c). Consider the following graph of $y = \frac{1}{1+x^4}$.



Using a similar line of logic presented in part (b), we have that B is the total area of the rectangles above. Hence,

$$\lim_{n\to\infty} B = \int_0^1 \frac{1}{1+x^4}\,\mathrm{d}x.$$

Let $f(x) = \frac{1}{1+x^4}$. Using the trapezium rule,

$$\lim_{n\to\infty} B \approx \frac{1}{2}\cdot\frac{1}{5}\Big[f(0) + 2\big[f(0.2) + f(0.4) + f(0.6) + f(0.8)\big] + f(1)\Big] = 0.86 \text{ (2 d.p.)}.$$

B11 Functions of Two Variables

Tutorial B11

Problem 1. Find the natural domain of the function f for the following:

(a) $f(x, y) = \sqrt{1 - x^2 - y^2}$

(b) $f(x, y) = \ln(x^2 - y)$

(c) $f(x, y) = \arcsin(x + y)$

(d) $f(x, y) = \frac{1}{x^2 - y^2}$

Solution.

Part (a). Observe that the argument of the square root must be non-negative. Hence, $1 - x^2 - y^2 \geq 0 \implies x^2 + y^2 \leq 1$. Thus,

$$D_f = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}.$$

Part (b). Observe that the argument of the natural log must be positive. Hence, $x^2 - y > 0 \implies y < x^2$. Thus,

$$D_f = \{(x, y) \in \mathbb{R}^2 : y < x^2\}.$$

Part (c). Observe that the argument of arcsin must be within the range of sin, i.e. between -1 and 1 inclusive. Hence, $-1 \leq x + y \leq 1$. Thus,

$$D_f = \{(x, y) \in \mathbb{R}^2 : -1 \leq x + y \leq 1\}.$$

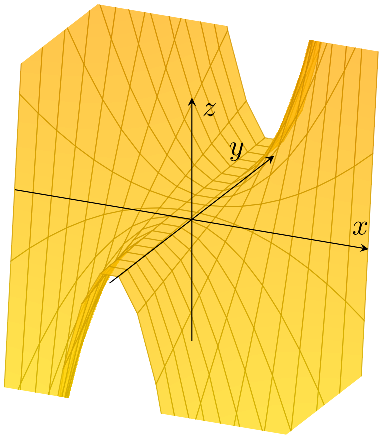
Part (d). Observe that the denominator must be non-zero. Hence, $x^2 - y^2 \neq 0 \implies y^2 \neq x^2 \implies y \neq x$ or $y \neq -x$. Thus,

$$D_f = \{(x, y) \in \mathbb{R}^2 : y \neq x \text{ or } y \neq -x\}.$$

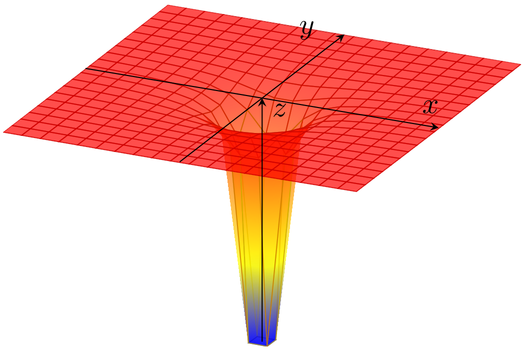
* * * * *

Problem 2. Identify the correct equations of the following surfaces in 3-D space.

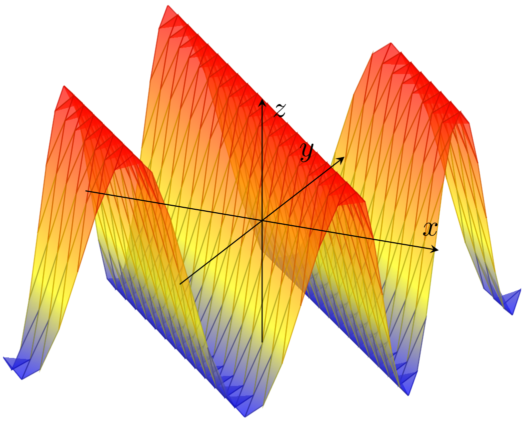
- $z = \cos(x + y)$
- $z = x^2 y + 1$
- $z = 3 - x + y$
- $z = -\frac{1}{\sqrt{x^2 + y^2}}$



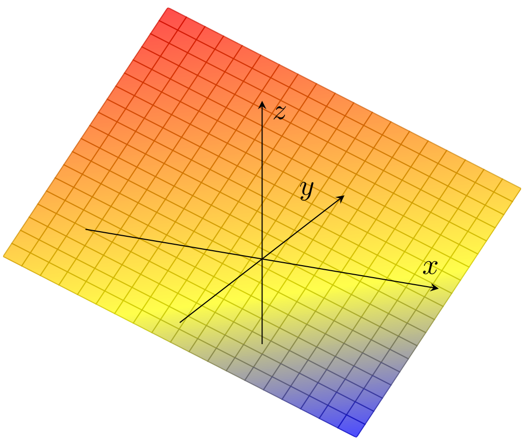
(a)



(b)



(c)



(d)

Solution.

Part (a). $z = x^2y + 1$

Part (b). $z = -\frac{1}{\sqrt{x^2+y^2}}$

Part (c). $z = \cos(x + y)$
Part (d). $z = 3 - x + y$

* * * * *

Problem 3. Let $f(x, y) = x^2 - 2x^3 + 3xy$. Find an equation of the level curve that passes through the point

(a) $(-1, 1)$

(b) $(2, -1)$

(c) $(1, 5)$

Solution.
Part (a). Note that $f(-1, 1) = 0$. Hence, the level curve is given by

$$x^2 - 2x^3 + 3xy = 0.$$

Part (b). Note that $f(2, -1) = -18$. Hence, the level curve is given by

$$x^2 - 2x^3 + 3xy = -18.$$

Part (c). Note that $f(1, 5) = 14$. Hence, the level curve is given by

$$x^2 - 2x^3 + 3xy = 14.$$

* * * * *

Problem 4. If $V(x, y)$ is the voltage or potential at a point (x, y) in the xy-plane, then the level curves of V are called equipotential curves. Along such a curve, the voltage remains constant. Given that

$$V(x, y) = \frac{8}{\sqrt{16 + x^2 + y^2}}$$

find an equation of the equipotential curves at which

(a) $V = 2.0$

(b) $V = 1.0$

(c) $V = 0.5$

Solution. Rearranging the given equation, we have

$$x^2 + y^2 = \frac{64}{V^2} - 16.$$

Part (a). When $V = 2.0$, we have $x^2 + y^2 = \frac{64}{2.0^2} - 16 = 0$, whence

$$x = 0 \text{ and } y = 0.$$

Part (b). When $V = 1.0$, we have

$$x^2 + y^2 = \frac{64}{1.0^2} - 16 = 48.$$

Part (c). When $V = 0.5$, we have

$$x^2 + y^2 = \frac{64}{0.5^2} - 16 = 240.$$

* * * * *

Problem 5. Given that $f(x,y) = x^4 \sin(xy^3)$, find $f_x(x,y)$, $f_y(x,y)$, $f_{xy}(x,y)$ and $f_{yx}(x,y)$.

Solution. Differentiating f with respect to x,

$$f_x(x,y) = 4x^3 \sin(xy^3) + x^4 y^3 \cos(xy^3) \, .$$

Differentiating f with respect to y,

$$f_y(x,y) = 3x^5 y^2 \cos(xy^3) \, .$$

Differentiating f_x with respect to y,

$$\begin{aligned} f_{xy}(x,y) &= 12x^4 y^2 \cos(xy^3) + x^4 \left[3y^2 \cos(xy^3) - 3xy^5 \sin(xy^3)\right] \\ &= 15x^4 y^2 \cos(xy^3) - 3x^5 y^5 \sin(xy^3) \, . \end{aligned}$$

Differentiating f_y with respect to x,

$$\begin{aligned} f_{yx}(x,y) &= 3y^2 \left[5x^4 \cos(xy^3) - x^5 y^3 \sin(xy^3)\right] \\ &= 15x^4 y^2 \cos(xy^3) - 3x^5 y^5 \sin(xy^3) \, . \end{aligned}$$

* * * * *

Problem 6. Given that $z = x^2 y$, $x = t^2$, $y = t^3$, use the chain rule to find $\frac{\mathrm{d}z}{\mathrm{d}t}$ in terms of t.

Solution.

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial z}{\partial x} \cdot \frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial z}{\partial y} \cdot \frac{\mathrm{d}y}{\mathrm{d}t} = 2xy\,(2t) + x^2\left(3t^2\right) = 2t^2 t^3 \cdot 2t + \left(t^2\right)^2 \cdot 3t^2 = 7t^6.$$

* * * * *

Problem 7. Find the gradient of $f(x,y) = 3x^2 y$ at the point $(1,2)$ and use it to calculate the directional derivative of f at $(1,2)$ in the direction of the vector $\mathbf{u} = 3\mathbf{i} + 4\mathbf{j}$.

Solution. Note that $f_x(x,y) = 6xy$ and $f_y(x,y) = 3x^2$. Hence, ∇f at $(1,2)$ is $(12,\, 3)^\mathsf{T}$.

Observe that the directional derivative of f in the direction of $\mathbf{u}$ at $(1,2)$ is given by

$$\nabla f \cdot \hat{\mathbf{u}} = \begin{pmatrix} 12 \\ 3 \end{pmatrix} \cdot \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \frac{48}{5}.$$

Thus, the instantaneous rate of change at $(1,2)$ in the direction of $\mathbf{u}$ is $48/5$.

* * * * *

Problem 8. Suppose that a point moves along the intersection of the sphere $x^2+y^2+z^2 = 1$ with the plane $x = \frac{2}{3}$. Find the rate of z with respect to y when the point is at $\left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)$.

Solution. Note that $x^2+y^2+z^2=1 \implies z=\pm\sqrt{1-x^2-y^2}$. Given that the object we want (the rate of change of z with respect to y) will later be evaluated when $z=\frac{2}{3}>0$, we consider only the positive branch. Let $f(x,y)=\sqrt{1-x^2-y^2}$. Then $f_y(x,y)=\frac{-y}{\sqrt{1-x^2-y^2}}$. Evaluating at the desired point, we get,

$$f_y\left(\frac{2}{3},\frac{1}{3}\right)=\frac{-1/3}{\sqrt{1-(2/3)^2-(1/3)^2}}=-\frac{1}{2}.$$

* * * * *

Problem 9.

(a) The Cauchy-Riemann equations are such that $\partial u/\partial x=\partial v/\partial y$ and $\partial u/\partial y=-\partial v/\partial x$ for $u(x,y)$ and $v(x,y)$. Show that $u=\mathrm{e}^x\cos y$, $v=\mathrm{e}^x\sin y$ satisfy the Cauchy-Riemann equations.

(b) Show that the function $f(x,y)=\mathrm{e}^x\sin y+\mathrm{e}^y\cos x$ satisfies that Laplace equation, i.e. $\partial^2 f/\partial x^2+\partial^2 f/\partial y^2=0$.

(c) If $u(x,y)$ and $v(x,y)$ satisfy the Cauchy-Riemann equations, state the conditions for both u and v to satisfy the Laplace equation.

Solution.

Part (a). Differentiating u with respect to x, we get $\partial u/\partial x=\mathrm{e}^x\cos y$. Differentiating v with respect to y, we get $\partial v/\partial y=\mathrm{e}^x\cos y$. Hence, $\partial u/\partial x=\partial v/\partial y$.

Differentiating u with respect to y, we get $\partial u/\partial y=-\mathrm{e}^x\sin y$. Differentiating v with respect to x, we get $\partial v/\partial x=\mathrm{e}^x\sin y$. Hence, $\partial u/\partial y=-\partial v/\partial x$.

Thus, u and v satisfy the Cauchy-Riemann equations.

Part (b). Differentiating f twice with respect to x,

$$\frac{\partial^2 f}{\partial x^2}=\frac{\partial}{\partial x}(\mathrm{e}^x\sin y-\mathrm{e}^y\sin x)=\mathrm{e}^x\sin y-\mathrm{e}^y\cos x.$$

Differentiating f twice with respect to y,

$$\frac{\partial^2 f}{\partial y^2}=\frac{\partial}{\partial y}(\mathrm{e}^x\cos y+\mathrm{e}^y\cos x)=-\mathrm{e}^x\sin y+\mathrm{e}^y\cos x.$$

Hence,

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}=(\mathrm{e}^x\sin y-\mathrm{e}^y\cos x)+(-\mathrm{e}^x\sin y+\mathrm{e}^y\cos x)=0.$$

Thus, $f(x,y)=\mathrm{e}^x\sin y+\mathrm{e}^y\cos x$ satisfies the Laplace equation.

Part (c). Suppose $u(x,y)$ and $v(x,y)$ satisfy the Cauchy-Riemann equations. This gives

$$\begin{cases} u_x=v_y \\ u_y=-v_x \end{cases}$$

Differentiating with respect to x and y, we obtain

$$\begin{cases} u_{xx}=v_{yx} \\ u_{yx}=-v_{xx} \end{cases} \quad\text{and}\quad \begin{cases} u_{xy}=v_{yy} \\ u_{yy}=-v_{xy} \end{cases}$$

This gives

$$\begin{cases} u_{xx}+u_{yy}=v_{yx}-v_{xy} \\ v_{xx}+v_{yy}=-u_{yx}+u_{xy} \end{cases}$$

Hence, if u and v both satisfy the Laplace equation, we require

$$\begin{cases} v_{yx} - v_{xy} = 0 \\ -u_{yx} + u_{xy} = 0 \end{cases}$$

which gives the conditions $u_{xy} = u_{yx}$ and $v_{xy} = v_{yx}$.

* * * * *

Problem 10. Find the equation of the tangent plane to the surface $z = x^2y$ at the point $(2, 1, 4)$. Hence, state the normal vector of the tangent plane.

Solution. Let $f(x, y) = x^2y$. Then $f_x(x, y) = 2xy$ and $f_y(x, y) = x^2$. Hence, the equation of the tangent plane at $(2, 1, 4)$ is given by

$$z = 4 + f_x(2, 1)(x - 2) + f_y(2, 1)(y - 1) = 4 + 4(x - 2) + 4(y - 1) = 4x + 4y - 8.$$

Rearranging,

$$4x + 4y - z = 8,$$

whence the normal vector of the tangent plane is $(4,\ 4,\ -1)^{\mathsf{T}}$.

* * * * *

Problem 11. The volume of a right-circular cone of radius r cm and height h cm is denoted by V. If h increases from 10 cm to 10.01 cm and r decreases from 12 cm to 11.95 cm, use a linear approximation to estimate the volume of the cone after the changes.

Solution. Let $V(r, h) = \frac{1}{3}\pi r^2 h$ be the volume of the cone. We have $V_r(r, h) = \frac{2}{3}\pi r h$ and $V_h(r, h) = \frac{1}{3}\pi r^2$. The equation of the tangent plane at $r = 12$ and $h = 10$ is given by

$$\begin{aligned} v &= V(12, 10) + V_r(12, 10)(r - 12) + V_h(12, 10)(h - 10) \\ &= \frac{1}{3}\pi(12^2)(10) + \frac{2}{3}\pi(12)(10)(r - 12) + \frac{1}{3}\pi(12^2)(h - 10) = 16\pi(5r + 3h - 60). \end{aligned}$$

Evaluating at $r = 11.95$ and $h = 10.01$, we have

$$v = 16\pi\,[5(11.95) + 3(10.01) - 60] = 476.48\pi.$$

The volume of the cone after the changes is hence approximately 476.48π cm^3.

* * * * *

Problem 12. The radius of a right-circular cylinder is measured with an error of at most 2%, and the height is measured with an error of at most 4%. Approximate the maximum possible percentage error in the volume of the cylinder calculated from these measurements.

Solution. Let the volume of the cylinder be $V = \pi r^2 h$. By the chain rule, we have

$$\mathrm{d}V = \frac{\partial V}{\partial r}\,\mathrm{d}r + \frac{\partial f}{\partial h}\,\mathrm{d}h = 2\pi r h\,\mathrm{d}r + \pi r^2\,\mathrm{d}h.$$

Dividing throughout by $V = \pi r^2 h$,

$$\frac{\mathrm{d}V}{V} = \frac{2\pi r h\,\mathrm{d}r + \pi r^2\,\mathrm{d}h}{\pi r^2 h} = 2\frac{\mathrm{d}r}{r} + \frac{\mathrm{d}h}{h}.$$

Note that $\mathrm{d}V/V$ measures the percentage error of the volume V, while $\mathrm{d}r/r$ and $\mathrm{d}h/h$ measure the percentage error of the radius and height respectively. Hence,

$$\max\frac{\mathrm{d}V}{V} = 2(2\%) + 4\% = 8\%.$$

Problem 13. On a certain mountain, the elevation z above a point (x, y) in a horizontal xy-plane that lies at sea level is $z = 2000 - 2x^2 - 4y^2$ ft. The positive x-axis points east, and the positive y-axis points north. A climber is at the point $(-20, 5, 1100)$.

(a) If the climber uses a compass reading to walk due northeast, will he ascend or descend? Find this rate.

(b) Find the direction where the climber should walk to travel a level path.

Solution.

Part (a). Let $f(x, y) = 2000 - 2x^2 - 4y^2$. Then $f_x(x, y) = -4x$ and $f_y(x, y) = -8y$. Hence,

$$\nabla f = \begin{pmatrix} -4x \\ -8y \end{pmatrix} = -4 \begin{pmatrix} x \\ 2y \end{pmatrix}$$

Note that the vector $(1,\ 1)^{\mathsf{T}}$ points northeast.

$$\nabla f \cdot \widehat{\begin{pmatrix} 1 \\ 1 \end{pmatrix}} = -4 \begin{pmatrix} x \\ 2y \end{pmatrix} \cdot \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -2\sqrt{2}\,(x + 2y)\,.$$

Evaluating at $(-20, 5, 1100)$, the instantaneous rate of change of the climber's altitude would be $-2\sqrt{2}\,(-20 + 2 \cdot 5) = 20\sqrt{2}$ ft/s. That is, the climber would ascend at a rate of $20\sqrt{2}$ feet per second.

Part (b). For a level path, the instantaneous rate of change of the climber's altitude should be 0. Let the direction of the climber be $u = (a,\ b)^{\mathsf{T}}$.

$$D_{\mathbf{u}} f(x, y)\big|_{(-20,5)} = -4 \begin{pmatrix} -20 \\ 10 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} \implies -2a + b = 0 \implies b = 2a.$$

We hence have

$$\mathbf{u} = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 2a \end{pmatrix} = a \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Thus, the climber should walk in the direction of $(1,\ 2)^{\mathsf{T}}$.

* * * * *

Problem 14. Find the absolute maximum and minimum values of $f(x, y) = 3xy - 6x - 3y + 7$ on the closed triangular region R with vertices $(0, 0)$, $(3, 0)$ and $(0, 5)$.

Solution. Note that $f_x(x, y) = 3y - 6$ and $f_y(x, y) = 3x - 3$, whence $f_{xx}(x, y) = f_{yy}(x, y) = 0$ and $f_{xy} = 3$. For stationary points,

$$\nabla f = \mathbf{0} \implies \begin{pmatrix} 3y - 6 \\ 3x - 3 \end{pmatrix} = \mathbf{0} \implies x = 1,\ y = 2.$$

Consider the nature of the stationary point at $(1, 2)$. We have

$$D = f_{xx}(1, 2) f_{yy}(1, 2) - [f_{xy}(1, 2)]^2 = -9 < 0$$

Hence, by the second derivative test, we see that $f(x, y)$ has a saddle point at $(1, 2)$. Thus, the extrema of $f(x, y)$ must occur along its boundary.

Note that the boundary of $f(x, y)$ is given by

- $x = 0,\ y \in [0, 5]$
- $x \in [0, 3],\ y = 0$

- $x \in [0,3]$, $y = 5 - \frac{5}{3}x$

Case 1: $x = 0$, $y \in [0,5]$. We have that $f(0,y) = -3y + 7$, which clearly attains a maximum of 7 at $y = 0$ and a minimum of -8 at $y = 5$.

Case 2: $x \in [0,3]$, $y = 0$. We have that $f(x,0) = -6x + 7$, which clearly attains a maximum of 7 at $x = 0$ and a minimum of -11 at $x = 3$.

Case 3: $x \in [0,3]$, $y = 5 - \frac{5}{3}x$. Observe that

$$f\left(x, 5 - \frac{5}{3}x\right) = 3x\left(5 - \frac{5}{3}x\right) - 6x - 3\left(5 - \frac{5}{3}x\right) + 7 = -(x-2)(5x-4)$$

is concave down and has a turning point at $x = 1.4$. Hence, the function clearly attains a maximum of 1.8 when $x = 1.4$ and a minimum of -11 when $x = 3$ (note that at $x = 0$, the function returns -8).

Hence, the maximum of $f(x,y)$ is 7, while the minimum is -11.

* * * * *

Problem 15. Find the dimensions of a rectangular box, open at the top, having a volume of 32 cm^3, and requiring the least amount of material for its construction.

Solution. Let the box have side lengths of x, y and z cm. Given that the volume of the box is fixed at 32 cm^3, we have

$$xyz = 32 \implies z = \frac{32}{xy}$$

Let the surface area of the box be measured by $f(x,y)$. Then

$$f(x,y) = xy + 2yz + 2xz = xy + 2y\left(\frac{32}{xy}\right) + 2x\left(\frac{32}{xy}\right) = xy + 64x^{-1} + 64y^{-1}.$$

Note that

$$\nabla f = \begin{pmatrix} f_x(x,y) \\ f_y(x,y) \end{pmatrix} = \begin{pmatrix} y - 64x^{-2} \\ x - 64y^{-2} \end{pmatrix}.$$

For stationary points, $\nabla f = \mathbf{0}$. We hence obtain

$$\begin{cases} y = 64x^{-2} \\ x = 64y^{-2} \end{cases} \implies \begin{cases} yx^2 = 64 \\ xy^2 = 64 \end{cases} \implies x^3y^3 = 64^2 \implies xy = 16$$

Hence,

$$x = \frac{x^2y}{xy} = \frac{64}{16} = 4,$$

whence $y = 4$ and $z = 2$. Thus, $f(x,y)$ has a stationary point at $(4,4,2)$.

We now consider the nature of this stationary point. Note that

$$f_{xx}(x,y) = 128x^{-3}, \quad f_{yy} = 128y^{-3}, \quad f_{xy} = 1.$$

Hence,

$$D = f_{xx}(4,4)f_{yy}(4,4) - [f_{xy}(4,4)]^2 = 3$$

Since $D > 0$ and $f_{xx}(4,4) = 2 > 0$, by the second derivative test, $f(x,y)$ attains a minimum at $(4,4,2)$. Thus, the amount of material required is lowest for a box of dimension $4 \times 4 \times 2$.

Problem 16. Find the quadratic approximation of $f(x, y) = x^2y + xy^2$ around the point $(1, 1)$.

Solution. Taking partial derivatives, we have

$$f_x(x, y) = 2xy + y^2, \quad f_y(x, y) = 2xy + x^2$$
$$f_{xx}(x, y) = 2y, \quad f_{xy}(x, y) = 2x + 2y, \quad f_{yy}(x, y) = 2x.$$

Hence, the required quadratic approximation $Q(x, y)$ is given by

$$\begin{aligned}
Q(x, y) &= f(1, 1) + f_x(1, 1)(x - 1) + f_y(1, 1)(y - 1) \\
&\qquad + \frac{1}{2} f_{xx}(1, 1)(x - 1)^2 + f_{xy}(1, 1)(x - 1)(y - 1) + \frac{1}{2} f_{yy}(1, 1)(y - 1)^2 \\
&= 2 + 3(x - 1) + 3(y - 1) + (x - 1)^2 + 4(x - 1)(y - 1) + (y - 1)^2 \\
&= 2 - 3x - 3y + 4xy + x^2 + y^2
\end{aligned}$$

Self-Practice B11

Problem 1. At what rate is the area of a rectangle changing if its length is 15 units and increasing at 3 units/s while its width is 6 units and increasing at 2 units/s?

Solution. Let $l(t)$ and $w(t)$ be the length and width of the rectangle respectively, where t is the time in seconds. Let $A = lw$ be the area of the rectangle. Note that

$$\frac{\partial A}{\partial l} = w \quad \text{and} \quad \frac{\partial A}{\partial w} = l.$$

Also,

$$\frac{\mathrm{d}l}{\mathrm{d}t} = 3 \quad \text{and} \quad \frac{\mathrm{d}w}{\mathrm{d}t} = 2.$$

Hence,

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\partial A}{\partial l}\frac{\mathrm{d}l}{\mathrm{d}t} + \frac{\partial A}{\partial w}\frac{\mathrm{d}w}{\mathrm{d}t} = 3w + 2l.$$

When $l = 15$ and $w = 6$, we have

$$\frac{\mathrm{d}A}{\mathrm{d}t} = 3(6) + 2(15) = 48.$$

Thus, the area of the rectangle is increasing at a rate of 48 units^2/s.

* * * * *

Problem 2. A particle moving along a metal plate in the xy-plane has the velocity $\mathbf{v} = \mathbf{i} - 4\mathbf{j}$ cm/s at the point $(3, 2)$. If the temperature of the plate at points in the xy-plane is $T(x, y) = y^2 \ln x$ where $x \geq 1$, in degrees Celsius, find $\mathrm{d}T/\mathrm{d}t$ at $(3, 2)$.

Solution. Note that

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{\partial T}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial T}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = \left(\frac{y^2}{x}\right)(1) + (2y\ln x)(-4) = \frac{y^2}{x} - 9y\ln x.$$

At $(3, 2)$, we have

$$\left.\frac{\mathrm{d}T}{\mathrm{d}t}\right|_{(3,2)} = \frac{2^2}{3} - 9(2)\ln 3 = -16.2 \text{ (3 d.p.)}.$$

* * * * *

Problem 3. Given that $f(x, y) = x^2 \mathrm{e}^y$, find the maximum value of a directional derivative at $(-2, 0)$ and give a unit vector in the direction in which the maximum value occurs.

Solution. At $(-2, 0)$, we have

$$\nabla f = \begin{pmatrix} 2x\mathrm{e}^y \\ x^2\mathrm{e}^y \end{pmatrix} = \begin{pmatrix} -4 \\ 4 \end{pmatrix}.$$

Note that

$$D_{\mathbf{u}} f(x, y) = \nabla f \cdot \mathbf{u} = |\nabla f| \cos\theta,$$

where θ is the angle between ∇f and $\mathbf{u}$. Hence, the maximum value of the directional derivative is

$$\left|\begin{pmatrix} -4 \\ 4 \end{pmatrix}\right| = \sqrt{(-4)^2 + 4^2} = 4\sqrt{2}.$$

This occurs when $\theta = 0$, i.e. when $\mathbf{u}$ is the same direction as ∇f. Hence,

$$\mathbf{u} = \frac{1}{4\sqrt{2}}\begin{pmatrix} -4 \\ 4 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Problem 4. If the electric potential at a point (x, y) in the xy-plane is $V(x, y)$, where $V(x, y) = \mathrm{e}^{-2x} \cos 2y$, find the direction where V decreases most rapidly at $(0, \pi/6)$.

Solution. At $(0, \pi/6)$, we have

$$\nabla V = \begin{pmatrix} -2\mathrm{e}^{-2x} \cos 2y \\ -2\mathrm{e}^{-2x} \sin 2y \end{pmatrix} = -\begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}.$$

Note that

$$D_{\mathbf{u}} V(x, y) = \nabla V \cdot \mathbf{u} = |\nabla V| \cos \theta,$$

where θ is the angle between ∇V and $\mathbf{u}$. Hence, V decreases the most when $\theta = \pi$, i.e. when $\mathbf{u}$ is in the opposite direction as ∇V. Thus, the desired direction is $\begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}$.

* * * * *

Problem 5. Find all local extrema and saddle points of $f(x, y) = 4xy - x^4 - y^4$.

Solution. Note that

$$\nabla f = \begin{pmatrix} 4y - 4x^3 \\ 4x - 4y^3 \end{pmatrix}.$$

Setting this equal to the zero vector, we have the system

$$\begin{cases} -4x^3 + \ \ 4y = 0 \\ \ \ \ \ 4x - 4y^3 = 0 \end{cases}.$$

From the first equation, we get $y = x^3$. Substituting this into the second equation yields

$$x - x^9 = x\left(1 - x^8\right) = 0.$$

Note that $x^8 - 1$ factors as $\left(x^4 + 1\right)\left(x^2 + 1\right)(x + 1)(x - 1)$. We thus have $x = -1, 0, 1$, which correspond to the points $(-1, -1)$, $(0, 0)$ and $(1, 1)$.

Let

$$D = f_{xx} f_{yy} - f_{xy}^2 = \left(-12x^2\right)\left(-12y^2\right) - (4)^2 = 144x^2y^2 - 16.$$

Case 1. At $(-1, -1)$, we have

$$D = 144(-1)^2(-1)^2 - 16 = 128 > 0.$$

Since $f_{xx} = -12(-1)^2 = -12 < 0$, by the second partial derivative test, $(-1, -1)$ is a maximum point.

Case 2. At $(0, 0)$, we have

$$D = 144(0)^2(0)^2 - 16 = -16 < 0.$$

By the second partial derivative test, $(0, 0)$ is a saddle point.

Case 3. At $(1, 1)$, we have

$$D = 144(1)^2(1)^2 - 16 = 128 > 0.$$

Since $f_{xx} = -12(1)^2 = -12 < 0$, by the second partial derivative test, $(1, 1)$ is a maximum point.

* * * * *

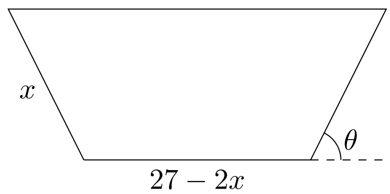
Problem 6. Find the absolute extrema of $f(x, y) = x^2 + 2y^2 - x$ such that the domain of this function f is the circular region $x^2 + y^2 \leq 4$.

Solution. Note that
$$\nabla f = \begin{pmatrix} 2x-1 \\ 4y \end{pmatrix}.$$
Setting this equal to the zero vector, we see that f has only one stationary point at $(1/2, 0)$, which is in the domain. At this point,
$$f\left(\frac{1}{2}, 0\right) = \left(\frac{1}{2}\right)^2 + 2(0)^2 - \frac{1}{2} = -\frac{1}{4}.$$
We now consider the points along the boundary of D_f, which is given by the equation $x^2 + y^2 = 4$. Substituting $y^2 = 4 - x^2$ into the definition of $f(x, y)$, we get the univariate function $g(x)$:
$$g(x) = x^2 + 2\left(4 - x^2\right) - x = -\left(x + \frac{1}{2}\right)^2 + \frac{33}{4}.$$
Also note that $x \in [-2, 2]$. Clearly, $g(x)$ attains a maximum of $33/4$ at $x = -1/2$ and a minimum of 2 at $x = 2$.

Thus, the absolute maximum of $f(x, y)$ is $33/4$, while the absolute minimum of $f(x, y)$ is $-1/4$.

* * * * *

Problem 7. A length of sheet metal 27 cm wide is to be made into a water trough by bending up two sides as shown in the figure below. Find the values of x and θ such that the trapezoid-shaped cross-section has a maximum area.



Solution 1. Take the mirror image of the figure and place it on top of the original image. We get a hexagon with perimeter 54, and we are tasked with maximizing its area. It is a well-known result that for n-sided polygons with fixed perimeters, the regular n-gon encloses the largest area. Hence, we have $x = 54/6 = 9$ and $\theta = 2\pi/6 = \pi/3$.

Solution 2. Observe that the longer side of the trapezium is given by $(27-2x)+2(x\cos\theta)$, while the height of the trapezium is given by $x\sin\theta$. The area A of the trapezium is thus given by
$$A = \frac{(27-2x) + (27-2x+2x\cos\theta)}{2}(x\sin\theta) = 27x\sin\theta + x^2\sin\theta(\cos\theta - 2).$$
Observe that
$$\nabla A = \begin{pmatrix} A_x \\ A_\theta \end{pmatrix} = \begin{pmatrix} 27\sin\theta + 2x\sin\theta(\cos\theta - 2) \\ 27x\cos\theta + x^2\left(\cos^2\theta - \sin^2\theta - 2\cos\theta\right) \end{pmatrix}.$$
Setting this equal to the zero vector, we get the following system:
$$\begin{cases} 27\sin\theta + 2x\sin\theta(\cos\theta - 2) & = 0, \\ 27x\cos\theta + x^2\left(\cos^2\theta - \sin^2\theta - 2\cos\theta\right) & = 0. \end{cases}$$
From the first equation, we get
$$x = \frac{-27}{2(\cos\theta - 2)}.$$

Substituting this into the second equation,

$$27\left(\frac{-27}{2\left(\cos\theta-2\right)}\right)\cos\theta+\left(\frac{-27}{2\left(\cos\theta-2\right)}\right)^2\left(\cos^2\theta-\sin^2\theta-2\cos\theta\right)=0.$$

Clearing denominators and simplifying, we get

$$-2\cos\theta\left(\cos\theta-2\right)+\left(\cos^2\theta-\sin^2\theta-2\cos\theta\right)=0.$$

Expanding, we have

$$-\cos^2\theta+2\cos\theta-\sin^2\theta=0,$$

from which we immediately get $\cos\theta=1/2$, whence $\theta=\pi/3$ and $x=9$.

We now calculate the second partial derivatives of A at $\theta=\pi/3$ and $x=9$. First, we have

$$A_{xx}=2\sin\theta\left(\cos\theta-2\right)=-2.5981\ (5\text{ s.f.}).$$

Secondly, we have

$$A_{\theta\theta}=-27x\sin\theta+x^2\left(-2\cos\theta\sin\theta-2\sin\theta\cos\theta+2\sin\theta\right)=-210.44\ (5\text{ s.f.}).$$

Lastly, we have

$$A_{\theta x}=27\cos\theta+2x\left(\cos^2\theta-\sin^2\theta-2\cos\theta\right)=-13.5.$$

Since

$$D=A_{xx}A_{\theta\theta}-A_{\theta x}^2=(-2.5981)(-210.44)-(-13.5)^2=364.5>0$$

and $A_{xx}=-2.5981<0$, by the second partial derivative test, A attains a maximum when $x=9$ and $\theta=\pi/3$.

* * * * *

Problem 8. A Further Maths student smiled when a question asked for him to find the quadratic approximation for the function of $f(x,y)=xy-3y-x$ around the point $(2,3)$. Explain why he is so delighted.

Solution. The function $f(x,y)=xy-3y-x$ is already a quadratic, hence the quadratic approximation to $f(x,y)$ is simply $f(x,y)$ itself.

* * * * *

Problem 9. A company produces two products, A and B, which require different amounts of two resources, Resource 1 and Resource 2. The profit generated by selling product A is \$10 per unit, and the profit from selling product B is \$15 per unit. Each unit of product A requires 2 units of Resource 1 and 1 unit of Resource 2. Each unit of product B requires 1 unit of Resource 1 and 3 units of Resource 2. The company has a total of 100 units of Resource 1 and 90 units of Resource 2. What should the company produce in order to maximize its profitability?

Solution. We use linear programming to solve this problem.

Let n and m be the amount of product A and B produced by the company. Let Π be the total revenue earned, i.e.

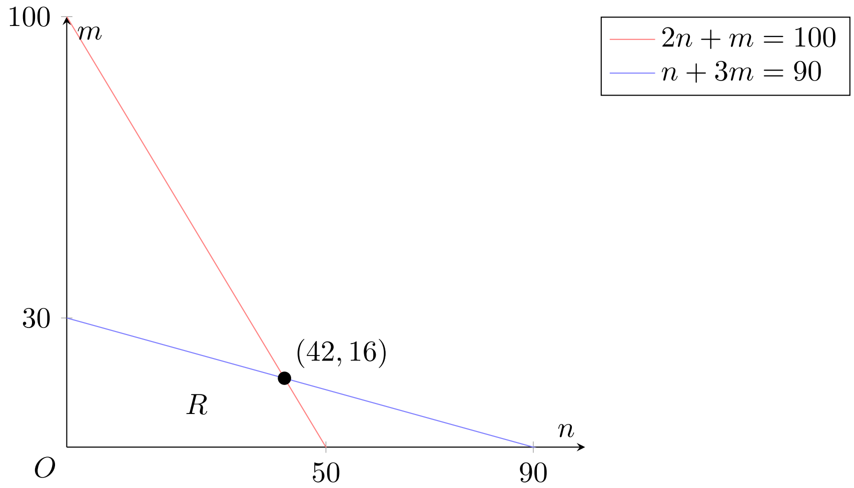
$$\Pi=10n+15m.$$

Due to resource constraints, we have the inequalities

$$\begin{cases}2n+m & \leq 100,\\ n+3m & \leq 90.\end{cases}$$

Additionally, we have $n, m \geq 0$.

We can visualize these inequalities graphically:

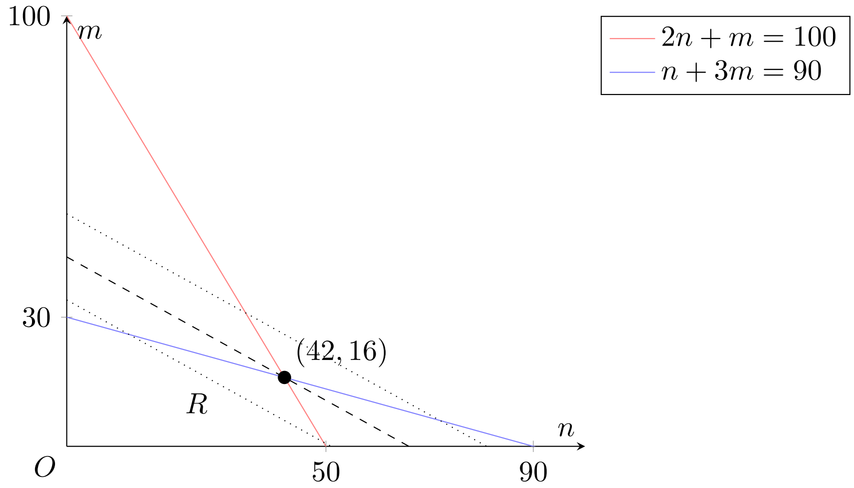


Here, R represents the feasible zone, i.e. the region where both inequalities are satisfied. This is the region where the company can produce. Also note that the two "boundaries" intersect at $(42, 16)$.

Now, recall that $\Pi = 10n + 15m$. Rearranging,

$$m = \frac{\Pi}{15} - \frac{2}{3}n.$$

If we plot this, we get a line with m-intercept $\Pi/15$ and gradient $-2/3$. Our goal of maximizing Π can be restated as "find the largest value $\Pi/15$ such that a line with gradient $-2/3$ intersects the region R once".



From the figure above, it is easy to see that the "optimal line" will only intersect the region R at only one point: $(42, 16)$. Hence, the company should produce 42 units of product A and 16 units of product B.

Assignment B11

Problem 1. Show that if f is differentiable at (x_0, y_0) and $\nabla f(x_0, y_0) \neq \mathbf{0}$, then $\nabla f(x_0, y_0)$ is perpendicular to the level curve of f through (x_0, y_0).

Solution. Let $f(x, y) = (x(t), y(t))$. Let the level curve at (x_0, y_0) have equation $f(x, y) = c$. Implicitly differentiating this with respect to t, we get

$$\frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = \begin{pmatrix} f_x \\ f_y \end{pmatrix} \cdot \begin{pmatrix} \mathrm{d}x/\mathrm{d}t \\ \mathrm{d}y/\mathrm{d}t \end{pmatrix} = \nabla f \cdot \mathbf{u} = 0,$$

where $\mathbf{u}$ is the tangent to the level curve at (x_0, y_0). Since both ∇f and $\mathbf{u}$ are non-zero vectors, they must be perpendicular to each other.

* * * * *

Problem 2. Find the quadratic approximation of $f(x, y) = \mathrm{e}^{x^2+y^2}$ around the point $\left(\frac{1}{2}, 0\right)$.

Solution. Observe that we have

$$f_x(x, y) = 2x\mathrm{e}^{x^2+y^2}, \quad f_y(x, y) = 2y\mathrm{e}^{x^2+y^2}$$
$$f_{xx}(x, y) = 2\mathrm{e}^{x^2+y^2}(2x^2 + 1), \quad f_{xy}(x, y) = 4xy\mathrm{e}^{x^2+y^2}, \quad f_{yy}(x, y) = 2\mathrm{e}^{x^2+y^2}(2y^2 + 1).$$

Evaluating $f(x, y)$ and the above partial derivatives at $\left(\frac{1}{2}, 0\right)$, we obtain

$$f(x, y) = \mathrm{e}^{1/4}, \quad f_x(x, y) = \mathrm{e}^{1/4}, \quad f_y(x, y) = 0$$
$$f_{xx}(x, y) = 3\mathrm{e}^{1/4}, \quad f_{xy}(x, y) = 0, \quad f_{yy}(x, y) = 2\mathrm{e}^{1/4}.$$

The quadratic approximation $Q(x, y)$ to $f(x, y)$ at $\left(\frac{1}{2}, 0\right)$ is hence

$$Q(x, y) = \mathrm{e}^{1/4} + \mathrm{e}^{1/4}\left(x - \frac{1}{2}\right) + 3\mathrm{e}^{1/4}\left(x - \frac{1}{2}\right)^2 + \mathrm{e}^{1/4}y^2.$$

* * * * *

Problem 3. A common problem in experimental work is to obtain a mathematical relationship between two variables x and y by "fitting" a curve to points in the plane corresponding to various experimentally determines values of x and y, say

$$(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_n, y_n).$$

Based on theoretical considerations, or simply on the pattern of the points, one decides on the general form of the curve to be fitted. Often, the "curve" to be fitted is a straight line, $y = ax + b$. One criterion for selecting a line of "best fit" is to choose a and b to minimize the function

$$f(a, b) = \sum_{k=1}^{n}(ax_k + b - y_k)^2.$$

Geometrically, $|ax_k + b - y_k|$ is the vertical distance between the data point (x_k, y_k) and the line $y = ax + b$, so in effect, minimizing $f(a, b)$ minimizes the sum of the squares of the vertical distances. This procedure is called the method of least squares.

(a) Show that the conditions $\partial f/\partial a = 0$ and $\partial f/\partial b = 0$ result in the equations

$$\left(\sum_{k=1}^{n} x_k^2\right) a + \left(\sum_{k=1}^{n} x_k\right) b = \sum_{k=1}^{n}(x_k y_k)$$
$$\left(\sum_{k=1}^{n} x_k\right) a + nb = \sum_{k=1}^{n} y_k$$

(b) Solve the equations for a and b to show that

$$a = \frac{n\sum_{k=1}^n (x_k y_k) - \left(\sum_{k=1}^n x_k\right)\left(\sum_{k=1}^n y_k\right)}{n\sum_{k=1}^n x_k^2 - \left(\sum_{k=1}^n x_k\right)^2}$$

and

$$b = \frac{\left(\sum_{k=1}^n x_k^2\right)\left(\sum_{k=1}^n y_k\right) - \left(\sum_{k=1}^n x_k\right)\left(\sum_{k=1}^n (x_k y_k)\right)}{n\sum_{k=1}^n x_k^2 - \left(\sum_{k=1}^n x_k\right)^2}.$$

(c) Given that $\bar{x} = \frac{1}{n}\sum_{k=1}^n x_k$, show that $n\sum_{k=1}^n x_k^2 - \left(\sum_{k=1}^n x_k\right)^2 > 0$.

(d) Find $f_{aa}(a,b)$, $f_{bb}(a,b)$ and $f_{ab}(a,b)$.

(e) Show that f has a relative minimum at the critical point found in (b).

Solution.

Part (a). Observe that

$$\frac{\partial f}{\partial a} = \frac{\partial}{\partial a}\sum_{k=1}^n (ax_k + b - y_k)^2 = \sum_{k=1}^n 2x_k(ax_k + b - y_k) = 2\sum_{k=1}^n (ax_k^2 + bx_k - x_k y_k).$$

Hence,

$$\frac{\partial f}{\partial a} = 2\sum_{k=1}^n (ax_k^2 + bx_k - x_k y_k) = 0 \implies a\sum_{k=1}^n x_k^2 + b\sum_{k=1}^n x_k = \sum_{k=1}^n x_k y_k.$$

Observe that

$$\frac{\partial f}{\partial b} = \frac{\partial}{\partial b}\sum_{k=1}^n (ax_k + b - y_k)^2 = \sum_{k=1}^n 2(ax_k + b - y_k) = 2\left[\sum_{k=1}^n (ax_k - y_k) + bn\right].$$

Hence,

$$\frac{\partial f}{\partial b} = 2\left[\sum_{k=1}^n (ax_k - y_k) + bn\right] = 0 \implies a\sum_{k=1}^n x_k + bn = \sum_{k=1}^n y_k.$$

Part (b). Let

$$A = \sum_{k=1}^n x_k^2, \quad B = \sum_{k=1}^n x_k, \quad C = \sum_{k=1}^n (x_k y_k), \quad D = n, \quad E = \sum_{k=1}^n y_k.$$

The above equations transform into

$$\begin{cases} Aa + Bb = C \\ Ba + Db = E \end{cases}.$$

One can easily solve the system for a and b, yielding

$$a = \frac{CD - BE}{AD - B^2}, \quad b = \frac{AE - BC}{AD - B^2}.$$

Thus,

$$a = \frac{n\sum_{k=1}^n (x_k y_k) - \left(\sum_{k=1}^n x_k\right)\left(\sum_{k=1}^n y_k\right)}{n\sum_{k=1}^n x_k^2 - \left(\sum_{k=1}^n x_k\right)^2}$$

and

$$b = \frac{\left(\sum_{k=1}^n x_k^2\right)\left(\sum_{k=1}^n y_k\right) - \left(\sum_{k=1}^n x_k\right)\left(\sum_{k=1}^n (x_k y_k)\right)}{n\sum_{k=1}^n x_k^2 - \left(\sum_{k=1}^n x_k\right)^2}.$$

Part (c). Observe that

$$\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k \implies \sum_{k=1}^{n} x_k = n\bar{x}.$$

Consider $n\sum_{k=1}^{n} x_k^2 - \left(\sum_{k=1}^{n} x_k\right)^2$.

$$\begin{aligned} n\sum_{k=1}^{n} x_k^2 - \left(\sum_{k=1}^{n} x_k\right)^2 &= n\left(\sum_{k=1}^{n} x_k^2 - n\bar{x}^2\right) = n\left(\sum_{k=1}^{n} x_k^2 - 2n\bar{x}^2 + n\bar{x}^2\right) \\ &= n\left[\sum_{k=1}^{n} x_k^2 - 2n\bar{x}\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right) + \sum_{k=1}^{n} \bar{x}^2\right] = n\left(\sum_{k=1}^{n} x_k^2 - \sum_{k=1}^{n} 2x_k\bar{x} + \sum_{k=1}^{n} \bar{x}^2\right) \\ &= n\sum_{k=1}^{n}\left(x_k^2 - 2x_k\bar{x} + \bar{x}^2\right) = n\sum_{k=1}^{n}\left(x_k - \bar{x}\right)^2. \end{aligned}$$

Given that the RHS is a sum of squares, it must be greater than or equal to 0. We thus have the inequality

$$n\sum_{k=1}^{n} x_k^2 - \left(\sum_{k=1}^{n} x_k\right)^2 \geq 0.$$

However, if $n\sum_{k=1}^{n} x_k^2 - \left(\sum_{k=1}^{n} x_k\right)^2 = 0$, then both a and b would be undefined. Thus, we must have a strict inequality, which gives

$$n\sum_{k=1}^{n} x_k^2 - \left(\sum_{k=1}^{n} x_k\right)^2 > 0.$$

Part (d). From (a), we have

$$f_a(a, b) = 2a\sum_{k=1}^{n} x_k^2 + 2b\sum_{k=1}^{n} x_k - 2\sum_{k=1}^{n} x_k y_k$$

and

$$f_b(a, b) = 2a\sum_{k=1}^{n} x_k + 2nb - 2\sum_{k=1}^{n} y_k.$$

Thus,

$$f_{aa}(a, b) = 2\sum_{k=1}^{n} x_k^2, \qquad f_{ab}(a, b) = 2\sum_{k=1}^{n} x_k, \qquad f_{bb}(a, b) = 2n.$$

Part (e). Let $D = f_{aa}(a, b)f_{bb}(a, b) - \left[f_{ab}(a, b)\right]^2$. From part (d), we have

$$D = 4\left[n\sum_{k=1}^{n} x_k^2 - \left(\sum_{k=1}^{n} x_k\right)^2\right],$$

which is clearly positive from part (c). Furthermore, $f_{aa}(a, b) = 2\sum_{k=1}^{n} x_k^2$ is clearly positive (note that we reject the equality for the reason stated in part (c)). Thus, by the second partial derivative test, the critical point found in part (b) must be a minimum point.

B12 Separable Differential Equations

Tutorial B12

Problem 1. Given that $y = 1$ when $x = 1$, find the particular solution of the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y^2}{x}$.

Solution.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y^2}{x} \implies \frac{1}{y^2}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{x} \implies \int \frac{1}{y^2}\,\mathrm{d}y = \int \frac{1}{x}\,\mathrm{d}x$$
$$\implies -\frac{1}{y} = \ln|x| + C_1 \implies y = \frac{1}{C - \ln|x|}, \quad C = -C_1.$$

Since $y(1) = 1$, we have

$$1 = \frac{1}{C - \ln|1|} \implies C = 1 \implies y = \frac{1}{1 - \ln|x|}.$$

* * * * *

Problem 2. Two variables x and t are connected by the differential equation $\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{kx}{10-x}$, where $0 < x < 10$ and where k is a constant. It is given that $x = 1$ when $t = 0$ and that $x = 2$ when $t = 1$. Find the value of t when $x = 5$, given your answer to three s.f.

Solution.

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{kx}{10-x} \implies \frac{10-x}{x}\frac{\mathrm{d}x}{\mathrm{d}t} = k$$
$$\implies \int \frac{10-x}{x}\,\mathrm{d}x = \int k\,\mathrm{d}t \implies 10\ln x - x = kt + C.$$

Evaluating at $x = 1$ and $t = 0$,

$$10\ln(1) - 1 = k(0) + C \implies C = -1.$$

Evaluating at $x = 2$ and $t = 1$,

$$10\ln(2) - 2 = k(1) - 1 \implies k = 10\ln 2 - 1.$$

Hence, evaluating at $x = 5$, we get

$$10\ln(5) - 5 = (10\ln 2 - 1)\,t - 1 \implies t = 2.04 \text{ (3 s.f.)}.$$

* * * * *

Problem 3. Use the substitution $y = u - 2x$ to find the general solution of the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{8x+4y+1}{4x+2y+1}$.

Solution. Note that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{8x+4y+1}{4x+2y+1} = -2 + \frac{1}{4x+2y+1}.$$

Also note that under the substitution $y = u - 2x$, we have

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}x} - 2.$$

Thus,

$$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{1}{4x+2y+1} = \frac{1}{4x+2(u-2x)+1} = \frac{1}{2u+1} \implies (2u+1)\frac{\mathrm{d}u}{\mathrm{d}x} = 1.$$

Integrating with respect to x,

$$\int (2u+1)\ \mathrm{d}u = \int 1\,\mathrm{d}x \implies u^2 + u = x + C \implies (y+2x)^2 + y + x = C.$$

* * * * *

Problem 4. By using the substitution $z = y\mathrm{e}^{2x}$, find the general solution of the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} + 2y = x\mathrm{e}^{-2x}$.

Find the particular solution of the differential equation given that $\frac{\mathrm{d}y}{\mathrm{d}x} = 1$ when $x = 0$.

Solution. Note that

$$z = ye^{2x} \implies \frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{e}^{2x} + 2y\mathrm{e}^{2x} = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{e}^{2x} + 2z \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}x}\mathrm{e}^{-2x} - 2y.$$

Substituting this into the given differential equation,

$$\frac{\mathrm{d}z}{\mathrm{d}x}\mathrm{e}^{-2x} - 2y + 2y = x\mathrm{e}^{-2x} \implies \frac{\mathrm{d}z}{\mathrm{d}x} = x.$$

Integrating with respect to x, we easily see that

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

Since $\frac{\mathrm{d}y}{\mathrm{d}x} = 1$ when $x = 0$, we have

$$1 + 2y = 0 \implies y = -\frac{1}{2} \implies C = -\frac{1}{2}.$$

The desired particular solution is hence

$$y = \frac{x^2-1}{2\mathrm{e}^{2x}}.$$

* * * * *

Problem 5. Find the general solution of the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = 6xy^3$.

Find its particular solution given that $y = 0.5$ when $x = 0$.

Determine the interval of validity for the particular solution.

Solution.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 6xy^3 \implies \frac{1}{y^3}\frac{\mathrm{d}y}{\mathrm{d}x} = 6x \implies \int \frac{1}{y^3}\,\mathrm{d}y = \int 6x\,\mathrm{d}x$$
$$\implies -\frac{1}{2y^2} = 3x^2 + C_1 \implies y^2 = \frac{1}{C - 6x^2}.$$

Since $y(0) = 0.5$, we have

$$(0.5)^2 = \frac{1}{C - 6(0)^2} \implies C = 4.$$

Thus, the particular solution is

$$y^2 = \frac{1}{4 - 6x^2}.$$

For the solution to be valid, we require $4 - 6x^2 > 0$, whence $x \in \left(-\sqrt{2/3}, \sqrt{2/3}\right)$.

* * * * *

Problem 6.

(a) Find the general solution of the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3x}{x^2+1}$.

(b) What can you say about the gradient of every solution as $x \to \pm\infty$?

(c) Find the particular solution of the differential equation for which $y = 2$ when $x = 0$. Hence, sketch the graph of this solution.

Solution.

Part (a).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3x}{x^2+1} = \frac{3}{2}\left(\frac{2x}{x^2+1}\right) \implies y = \frac{3}{2}\ln\left(x^2+1\right) + C.$$

Part (b).

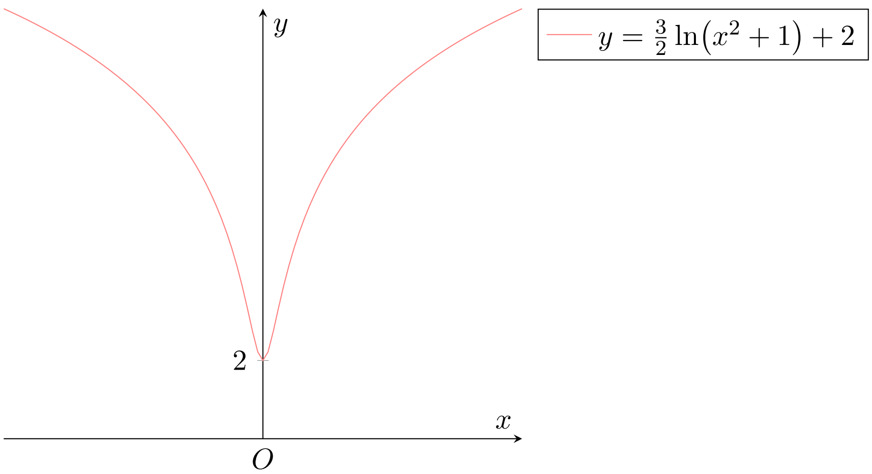
$$\lim_{x\to\pm\infty}\frac{\mathrm{d}y}{\mathrm{d}x} = \lim_{x\to\pm\infty}\frac{3x}{x^2+1} = 0.$$

Part (c). Evaluating the general solution at $x = 0$ and $y = 2$, we get

$$2 = \frac{3}{2}\ln\left(0^2+1\right) + C \implies C = 2.$$

Thus, the particular solution is

$$y = \frac{3}{2}\ln\left(x^2+1\right) + 2.$$



Problem 7. The variables x, y and z are connected by the following differential equations.

$$\frac{\mathrm{d}z}{\mathrm{d}x} = 3 - 2z \qquad (*)$$
$$\frac{\mathrm{d}y}{\mathrm{d}x} = z$$

(a) Given that $z < \frac{3}{2}$, solve equation $(*)$ to find z in terms of x.

(b) Hence, find y in terms of x.

(c) Use the result in part (b) to show that

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

for constants a and b to be determined.

(d) The curve of the solution in part (b) passes through the points $(0, 1)$ and $\left(2, 3 + \mathrm{e}^{-4}\right)$. Sketch this curve, indicating its axial intercept and asymptote (if any).

Solution.

Part (a).

$$\frac{\mathrm{d}z}{\mathrm{d}x} = 3 - 2z \implies \frac{1}{3-2z}\frac{\mathrm{d}z}{\mathrm{d}x} = 1 \implies \int \frac{1}{3-2z}\,\mathrm{d}z = \int 1\,\mathrm{d}x$$
$$\implies -\frac{1}{2}\ln(3-2z) = x + C_1 \implies z = \frac{3}{2} - A\mathrm{e}^{-2x}, \quad A = \frac{\mathrm{e}^{-2C_1}}{2}.$$

Thus, the general solution is

$$z = \frac{3}{2} - A\mathrm{e}^{-2x}, \quad A \in \mathbb{R}^+.$$

Part (b).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3}{2} - A\mathrm{e}^{-2x} \implies y = \int \left(\frac{3}{2} - A\mathrm{e}^{-2x}\right)\mathrm{d}x = \frac{3}{2}x + \frac{A}{2}\mathrm{e}^{-2x} + B, \quad B \in \mathbb{R}.$$

Part (c).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3}{2} - A\mathrm{e}^{-2x} \implies \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 2A\mathrm{e}^{-2x} = 2\left(\frac{3}{2} - \frac{\mathrm{d}y}{\mathrm{d}x}\right) = -2\frac{\mathrm{d}y}{\mathrm{d}x} + 3.$$

Hence, $a = -2$ and $b = 3$.

Part (d). Evaluating the general solution at $(0, 1)$, we obtain

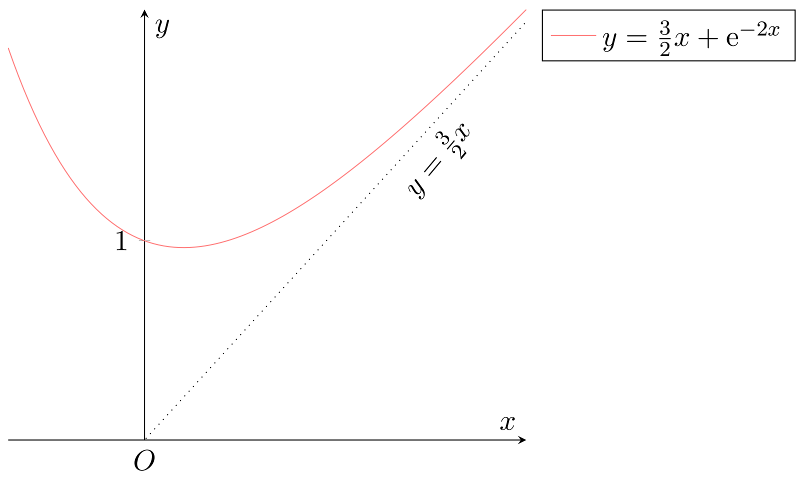
$$1 = \frac{3}{2}(0) + \frac{A}{2}\mathrm{e}^{-2(0)} + B \implies B = 1 - \frac{A}{2}.$$

Evaluating the general solution at $\left(2, 3 + \mathrm{e}^{-4}\right)$, we obtain

$$3 + \mathrm{e}^{-4} = \frac{3}{2}(2) + \frac{A}{2}\mathrm{e}^{-2(2)} + \left(1 - \frac{A}{2}\right) \implies A = 2.$$

The curve thus has equation

$$y = \frac{3}{2}x + \mathrm{e}^{-2x}.$$



* * * * *

Problem 8. A bottle containing liquid is taken from a refrigerator and placed in a room where the temperature is a constant 20 °C. As the liquid warms up, the rate of increase of its temperature θ °C after time t minutes is proportional to the temperature difference $(20-\theta)$ °C. Initially the temperature of the liquid is 10 °C and the rate of increase of the temperature is 1 °C per minute. By setting up and solving a differential equation, show that $\theta = 20 - 10e^{-t/10}$.

Find the time it takes the liquid to reach a temperature of 15 °C, and state what happens to θ for large values of t. Sketch a graph of θ against t.

Solution. Since $\frac{d\theta}{dt} \propto (20-\theta)$, we have $\frac{d\theta}{dt} = k(20-\theta)$, where k is a constant. We now solve for θ.

$$\frac{d\theta}{dt} = k(20-\theta) \implies \frac{1}{20-\theta}\frac{d\theta}{dt} = k \implies \int \frac{1}{20-\theta}\, d\theta = \int k\, dt$$
$$\implies -\ln(20-\theta) = kt + C_1 \implies \theta = 20 - Ce^{-kt}, \quad C = e^{-C_1}.$$

Evaluating at $\theta = 0$ and $\theta = 10$, we get

$$10 = 20 - Ce^{-0} \implies C = 10.$$

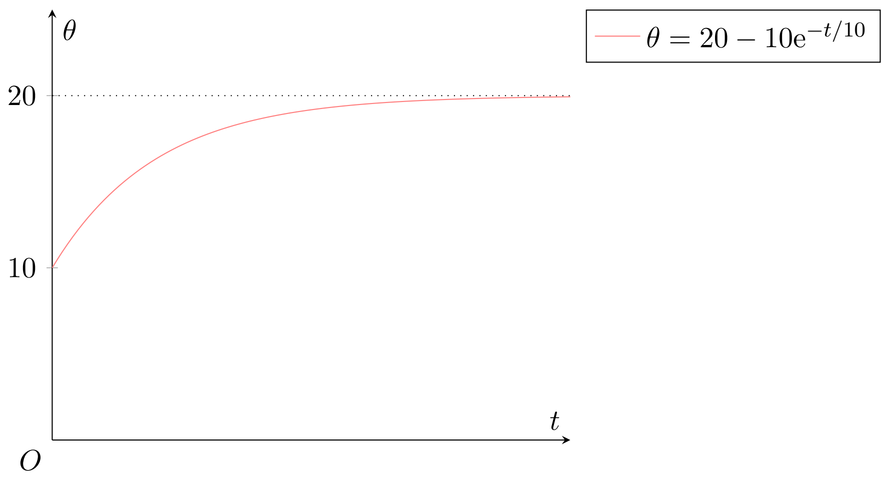
Additionally, since $\frac{d\theta}{dt} = 1$ when $t = 0$, we have

$$1 = k\left[20 - (20 - 10e^{0})\right] = 10k \implies k = \frac{1}{10}.$$

Thus,

$$\theta = 20 - 10e^{-t/10}.$$

Using G.C., when $\theta = 15$, we have $t = 6.93$. Thus, it takes 6.93 minutes for the liquid to reach a temperature of 15°C. As t tends to infinity, θ tends towards 20.



* * * * *

Problem 9.

(a) Find $\int \frac{1}{100-v^2}\,\mathrm{d}x$.

(b) A stone is dropped from a stationary balloon. It leaves the balloon with zero speed, and t seconds later its speed v metres per second satisfies the differential equation

$$\frac{\mathrm{d}v}{\mathrm{d}t} = 10 - 0.1v^2.$$

(i) Find t in terms of v. Hence, find the exact time the stone takes to reach a speed of 5 metres per second.

(ii) Find the speed of the stone after 1 second.

(iii) What happens to the speed of the stone for large values of t?

Solution.

Part (a).

$$\int \frac{1}{100-v^2}\,\mathrm{d}v = \frac{1}{2(10)}\ln\left(\frac{10+v}{10-v}\right) + C = \frac{1}{20}\ln\left(\frac{10+v}{10-v}\right) + C.$$

Part (b).

Part (b)(i).

$$\frac{\mathrm{d}v}{\mathrm{d}t} = 10 - 0.1v^2 = \frac{100-v^2}{10} \implies \frac{1}{100-v^2}\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{1}{10} \implies \int \frac{1}{100-v^2}\,\mathrm{d}v = \int \frac{1}{10}\,\mathrm{d}t$$
$$\implies \frac{1}{20}\ln\left(\frac{10+v}{10-v}\right) + C_1 = \frac{t}{10} \implies t = \frac{1}{2}\ln\left(\frac{10+v}{10-v}\right) + C, \quad C = 10C_1.$$

Evaluating the solution at $t = 0$ and $v = 0$, we get

$$0 = \frac{1}{2}\ln\left(\frac{10+0}{10+0}\right) + C \implies C = 0.$$

Thus, the general solution is

$$t = \frac{1}{2}\ln\left(\frac{10+v}{10-v}\right).$$

Consider $v = 5$, we have

$$t = \frac{1}{2}\ln\left(\frac{10+5}{10-5}\right) = \frac{1}{2}\ln 3.$$

It thus takes $\frac{1}{2}\ln 3$ seconds for the stone to reach a speed of 5 m/s.

Part (b)(ii). Consider $t = 1$. Using G.C., we get $v = 7.62$. Thus, after 1 second, the stone has a speed of 7.62 m/s.

Part (b)(iii). As $t \to \infty$, we have $\ln\left(\frac{10+v}{10-v}\right) \to \infty \implies \frac{10+v}{10-v} \to \infty$. Thus, $v \to 10$. Hence, for large values of t, the speed of the stone approaches 10 m/s.

* * * * *

Problem 10. Two scientists are investigating the change of a certain population of an animal species of size n thousand at time t years. It is known that due to its inability to reproduce effectively, the species is unable to replace itself in the long run.

(a) One scientist suggests that n and t are related by the differential equation $\frac{\mathrm{d}^2 n}{\mathrm{d}t^2} = 10 - 6t$. Given that $n = 100$ when $t = 0$, show that the general solution of this differential equation is $n = 5t^2 - t^3 + Ct + 100$, where C is a constant. Sketch the solution curve of the particular solution when $C = 0$, stating the axial intercepts clearly.

(b) The other scientist suggests that n and t are related by the differential equation $\frac{\mathrm{d}n}{\mathrm{d}t} = 3 - 0.02n$. Find n in terms of t, given again that $n = 100$ when $t = 0$. Explain in simple terms what will eventually happen to the population using this model.

Which is a more appropriate model in modelling the population of the animal species?

Solution.

Part (a).

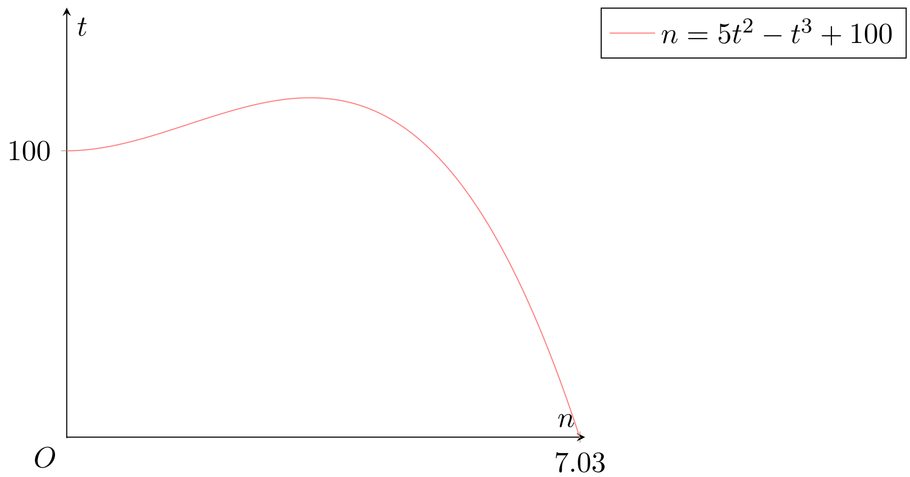
$$\frac{\mathrm{d}^2 n}{\mathrm{d}t^2} = 10 - 6t \implies \frac{\mathrm{d}n}{\mathrm{d}t} = \int (10 - 6t)\,\mathrm{d}\theta = 10t - 3t^2 + C$$
$$\implies n = \int \left(10t - 3t^2 + C\right)\mathrm{d}t = 5t^2 - t^3 + Ct + D.$$

Evaluating the solution at $t = 0$ and $n = 100$, we obtain $D = 100$. Thus,

$$n = 5t^2 - t^3 + Ct + 100.$$

Hence, when $C = 0$,

$$n = 5t^2 - t^3 + 100.$$



Part (b).

$$\frac{\mathrm{d}n}{\mathrm{d}t} = 3 - 0.02n = \frac{150 - n}{50} \implies \frac{1}{150 - n}\frac{\mathrm{d}n}{\mathrm{d}t} = \frac{1}{50} \implies \int \frac{1}{150 - n}\,\mathrm{d}n = \int \frac{1}{50}\,\mathrm{d}t$$
$$\implies -\ln(150 - n) = \frac{1}{50}t + C_1 \implies n = 150 - C\mathrm{e}^{-t/50}, \quad C = \mathrm{e}^{-C_1}.$$

When $t = 0$ and $n = 100$, we have $C = 50$. Hence,

$$n = 150 - 50\mathrm{e}^{-t/50}.$$

As $t \to \infty$, $n \to 150$. Hence, the population will decrease before plateauing at 150 thousand.

The first model is more appropriate, as it account for the fact that the species will eventually go extinct ($n = 0$) due to the fact that they cannot replace itself in the long run.

* * * * *

Problem 11. A rectangular tank has a horizontal base. Water is flowing into the tank at a constant rate, and flows out at a rate which is proportional to the depth of water in the tank. At time t seconds, the depth of water in the tank is x metres. If the depth is 0.5 m, it remains at this constant value. Show that $\frac{\mathrm{d}x}{\mathrm{d}t} = -k(2x-1)$, where k is a positive constant. When $t = 0$, the depth of water in the tank is 0.75 m and is decreasing at a rate of 0.01 m s^{-1}. Find the time at which the depth of water is 0.55 m.

Solution. Let V_i m^3/s be the rate at which water is flowing into the tank. Note that $V_i \geq 0$. Let the rate at which water is flowing out of the tank be $V_o x$ m^3/s. Let the base of the container be A m^2. Then

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{V_i - V_o x}{A}.$$

At $x = 0.5$, the volume of water in the tank is constant. Thus,

$$\left.\frac{\mathrm{d}x}{\mathrm{d}t}\right|_{x=0.5} = 0 \implies V_i - 0.5V_o = 0 \implies V_o = 2V_i \implies \frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{V_i(2x-1)}{A}.$$

Letting $k = V_i/A$, we have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -k(2x-1).$$

We now solve for t.

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -k(2x-1) \implies \frac{1}{2x-1}\frac{\mathrm{d}x}{\mathrm{d}t} = -k \implies \int \frac{1}{2x-1}\,\mathrm{d}x = \int -k\,\mathrm{d}t$$
$$\implies \frac{\ln(2x-1)}{2} + C_1 = -kt \implies t = -\frac{\ln(2x-1)+C}{2k}, \quad C = 2C_1$$

Evaluating the solution at $t = 0$ and $x = 0.75$, we get

$$0 = -\frac{\ln(2(0.75)-1)+C}{2k} \implies C_2 = \ln 2.$$

Additionally,

$$\left.\frac{\mathrm{d}x}{\mathrm{d}t}\right|_{t=0} = -0.01 \implies -0.01 = -k[2(0.75)-1] \implies k = 0.02.$$

Thus,

$$t = -\frac{\ln(2x-1)+\ln 2}{2(0.02)} = -25\ln(4x-2)\,.$$

Consider $x = 0.55$. Then

$$t = -25\ln(4(0.55)-2) = 25\ln 5.$$

Thus, when $t = 25\ln 5$, the depth of the water is 0.55 m.

Problem 12. In a model of mortgage repayment, the sum of money owned to the Building Society is denoted by x and the time is denoted by t. Both x and t are taken to be continuous variables. The sum of money owned to the Building Society increases, due to interest, at a rate proportional to the sum of money owed. Money is also repaid at a constant rate r.

When $x = a$, interest and repayment balance. Show that, for $x > 0$, $\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{r}{a}(x - a)$.

Given that, when $t = 0$, $x = A$, find x in terms of t, r, a and A.

On a single, clearly labelled sketch, show the graph of x against t in the two cases:

(a) $A > a$.

(b) $A < a$.

State the circumstances under which the loan is repaid in a finite time T and show that, in this case, $T = \frac{a}{r}\ln\frac{a}{a-A}$.

Solution. Let the rate at which money is owned to the Building Society be kx. Then

$$\frac{\mathrm{d}x}{\mathrm{d}t} = kx - r.$$

At $x = a$, interest and repayment balance. Hence,

$$\mathrm{d}x/\mathrm{d}t\big|_a = ka - r = 0 \implies k = \frac{r}{a}.$$

Thus,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{r}{a}x - r = \frac{r}{a}(x - a).$$

We now solve for x.

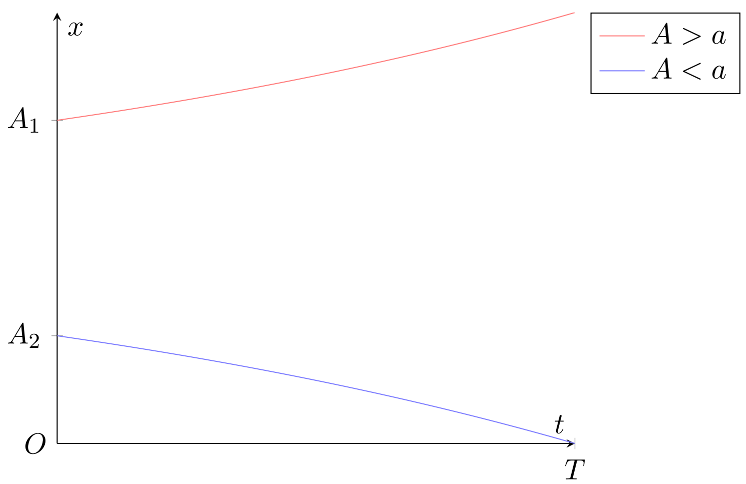
$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{r}{a}(x-a) \implies \frac{1}{x-a}\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{r}{a} \implies \int \frac{1}{x-a}\,\mathrm{d}x = \int \frac{r}{a}\,\mathrm{d}t$$
$$\implies \ln|x-a| = \frac{r}{a}t + C_1 \implies x = C\mathrm{e}^{rt/a} + a.$$

When $t = 0$, we have $x = A$. Substituting this into the solution, we obtain

$$A = C + a \implies C = A - a.$$

Thus,

$$x = (A-a)\mathrm{e}^{rt/a} + a.$$



For the loan to be repaid in finite time, $A < a$. At time T, the loan has been repaid, i.e. $x = 0$. Thus,

$$(A-a)\mathrm{e}^{rt/a} + a = 0 \implies \mathrm{e}^{rt/a} = \frac{a}{a-A} \implies \frac{rt}{a} = \ln\left(\frac{a}{a-A}\right) \implies t = \frac{a}{r}\ln\left(\frac{a}{a-A}\right).$$

Self-Practice B12

Problem 1. Show that the differential equation $x^2\frac{\mathrm{d}y}{\mathrm{d}x} - 2xy + 3 = 0$ may be reduced by means of the substitution $y = ux^2$ to $\frac{\mathrm{d}u}{\mathrm{d}x} = -\frac{3}{x^4}$. Hence, other otherwise, show that the general solution for y in terms of x is $y = Cx^2 + \frac{1}{x}$, where C is an arbitrary constant.

Solution. Since $y = ux^2$, by the chain rule, one has

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}x}x^2 + 2ux.$$

Substituting this into the given DE, we obtain

$$x^2\left(\frac{\mathrm{d}u}{\mathrm{d}x}x^2 + 2ux\right) - 2x\left(ux^2\right) + 3 = 0 \implies \frac{\mathrm{d}u}{\mathrm{d}x} = -\frac{3}{x^4}.$$

Integrating both sides with respect to x,

$$\int \mathrm{d}u = \int -\frac{3}{x^4}\,\mathrm{d}x \implies u = x^{-3} + C.$$

Since $u = y/x^2$, we have the general solution $y = Cx^2 + 1/x$.

* * * * *

Problem 2. Use the substitution $z = y\mathrm{e}^x$ to find the general solution of the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} + y = 2x + 3$. Sketch on one diagram, the curve of a particular solution for which $y \to \infty$ as $x \to -\infty$, labelling the equation of this particular solution.

Solution. Since $z = y\mathrm{e}^x$, by the chain rule, one has

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{e}^x + y\mathrm{e}^x \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}x}\mathrm{e}^{-x} - y.$$

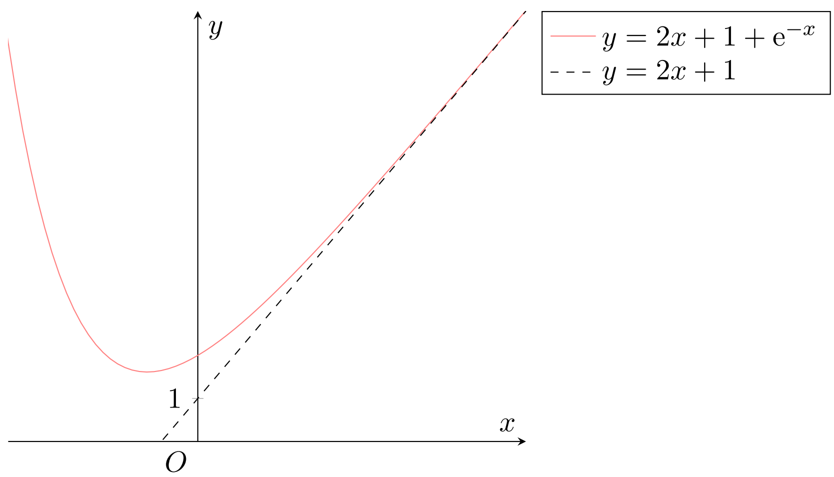
Substituting this into the given DE, we have

$$\left(\frac{\mathrm{d}z}{\mathrm{d}x}\mathrm{e}^{-x} - y\right) + y = 2x + 3 \implies \frac{\mathrm{d}z}{\mathrm{d}x} = (2x+3)\mathrm{e}^x.$$

Integrating both sides with respect to x,

$$\int \mathrm{d}z = \int (2x+3)\,\mathrm{e}^x\,\mathrm{d}x \implies y\mathrm{e}^x = z = (2x+1)\mathrm{e}^x + C.$$

Thus, $y = 2x + 1 + C\mathrm{e}^{-x}$.



Problem 3.

(a) Find the general solution of the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = (x+2)(y-3),$$

giving your answer in the form $y = f(x)$.

(b) Given that u and t are related by

$$\frac{\mathrm{d}u}{\mathrm{d}t} = 16 - 9u^2,$$

and that $u = 1$ when $t = 0$, find t in terms of u, simplifying your answer.

Solution.

Part (a). Manipulating the given DE, we have

$$\frac{1}{y-3}\frac{\mathrm{d}y}{\mathrm{d}x} = x+2 \implies \int \frac{1}{y-3}\,\mathrm{d}y = \int (x+2)\,\mathrm{d}x$$
$$\implies \ln|y-3| + A = \frac{1}{2}x^2 + 2x \implies B(y-3) = \mathrm{e}^{\frac{1}{2}x^2+2x} \implies y = C\mathrm{e}^{\frac{1}{2}x^2+2x} + 3.$$

Part (b). Manipulating the given DE, we have

$$\frac{1}{16-9u^2}\frac{\mathrm{d}u}{\mathrm{d}t} = 1 \implies \int \frac{1}{4^2-(3u)^2}\,\mathrm{d}u = \int \mathrm{d}t$$
$$\implies t = \frac{1}{3}\cdot\frac{1}{2(4)}\ln\left(\frac{4+3u}{4-3u}\right) + C = \frac{1}{24}\ln\left(\frac{4+3u}{4-3u}\right) + C.$$

At $t = 0$, $u = 1$. Hence,

$$0 = \frac{1}{24}\ln\left(\frac{4+3}{4-3}\right) + C \implies C = -\frac{1}{24}\ln 7.$$

Thus,

$$t = \frac{1}{24}\ln\left(\frac{4+3u}{4-3u}\right) - \frac{1}{24}\ln 7 = \frac{1}{24}\ln\left(\frac{4+3u}{7(4-3u)}\right).$$

* * * * *

Problem 4. At each instant of time the rate of increase of money in a savings account is proportional to the amount in the account at that instant. The constant of proportionality does not vary with time. Denote the amount in the account at time t years by \$$x$. When $x = 1000$, the rate of increase is \$50 per year. Obtain a differential equation relating x and t.

(a) Initially, when $t = 0$, the account contained \$900. Find the amount in the account exactly 3 years later.

(b) Find, in years correct to 2 places of decimals, the time when the account contains \$1800.

(c) Comment on whether the model can be regarded as a good model of the situation in the real world.

Solution. We have $\frac{\mathrm{d}x}{\mathrm{d}t} = kx$ for some $k \in \mathbb{R}^+$. Thus,

$$\left.\frac{\mathrm{d}x}{\mathrm{d}t}\right|_{x=1000} = 50 \implies k(10000) = 50 \implies k = \frac{1}{20}.$$

Hence,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{20}x.$$

Part (a). Solving for x, we get

$$\begin{aligned} &\frac{1}{x}\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{20} \implies \int \frac{1}{x}\,\mathrm{d}x = \int \frac{1}{20}\,\mathrm{d}t \\ \implies& \ln|x| + A = \frac{1}{20}t \implies Bx = \mathrm{e}^{t/20} \implies x = C\mathrm{e}^{t/20}. \end{aligned}$$

When $t = 0$, $x = 900$. Hence, $C = 900$ and

$$x = 900\mathrm{e}^{t/20}.$$

At $t = 3$,

$$x = 900\mathrm{e}^{3/20} = 1045.65 \text{ (2 d.p.)}.$$

Hence, there will be \$1045.65 in the account 3 years later.

Part (b). Let $x = 900\mathrm{e}^{t/20} = 1800$. Using G.C., $t = 13.86$ (2 d.p.). Hence, it will take 13.86 years for the account to contain \$1800.

Part (c). The model is not a good model of the situation in the real world as there is finite money in the world, but the model predicts that the amount of money in the bank account will grow forever.

* * * * *

Problem 5. Salt is dissolved in a tank filled with 120 litres of water. Salt water containing 20 g of salt per litre is poured in at a rate of 3 litres per minute and the mixture flows out at a constant rate of 3 litres per minute. The contents of the tank are kept well mixed at all times. Let the amount of salt in the tank (in grams) be denoted by S and the time (in minutes) be denoted by t.

(a) Show that $\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{2400-S}{40}$.

(b) Given that 400g of salt was dissolved in the tank initially, find the amount of salt in the tank after 1 hour, giving your answer to the nearest grams.

Solution.

Part (a). At any instant, the amount of salt entering the tank is $3(20/1)$ g, while the amount of salt leaving the tank is $3(S/120)$. Thus,

$$\frac{\mathrm{d}S}{\mathrm{d}t} = 3\left(\frac{20}{1}\right) - 3\left(\frac{S}{120}\right) = 60 - \frac{1}{40}S = \frac{2400 - S}{40}.$$

Part (b). Note that

$$\frac{\mathrm{d}S}{\mathrm{d}t} + \frac{1}{40}S = 60.$$

Multiplying through by $\mathrm{e}^{t/40}$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(S\mathrm{e}^{t/40}\right) = \frac{\mathrm{d}S}{\mathrm{d}t}\mathrm{e}^{t/40} + \frac{1}{40}S\mathrm{e}^{t/40} = 60\mathrm{e}^{t/40}.$$

Integrating both sides with respect to t,

$$Se^{t/40} = \int 60e^{t/40}\,dt = 2400e^{t/40} + C \implies S = 2400 + Ce^{-t/40}.$$

At $t = 0$, $S = 400$. Hence,

$$400 = 2400 + C \implies C = -2000.$$

At $t = 60$,

$$S = 2400 - 2000e^{60/40} = 1954.$$

Thus, there will be 1954 g of salt in the tank after 1 hour.

* * * * *

Problem 6. In a certain country, the price of a brand-new car of a particular make, manufactured on 1 January 1996, is \$32,000. According to a model of car pricing, the price P of the car (in \$) depreciates at a rate proportional to P when the car is t years old (as from 1 January 1996). Write down a differential equation relating P and t.

By solving this differential equation, show that $P = 32000e^{-kt}$ where k is a positive constant.

A man purchased a used car of this particular make for \$2000, at the price predicted by the model, on 1 January 2006. Subsequently on 1 January 2007, the man sold the used car for \$800. Determine if the man sold his car below the price predicted by the model.

Solution. We have

$$\frac{dP}{dt} = -kP,$$

where k is a positive real number. Solving, we have

$$\frac{1}{P}\frac{dP}{dt} = -k \implies \int \frac{1}{P}\,dP = -k\int dt \implies \ln P = -kt + C \implies P = Ce^{-kt}.$$

At $t = 0$, $P = 32000$. Hence, $C = 32000$, whence

$$P = 32000e^{-kt}.$$

When $t = 10$, $P = 2000$. Thus,

$$2000 = 32000e^{-10k} \implies k = \frac{1}{10}\ln 16.$$

Hence, at $t = 11$, the model predicts P to be

$$P = 32000e^{-11\left(\frac{1}{10}\ln 16\right)} = 1515 > 800.$$

Thus, the man sold his car below the price predicted by the model.

Assignment B12

Problem 1. The curve $y = f(x)$ passes through the origin and has gradient given by

$$\frac{dy}{dx} = \frac{3x^2 - 4x + 1}{2y - 5}.$$

(a) Find $f(x)$.

(b) By considering $\frac{dy}{dx}$, deduce the coordinates of the point on the curve where it is tangent to the x-axis.

(c) Determine the interval of validity for the solution.

Solution.

Part (a).

$$\frac{dy}{dx} = \frac{3x^2 - 4x + 1}{2y - 5} \implies (2y-5)\frac{dy}{dx} = 3x^2 - 4x + 1$$
$$\implies \int (2y-5)\,dy = \int \left(3x^2 - 4x + 1\right)\,dx \implies y^2 - 5y = x^3 - 2x^2 + x + C_1.$$

Note that $x^3 - 2x^2 + x = x(x-1)^2$. Hence,

$$y^2 - 5y - x(x-1)^2 + C_2 = 0 \implies y = \frac{5 \pm \sqrt{4x(x-1)^2 + C_3}}{2}.$$

Since the curve passes through the origin $(0, 0)$, we have

$$0 = \frac{5 - \sqrt{C_3}}{2} \implies C_3 = 25.$$

Thus,

$$f(x) = \frac{5 - \sqrt{4x(x-1)^2 + 25}}{2}.$$

Note that we reject the positive branch since $f(x) > 0$ in that case.

Part (b). When the curve is tangent to the x-axis, we have $\frac{dy}{dx} = 0$ and $y = 0$. Note that

$$\frac{dy}{dx} = 0 \implies 3x^2 - 4x + 1 = 0 \implies x = \frac{1}{3} \text{ or } 1.$$

Also note that

$$y = 0 \implies 4x(x-1)^2 = 0 \implies x = 0 \text{ or } 1.$$

Hence, the required point is $(1, 0)$.

Part (c). Since the square root function is defined only on the non-negative reals, we require

$$4x(x-1)^2 + 25 \geq 0 \implies x \geq -1.24.$$

Thus, the interval of validity is $[-1.24, \infty)$.

* * * * *

Problem 2.

(a) Using the substitution $y = ux$, find the general solution of the differential equation

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

where $x > 0$.

(b) Find the particular solution of the differential equation for which $y = -1$ when $x = 1$.

(c) Without sketching the curve of the solution in (b), determine the number of stationary points the solution curve has.

Solution.

Part (a). Note that

$$y = ux \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}x}x + u.$$

Substituting this into the given differential equation,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x+y}{x} \implies \frac{\mathrm{d}u}{\mathrm{d}x}x + u = \frac{x+ux}{x} \implies \frac{\mathrm{d}u}{\mathrm{d}x} = \frac{1}{x}$$
$$\implies u = \ln(x) + C \implies y = x\ln x + Cx.$$

Part (b). Evaluating the solution at $x = 1$ and $y = -1$, we get

$$-1 = 1\ln 1 + C(1) \implies C = -1.$$

Thus,

$$y = x\ln x - x.$$

Part (c). Note that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x+y}{x} = \frac{x + x\ln x - x}{x} = \ln x.$$

Since $\ln x$ only has one root (at $x = 1$), the solution curve has only 1 stationary point.

* * * * *

Problem 3. As a tree grows, the rate of increase of its height, h m, with respect to time, t years after planting, is modelled by the differential equation

$$\frac{\mathrm{d}h}{\mathrm{d}t} = \frac{1}{10}\sqrt{16 - \frac{1}{2}h}.$$

The tree is planted as a seedling of negligible height, so that $h = 0$ when $t = 0$.

(a) State the maximum height of the tree, according to this model.

(b) Find an expression for t in terms of h, and hence find the time the tree takes to reach half of its maximum height.

Solution.

Part (a). Note that $\frac{\mathrm{d}h}{\mathrm{d}t} \geq 0 \implies h \leq 32$. Thus, the maximum height of the tree is 32 m.

Part (b).

$$\frac{\mathrm{d}h}{\mathrm{d}t} = \frac{1}{10}\sqrt{16 - \frac{1}{2}h} \implies 10\left(16 - \frac{1}{2}h\right)^{-1/2}\frac{\mathrm{d}h}{\mathrm{d}t} = 1$$
$$\implies 10\int\left(16 - \frac{1}{2}h\right)^{-1/2}\mathrm{d}h = \int 1\,\mathrm{d}t \implies -10\sqrt{16 - \frac{1}{2}h} + C = t.$$

Since $h = 0$ when $t = 0$, we have

$$-10\sqrt{16} + C = 0 \implies C = 40.$$

Thus,

$$t = 40 - 10\sqrt{16 - \frac{1}{2}h}.$$

When $h = \frac{32}{2} = 16$, we have

$$t = 40 - 10\sqrt{16 - \frac{1}{2}(16)} = 11.7 \text{ (3 s.f.)}.$$

Thus, it takes 11.7 years for the tree to reach half its maximum height.

* * * * *

Problem 4.

(a) Find $\int \frac{1}{x(1000-x)} \, \mathrm{d}x$.

(b) A communicable disease is spreading within a small community with a population of 1000 people. A scientist found out that the rate at which the disease spreads is proportional to the product of the number of people who are infected with the disease and the number of people who are not infected with the disease. It is known that one person in this community is infected initially and five days later, 12% of the population is infected.

Given that the infected population is x at time t days after the start of the spread of the disease, show that it takes less than 8 days for half the population to contract the disease.

(c) State an assumption made by the scientist.

Solution.

Part (a).

$$\begin{aligned}\int \frac{1}{x(1000-x)} \, \mathrm{d}x &= \int \frac{1}{1000}\left(\frac{1}{x} + \frac{1}{1000-x}\right) \mathrm{d}x \\ &= \frac{\ln|x| - \ln|1000-x|}{1000} + C = \frac{1}{1000}\ln\left|\frac{x}{1000-x}\right| + C.\end{aligned}$$

Part (b). Note that $\frac{\mathrm{d}x}{\mathrm{d}t} \propto x(1000-x) \implies \frac{\mathrm{d}x}{\mathrm{d}t} = kx(1000-x)$ for some $k \in \mathbb{R}^+$.

$$\begin{aligned}&\frac{\mathrm{d}x}{\mathrm{d}t} = kx(1000-x) \implies \frac{1}{x(1000-x)}\frac{\mathrm{d}x}{\mathrm{d}t} = k \\ \implies &\int \frac{1}{x(1000-x)} \, \mathrm{d}x = \int k \, \mathrm{d}t \implies \frac{1}{1000}\ln\left(\frac{x}{1000-x}\right) + C = kt.\end{aligned}$$

Note that when $t = 0$, we have $x = 1$. Thus,

$$\frac{1}{1000}\ln\left(\frac{1}{999}\right) + C = 0 \implies C = \frac{\ln 999}{1000}.$$

When $t = 5$, $x = 120$. Hence,

$$\frac{1}{1000}\ln\left(\frac{120}{880}\right) + \frac{\ln 999}{1000} = 5k \implies k = \frac{1}{5000}\left(\ln\frac{3}{22} + \ln 999\right).$$

Thus,

$$\begin{aligned} t &= \left[\frac{1}{5000}\left(\ln\frac{3}{22} + \ln 999\right)\right]^{-1}\left[\frac{1}{1000}\ln\left(\frac{x}{1000-x}\right) + \frac{\ln 999}{1000}\right] \\ &= \frac{5}{\ln(3/22) + \ln 999}\left[\ln\left(\frac{x}{1000-x}\right) + \ln 999\right] \end{aligned}$$

Hence, when half the population is infected, i.e. $x = 500$, we have $t = 7.03 < 8$. Thus, it takes less than 8 days for half the population to contract the disease.

Part (c). The assumption is that there are no measures taken by the population to limit the spread of the disease (e.g. quarantine).

B13 Linear First Order Differential Equations

Tutorial B13

Problem 1. Solve the following differential equations:

(a) $2\sec x\frac{\mathrm{d}y}{\mathrm{d}x} = \sqrt{1-y^2}$

(b) $\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{t}{y-t^2y}$, given $y=4$ when $t=0$

Solution.

Part (a).

$$2\sec x\frac{\mathrm{d}y}{\mathrm{d}x} = \sqrt{1-y^2} \implies \frac{2}{\sqrt{1-y^2}}\frac{\mathrm{d}y}{\mathrm{d}x} = \cos x \implies \int \frac{2}{\sqrt{1-y^2}}\,\mathrm{d}y = \int \cos x\,\mathrm{d}x$$
$$\implies 2\arcsin y = \sin x + C_1 \implies y = \sin\left(\frac{1}{2}\sin x + C\right), \quad C = \frac{C_1}{2}.$$

Part (b).

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{t}{y-t^2y} \implies y\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{t}{1-t^2} \implies \int y\,\mathrm{d}y = \int \frac{t}{1-t^2}\,\mathrm{d}t$$
$$\implies \frac{1}{2}y^2 = -\frac{1}{2}\ln\left|1-t^2\right| + C_1 \implies y^2 = C_2 - \ln\left|1-t^2\right|, \quad C_2 = 2C_1.$$

Since $y = 4 \geq 0$ when $t = 0$, we have

$$4^2 = C_2 - \ln|1-0| \implies C_2 = 16.$$

Hence,

$$y = \sqrt{16 - \ln|1-t^2|}.$$

* * * * *

Problem 2. Solve the following differential equations:

(a) $xy' + (2x-3)y = 4x^4$

(b) $(1+x)y' + y = \cos x$, $y(0) = 1$

(c) $(1+t^2)\frac{\mathrm{d}y}{\mathrm{d}t} = 2ty + 2$

(d) $(x+1)\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{y}{\ln(x+1)} = x^2 + x$, where $x > 0$

Solution.

Part (a). Note that

$$xy' + (2x-3)y = 4x^4 \implies y' + \left(2 - \frac{3}{x}\right) y = 4x^3.$$

Hence, the integrating factor is

$$\text{I. F.} = \exp \int \left(2 - \frac{3}{x}\right) \mathrm{d}x = \exp(2x - \ln 3) = \frac{\mathrm{e}^{2x}}{x^3}.$$

Multiplying through by the integrating factor, we get

$$\frac{\mathrm{e}^{2x}}{x^3} y' + \frac{\mathrm{e}^{2x}}{x^3}\left(2 - \frac{3}{x}\right) y = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^{2x}}{x^3} y\right) = 4\mathrm{e}^{2x} \implies \frac{\mathrm{e}^{2x}}{x^3} y = \int 4\mathrm{e}^{2x}\,\mathrm{d}x = 2\mathrm{e}^{2x} + C$$

$$\implies y = \frac{x^3}{\mathrm{e}^{2x}}\left(2\mathrm{e}^{2x} + C\right) = 2x^3 + Cx^3\mathrm{e}^{-2x}.$$

Part (b).

$$(1+x)y' + y = \frac{\mathrm{d}}{\mathrm{d}x}[(1+x)y] = \cos x \implies (1+x)y = \int \cos x\,\mathrm{d}x = \sin x + C$$

$$\implies y = \frac{\sin x + C}{x+1}$$

Since $y(0) = 1$,

$$1 = \frac{\sin 0 + C}{0+1} \implies C = 1 \implies y = \frac{\sin x + 1}{x+1}.$$

Part (c). Let $t = \tan\theta$. Observe that

$$\frac{\mathrm{d}t}{\mathrm{d}\theta} = \sec^2\theta = 1 + t^2.$$

Hence,

$$(1+t^2)\frac{\mathrm{d}y}{\mathrm{d}t} = (1+t^2)\frac{\mathrm{d}y}{\mathrm{d}\theta}\cdot\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{1+t^2}{1+t^2}\cdot\frac{\mathrm{d}y}{\mathrm{d}\theta} = \frac{\mathrm{d}y}{\mathrm{d}\theta}.$$

Substituting this into the given differential equation,

$$\frac{\mathrm{d}y}{\mathrm{d}\theta} = 2y\tan\theta + 2 \implies \cos^2\theta\frac{\mathrm{d}y}{\mathrm{d}\theta} - 2y\sin\theta\cos\theta = 2\cos^2\theta \implies \frac{\mathrm{d}}{\mathrm{d}\theta}\left(y\cos^2\theta\right) = 2\cos^2\theta$$

$$\implies y\cos^2\theta = \int 2\cos^2\theta\,\mathrm{d}t = \int (1+\cos 2\theta)\,\mathrm{d}\theta = \theta + \frac{\sin 2\theta}{2} + C = \theta + \sin\theta\cos\theta + C$$

$$\implies y = (\theta + C)\sec^2\theta + \tan\theta = (\arctan t + C)\left(1+t^2\right) + t.$$

Part (d). Note that

$$(x+1)\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{y}{\ln(x+1)} = x^2 + x \implies \frac{\mathrm{d}y}{\mathrm{d}x} + \frac{y}{(x+1)\ln(x+1)} = x.$$

Hence, the integrating factor is

$$\text{I. F.} = \exp\int \frac{1/(x+1)}{\ln(x+1)}\,\mathrm{d}x = \exp\ln\ln(x+1) = \ln(x+1)\,.$$

Multiplying through by the integrating factor, we get

$$\ln(x+1)\,\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{y}{(x+1)} = \frac{\mathrm{d}}{\mathrm{d}x}\left(y\ln(x+1)\right) = x\ln(x+1)$$

$$\implies y\ln(x+1) = \int x\ln(x+1)\,\mathrm{d}x.$$

Integrating by parts, we get

$$\begin{aligned} y\ln(x+1) &= \frac{x^2}{2}\ln(x+1) - \frac{1}{2}\int \frac{x^2}{x+1}\,\mathrm{d}x \\ &= \frac{x^2}{2}\ln(x+1) - \frac{1}{2}\int \left(x - 1 + \frac{1}{x+1}\right)\mathrm{d}x \\ &= \frac{x^2}{2}\ln(x+1) - \frac{x^2}{4} + \frac{x}{2} - \frac{\ln(x+1)}{2} + C. \end{aligned}$$

Thus,

$$y = \frac{x^2}{2} - \frac{x^2}{4\ln(x+1)} + \frac{x}{2\ln(x+1)} - \frac{1}{2} + \frac{C}{\ln(x+1)}.$$

* * * * *

Problem 3. Given a general first order differential equation, $\frac{\mathrm{d}y}{\mathrm{d}x} = f(x,y)$, if $f(x,y)$ is such that $f(kx,ky) = f(x,y)$, then the equation may be reduced to a separable equation by means of the substitution $y = ux$. Hence, solve the following differential equation: $(x+y)y' = x - y$.

Solution. Note that

$$(x+y)y' = x - y \implies y' = \frac{x-y}{x+y}.$$

Let $f(x,y) = \frac{x-y}{x+y}$. Then

$$f(kx,ky) = \frac{kx-ky}{kx+ky} = \frac{x-y}{x+y} = f(x,y).$$

Hence, the differential equation can be solved with the substitution $y = ux$, whence $y' = u'x + u$. Substituting this into the differential equation, we get

$$\begin{aligned} u'x + u = \frac{x-ux}{x+ux} = \frac{1-u}{1+u} &\implies \left(\frac{1+u}{1-2u-u^2}\right)u' = \frac{1}{x} \\ \implies \int \frac{1+u}{1-2u-u^2}\,\mathrm{d}u &= \int \frac{1+u}{2-(1+u)^2}\,\mathrm{d}u = \int \frac{1}{x}\,\mathrm{d}x \\ \implies -\frac{1}{2}\ln\left|2-(1+u)^2\right| = -\frac{1}{2}\ln\left|2-\left(1+\frac{y}{x}\right)^2\right| &= \ln x + C_1 \implies (x+y)^2 = 2x^2 + C, \end{aligned}$$

where $C = -\mathrm{e}^{-2C_1} \in \mathbb{R}^-$.

* * * * *

Problem 4. Using the substitution $u = \frac{1}{y}$, solve $\frac{\mathrm{d}y}{\mathrm{d}x} + 2y = e^x y^2$.

Solution. Note that

$$u = \frac{1}{y} \implies y = \frac{1}{u} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{1}{u^2}\cdot\frac{\mathrm{d}u}{\mathrm{d}x}.$$

Substituting this into the differential equation,

$$\begin{aligned} -\frac{1}{u^2}\cdot\frac{\mathrm{d}u}{\mathrm{d}x} + \frac{2}{u} = \frac{\mathrm{e}^x}{u^2} &\implies \frac{\mathrm{d}u}{\mathrm{d}x} - 2u = -\mathrm{e}^x \implies \mathrm{e}^{-2x}\frac{\mathrm{d}u}{\mathrm{d}x} - 2\mathrm{e}^{-2x}u = \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-2x}u\right) = -\mathrm{e}^{-x} \\ \implies \mathrm{e}^{-2x}u = \int -\mathrm{e}^{-x}\,\mathrm{d}x = \mathrm{e}^{-x} + C &\implies u = \mathrm{e}^x + C\mathrm{e}^{2x} \implies y = \frac{1}{\mathrm{e}^x + C\mathrm{e}^{2x}}. \end{aligned}$$

* * * * *

Problem 5. Assuming that $p(x) \neq 0$, state conditions under which the linear equation $y' + p(x)y = f(x)$ is separable. If the equation satisfies these conditions, solve it by separation of variables and by the method of integrating factor.

Solution. The linear equation $y' + p(x)y = f(x)$ is separable if $p(x)$ is a scalar multiple of $f(x)$, i.e. $p(x) = \lambda f(x)$ for some $\lambda \in \mathbb{R}$.

We begin by solving using separation of variables. Note that

$$y' = f(x) - p(x)y = f(x) - \lambda f(x)y = f(x)(1 - \lambda y).$$

Hence,

$$\begin{aligned}\frac{1}{1-\lambda y}y' = f(x) &\implies \int \frac{1}{1-\lambda y}\,\mathrm{d}y = -\frac{\ln|1-\lambda y|}{\lambda} = \int f(x)\,\mathrm{d}x \\ &\implies \ln|1-\lambda y| = -\int \lambda f(x)\,\mathrm{d}x = -\int p(x)\,\mathrm{d}x \\ &\implies y = \frac{1}{\lambda}\left[1 - C_1 \mathrm{e}^{-\int p(x)\,\mathrm{d}x}\right] = \frac{1}{\lambda} + C\mathrm{e}^{-\int p(x)\,\mathrm{d}x}.\end{aligned}$$

Integrating Factor. Note that the integrating factor is $\mathrm{e}^{\int p(x)\,\mathrm{d}x}$. Multiplying through, we get

$$\begin{aligned}&\mathrm{e}^{\int p(x)\,\mathrm{d}x}y' + \mathrm{e}^{\int p(x)\,\mathrm{d}x}p(x)y = \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{\int p(x)\,\mathrm{d}x}y\right) = \mathrm{e}^{\int p(x)\,\mathrm{d}x}f(x) = \frac{1}{\lambda}e^{\int p(x)\,\mathrm{d}x}p(x) \\ \implies &\mathrm{e}^{\int p(x)\,\mathrm{d}x}y = \int \frac{1}{\lambda}\mathrm{e}^{\int p(x)\,\mathrm{d}x}p(x)\,\mathrm{d}x = \frac{1}{\lambda}\mathrm{e}^{\int p(x)\,\mathrm{d}x} + C \implies y = \frac{1}{\lambda} + Ce^{-\int p(x)\,\mathrm{d}x}.\end{aligned}$$

* * * * *

Problem 6. The variables x and y are related by the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{y}{x} = y^3$.

(a) State clearly why the integrating factor method cannot be used to solve this equation.

(b) The variables y and z are related by the equation $\frac{1}{y^2} = -2z$. Show that $\frac{\mathrm{d}z}{\mathrm{d}x} - \frac{2z}{x} = 1$.

(c) Find the solution of the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{y}{x} = y^3$, given that $y = 2$ when $x = 1$.

Solution.

Part (a). The differential is non-linear due to the presence of the y^3 term.

Part (b).

$$\frac{1}{y^2} = -2z \implies \frac{\mathrm{d}z}{\mathrm{d}x} = \frac{1}{y^3}\cdot\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{y^3}\left(y^3 - \frac{y}{x}\right) = 1 - \frac{1}{y^2}\cdot\frac{1}{x} = 1 + \frac{2z}{x} \implies \frac{\mathrm{d}z}{\mathrm{d}x} - \frac{2z}{x} = 1.$$

Part (c).

$$\begin{aligned}\frac{\mathrm{d}z}{\mathrm{d}x} - \frac{2z}{x} = 1 &\implies \frac{1}{x^2}\frac{\mathrm{d}z}{\mathrm{d}x} - \frac{2z}{x^3} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{z}{x^2}\right) = \frac{1}{x^2} \implies \frac{z}{x^2} = \int \frac{1}{x^2}\,\mathrm{d}x = -\frac{1}{x} + C_1 \\ &\implies z = -\frac{1}{2y^2} = -x + C_1x^2 \implies y^2 = \frac{1}{2x + C_2x^2}.\end{aligned}$$

Since $y(1) = 2$, we have

$$2^2 = \frac{1}{2 + C_2} \implies C_2 = -\frac{7}{4}.$$

Thus,

$$y^2 = \frac{1}{2x - 7x^2/4} = \frac{4}{8x - 7x^2} \implies y = \frac{2}{\sqrt{8x - 7x^2}}.$$

Note that we reject the negative branch since $y(1) = 2 \geq 0$.

* * * * *

Problem 7. Show that the substitution $v = \ln y$ transforms the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x)(y\ln y)$ into the linear equation $\frac{\mathrm{d}v}{\mathrm{d}x} + P(x) = Q(x)v(x)$. Hence, solve the equation $x\frac{\mathrm{d}y}{\mathrm{d}x} - 4x^2y + 2y\ln y = 0$.

Solution. Note that

$$v = \ln y \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{y} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = y\frac{\mathrm{d}v}{\mathrm{d}x}.$$

Substituting this into the differential equation, we get

$$\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x)(y \ln y) \implies y\frac{\mathrm{d}v}{\mathrm{d}x} + P(x)y = Q(x)(yv) \implies \frac{\mathrm{d}v}{\mathrm{d}x} + P(x) = Q(x)v.$$

We now solve the given differential equation. Note that

$$x\frac{\mathrm{d}y}{\mathrm{d}x} - 4x^2y + 2y\ln y = 0 \implies \frac{\mathrm{d}y}{\mathrm{d}x} - 4xy = -\frac{2}{x}(y \ln y).$$

Hence,

$$P(x) = -4x, \quad Q(x) = -\frac{2}{x}.$$

Now, from the previous part, we can rewrite the differential equation as

$$\frac{\mathrm{d}v}{\mathrm{d}x} - 4x = -\frac{2v}{x},$$

where $v(x) = \ln y$. Thus,

$$x^2\frac{\mathrm{d}v}{\mathrm{d}x} + 2xv = \frac{\mathrm{d}}{\mathrm{d}x}\left(x^2v\right) = 4x^3 \implies x^2v = x^2\ln y = \int 4x^3 \,\mathrm{d}x = x^4 + C$$
$$\implies y = \exp\left(x^2 + Cx^{-2}\right).$$

* * * * *

Problem 8. The normal at any point on a certain curve always passes through the point $(2, 3)$. Form a differential equation to express this property. Without solving the differential equation, find the equation of the curve where the stationary points of the family of curves will lie on. Which family of standard curves will have their stationary points lying along a curve with such an equation found earlier?

Solution. Clearly,

$$y - 3 = \frac{-1}{\mathrm{d}y/\mathrm{d}x}(x - 2).$$

Note that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{x-2}{y-3}.$$

For stationary points, $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$, whence $x = 2$ and $y \neq 3$.

Also note that

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -\frac{(y-3) - (x-2)\frac{\mathrm{d}y}{\mathrm{d}x}}{(y-3)^2}.$$

At stationary points, $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$, giving

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -\frac{1}{y-3}.$$

Hence, when $y > 3$, we have $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} < 0$, giving a maximum. Likewise, when $y < 3$, we have $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} > 0$, giving a minimum. This suggests that the required family of standard curves is the family of circles with centre $(2, 3)$.

Problem 9. Obtain the general solution of the differential equation $x\frac{\mathrm{d}y}{\mathrm{d}x} - y = x^2 + 1$ in the form $y = x^2 + Cx - 1$, where C is an arbitrary constant.

Show that each solution curve of the differential equation has one minimum point.

Find the equation of the curve of which all these minimum points lie.

Sketch some of the family of solution curves including those corresponding to some negative values of C, some positive values of C, and $C = 0$.

Solution.

$$x\frac{\mathrm{d}y}{\mathrm{d}x} - y = x^2 + 1 \implies \frac{1}{x}\frac{\mathrm{d}y}{\mathrm{d}x} - \frac{y}{x^2} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{y}{x} = 1 + \frac{1}{x^2}$$
$$\implies \frac{y}{x} = \int \left(1 + \frac{1}{x^2}\right)\,\mathrm{d}x = x - \frac{1}{x} + C \implies y = x^2 + Cx - 1.$$

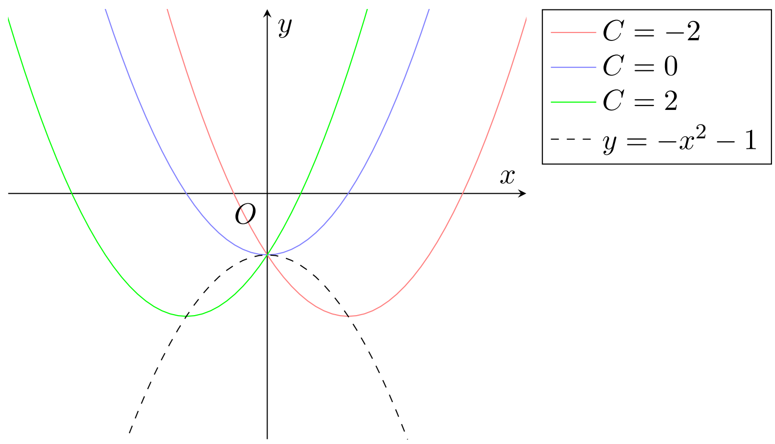
Note that

$$y = x^2 + Cx - 1 = \left(x + \frac{C}{2}\right)^2 - \left(1 + \frac{C^2}{4}\right).$$

Thus, y has a unique minimum point at $\left(-\frac{C}{2}, -\left(1 + \frac{C^2}{4}\right)\right)$.

For stationary points, $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$. Thus, the minimum points lie on the curve with equation

$$-y = x^2 + 1 \implies y = -x^2 - 1.$$



* * * * *

Problem 10. Show that the general solution of the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} + 2xy - 2x\left(x^2 + 1\right) = 0$$

can be expressed in the form $y = x^2 + Ce^{-x^2}$, where C is an arbitrary constant.

Deduce, with reasons, the number of stationary points of the solution curves of the equation when

(a) $C \leq 1$;

(b) $C > 1$.

Solution. Note that the integrating factor is $\exp\left(\int 2x\,\mathrm{d}x\right) = \mathrm{e}^{x^2}$. Multiplying the integrating factor throughout the differential equation, we get

$$\mathrm{e}^{x^2}\frac{\mathrm{d}y}{\mathrm{d}x} + 2x\mathrm{e}^{x^2}y = \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{x^2}y\right) = 2x\mathrm{e}^{x^2}\left(x^2 + 1\right)$$
$$\implies \mathrm{e}^{x^2}y = \int 2x\mathrm{e}^{x^2}\left(x^2 + 1\right)\,\mathrm{d}x = \mathrm{e}^{x^2}(x^2 + 1) - \int 2x\mathrm{e}^{x^2}\,\mathrm{d}x = \mathrm{e}^{x^2}(x^2 + 1) - \mathrm{e}^{x^2} + C$$
$$\implies y = \left(x^2 + 1\right) - 1 + C\mathrm{e}^{-x^2} = x^2 + C\mathrm{e}^{-x^2}.$$

For stationary points, $\frac{dy}{dx} = 0$. Hence,

$$2xy - 2x\left(x^2+1\right) = 0 \implies x = 0 \text{ or } y - \left(x^2+1\right) = 0.$$

Consider $y - \left(x^2+1\right) = 0$.

$$y - \left(x^2+1\right) = \left(x^2 + C\mathrm{e}^{-x^2}\right) - \left(x^2+1\right) = C\mathrm{e}^{-x^2} - 1 = 0 \implies x^2 = \ln C.$$

Part (a). When $C < 1$, we have $\ln C < 0$. Hence, there are no solutions to $x^2 = \ln C$, whence there is only 1 stationary point (at $x = 0$).

When $C = 1$, we have $\ln C = 0$, whence the only solution to $x^2 = \ln C$ is $x = 0$. Thus, there is still only 1 stationary point (at $x = 0$).

Part (b). When $C > 1$, we have $\ln C > 0$, whence there are two solutions to $x^2 = \ln C$, namely $x = \pm \ln C$. Thus, there are 3 stationary points (at $x = 0$ and $x = \pm \ln C$).

* * * * *

Problem 11. Using the substitution $y = x^2 \ln t$, where $t > 0$, show that the differential equation

$$2xt \ln t \frac{\mathrm{d}x}{\mathrm{d}t} + (3\ln t + 1)x^2 = \frac{\mathrm{e}^{-2t}}{t} \tag{$*$}$$

can be reduced to a differential equation of the form

$$\frac{\mathrm{d}y}{\mathrm{d}t} + P(t)y = \frac{\mathrm{e}^{-2t}}{t^2},$$

where $P(t)$ is some function of t to be determined.

Hence, find x^2 in terms of t.

Sketch, on a single diagram, two solution curves for the differential equation $(*)$, C_1 and C_2, of which only C_1 has stationary point(s). Label the equations of any asymptotes in your diagram.

Solution. Note that

$$y = x^2 \ln t \implies \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{x^2}{t} + 2x\ln t\frac{\mathrm{d}x}{\mathrm{d}t} \implies 2xt\ln t\frac{\mathrm{d}x}{\mathrm{d}t} = t\frac{\mathrm{d}y}{\mathrm{d}t} - x^2.$$

Hence,

$$2xt\ln t\frac{\mathrm{d}x}{\mathrm{d}t} + (3\ln t + 1)x^2 = \left(t\frac{\mathrm{d}y}{\mathrm{d}t} - x^2\right) + (3\ln t + 1)x^2 = t\frac{\mathrm{d}y}{\mathrm{d}t} + 3x^2\ln t = t\frac{\mathrm{d}y}{\mathrm{d}t} + 3y.$$

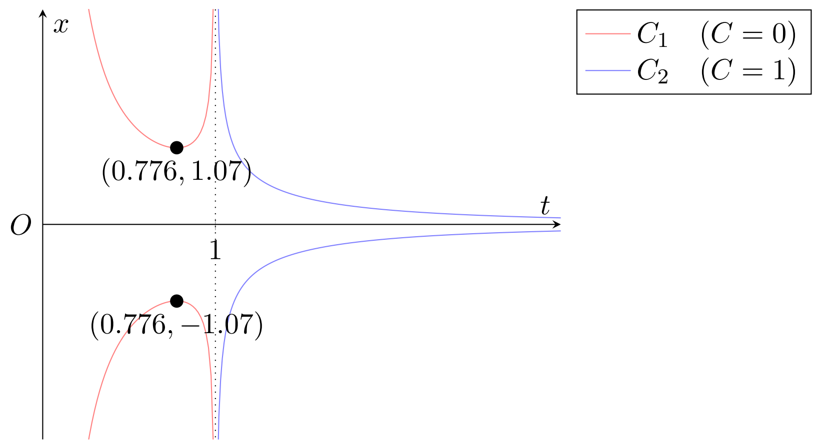
Our differential equation thus becomes

$$t\frac{\mathrm{d}y}{\mathrm{d}t} + 3y = \frac{\mathrm{e}^{-2t}}{t} \implies \frac{\mathrm{d}y}{\mathrm{d}t} + \frac{3y}{t} = \frac{\mathrm{e}^{-2t}}{t^2},$$

whence $P(t) = 3/t$. We now solve the differential equation. Observe that

$$t^3\frac{\mathrm{d}y}{\mathrm{d}t} + 3t^2y = \frac{\mathrm{d}}{\mathrm{d}t}\left(t^3y\right) = t\mathrm{e}^{-2t} \implies t^3y = \int t\mathrm{e}^{-2t}\,\mathrm{d}t = -\frac{1}{2}t\mathrm{e}^{-2t} - \frac{1}{4}\mathrm{e}^{-2t} + C_1$$

$$\implies t^3x^2\ln t = -\frac{(2t+1)\,\mathrm{e}^{-2t}}{4} + \frac{C}{4} \implies x^2 = \frac{C - (2t+1)\,\mathrm{e}^{-2t}}{4t^3\ln t}.$$



* * * * *

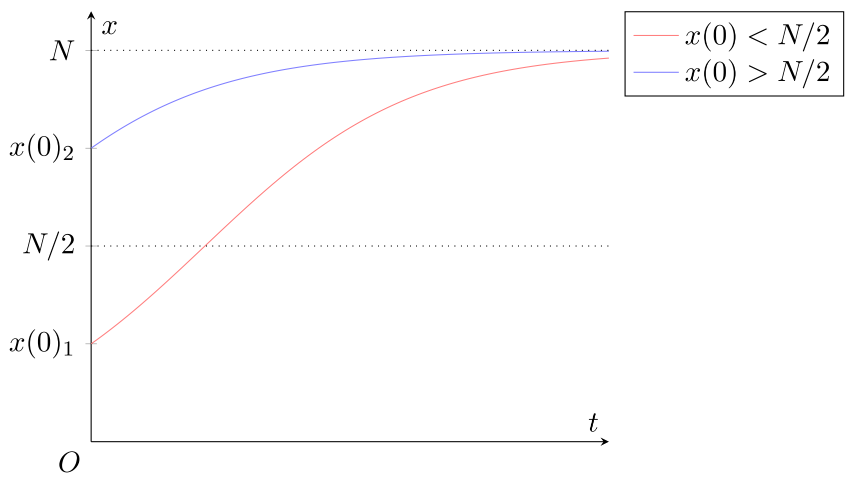
Problem 12. It is suggested that the spread of a highly contagious disease on an isolated island with a population of N may be modelled by the differential equation $\frac{\mathrm{d}x}{\mathrm{d}t} = kx(N-x)$, where k is a positive constant, and $x(t)$ is the number of individuals infected with the disease at time t.

(a) Without solving the differential equation, sketch the graph of $x(t)$ against t for cases when $x(0) < \frac{N}{2}$ and $x(0) > \frac{N}{2}$.

(b) Given that $x(0) = x_0$, solve the differential equation for an explicit expression of $x(t)$.

Solution.

Part (a).



Part (b).

$$\frac{\mathrm{d}x}{\mathrm{d}t} = kx(N-x) \implies \frac{1}{x(N-x)}\frac{\mathrm{d}x}{\mathrm{d}t} = k \implies \int \frac{1}{x(N-x)}\,\mathrm{d}x = \int k\,\mathrm{d}t$$
$$\implies \int \frac{1}{N}\left(\frac{1}{x} - \frac{1}{N-x}\right) = \int k\,\mathrm{d}t \implies \frac{\ln x - \ln(N-x)}{N} = kt + C_1$$
$$\implies \ln\frac{x}{N-x} = Nkt + C_2 \implies x = \frac{C_3 N\mathrm{e}^{Nkt}}{1 + C_3\mathrm{e}^{Nkt}}.$$

At $t = 0$, we have $x = x_0$. Hence,

$$x_0 = \frac{C_3 N \mathrm{e}^0}{1 + C_3 \mathrm{e}^0} \implies C_3 = \frac{x_0}{N - x_0}.$$

This gives

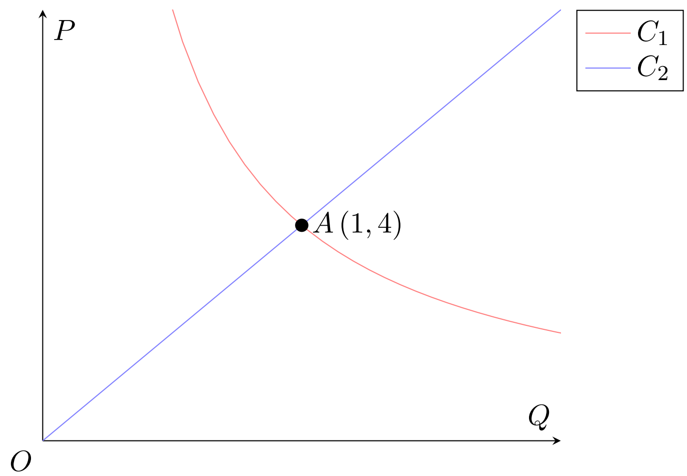
$$x = \frac{\frac{x_0}{N-x_0} N \mathrm{e}^{Nkt}}{1 + \frac{x_0}{N-x_0} \mathrm{e}^{Nkt}} = \frac{N x_0 \mathrm{e}^{Nkt}}{N - x_0 + x_0 \mathrm{e}^{Nkt}}.$$

* * * * *

Problem 13. In the diagram below, the curve C_1 and the line C_2 illustrate the relationship between price (P dollars per kg) and quantity (Q tonnes) for consumers and producers respectively.

The curve C_1 shows the quantity of rice that consumers will buy at each price level while the line C_2 shows the quantity of rice that producers will produce at each price level. C_1 and C_2 intersect at point A, which has the coordinates $(1, 4)$.

The quantity of rice that consumers will buy is inversely proportional to the price of the rice. The quantity of rice that producers will produce is directly proportional to the price.



(a) Interpret the coordinates of A in the context of the question.

(b) Solve for the equations of C_1 and C_2, expressing Q in terms of P.

Shortage occurs when the quantity of rice consumers will buy exceeds the quantity of rice producers will produce. It is known that the rate of increase of P after time t months is directly proportional to the quantity of rice in shortage.

(c) Given that the initial price is \$3 and that after 1 month, the price is \$3.65, find P in terms of t and sketch this solution curve, showing the long-term behaviour of P.

Suggest a reason why producers might use $P = aQ + b$, where $a, b \in \mathbb{R}^+$, instead of C_2 to model the relationship between price and quantity of rice produced.

Solution.

Part (a). The coordinates of A represent the equilibrium price and quantity of rice. That is, 1 tonne of rice will be transacted at a price of \$4 per kg.

Part (b). Note that $C_1 : P = \frac{k_1}{Q}$ and $C_2 : P = k_2 Q$ for some constants k_1 and k_2. At $A(1, 4)$, we obtain $k_1 = k_2 = 4$. Thus,

$$C_1 : Q = \frac{4}{P}, \quad C_2 : Q = \frac{P}{4}.$$

Part (c). At a given price $P < 4$, the difference in the amount of rice demanded and produced is given by $\frac{4}{P} - \frac{P}{4} = \frac{16-P^2}{4P}$. Hence, $\frac{\mathrm{d}P}{\mathrm{d}t} = k \cdot \frac{16-P^2}{4P}$.

$$\frac{\mathrm{d}P}{\mathrm{d}t} = k \cdot \frac{16 - P^2}{4P} \implies \frac{2P}{16 - P^2}\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{k}{2} \implies \int \frac{2P}{16 - P^2}\,\mathrm{d}P = \int \frac{k}{2}\,\mathrm{d}t$$
$$\implies -\ln\left(16 - P^2\right) = \frac{kt}{2} + C_1 \implies P = \sqrt{16 - C_2\mathrm{e}^{-kt/2}}.$$

Note that we used the fact that $0 < P < 4$ when solving for P.

At $t = 0$, $P = 3$. Hence,

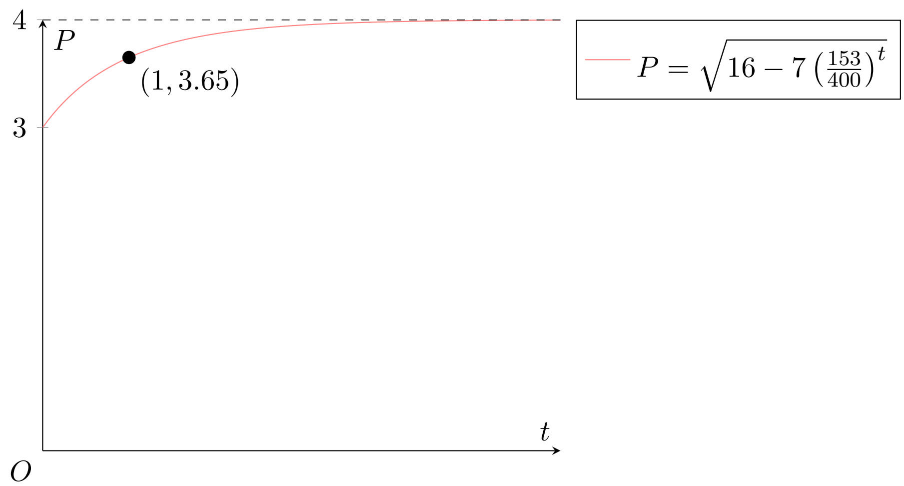
$$3 = \sqrt{16 - C_2\mathrm{e}^0} \implies C_2 = 7.$$

At $t = 1$, $P = 3.65$. Hence,

$$3.65 = \sqrt{16 - 7\mathrm{e}^{-k(1)/2}} \implies \mathrm{e}^{-k/2} = \frac{153}{400}.$$

Thus,

$$P = \sqrt{16 - 7\left(\frac{153}{400}\right)^t}.$$



The model $P = aQ + b$ accounts for the fixed cost involved in producing rice.

* * * * *

Problem 14. A rectangular tank contains 100 litres of salt solution at a concentration of 0.01 kg/litre. A salt solution with a concentration of 0.5 kg/litre flows into the tank at the rate of 6 litres/min. The mixture in the tank is kept uniform by stirring the mixture and the mixture flows out at the rate of 4 litres/min. If y kg is the mass of salt in the solution in the tank after t minutes, show that y satisfies the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}t} = 3 - \frac{ky}{100 + mt},$$

where k and m are constants to be determined.

Find the particular solution of the differential equation.

Solution. Note that

$$\frac{\mathrm{d}y}{\mathrm{d}t} = 0.5(6) - 4\left(\frac{y}{100+(6-4)t}\right) = 3 - \frac{4y}{100+2t}.$$

Hence, $k = 4$ and $m = 2$.

We now solve the differential equation. Multiplying throughout by $(100+2t)^2$, we get

$$\begin{aligned} 2(100+2t)^2\frac{\mathrm{d}y}{\mathrm{d}t} + 4(100+2t)y &= \frac{\mathrm{d}}{\mathrm{d}t}\left[(100+2t)^2 y\right] = 3(100+2t)^2 \\ \implies (100+2t)^2 y &= \int 3(100+2t)^2\,\mathrm{d}t = \frac{1}{2}(100+2t)^3 + C_1 = 4(50+t)^3 + C_1 \\ \implies y &= \frac{4(50+t)^3 + C_1}{(100+2t)^2} = \frac{4(50+t)^3 + C_1}{4(50+t)^2} = 50 + t + \frac{C}{(50+t)^2}. \end{aligned}$$

At $t = 0$, $y = 100(0.01) = 1$. Hence,

$$1 = 50 + 0 + \frac{C}{(50+0)^2} \implies C = -122500.$$

Thus,

$$y = 50 + t - \frac{122500}{(50+t)^2}.$$

* * * * *

Problem 15. A first order differential equation of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = q(x)y^n, \quad n \neq 0, 1$$

is called a Bernoulli equation. Show that the substitution $u = y^{1-n}$ reduces the Bernoulli equation into the linear equation $\frac{\mathrm{d}u}{\mathrm{d}x} + (1-n)p(x)u(x) = (1-n)q(x)$.

A cardiac pacemaker is designed to provide electrical impulses I amps such that as time t increases, I oscillates with a fixed amplitude of one amp. It is proposed that the following differential equation $\frac{\mathrm{d}I}{\mathrm{d}t} + (\tan t)I = (I \sin t)^2$ can be used to describe how I changes with t.

By using a substitution of the form $u = I^{1-n}$, find I in terms of t.

State one limitation of this model.

Solution. Note that

$$u = y^{1-n} \implies \frac{\mathrm{d}u}{\mathrm{d}x} = (1-n)y^n\frac{\mathrm{d}y}{\mathrm{d}x} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}x}\cdot\frac{y^n}{1-n}.$$

Substituting this into the given differential equation, we get

$$\begin{aligned} \frac{\mathrm{d}u}{\mathrm{d}x}\cdot\frac{y^n}{1-n} + p(x)y = q(x)y^n &\implies \frac{\mathrm{d}u}{\mathrm{d}x}\cdot\frac{1}{1-n} + p(x)u = q(x) \\ &\implies \frac{\mathrm{d}u}{\mathrm{d}x} + (1-n)p(x)u = (1-n)q(x) \end{aligned}$$

Let $n = 2$. Then $u = I^{-1}$. We also have $p(x) = \tan t$ and $q(x) = \sin^2 t$.

$$\begin{aligned} &\frac{\mathrm{d}I}{\mathrm{d}t} + (\tan t)I = (I\sin t)^2 \implies \frac{\mathrm{d}u}{\mathrm{d}t} - (\tan t)u = -\sin^2 t \\ \implies &\cos t\frac{\mathrm{d}u}{\mathrm{d}t} - (\sin t)u = -\cos t\sin^2 t \implies \frac{\mathrm{d}}{\mathrm{d}t}(u\cos t) = -\cos t\sin^2 t \\ \implies &u\cos t = \int -\cos t\sin^2\theta\,\mathrm{d}t = -\frac{1}{3}\sin^3 t + C \\ \implies &u = \frac{-1/3\cdot\sin^3 t + C}{\cos t} \implies I = \frac{\cos t}{-1/3\cdot\sin^3 t + C} = \frac{3\cos t}{3C - \sin^3 t} \end{aligned}$$

Consider the stationary points of I. For stationary points, we have $\frac{\mathrm{d}I}{\mathrm{d}t} = 0$. Hence,

$$\frac{\sin t}{\cos t} I = I^2 \sin^2 t \implies I \sin t \left(I \sin t - \frac{1}{\cos t} \right) = 0.$$

Hence, $\sin t = 0$ or $I \sin t - \frac{1}{\cos t} = 0$. Note that $I = \frac{3 \cos t}{3C - \sin^3 t} \neq 0$ since $\cos t \neq 0$. We now consider the latter case.

$$I \sin t - \frac{1}{\cos t} = 0 \implies I \sin t \cos t = 1 \implies I \sin 2t = 2.$$

Since I has an amplitude of 1, we have that $I \in [-1, 1]$. Since $\sin 2t \in [-1, 1]$, we have that $I \sin 2t \in [-1, 1]$. Thus, $I \sin 2t$ can never be 2. Hence, stationary points only occur when $\sin t = 0$, implying $t = k\pi$. Thus, $I(0) = \pm 1$. This gives

$$I(0) = \frac{3 \cos 0}{3C - \sin^3 0} = 1 \implies C = 1,$$

whence

$$I = \frac{3 \cos t}{3 - \sin^3 t}.$$

A limitation of this model is that it does not reflect the fact that the oscillations may gradually get weaker.

Self-Practice B13

Problem 1. Food energy taken in by a man goes partly to maintain the healthy functioning of his body and partly to increase body mass. The total food energy intake of the man per day is assumed to be a constant denoted by I (in joules). The food energy required to maintain the healthy functioning of his body is proportional to his body mass M (in kg). The increase of M with respect to time t (in days) is proportional to the energy not used by his body. If the man does not eat for one day, his body mass will be reduced by 1%.

(a) Show that I, M and t are related by the following differential equation:

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \frac{I - aM}{100a},$$

where a is a constant. State an assumption for this model to be valid.

(b) Find the total food energy intake per day, I, of the man in terms of a and M if he wants to maintain a constant body mass.

It is given that the man's initial mass is 100 kg.

(c) Solve the differential equation in part (a), giving M in terms of I, a and t.

(d) Sketch the graph of M against t for the case where $I > 100a$. Interpret the shape of the graph with regard to the man's food energy intake.

(e) If the man's total food energy intake per day is $50a$, find the time taken in days for the man to reduce his body mass from 100 kg to 90 kg.

* * * * *

Problem 2. Find the general solution of the differential equation

$$x\frac{\mathrm{d}y}{\mathrm{d}x} - 3y = x^5 \mathrm{e}^{2x}.$$

Sketch the family of solution curves, showing clearly all the essential features sufficiently.

* * * * *

Problem 3. Let the variables x and y be related by the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{y}{x} = xy^n,$$

where n is a real number. Find the general solution for y in terms of x for the following cases:

(a) $n = 0$;

(b) $n = 1$;

(c) $n \geq 2$, using the substitution $u = y^{1-n}$.

* * * * *

Problem 4. *Orthogonal trajectories* are a family of curves that intersect another family of curves perpendicularly.

The electrostatic field created by a single positive charge is a collection of straight lines that radiate away from the charge. Equipotential lines are where the electric potentials are equal on a 2-dimensional surface (**these lines can be curves**). It is given that the equipotential lines are orthogonal trajectories of the electric field lines.

(a) By forming a differential equation satisfied by equipotential lines and solving it, show that the equipotential line of a point charge forms a family of circles with centre at the origin, taking the point charge to be at the origin.

When a point charge is placed at $(0, h_1)$, there is an equipotential line tangential to the x-axis. The collection of these equipotential lines for all $h_1 \in \mathbb{R}$, $h_1 \neq 0$ forms a family of circles denoted by C.

(b) By first writing the Cartesian equation of a circle tangential to the x-axis and with centre $(0, h_1)$, show that the orthogonal trajectories of the family of circles C satisfy

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y^2 - x^2}{2xy}.$$

Hence, by using the substitution $Y = y^2$ show that the orthogonal trajectories of the family of circles C, form a family of circles that are tangential to the y-axis at the origin.

Assignment B13

Problem 1. Two biological cultures, X and Y, react with each other, and their volumes at time t are x and y respectively, in appropriate units. Their rates of growth are modelled by the simultaneous equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = (2-x)y,$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{y^2}{x}$$

When $t = 0$, $x = y = 1$.

(a) Show that $x = \frac{2y^2}{1+y^2}$.

(b) Find and simplify expressions for y and x in terms of t.

(c) Sketch the graph of y against x for $0 < t < \frac{\pi}{2}$.

Solution.

Part (a). Note that $x, y > 0$ since they represent volume. Also, for $x \in (0, 2)$, we have $\frac{\mathrm{d}x}{\mathrm{d}t} = (2-x)y > 0$. When $x = 2$, we have $\frac{\mathrm{d}x}{\mathrm{d}t} = 0$. Hence, $0 < x \leq 2$. Now observe that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{y^2/x}{(2-x)y} = \frac{y}{x(2-x)} \implies \frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{x(2-x)}.$$

Integrating both sides with respect to x, we get

$$\implies \int \frac{1}{y}\,\mathrm{d}y = \int \frac{1}{x(2-x)}\,\mathrm{d}x = \frac{1}{2}\int \left(\frac{1}{x} + \frac{1}{2-x}\right)\mathrm{d}x$$
$$\implies \ln y = \frac{1}{2}\left[\ln x - \ln(2-x)\right] + C_1 \implies y = C_2\sqrt{\frac{x}{2-x}}.$$

At $t = 0$, $x = y = 1$. Hence,

$$1 = C_2\sqrt{\frac{1}{2-1}} \implies C_2 = 1.$$

Thus,

$$y = \sqrt{\frac{x}{2-x}} \implies x = \frac{2y^2}{1+y^2}.$$

Part (b). Observe that

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{y^2}{x} = \frac{y^2}{2y^2/(1+y^2)} = \frac{1}{2}(1+y^2) \implies \frac{1}{1+y^2}\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{2}.$$

Integrating both sides with respect to t, we get

$$\int \frac{1}{1+y^2}\,\mathrm{d}y = \int \frac{1}{2}\,\mathrm{d}t \implies \arctan y = \frac{t}{2} + C \implies y = \tan\left(\frac{t}{2} + C\right).$$

At $t = 0$, $y = 1$. Hence,

$$1 = \tan C \implies C = \frac{\pi}{4},$$

whence

$$y = \tan\left(\frac{t}{2} + \frac{\pi}{4}\right) = \frac{1 - \cos(t + \pi/2)}{\sin(t + \pi/2)} = \frac{1 + \sin t}{\cos t} = \sec t + \tan t.$$

Observe that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = (2-x)y = (2-x)\sqrt{\frac{x}{2-x}} = \sqrt{x(2-x)} \implies \frac{1}{\sqrt{x(2-x)}}\frac{\mathrm{d}x}{\mathrm{d}t} = 1.$$

Integrating both sides with respect to t, we get

$$\int \frac{1}{\sqrt{x(2-x)}}\,\mathrm{d}x = \int 1\,\mathrm{d}t \implies 2\arcsin\left(\sqrt{\frac{x}{2}}\right) = t + C_1 \implies x = 2\sin^2\left(\frac{t}{2} + C_2\right).$$

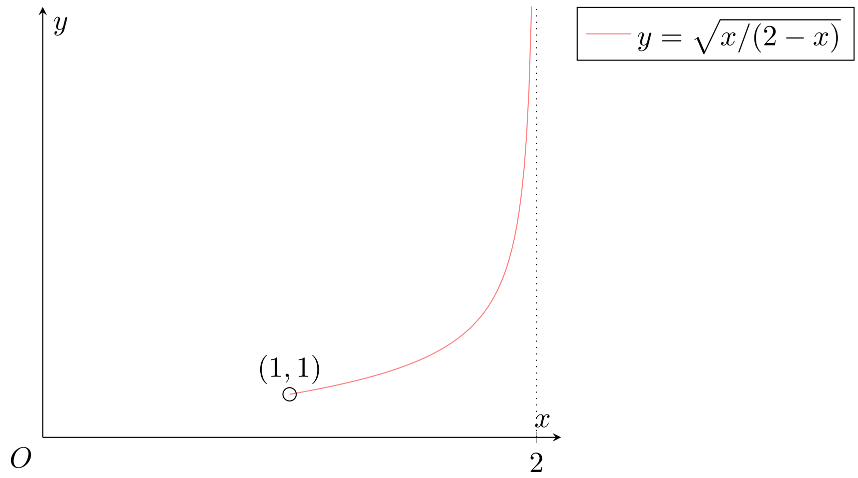
At $t=0$, $x=1$. Hence,

$$1 = 2\sin^2 C_2 \implies C_2 = \frac{\pi}{4}.$$

Thus,

$$x = 2\sin^2\left(\frac{1}{2}t + \frac{\pi}{4}\right) = 1 - \cos\left(t + \frac{\pi}{2}\right) = 1 + \sin t.$$

Part (c). Note that $0 < t < \frac{\pi}{2} \implies 1 < x < 2$.



* * * * *

Problem 2. Find the general solution of the differential equation

$$x\frac{\mathrm{d}y}{\mathrm{d}x} + 4y - 10x = 0.$$

Find the particular solution such that $y \to 0$ as $x \to 0$.

Show, on a single diagram, a sketch of this particular solution and one typical member of the family, F of solution curves for which $\frac{\mathrm{d}y}{\mathrm{d}x}$ is positive whenever x is positive.

Show that there is a straight line which passes through the maximum point of every member of F and find its equation.

Solution.

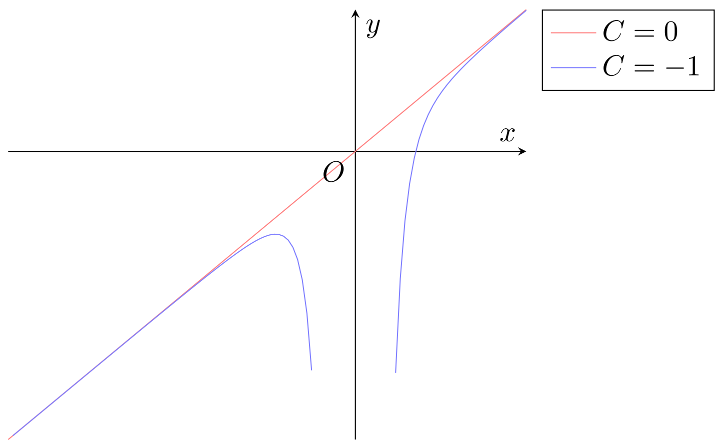
$$x\frac{\mathrm{d}y}{\mathrm{d}x} + 4y - 10x = 0 \implies x^4\frac{\mathrm{d}y}{\mathrm{d}x} + 4x^3y = \frac{\mathrm{d}}{\mathrm{d}x}\left(x^4y\right) = 10x^4$$
$$\implies x^4y = \int 10x^4\,\mathrm{d}x = 2x^5 + C \implies y = 2x + Cx^{-4}$$

As $x \to 0$, $x^{-4} \to \infty$. Hence, C must be 0, whence the particular solution is $y = 2x$.

Note that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2 - 4Cx^{-5} > 0 \implies C < \frac{x^5}{2}.$$

Since $x > 0$, we hence have the constraint $C \leq 0$ for members of F.



Consider the stationary points of members of F. For stationary points, $\frac{dy}{dx} = 0$. Hence,

$$x\frac{dy}{dx} + 4y - 10x = 0 \implies 4y - 10x = 0 \implies y = \frac{5}{2}x.$$

Differentiating the original differential equation with respect to x, we obtain

$$x\frac{dy}{dx} + 4y - 10x = 0 \implies \left(x\frac{d^2y}{dx^2} + \frac{dy}{dx}\right) + 4\frac{dy}{dx} - 10 = 0 \implies \frac{d^2y}{dx^2} = \frac{10}{x}.$$

Note that for members of F, we have that $\frac{dy}{dx} > 0$ for $x > 0$. Hence, there are no stationary points when $x > 0$. That is, any stationary point must occur when $x < 0$ (indeed, there is a stationary point when $x = \sqrt[5]{2C} < 0$). Furthermore, when $x < 0$, $\frac{d^2y}{dx^2} < 0$. Hence, all stationary points must be a maximum point. Thus, $y = \frac{5}{2}x$ passes through the maximum point of every member of F.

* * * * *

Problem 3.

(a) The variables x and y are related by the differential equation

$$x^2\frac{dy}{dx} - 2xy + y = 0.$$

(i) Find the general solution of this differential equation, expressing y in terms of x.

(ii) Find the particular solution for which $y = -\mathrm{e}$ when $x = 1$. Obtain the coordinates of the turning point of the solution curve of this particular solution and sketch the curve for $x > 0$.

(b) Find the general solution of the differential equation

$$\frac{dy}{dx} + xy = \mathrm{e}^x x^2,$$

expressing y in terms of x.

Solution.

Part (a).

Part (a)(i). Note that

$$x^2\frac{dy}{dx} - 2xy + y = 0 \implies \frac{1}{y}\frac{dy}{dx} = \frac{2}{x} - \frac{1}{x^2}.$$

Integrating with respect to x on both sides, we get

$$\int \frac{1}{y}\,\mathrm{d}y = \int \left(\frac{2}{x} - \frac{1}{x^2}\right)\,\mathrm{d}x \implies \ln|y| = 2\ln|x| + \frac{1}{x} + C_1 \implies y = C_2 x^2 \mathrm{e}^{1/x}.$$

Part (a)(ii). When $x = 1$, $y = -e$. Hence,

$$-e = C_2\left(1^2\right)\left(\mathrm{e}^1\right) \implies C_2 = -1 \implies y = -x^2 e^{1/x}.$$

For stationary points, $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$. Hence, $y(2x-1) = 0$, whence $x = \frac{1}{2}$. Note that we reject $y = 0$ since $e^{1/x} \neq 0$ and $x \neq 0$ due to the presence of a $\frac{1}{x}$ term. Hence, y has a stationary point at $(1/2, -\mathrm{e}^2/4)$.

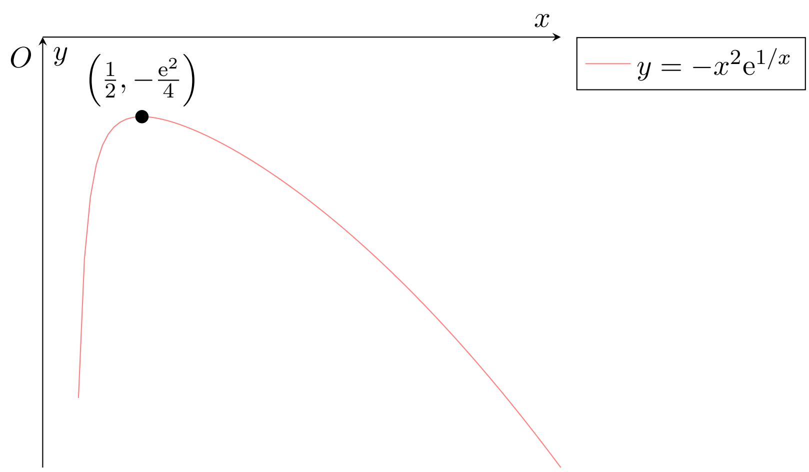
Differentiating the original differential equation with respect to x, we obtain

$$x^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 2y = 0 \implies \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{2y}{x^2}.$$

Hence, at $(1/2, -\mathrm{e}^2/4)$, we have

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{-e^2/2}{1/4} < 0,$$

whence it is a turning point.



Part (b). Observe that

$$\frac{\mathrm{d}y}{\mathrm{d}x} + xy = \mathrm{e}^x x^2 \implies e^{\frac{1}{2}x^2}\frac{\mathrm{d}y}{\mathrm{d}x} + x\mathrm{e}^{\frac{1}{2}x^2}y = \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{\frac{1}{2}x^2}y\right) = \mathrm{e}^{\frac{1}{2}x^2+x}x^2.$$

Thus,

$$\mathrm{e}^{\frac{1}{2}x^2}y = \int \mathrm{e}^{\frac{1}{2}x^2+x}x^2\,\mathrm{d}x.$$

Suppose $\int e^{\frac{1}{2}x^2+x}x^2\,\mathrm{d}x = P(x)e^{\frac{1}{2}x^2+x} + C$ for some function $P(x)$. Differentiating both sides with respect to x, we obtain

$$x^2e^{\frac{1}{2}x^2+x} = e^{\frac{1}{2}x^2+x}\left[(x+1)P(x) + P'(x)\right],$$

whence

$$x^2 = (x+1)P(x) + P'(x).$$

Thus, $P(x)$ is a polynomial of degree 1. Let $P(x) = ax + b$. For some constants a and b. Then

$$x^2 = ax^2 + (a+b)x + (a+b).$$

Comparing coefficients of x^2, x and constant terms, we have $a = 1$ and $a + b = 0 \implies b = -1$. Thus,

$$\int x^2 e^{\frac{1}{2}x^2+x}\,\mathrm{d}x = (x-1)e^{\frac{1}{2}x^2+x} + C.$$

Hence, we have

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

B14 Euler Method and Improved Euler Method

Tutorial B14

Problem 1. Consider the initial value problem

$$\frac{\mathrm{d}y}{\mathrm{d}t} = 4y - 1, \qquad y(0) = 1.$$

Use the Euler method with step size $\Delta t = 0.1$ to estimate $y(0.5)$.

Explain whether the approximation is an underestimate or an overestimate of the actual value.

Solution. Let $f(y) = 4y - 1$ and $y_0 = 1$. By the Euler method ($y_{n+1} = y_n + \Delta t f(y_n)$),

t	n	y_n
0.0	0	1
0.1	1	1.3
0.2	2	1.72
0.3	3	2.308
0.4	4	3.1312
0.5	5	4.28368

Hence, $y(0.5) \approx 4.28$.

Observe that $\frac{\mathrm{d}^2y}{\mathrm{d}t^2} = 4\frac{\mathrm{d}y}{\mathrm{d}t} > 0$ for $y > 0$. Hence, y is concave upward. Thus, the approximation is an underestimate.

* * * * *

Problem 2. A solution to the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = y - x$ has $y = 0.5$ at $x = 0$.

(a) Use the Euler method with step size 0.2 to estimate y at $x = 1$. State with a reason whether this value of y is an underestimate or an overestimate.

(b) Find the exact value of y at $x = 1$.

(c) By changing the step size to $\Delta x = 0.1$, comment on the accuracy of the approximations. What are the trade-offs, if any?

Solution.

Part (a). Let $f(x, y) = y - x$, $x_0 = 0$, $y_0 = 0.5$ and $\Delta x = 0.2$. By the Euler method ($y_{n+1} = y_n + \Delta t f(y_n)$),

x	n	y_n
0.0	0	0.5
0.2	1	0.6
0.4	2	0.68
0.6	3	0.736
0.8	4	0.7632
1.0	5	0.75584

Hence, $y(1) \approx 0.756$.
Observe that for $x \in [0,1]$,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y - x < 1 \implies \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\mathrm{d}y}{\mathrm{d}x} - 1 < 0.$$

Thus, y is concave downward near $x = 1$, whence the approximation is an overestimate.

Part (b).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y - x \implies \mathrm{e}^{-x}\frac{\mathrm{d}y}{\mathrm{d}x} - \mathrm{e}^{-x}y = \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-x}y\right) = -x\mathrm{e}^{-x}$$
$$\implies \mathrm{e}^{-x}y = \int -x\mathrm{e}^{-x}\,\mathrm{d}x = x\mathrm{e}^{-x} + \mathrm{e}^{-x} + C \implies y = x + 1 + C\mathrm{e}^{x}.$$

At $x = 0$, $y = \frac{1}{2}$. Hence,

$$\frac{1}{2} = 1 + C \implies C = -\frac{1}{2} \implies y = x + 1 - \frac{\mathrm{e}^x}{2}.$$

Evaluating y at $x = 1$ yields,

$$y = 1 + 1 - \frac{e^1}{2} = 2 - \frac{\mathrm{e}}{2}.$$

Part (c). The accuracy of the approximations will improve. However, more calculations will need to be done.

* * * * *

Problem 3. Consider the initial value problem

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

(a) Use the Euler method with step size $\Delta t = 0.2$ to estimate y at $t = 0.6$. Compare the approximated results with the exact solution.

(b) Use the improved Euler method with step size $\Delta t = 0.2$ to estimate y at $t = 0.6$. Compare the approximated results with the exact solution.

Solution. We begin by finding the exact solution to the differential equation. Observe that

$$\frac{\mathrm{d}y}{\mathrm{d}t} = t + y \implies \mathrm{e}^{-t}\frac{\mathrm{d}y}{\mathrm{d}t} - \mathrm{e}^{-t}y = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{-t}y\right) = t\mathrm{e}^{-t}$$
$$\implies \mathrm{e}^{-t}y = \int t\mathrm{e}^{-t}\,\mathrm{d}t = -t\mathrm{e}^{-t} - \mathrm{e}^{-t} + C \implies y = -t - 1 + C\mathrm{e}^{t}.$$

At $t = 0$, $t = 1$. Hence,

$$1 = -1 + C \implies C = 2 \implies y = -t - 1 + 2\mathrm{e}^t.$$

Evaluating at $t = 0.6$, we get

$$y = -0.6 - 1 + 2\mathrm{e}^{0.6} = 2.044.$$

Part (a). Let $f(t, y) = t + y$, $t_0 = 0$, $y_0 = 1$ and $\Delta t = 0.2$. By the Euler method,

t	n	y_n
0.0	0	1
0.2	1	1.2
0.4	2	1.48
0.6	3	1.856

Hence, $y(0.6) \approx 1.856$.

The approximation is not very close to the exact solution, with a percentage error of 9.20%.

Part (b). By the improved Euler method,

$$\widetilde{y_1} = y_0 + \Delta t\, f(t_0, y_0) = 1.2$$
$$y_1 = y_0 + \frac{1}{2}\Delta t\Big[f(t_0, y_0) + f(t_1, \widetilde{y_1})\Big] = 1.24$$

$$\widetilde{y_2} = y_1 + \Delta t\, f(t_1, y_1) = 1.528$$
$$y_2 = y_1 + \frac{1}{2}\Delta t\Big[f(t_1, y_1) + f(t_2, \widetilde{y_2})\Big] = 1.5768$$

$$\widetilde{y_3} = y_2 + \Delta t\, f(t_2, y_2) = 1.97216$$
$$y_3 = y_2 + \frac{1}{2}\Delta t\Big[f(t_2, y_2) + f(t_3, \widetilde{y_3})\Big] = 2.031696$$

Hence, $y(0.6) \approx 2.032$.

The approximation is very close to the exact solution, with a percentage error of 0.602%.

* * * * *

Problem 4. Consider the initial value problem

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -y^2, \qquad y(0) = \frac{1}{2}, \qquad 0 \le t \le 2.$$

(a) Determine an analytic solution for the problem.

(b) Using the improved Euler method with a step size of 0.5, determine an approximate value for $y(2)$ and its error.

Solution.

Part (a).

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -y^2 \implies -y^{-2}\frac{\mathrm{d}y}{\mathrm{d}t} = 1 \implies \int -y^2\,\mathrm{d}y = \int \mathrm{d}t$$
$$\implies y^{-1} = t + C \implies y = \frac{1}{t+C}.$$

Note that

$$y(0) = \frac{1}{2} \implies \frac{1}{2} = \frac{1}{C} \implies C = 2.$$

Hence,

$$y = \frac{1}{t+2}.$$

Part (b). Let $f(y) = -y^2$, $t_0 = 0$, $y_0 = \frac{1}{2}$ and $\Delta t = 0.5$. By the improved Euler method,

$$\widetilde{y_1} = y_0 + \Delta t\, f(y_0) = 0.375$$
$$y_1 = y_0 + \frac{1}{2}\Delta t\Big[f(y_0) + f(\widetilde{y_1})\Big] = 0.4023438$$

$$\widetilde{y_2} = y_1 + \Delta t\, f(y_1) = 0.3214035$$
$$y_2 = y_1 + \frac{1}{2}\Delta t\Big[f(y_1) + f(\widetilde{y_2})\Big] = 0.3360486$$

$$\widetilde{y_3} = y_2 + \Delta t\, f(y_2) = 0.2795843$$
$$y_3 = y_2 + \frac{1}{2}\Delta t\Big[f(y_2) + f(\widetilde{y_3})\Big] = 0.2882746$$

$$\widetilde{y_4} = y_3 + \Delta t\, f(y_3) = 0.2467235$$
$$y_4 = y_3 + \frac{1}{2}\Delta t\Big[f(y_3) + f(\widetilde{y_4})\Big] = 0.2522809$$

Hence, $y(2) \approx 0.252$. Thus, the error is

$$\text{Error} = 0.2522809 - \frac{1}{2+2} = 0.0022809$$

* * * * *

Problem 5. It is given that $\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^y + x$, and that a particular solution curve passes through the point $(0, 1)$.

(a) Use the Euler method with a step size of 0.1 to estimate the value of y at $x = 0.5$.

(b) If the estimate for y at $x = 0.5$ is calculated using the improved Euler method with a step size of 0.1, determine whether this estimate will be greater or less than the value you have calculated in (a). Justify your answer.

Solution.

Part (a). Let $f(x, y) = \mathrm{e}^y + x$, $x_0 = 0$, $y_0 = 1$ and $\Delta x = 0.1$. By the Euler method,

t	n	y_n
0.0	0	1
0.1	1	1.27183
0.2	2	1.63857
0.3	3	2.17334
0.4	4	3.08210
0.5	5	5.30253

Hence, $y(0.5) \approx 5.30$.

Part (b). Observe that for $x > 0$,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^y + x > 0 \implies \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \mathrm{e}^y\frac{\mathrm{d}y}{\mathrm{d}x} + 1 > 0.$$

Thus, y is concave upwards, whence the estimates are underestimates. Since the improved Euler method is more accurate than the Euler method, it will be greater than the value calculated in (a).

Problem 6. A differential equation is given by

$$\frac{\mathrm{d}y}{\mathrm{d}t} = y^2 - 2y + 1, \qquad y(0) = 2, \qquad 0 \le t \le 2.$$

Copy and complete the table showing the use of the improved Euler's method with step size 0.5 to estimate y at $t = 2$.

n	t_n	Euler (y_n)	$\widetilde{y_n}$	Improved Euler (y_n)	Actual y_n
0	0.0	2		2	2
1	0.5	2.5	2.5	2.8125	
2	1.0				
3	1.5				
4	2.0				

Compare and comment on your values obtained using the improved Euler method with the values obtained from the Euler method and the actual solution.

Solution. Let $f(y) = y^2 - 2y + 1$ and $\Delta t = 0.5$.

Euler Method.

$$\begin{aligned} y_2 &= y_1 + \Delta t\, f(y_1) = 3.625 \\ y_3 &= y_2 + \Delta t\, f(y_2) = 7.070313 \\ y_4 &= y_3 + \Delta t\, f(y_3) = 25.49466 \end{aligned}$$

Improved Euler Method.

$$\begin{aligned} \widetilde{y_2} &= y_1 + \Delta t\, f(y_1) = 4.455078 \\ y_2 &= y_1 + \frac{1}{2}\Delta t\big[f(y_1) + f(\widetilde{y_2})\big] = 6.618180 \end{aligned}$$

$$\begin{aligned} \widetilde{y_3} &= y_2 + \Delta t\, f(y_2) = 22.40016 \\ y_3 &= y_2 + \frac{1}{2}\Delta t\big[f(y_2) + f(\widetilde{y_3})\big] = 129.0008 \end{aligned}$$

$$\begin{aligned} \widetilde{y_4} &= y_3 + \Delta t\, f(y_3) = 8321.107 \\ y_4 &= y_3 + \frac{1}{2}\Delta t\big[f(y_3) + f(\widetilde{y_4})\big] = 17310270 \end{aligned}$$

Actual Value.

$$\frac{\mathrm{d}y}{\mathrm{d}t} = y^2 - 2y + 1 = (y-1)^2 \implies \frac{1}{(y-1)^2}\frac{\mathrm{d}y}{\mathrm{d}t} = 1 \implies \int \frac{1}{(y-1)^2}\,\mathrm{d}y = \int 1\,\mathrm{d}t$$
$$\implies -\frac{1}{y-1} = x + C \implies y = 1 - \frac{1}{x+C}.$$

Note that

$$y(0) = 2 \implies 2 = 1 - \frac{1}{C} \implies C = -1.$$

Hence,

$$y = 1 - \frac{1}{x-1}.$$

n	t_n	Euler (y_n)	$\widetilde{y_n}$	Improved Euler (y_n)	Actual y_n
0	0.0	2		2	2
1	0.5	2.5	2.5	2.8125	3
2	1.0	3.625	4.455078	6.618180	-
3	1.5	7.070313	22.40016	129.0008	-1
4	2.0	25.49466	8321.107	17310270	0

The values obtained using the improved Euler method deviate significantly from that obtained using the Euler method and the actual solution. This is because y has a discontinuity at $t = 1$, making both Euler methods inappropriate to use.

* * * * *

Problem 7. A solution of the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y - x$$

has $y = 2$ at $x = 0$.

(a) Use the Euler method with step size 0.5 to estimate y at $x = 1$. Explain whether you expect this value of y to be an underestimate or overestimate of the true value.

(b) Copy and complete the table showing the use of the improved Euler method with step size 0.5 to estimate y at $x = 1$.

x	y	$y - x$	$\widetilde{y}$	$\Delta y / \Delta x$
0	2	2	3	$(2 + 2.5)/2$
0.5	3.125	2.625	4.438	
1				

(c) Show that the exact value of y at $x = 1$ is $2 + \mathrm{e}$.

Solution.

Part (a). Let $f(x, y) = y - x$, $x_0 = 0$, $y_0 = 2$, $\Delta x = 0.5$. By the Euler method,

x	n	y_n
0.0	0	2
0.5	1	3
1.0	2	4.25

Hence, $y(1) \approx 4.25$.
Note that for $x > 0$, we get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y - x > \implies \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \frac{\mathrm{d}y}{\mathrm{d}x} - 1 > 0.$$

Thus, y is concave upwards, whence the value of y is an underestimate.

Part (b). From the improved Euler method, one has

$$y_2 = y_1 + \frac{1}{2}\Delta x\Big[f(x_1, y_1) + f(x_2, \widetilde{y_2})\Big].$$

Thus,

$$\frac{y_2 - y_1}{\Delta x} = \frac{\Delta y}{\Delta x} = \frac{1}{2}\Big[f(x_1, y_1) + f(x_2, \widetilde{y_2})\Big] = \frac{1}{2}\left[2.625 + (4.438 - 1)\right] = 3.031.$$

Also,

$$y_2 = y_1 + \Delta x \left(\frac{\Delta y}{\Delta x} \right) = 3.125 + 0.5\,(3.031) = 4.64.$$

x	y	$y - x$	$\widetilde{y}$	$\Delta y / \Delta x$
0	2	2	3	$(2 + 2.5)/2$
0.5	3.125	2.625	4.438	3.031
1	4.64			

Part (c).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y - x \implies \mathrm{e}^{-x}\frac{\mathrm{d}y}{\mathrm{d}x} - \mathrm{e}^{-x}y = \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-x}y\right) = -x\mathrm{e}^{-x}$$

$$\implies \mathrm{e}^{-x}y = \int \left(-x\mathrm{e}^{-x}\right)\,\mathrm{d}x = x\mathrm{e}^{-x} + \mathrm{e}^{-x} + C \implies y = x + 1 + C\mathrm{e}^{x}.$$

At $x = 0$, $y = 2$, whence

$$2 = 1 + C \implies C = 1 \implies y = x + 1 + \mathrm{e}^{x}.$$

Thus, at $x = 1$,

$$y = 1 + 1 + \mathrm{e}^{1} = 2 + \mathrm{e}.$$

* * * * *

Problem 8. Initially, a tank is fully filled with 100 litres of pure water. There exists a tap at the top of the tank. This tap supplies brine, containing 1 g of salt per litre, into the tank at a rate of 1 litre per minute. There also exists another tap at the bottom of the tank which allows the mixture to flow out at a constant rate of 2 litres per minute. At time T (in minutes), the amount of salt and the volume of the mixture in the tank are denoted by S (in grams) and V (in litres) respectively. Both taps are turned on simultaneously at time $T = 0$. The tap at the bottom of the tank is turned off at time $T = 75$. The mixture in the tank is assumed to be well-stirred and homogenous at all times.

(a) Show that $\frac{\mathrm{d}S}{\mathrm{d}T} = \frac{100 - T - 2S}{100 - T}$, $0 < T < 75$.

(b) By solving the differential equation, show that the amount of salt in the tank after 75 minutes is 18.75 grams.

At the instance when the tap at the bottom is turned off, a crack is accidentally created at the bottom of the tank. According to Torricelli's law, the mixture flows out from the crack at a rate proportional to the square-root of its volume. It can be assumed that the mixture flow obeys Torricelli's law, regardless of its viscosity. Let the amount of salt and the volume of the mixture in the tank be denoted by s (in grams) and v (in litres) respectively, t minutes after the crack has been accidentally created. It has been observed that the volume of the mixture in the tank stays constant at 36 litres after a long period of time.

(c) Show that $\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{6 - \sqrt{v}}{6}$. Estimate the time taken for the mixture in the tank to rise to 26 litres after the crack has been created, by using

  (i) Euler's Method with two iterations,

  (ii) Simpson's Rule with two strips.

(d) Show that $\frac{\mathrm{d}s}{\mathrm{d}v} = \frac{6\sqrt{v}-s}{6\sqrt{v}-v}$. Use the improved Euler method with one iteration to estimate the amount of salt in the tank at the instant when the mixture in the tank rises to 26 litres after the crack has been created. Given your answer to 4 decimal places.

Solution.

Part (a). The concentration of salt in the tank is given by $\frac{S}{V}$. Let V_i and V_o be the volume of liquid entering and leaving the tank in litres, respectively. Then

$$\frac{\mathrm{d}V}{\mathrm{d}T} = \frac{\mathrm{d}V_i}{\mathrm{d}T} - \frac{\mathrm{d}V_o}{\mathrm{d}T} = 1 - 2 = -1 \implies V = 100 - T,$$

since $V = 100$ initially. Thus,

$$\frac{\mathrm{d}S}{\mathrm{d}T} = 1\left(\frac{\mathrm{d}V_i}{\mathrm{d}T}\right) - \frac{S}{V}\left(\frac{\mathrm{d}V_o}{\mathrm{d}T}\right) = 1 - \frac{2S}{100-T} = \frac{100-T-2S}{100-T}.$$

Part (b).

$$\frac{\mathrm{d}S}{\mathrm{d}T} = 1 - \frac{2S}{100-T} \implies \frac{\mathrm{d}S}{\mathrm{d}T} + \frac{2S}{100-T} = 1$$
$$\implies \frac{1}{(100-T)^2}\frac{\mathrm{d}S}{\mathrm{d}T} + \frac{2S}{(100-T)^3} = \frac{\mathrm{d}}{\mathrm{d}T}\left(\frac{S}{(100-T)^2}\right) = \frac{1}{(100-T)^2}$$
$$\implies \frac{S}{(100-T)^2} = \int \frac{\mathrm{d}T}{(100-T)^2} = \frac{1}{100-T} + C \implies S = 100 - T + C(100-T)^2.$$

When $T = 0$, $S = 0$, whence

$$0 = 100 + C\left(100^2\right) \implies C = -\frac{1}{100} \implies S = 100 - T - \frac{(100-T)^2}{100}.$$

When $T = 75$,

$$S = 100 - 75 - \frac{(100-75)^2}{100} = 18.75.$$

Hence, the amount of salt in the tank after 75 minutes is 18.75 grams.

Part (c). Let v_i and v_o be the volume of liquid entering and leaving the tank in litres, respectively. Since the top tap is still open, we have $\frac{\mathrm{d}v_i}{\mathrm{d}t} = 1$. By Torricelli's law, we also have $\frac{\mathrm{d}v_0}{\mathrm{d}t} = k\sqrt{v}$. Thus,

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\mathrm{d}v_i}{\mathrm{d}t} - \frac{\mathrm{d}v_o}{\mathrm{d}t} = 1 - k\sqrt{v}.$$

Since the volume of the tank remains constant at 36 litres eventually, we have

$$1 - k\sqrt{36} = 0 \implies k = \frac{1}{6}.$$

Thus,

$$\frac{\mathrm{d}v}{\mathrm{d}t} = 1 - \frac{\sqrt{v}}{6} = \frac{6-\sqrt{v}}{6}.$$

Observe that $\frac{\mathrm{d}t}{\mathrm{d}v} = \frac{6}{6-\sqrt{v}}$.

Part (c)(i). Let $f(v) = \frac{6}{6-\sqrt{v}}$, $v_0 = 25$, $t_0 = 0$ and $\Delta v = 0.5$. By the Euler method,

$$t_1 = t_0 + \Delta v\, f(v_0) = 3, \qquad t_2 = t_1 + \Delta v\, f(v_1) = 6.157.$$

Hence, when $t \approx 6.16$, the mixture in the tank has risen to 26 litres.

Part (c)(ii). Note that $t = \int \frac{6}{6-\sqrt{v}}\,\mathrm{d}v$. Hence, the desired time is given by $\int_{25}^{26} \frac{6}{6-\sqrt{v}}\,\mathrm{d}v$. By Simpson's rule,

$$\int_{25}^{26} \frac{6}{6-\sqrt{v}}\,\mathrm{d}v \approx \frac{1}{3}\cdot\frac{26-25}{2}\Big[f(25)+4f(25.5)+f(26)\Big] = 6.319.$$

Hence, when $t \approx 6.32$, the mixture in the tank has risen to 26 litres.

Part (d). Observe that

$$\frac{\mathrm{d}s}{\mathrm{d}t} = 1\left(\frac{\mathrm{d}v_i}{\mathrm{d}t}\right) - \frac{s}{v}\left(\frac{\mathrm{d}v_o}{\mathrm{d}t}\right) = 1 - \frac{s}{v}\frac{\sqrt{v}}{6} = \frac{6\sqrt{v}-s}{6\sqrt{v}}.$$

By the chain rule,

$$\frac{\mathrm{d}s}{\mathrm{d}v} = \frac{\mathrm{d}s}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}v} = \frac{6\sqrt{v}-s}{6\sqrt{v}}\cdot\frac{6}{6-\sqrt{v}} = \frac{6\sqrt{v}-s}{6\sqrt{v}-v}.$$

Let $f(s,v) = \frac{6\sqrt{v}-s}{6\sqrt{v}-v}$, $v_0 = 25$, $s_0 = 18.75$ and $\Delta v = 1$. By the improved Euler method,

$$\begin{aligned}\widetilde{s_1} &= s_0 + \Delta v\, f(s_0, v_0) = 21\\ s_1 &= s_0 + \frac{1}{2}\Delta v\Big[f(s_0,v_0) + f(\widetilde{s_1}, v_1)\Big] = 20.9192 \text{ (4 d.p.)}\end{aligned}$$

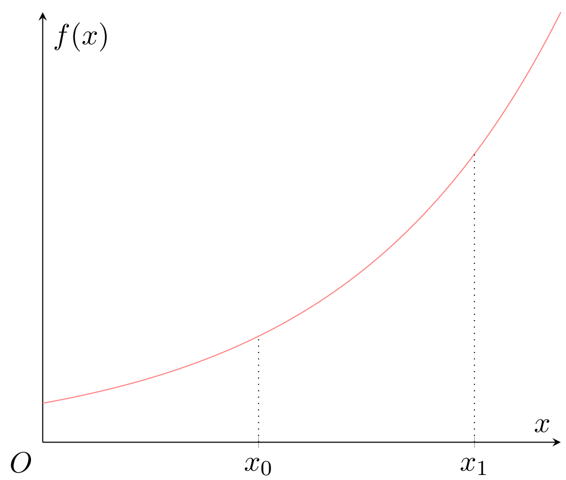
Hence, there is approximately 20.9192 grams of salt in the tank.

* * * * *

Problem 9. A solution to the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)$ has $y = y_0$ at $x = x_0$. It is required to estimate the value of y at $x = x_1$ using a numerical method with one step.

(a) Write down expressions for the value of y at $x = x_1$ obtained by using the Euler method and by using the improved Euler method.

(b) The graph of f is as below. Copy the graph and use it to illustrate the errors in the two estimates of y obtained by using the methods of part (a). State clearly whether the errors correspond to overestimates or underestimates.



(c) Given that $x_0 = 0$ and $f(x) = a + bx + cx^2$, where a, b and c are constants, find the error in using the improved Euler method with a single step of size h.

Solution.

Part (a). Let $\Delta x = x_1 - x_0$.

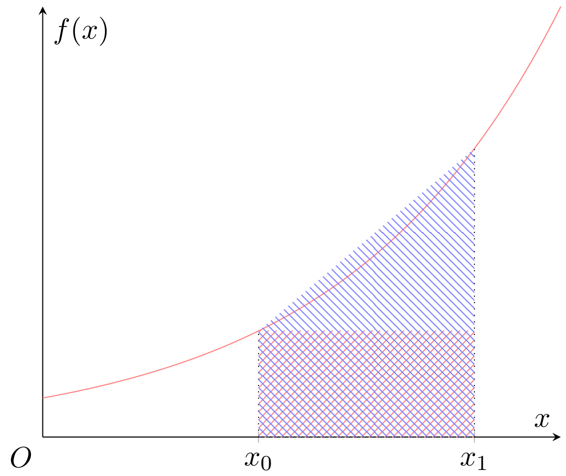
Euler Method.

$$y_1 = y_0 + \Delta x\, f(x_0)$$

Improved Euler Method.

$$y_1 = y_0 + \frac{1}{2}\Delta x\Big[f(x_0) + f(x_1)\Big]$$

Part (b).



By the fundamental theorem of calculus, the area under the graph of $f(x)$ between x_0 and x_1 is precisely $y_1 - y_0$. That is,

$$\int_{x_0}^{x_1} f(x)\,\mathrm{d}x = y_1 - y_0 = \Delta y.$$

Hence, the better the approximation of the integral, the better the approximation of Δy and thus y_1.

The Euler method gives the approximation $\Delta y = \Delta x\, f(x_0)$. This is represented by the area of the red-shaded rectangle with base Δx and height $f(x_0)$.

The improved Euler method gives the approximation $\Delta y = \Delta x\,\frac{1}{2}\big[f(x_0) + f(x_1)\big]$. This is represented by the area of the blue-shaded trapezium with base Δx and heights $f(x_0)$ and $f(x_1)$.

Thus, the improved Euler method gives a better approximation for the integral of $f(x)$ than the Euler method. Thus, the error of the estimate given by the improved Euler method is smaller than that of the Euler method.

The Euler method underestimates the integral, hence y_1 is also underestimated. Similarly, the improved Euler method overestimates the integral, hence y_1 is also overestimated.

Part (c). We have $f(x) = a + bx + cx^2$, $x_0 = 0$ and $\Delta x = h$. Let $y_0 = d$.

Improved Euler Method.

$$y_1 = d + \frac{1}{2}h\Big[(a) + \big(a + bh + ch^3\big)\Big] = d + ah + \frac{1}{2}bh^2 + \frac{1}{2}ch^3.$$

Actual Value.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = a + bx + cx^2 \implies y = \int \left(a + bx + cx^2\right)\mathrm{d}x = ax + \frac{1}{2}bx^2 + \frac{1}{3}cx^3 + C.$$

When $x = 0$, $y = d$. Hence, $C = d$, thus $y = d + ax + \frac{1}{2}bx^2 + \frac{1}{3}cx^3$. At $x = h$,

$$y_1 = d + ah + \frac{1}{2}bh^2 + \frac{1}{3}ch^3$$

Error.

$$\text{Error} = \left(d + ah + \frac{1}{2}bh^2 + \frac{1}{3}ch^3\right) - \left(d + ah + \frac{1}{2}bh^2 + \frac{1}{2}ch^3\right) = \frac{1}{6}ch^3.$$

* * * * *

Problem 10. The differential equation

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

where $y = 1$ when $x = 1$, is to be solved numerically.

(a) Carry out two steps of Euler's method with step length 0.1 to estimate the value of y when $x = 1.2$, giving your answer to 4 decimal places.

(b) The method in part (a) is now replaced by the improved Euler method. The estimate obtained is 2.0156, given to 4 decimal places. State, with a reason, whether this estimate and the one found in part (a) are likely to be overestimates or underestimates of the actual value of y when $x = 1.2$.

(c) Explain why it would be inappropriate to continue this process in part (a) to estimate the value of y when $x = 1.6$.

Solution.

Part (a). Let $f(x, y) = \mathrm{d}y/\mathrm{d}x = 1+y^2 \tan x$, $x_0 = 1$, $y_0 = 1$, $\Delta x = 0.1$ and $x_n = x_0+n\Delta x$. By the Euler method,

$$y_1 = y_0 + \Delta x\, f(x_0, y_0) = 1.2557408, \qquad y_2 = y_1 + \Delta x\, f(x_1, y_1) = 1.6655608.$$

Hence, $y(1.2) \approx 1.6656$ (4 d.p.).

Part (b). Observe that on the interval $x \in I = [1, 1.2]$, we have

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 1 + y^2 \tan x > 0.$$

Since y is continuous on I, and $y(1) = 1$, we also have $y > 0$ on I. Thus,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 2y\frac{\mathrm{d}y}{\mathrm{d}x} \tan x + y^2 \sec^2 x > 0,$$

whence y is concave upwards. Thus, the estimates are likely to be underestimates.

Part (c). $f(x, y)$ has a vertical asymptote at $x = \pi/2 \in (1.5, 1.6)$. Thus, the Euler method will fail.

Self-Practice B14

Problem 1. A solution of the differential equation $\frac{\mathrm{d}x}{\mathrm{d}t} + x\cot t = \csc t$ has $x = 0$ when $t = 1$.

(a) Use the improved Euler method with a step size of 0.5 to estimate x at $t = 2$.

(b) Solve the differential equation to find the exact value of x when $t = 2$.

(c) By considering the gradients of the curve at $t = 1.5$, $t = 2$ and $t = 2.5$, comment on the accuracy of the Euler method to estimate x at $t = 2$ and $t = 3$.

* * * * *

Problem 2. For the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{y}{x} = 3x$, consider the solution curve passing through the point $(1, 2)$.

(a) Compute the Euler approximation to $y(1.1)$ using step-size $h = 0.1$.

(b) State, giving a reason, if you would expect the estimate in part (a) to be an underestimate or an overestimate of the true value.

(c) Find the general solution of the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{y}{x} = 3x$.

(d) Hence, find the solution curve that passes through the point $(1, 2)$ and calculate the percentage error of your estimate in (a).

* * * * *

Problem 3. A differential equation is of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y}{x} f\left(\frac{y}{x}\right).$$

(a) By using the substitution $z = y/x$, show that

$$\ln|x| = \int \frac{1}{z\,[f(z) - 1]}\,\mathrm{d}z.$$

(b) A particular solution of the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y}{x} \ln\left(\frac{x}{y}\right)$$

has $y = 1$ when $x = 1$. Find y in terms of x.

(c) Copy and complete the table below, using the improved Euler method with step size 0.5 to estimate y at $x = 2$.

x	y	$\mathrm{d}y/\mathrm{d}x$	$\widetilde{y}$	$\Delta y/\Delta x$
1	1	0	1	0.13516
1.5	1.06758			
2				

(d) Use a graphical method to explain whether the estimated value of y found in part (c) is an under-estimate or over-estimate.

Problem 4. The variables P and t are related by the "modified" logistic equation

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{1}{10}P\left(1 - \frac{P}{10}\right)(P-1).$$

The differential equation is used to model the size, P (in thousands), of species of wolves in time, t (in years) in a given habitat.

(a) Biologists observe that if the population of wolves is "too small", adults run the risk of being unable to find a mate, resulting in a decrease to the population.

  (i) Explain how the model accounts for this observation.

  (ii) State the maximum population that the resources in the habitat can support.

  (iii) Find the equilibrium solutions.

  (iv) Sketch the possible solution curves for P as a function of t.

(b) You are given that the population is now 2000.

  (i) Copy and complete the table showing the use of the improved Euler method with step size 0.5 to estimate the population in a year's time.

t	P	$\mathrm{d}P/\mathrm{d}t$	$\widetilde{P}$	$\Delta P/\Delta t$
0	2	0.16	2.08	0.16896
0.5	2.08448			
1				

  (ii) How could the accuracy of the numerical method in part (b)(i) be improved?

(c) Suppose the wolves are being hunted at a fixed rate E (in thousands per year). Write down the new model for the population.

Assignment B14

Problem 1.

(a) Explain why the Euler method will fail for the initial-value problem

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y\cos\sqrt{x}, \qquad y(0) = 0,$$

where $y = y(x)$ satisfies that differential equation and is not a constant.

(b) Suppose the initial condition for the problem in part (a) is now $y(0) = 10$. Use the improved Euler method with a step size of 0.1 to find, to three decimal places, an estimate for $y(0.1)$.

(c) Solve the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{2}y(2-x), \qquad y > 0, \quad y(0) = 10,$$

expressing y in terms of x, and simplifying your answer as far as possible.

(d) Explain why the solution found in part (c) will give a reasonable estimate for $y(0.1)$ in part (b).

Solution.

Part (a). By the Euler method,

$$y_1 = y_0 + \Delta x(y_0\cos\sqrt{x_0}) = 0 + \Delta x(0\cos 0) = 0.$$

It follows that $y_n = 0$ for all $n \in \mathbb{N}$, whence y is the zero function. However, because y is not a constant function, y cannot be the zero function, a contradiction. Hence, the Euler method fails.

Part (b). Let $\Delta x = 0.1$, $y_0 = 10$ and $x_n = n\Delta x$.

$$\begin{aligned} \widetilde{y_1} &= y_0 + \Delta x(y_0\cos\sqrt{x_0}) = 11 \\ y_1 &= y_0 + \frac{1}{2}\Delta x\left[y_0\cos\sqrt{x_0} + \widetilde{y_1}\cos\sqrt{x_1}\right] = 11.023 \text{ (3 d.p.)} \end{aligned}$$

Hence, $y(0.1) \approx 11.023$.

Part (c).

$$\begin{aligned} &\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{2}y(2-x) \implies \frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{2}(2-x) \implies \int\frac{1}{y}\,\mathrm{d}y = \int\frac{1}{2}(2-x)\,\mathrm{d}x \\ &\implies \ln y = \frac{1}{2}\left[2x - \frac{1}{2}x^2\right] + C_1 = x - \frac{1}{4}x^2 + C_1 \implies y = C\exp\left(x - \frac{1}{4}x^2\right). \end{aligned}$$

Since $y(0) = 10$, we have $C = 10$. Thus,

$$y = 10\exp\left(x - \frac{1}{4}x^2\right).$$

Part (d). For small x, we have that $\cos\sqrt{x} \approx 1 - \frac{1}{2}(\sqrt{x})^2 = \frac{1}{2}(2-x)$. Thus,

$$y\cos\sqrt{x} \approx \frac{1}{2}y(2-x),$$

whence the two differential equations and thus their solutions are approximately equal. Since $x = 0.1$ is small, the solution found in part (c) will give a reasonable estimate for $y(0.1)$ in part (b).

* * * * *

Problem 2. Rewriting the given differential equation, we obtain

$$\frac{\mathrm{d}v}{\mathrm{d}x} = -\frac{7x}{v} - 24.$$

Let $f(x,v) = -\frac{7x}{v} - 24$, $\Delta x = 1$, $v_0 = 121$, and $x_n = n\Delta x$.

Solution.

Part (a). By the Euler method,

$$v_1 = v_0 + \Delta x f(x_0, v_0) = 97.$$

Thus, $y(1) \approx 97$.

Part (b). By the improved Euler method,

$$\begin{aligned}\widetilde{v_1} &= v_0 + \Delta x f(x_0, v_0) = 97\\ v_1 &= v_0 + \frac{1}{2}\Delta x\left[f(x_0, v_0) + f(x_1, \widetilde{v_1})\right] = 96.964\end{aligned}$$

Thus, $y(1) \approx 96.964$.

The gradient of v at $x = 0$ is $f(x_0, v_0) = -24$, which is very close to the gradient of v at $x = 1$, $f(x_1, \widetilde{v_1}) = -24.072$. Since the gradient of v is approximately constant for $0 \le x \le 1$, we have that v is approximately a linear function on that interval.

Observe that for $0 \le x \le 1$, $x/v \approx 0$ since $x \in [0,1]$, while $v \ge 96$. Thus, $\mathrm{d}v/\mathrm{d}x \approx -24$, whence $v = -24x + C$. Since $v = 121$ when $x = 0$, we have $v \approx -24x + 121$.

* * * * *

Problem 3. The function $y = y(x)$ satisfies

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{5}(\tan x + x^3 y).$$

The value of $y(h)$ is to be found, where h is a small positive number, and $y(0) = 0$.

(a) Use one step of the improved Euler method to find an alternative approximation to $y(h)$ in terms of h.

(b) It can be shown that $y = y(x)$ satisfies

$$y(h) = e^{0.05h^4}\int_0^h \frac{\tan x}{5}\mathrm{e}^{-0.05x^4}\,\mathrm{d}x.$$

Assume that h is small and hence find another approximation to $y(h)$ in terms of h.

(c) Discuss the relative merits of these two methods employed to obtain these approximations.

Solution.

Part (a). Let $f(x,y) = \frac{1}{5}(\tan x + x^3 y)$, $\Delta x = h$ and $y_0 = 0$. By the improved Euler method,

$$\begin{aligned}
\widetilde{y_1} &= y_0 + \Delta x f(x_0, y_0) = 0 \\
y_1 &= y_0 + \frac{1}{2}\Delta x\left[f(x_0,y_0) + f(x_1,\widetilde{y_1})\right] = 0 + \frac{1}{2}h\left[0 + \frac{1}{5}(\tan h + 0)\right] = \frac{h\tan h}{10}.
\end{aligned}$$

Hence,

$$y(h) \approx \frac{h\tan h}{10}.$$

Part (b). Since h is small, we have that $\mathrm{e}^{0.05h^4} \approx 1$. Furthermore, since we are integrating over the interval $x \in [0,h]$, the integrand $\frac{\tan x}{5}\mathrm{e}^{-0.05x^4}$ can likewise be approximated by $\frac{x}{5}$. Our integral hence transforms to

$$y(h) = \mathrm{e}^{0.05h^4}\int_0^h \frac{\tan x}{5}\mathrm{e}^{-0.05x^4}\,\mathrm{d}x \approx \int_0^h \frac{x}{5}\,\mathrm{d}x = \frac{h^2}{10}.$$

Part (c). The improved Euler method involves more steps, while the approximation in (b) is more direct.

B15 Modelling Populations with First Order Differential Equations

Assignment B15

Problem 1. In response to a massive ecosystem-wide destruction by goats on the island of Isabela in Ecuador, Project Isabela was started on the first day of 1997 to eliminate all goats on the island. Goat elimination was done by hunting at a constant rate. Suppose that the goat population, P (in thousands), can be modelled by the differential equation

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{P}{4}\left(1 - \frac{P}{150}\right) - H,$$

where t is measured in months and H is measured in thousands.

(a) State, in context, the significance of the term H.

(b) Find the greatest integer value of H for which it is still possible for some goats to survive in the long run.

(c) Based on the answer from part (b), discuss the long-term behaviour of the goat population for different initial populations.

The hunters involved in Project Isabela finally managed to eliminate all the goats on the island of Isabela on the first day of 2006.

(d) State an inequality that must be satisfied by H.

(e) Given that the initial goat population was 100 thousand, find the value of H, correct to 3 decimal places.

Solution.

Part (a). H represents the number of goats killed (in thousands) per month.

Part (b). Consider the equilibrium points of the differential equation.

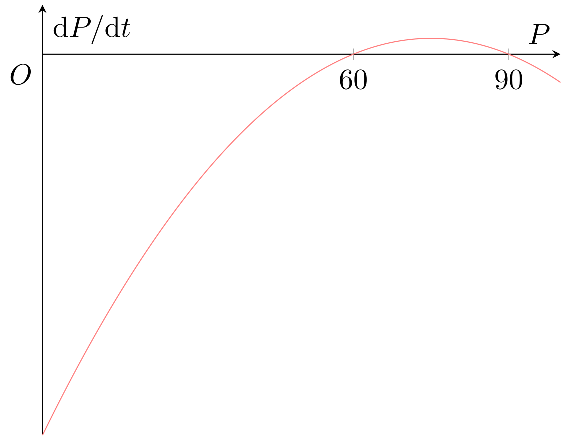
$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{P}{4}\left(1 - \frac{P}{150}\right) - H = -\frac{1}{600}\left(P^2 - 150P + 600H\right) = 0.$$

By the quadratic formula,

$$P = 75 \pm 5\sqrt{225 - 24H}.$$

For it to be possible for goats to survive in the long term, there must be at least one equilibrium point. That is, $\sqrt{225 - 24H} \geq 0 \implies H \leq 9.375$. Thus, the maximum integer value of H is 9.

Part (c). When $H = 9$, the equilibrium points are $P = 75 \pm 5\sqrt{225 - 24(9)} = 60$ or 90. Let the initial population be P_0.



When $P_0 = 0$, there are no goats initially. Hence, the population will remain at 0.

When $0 < P_0 < 60$, $\mathrm{d}P/\mathrm{d}t < 0$. Hence, the population of goats will decrease towards 0.

When $P_0 = 60$, $\mathrm{d}P/\mathrm{d}t = 0$. Hence, the population of goats will remain at 60 thousand.

When $60 < P_0 < 90$, $\mathrm{d}P/\mathrm{d}t > 0$. Hence, the population of goats will increase towards 90 thousand.

When $P_0 = 90$, $\mathrm{d}P/\mathrm{d}t = 0$. Hence, the population of goats will remain at 90 thousand.

When $P_0 > 90$, $\mathrm{d}P/\mathrm{d}t < 0$. Hence, the population of goats will decrease towards 90 thousand.

Part (d). H must satisfy the inequality $H > 9.375$.

Part (e). Note that $t = 120$, $P(0) = 100$ and $P(108) = 0$. Now observe that

$$\begin{aligned}\frac{\mathrm{d}P}{\mathrm{d}t} &= -\frac{1}{600}\left(P^2 - 150P + 600H\right) = -\frac{1}{600}\left[(P-75)^2 + \left(600H - 75^2\right)\right] \\ \implies &\frac{1}{(P-75)^2 + (600H - 75^2)}\frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{1}{600}.\end{aligned}$$

Integrating both sides with respect to t,

$$\begin{aligned}\int \frac{1}{(P-75)^2 + (600H-75^2)}\,\mathrm{d}P &= \int -\frac{1}{600}\,\mathrm{d}t \\ \implies \frac{1}{\sqrt{600H-75^2}}\arctan\left(\frac{P-75}{\sqrt{600H-75^2}}\right) &= -\frac{1}{600}t + C.\end{aligned}$$

Let $X = \frac{1}{\sqrt{600H-75^2}}$. This simplifies the above result to

$$X\arctan((P-75)X) = -\frac{1}{600}t + C.$$

When $t = 0$, $P = 100$. Hence,

$$C = X\arctan(25X)$$

When $t = 108$, $P = 0$. Hence,

$$X\arctan(-75X) = -\frac{108}{600} + X\arctan(25X)\,,$$

which has the solution $X = 0.073145$. Note that we reject $X = -0.073145$ since $X \geq 0$. We thus have

$$H = \frac{1}{600}\left(\frac{1}{0.073145^2} + 75^2\right) = 9.377 \text{ (3 d.p.)}.$$

B16 Second Order Differential Equations

Tutorial B16

Problem 1. Find the general solution of $3\frac{\mathrm{d}^2x}{\mathrm{d}t^2}+4\frac{\mathrm{d}x}{\mathrm{d}t}-7x=0$.

Solution. Consider the characteristic equation of the DE:

$$3m^2+4m-7=(3m+7)(m-1)=0.$$

We hence have $m=-7/3$ or $m=1$, whence

$$x=A\mathrm{e}^{-\frac{7}{3}t}+B\mathrm{e}^{t}.$$

* * * * *

Problem 2. Solve the following homogeneous second-order linear differential equations.

(a) $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+4\frac{\mathrm{d}y}{\mathrm{d}x}+3y=0$, given that $y=0$ and $\frac{\mathrm{d}y}{\mathrm{d}x}=-4$ when $x=0$.

(b) $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+6\frac{\mathrm{d}y}{\mathrm{d}x}+9y=0$, given that $y=1$ and $\frac{\mathrm{d}y}{\mathrm{d}x}=1$ when $x=0$.

(c) $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+\sqrt{3}\frac{\mathrm{d}y}{\mathrm{d}x}+3y=0$, given that $y=0$ and $\frac{\mathrm{d}y}{\mathrm{d}x}=-4$ when $x=0$.

Solution.

Part (a). Consider the characteristic equation of the DE:

$$m^2+4m+3=(m+1)(m+3)=0.$$

We hence have $m=-1$ or $m=-3$, whence

$$y=A\mathrm{e}^{-x}+B\mathrm{e}^{-3x} \implies \frac{\mathrm{d}y}{\mathrm{d}x}=-A\mathrm{e}^{-x}-3B\mathrm{e}^{-3x}.$$

Using the given conditions, we obtain the system

$$\begin{cases} A+B=0 \\ -A-3B=-4 \end{cases},$$

which has solution $A=-2$ and $B=2$. Thus,

$$y=-2\mathrm{e}^{-x}+2\mathrm{e}^{-3x}.$$

Part (b). Consider the characteristic equation of the DE:

$$m^2+6m+9=(m+3)^2=0.$$

We have a repeated root $m=-3$, whence

$$y=(A+Bx)\mathrm{e}^{-3x} \implies \frac{\mathrm{d}y}{\mathrm{d}x}=-3(A+Bx)\mathrm{e}^{-3x}+B\mathrm{e}^{-3x}.$$

Using the given conditions, we obtain the system

$$\begin{cases} A &= 1 \\ -3A + B &= 1 \end{cases},$$

which has solution $A = 1$ and $B = 4$. Thus,

$$y = (1 + 4x)\mathrm{e}^{-3x}.$$

Part (c). Consider the characteristic equation of the DE:

$$m^2 + \sqrt{3}m + 3 = 0.$$

Solving, we get

$$m = \frac{-\sqrt{3}}{2} \pm \frac{3}{2}\mathrm{i},$$

whence

$$y = \mathrm{e}^{-\frac{\sqrt{3}}{2}x}\left(A\cos\frac{3}{2}x + B\sin\frac{3}{2}x\right).$$

Differentiating, we get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^{-\frac{\sqrt{3}}{2}x}\left[\left(-\frac{\sqrt{3}}{2}A + \frac{3}{2}B\right)\cos\frac{3}{2}x + \left(-\frac{\sqrt{3}}{2}B - \frac{3}{2}A\right)\sin\frac{3}{2}x\right].$$

Using the given conditions, we obtain the system

$$\begin{cases} A &= 0 \\ -\frac{\sqrt{3}}{2}A + \frac{3}{2}B &= -4 \end{cases},$$

whence $A = 0$ and $B = -8/3$. Thus,

$$y = -\frac{8}{3}\mathrm{e}^{-\frac{\sqrt{3}}{2}x}\sin\frac{3}{2}x.$$

* * * * *

Problem 3. Find the general solution of

(a) $2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 3\frac{\mathrm{d}y}{\mathrm{d}x} - 5y = 10x^2 + 1$,

(b) $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 2\frac{\mathrm{d}y}{\mathrm{d}x} + 3y = 22\mathrm{e}^{4x}$,

(c) $\frac{\mathrm{d}^2s}{\mathrm{d}t^2} - 2\frac{\mathrm{d}s}{\mathrm{d}t} + s = 4\mathrm{e}^{t}$,

(d) $\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 16x = 3\cos 4t$.

Solution.

Part (a). Consider the characteristic equation of the DE:

$$2m^2 - 3m - 5 = (2m - 5)(m + 1) = 0.$$

The roots are $m = 5/2$ and $m = -1$, whence the complementary function is

$$y_c = A\mathrm{e}^{\frac{5}{2}x} + B\mathrm{e}^{-x}.$$

For the particular solution, we try

$$y_p = Cx^2 + Dx + E.$$

Note that

$$y_p' = 2Cx + D \quad \text{and} \quad y_p'' = 2C.$$

Substituting this into the DE,

$$2(2C) + -3(2Cx + D) - 5\left(Cx^2 + Dx + E\right) = 10x^2 + 1.$$

Comparing coefficients, we get the system

$$\begin{cases} -5C & = 10 \\ -6C - 5D & = 0 \\ \ \ 4C - 3D - 5E & = 1 \end{cases},$$

which has solution $C = -2$, $D = 12/5$ and $E = -81/25$. The general solution is thus

$$y = y_c + y_p = A\mathrm{e}^{\frac{5}{2}x} + B\mathrm{e}^{-x} - 2x^2 + \frac{12}{5}x - \frac{81}{25}.$$

Part (b). Consider the characteristic equation of the DE:

$$m^2 - 2m + 3 = 0 \implies m = 1 \pm \sqrt{2}\mathrm{i}.$$

The complementary function is hence

$$y_c = \mathrm{e}^x\left(A\cos\sqrt{2}x + B\sin\sqrt{2}x\right).$$

For the particular solution, we try

$$y_p = C\mathrm{e}^{4x}.$$

Note that

$$y_p' = 4C\mathrm{e}^{4x} \quad \text{and} \quad y_p'' = 16C\mathrm{e}^{4x}.$$

Substituting this into the DE,

$$16C\mathrm{e}^{4x} - 2\left(4C\mathrm{e}^{4x}\right) + 3C\mathrm{e}^{4x}22\mathrm{e}^{4x} \implies C = 2.$$

The general solution is thus

$$y = y_c + y_p = \mathrm{e}^x\left(A\cos\sqrt{2}x + B\sin\sqrt{2}x\right) + 2\mathrm{e}^{4x}.$$

Part (c). Consider the characteristic equation of the DE:

$$m^2 - 2m + 1 = (m-1)^2 = 0.$$

The only root is $m = 1$, whence the complementary function is

$$s_c = (A + Bt)\mathrm{e}^t.$$

For the particular solution, we try

$$s_p = Ct^2\mathrm{e}^t.$$

Note that

$$s_p' = C\mathrm{e}^t\left(t^2 + 2t\right) \quad \text{and} \quad s_p'' = C\mathrm{e}^t\left(t^2 + 4t + 2\right).$$

Substituting this into the DE,

$$Ce^t\left(t^2+4t+2\right)-2Ce^t\left(t^2+2t\right)+Ce^t\left(t^2\right)=4e^t \implies C=2.$$

The general solution is thus

$$s=s_c+s_p=\left(A+Bt+2t^2\right)e^t.$$

Part (d). Consider the characteristic equation of the DE:

$$m^2+16m=0 \implies m=\pm 4i.$$

The complementary function is hence

$$x_c=A\cos 4t+B\sin 4t.$$

For the particular solution, we try

$$x_p=t\left(C\cos 4t+D\sin 4t\right).$$

Note that

$$x_p'=4t\left(-C\sin 4t+D\cos 4t\right)+\left(C\cos 4t+D\sin 4t\right)$$

and

$$x_p''=16t\left(-C\cos 4t-D\sin 4t\right)+8\left(-C\sin 4t+D\cos 4t\right).$$

Substituting this into the DE,

$$\begin{aligned}16t\left(-C\cos 4t-D\sin 4t\right)+8\left(-C\sin 4t+D\cos 4t\right)\\+16t\left(C\cos 4t+D\sin 4t\right)=3\cos 4\theta.\end{aligned}$$

Simplifying, we get

$$-8C\sin 4t+8D\cos 4t=3\cos 4t,$$

whence $C=0$ and $D=3/8$. Thus, the general solution is

$$x=A\cos 4t+B\sin 4t+\frac{3}{8}t\sin 4t.$$

* * * * *

Problem 4.

(a) Find the general solution of the differential equation $\frac{d^2y}{dx^2}-4y=10e^{3x}$.

(b) Hence, find the solution for which $y=-2$ and $\frac{dy}{dx}=-6$ when $x=0$.

Solution.

Part (a). Observe that the characteristic equation of the DE is $m^2-4=0$, whence the roots are $m=\pm 2$. Hence, the complementary function is

$$y_c=Ae^{2x}+Be^{-2x}.$$

For the particular solution, we try $y_p=Ce^{3x}$. Note that

$$y_p'=3Ce^{3x} \quad \text{and} \quad y_p''=9Ce^{3x}.$$

Substituting this into the DE, get

$$9Ce^{3x} - 4Ce^{3x} = 10e^{3x} \implies C = 2,$$

whence the general solution is

$$y = y_c + y_p = Ae^{2x} + Be^{-2x} + 2e^{3x}.$$

Part (b). Note that

$$\frac{dy}{dx} = 2Ae^{2x} - 2Be^{-2x} + 6e^{3x}.$$

The given conditions thus give the system

$$\begin{cases} A + \ \ B + 2 = -2 \\ 2A - 2B + 6 = -6 \end{cases},$$

whence $A = -5$ and $B = 1$. Hence,

$$y = -5e^{2x} + e^{-2x} + 2e^{3x}.$$

* * * * *

Problem 5.

(a) Find the general solution of the differential equation $\frac{d^2y}{dx^2} = \sin x$. Find the particular solution that passes through the points $(0, \sqrt{2})$ and $\left(\frac{\pi}{4}, -\sqrt{2}\right)$.

(b) Find the general solution of the differential equation

(i) $\frac{d^2y}{dx^2} = 16 - 9x^2$,

(ii) $\left(9 - x^2\right)^2 \frac{d^2y}{dx^2} - x = 0$,

giving your answer in the form $y = f(x)$.

Solution.

Part (a). Integrating the DE with respect to x,

$$\frac{dy}{dx} = \int \sin x \, dx = -\cos x + A.$$

Integrating once more,

$$y = \int (-\cos x + A) \, dx = -\sin x + Ax + B.$$

At $(0, \sqrt{2})$, we have $B = \sqrt{2}$. At $(\frac{\pi}{4}, -\sqrt{2})$, we have

$$-\frac{\sqrt{2}}{2} + A\left(\frac{\pi}{4}\right) + B = -\sqrt{2} \implies A = -\frac{6\sqrt{2}}{\pi}.$$

Thus, the particular solution is

$$y = -\sin x - \frac{6\sqrt{2}}{\pi}x + \sqrt{2}.$$

Part (b).

Part (b)(i). Integrating with respect to x,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \int \left(16 - 9x^2\right) \mathrm{d}x = 16x - 3x^2 + A.$$

Integrating once more,

$$y = \int \left(16x - 3x^2 + A\right) \mathrm{d}x = 8x^2 - \frac{3}{4}x^4 + Ax + B.$$

Part (b)(ii). Rewriting, we get

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{x}{\left(9 - x^2\right)^2}.$$

Integrating with respect to x,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \int \frac{x}{\left(9 - x^2\right)^2} \mathrm{d}x.$$

Using the substitution $x = 3\sin\theta$, we have

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} &= \int \frac{3\sin\theta}{81\cos^4\theta} 3\cos\theta \,\mathrm{d}\theta = \frac{1}{9}\int \tan\theta\sec^2\theta \,\mathrm{d}\theta \\ &= \frac{1}{9}\left(\frac{\tan^2\theta}{2}\right) + C = \frac{1}{18}\left(\frac{(x/3)^2}{1 - (x/3)^2}\right) + C = \frac{1}{18}\left(\frac{x^2}{9 - x^2}\right) + C.\end{aligned}$$

Integrating once more,

$$\begin{aligned}y &= \int \left[\frac{1}{18}\left(\frac{x^2}{9 - x^2}\right) + C\right] \mathrm{d}x = \int \left[\frac{1}{18}\left(\frac{9}{9 - x^2} - 1\right) + C\right] \mathrm{d}x \\ &= \frac{1}{18}\left[\frac{3}{2}\ln\left|\frac{3 + x}{3 - x}\right| - x\right] + Cx + D = \frac{1}{12}\ln\left|\frac{3 + x}{3 - x}\right| + Ex + D,\end{aligned}$$

where $E = -1/18 + C$.

* * * * *

Problem 6.

(a) Find the particular solution of $\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 16x = 0$, given that $x = 3$ and $\frac{\mathrm{d}x}{\mathrm{d}t} = -8$ when $t = 0$.

(b) By writing the particular solution as $R\cos(4t + \alpha)$, find the first positive value of t for which x is maximum.

Solution.

Part (a). Note that the characteristic equation of the DE is $m^2 + 16 = 0$, whence the roots are $m = \pm 4\mathrm{i}$. Hence,

$$x = A\cos 4t + B\sin 4\theta.$$

Differentiating with respect to t, we obtain

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -4A\sin 4t + 4B\cos 4t.$$

When $x = 3$ and $t = 0$, we have $A = 3$. When $\frac{\mathrm{d}x}{\mathrm{d}t} = -8$ and $t = 0$, we have $B = -2$. Thus,

$$x = 3\cos 4t - 2\sin 4t.$$

Part (b). We have

$$x = 3\cos 4t - 2\sin 4t = \sqrt{3^2+2^2}\cos\left(4t - \arctan\frac{-2}{3}\right) = \sqrt{13}\cos(4t + 0.58800)\,.$$

x attains a maximum whenever $\cos(4t + 0.58800) = 1$. Thus,

$$4t + 0.58800 = 2\pi n \implies t = \frac{2\pi n - 0.58800}{4},$$

where n is an integer. The first positive value of t is hence

$$t = \frac{2\pi - 0.58800}{4} = 1.42 \text{ (3 s.f.)},$$

which occurs when $n = 1$.

* * * * *

Problem 7. Using the substitution $x = \mathrm{e}^u$, find the general solution of

(a) $x^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 2x\frac{\mathrm{d}y}{\mathrm{d}x} - 2y = 0$,

(b) $x^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 5x\frac{\mathrm{d}y}{\mathrm{d}x} - 6y = 0$.

Solution. Note that

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

and

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\mathrm{d}^2y}{\mathrm{d}u^2}\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2 + \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}^2u}{\mathrm{d}x^2}.$$

Since $u = \ln x$, we have $\mathrm{d}u/\mathrm{d}x = 1/x$ and $\mathrm{d}^2u/\mathrm{d}x^2 = -1/x^2$. Thus,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{x}\frac{\mathrm{d}y}{\mathrm{d}u} \quad\text{and}\quad \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{1}{x^2}\frac{\mathrm{d}^2y}{\mathrm{d}u^2} - \frac{1}{x^2}\frac{\mathrm{d}y}{\mathrm{d}u}.$$

Part (a). Substituting the above expressions into the DE, we have

$$x^2\left(\frac{1}{x^2}\frac{\mathrm{d}^2y}{\mathrm{d}u^2} - \frac{1}{x^2}\frac{\mathrm{d}y}{\mathrm{d}u}\right) + 2x\left(\frac{1}{x}\frac{\mathrm{d}y}{\mathrm{d}u}\right) - 2y = 0.$$

Simplifying, we get

$$\frac{\mathrm{d}^2y}{\mathrm{d}u^2} + \frac{\mathrm{d}y}{\mathrm{d}u} - 2y = 0.$$

The characteristic equation $m^2 + m - 2 = (m+2)(m-1) = 0$ has roots $m = -2$ and $m = 1$. Thus,

$$y = A\mathrm{e}^{-2u} + B\mathrm{e}^{u} = Ax^{-2} + Bx.$$

Part (b). Substituting the above expressions into the DE, we have

$$x^2\left(\frac{1}{x^2}\frac{\mathrm{d}^2y}{\mathrm{d}u^2} - \frac{1}{x^2}\frac{\mathrm{d}y}{\mathrm{d}u}\right) - 5x\left(\frac{1}{x}\frac{\mathrm{d}y}{\mathrm{d}u}\right) - 6y = 0.$$

Simplifying, we get

$$\frac{\mathrm{d}^2y}{\mathrm{d}u^2} - 6\frac{\mathrm{d}y}{\mathrm{d}u} - 6y = 0.$$

The characteristic equation $m^2 - 6m - 6 = 0$ has roots $m = 3 \pm \sqrt{15}$. Thus,

$$y = A\mathrm{e}^{(3+\sqrt{15})u} + B\mathrm{e}^{(3-\sqrt{15})u} = Ax^{3+\sqrt{15}} + Bx^{3-\sqrt{15}}.$$

Problem 8. Show, by means of the substitution $y = x^{-4}z$, that the differential equation

$$x^2 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + \left(4x^2 + 8x\right) \frac{\mathrm{d}y}{\mathrm{d}x} + \left(3x^2 + 16x + 12\right) y = 0$$

can be reduced to the form

$$\frac{\mathrm{d}^2 z}{\mathrm{d}x^2} + a\frac{\mathrm{d}z}{\mathrm{d}x} + bz = 0,$$

where a and b are constants to be determined. Hence, find the general solution of the above differential equation.

Solution. Note that $z = x^4 y$. Differentiating with respect to x,

$$\frac{\mathrm{d}z}{\mathrm{d}x} = x^4 \frac{\mathrm{d}y}{\mathrm{d}x} + 4yx^3.$$

Differentiating with respect to x again,

$$\frac{\mathrm{d}^2 z}{\mathrm{d}x^2} = x^4 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 8x^3 \frac{\mathrm{d}y}{\mathrm{d}x} + 12yx^2.$$

Consider the DE in question. Multiplying through by x^2,

$$x^4 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + \left(4x^4 + 8x^3\right) \frac{\mathrm{d}y}{\mathrm{d}x} + \left(3x^4 + 16x^3 + 12x^2\right) y = 0.$$

Now observe that we can split the LHS as

$$\left(x^4 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 8x^3 \frac{\mathrm{d}y}{\mathrm{d}x} + 12yx^2\right) + 4\left(x^4 \frac{\mathrm{d}y}{\mathrm{d}x} + 4yx^3\right) + 3x^4 y.$$

Thus,

$$\frac{\mathrm{d}^2 z}{\mathrm{d}x^2} + 4\frac{\mathrm{d}z}{\mathrm{d}x} + 3z = 0.$$

Hence, $a = 4$ and $b = 3$.

Note that the characteristic equation of this new DE is $m^2 + 4m + 3 = (m+3)(m+1) = 0$. Thus, the roots are $m = -3$ and $m = -1$, whence

$$z = A\mathrm{e}^{-3x} + B\mathrm{e}^{-x} \implies y = x^{-4}\left(A\mathrm{e}^{-3x} + B\mathrm{e}^{-x}\right).$$

* * * * *

Problem 9. By letting $x = \sqrt{t}$, show that the differential equation

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + \left(2x - \frac{1}{x}\right) \frac{\mathrm{d}y}{\mathrm{d}x} + 24x^2 = 0$$

where $x > 0$, may be transformed to

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

where a and b are constants to be determined. Hence, find the general solution of y in terms of x.

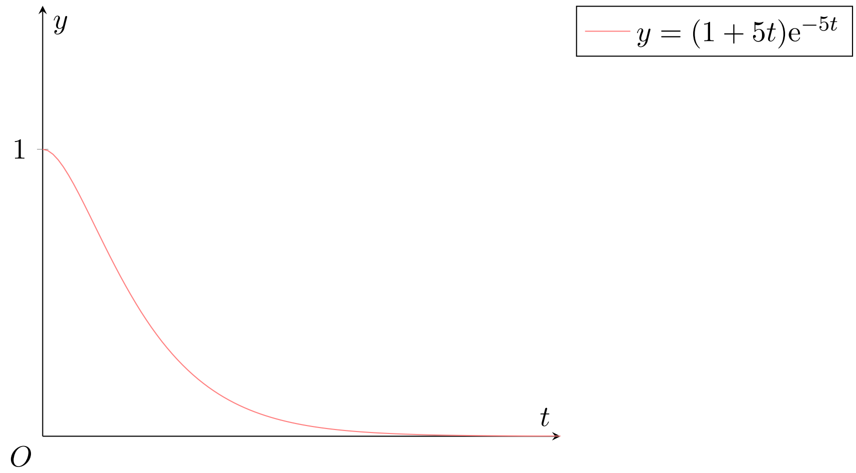
Solution. Note that
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x}$$
and
$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\mathrm{d}^2y}{\mathrm{d}t^2}\left(\frac{\mathrm{d}t}{\mathrm{d}x}\right)^2 + \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}^2t}{\mathrm{d}x^2}.$$
Since $t = x^2$, we have $\mathrm{d}t/\mathrm{d}x = 2x$ and $\mathrm{d}^2u/\mathrm{d}x^2 = 2$. Thus,
$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2x\frac{\mathrm{d}y}{\mathrm{d}t} \quad \text{and} \quad \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 4x^2\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 2\frac{\mathrm{d}y}{\mathrm{d}u}.$$
Substituting this into the given DE,
$$\left(4x^2\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 2\frac{\mathrm{d}y}{\mathrm{d}u}\right) + \left(2x - \frac{1}{x}\right)\left(2x\frac{\mathrm{d}y}{\mathrm{d}t}\right) + 24x^2 = 0.$$
Simplifying, we get
$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + \frac{\mathrm{d}y}{\mathrm{d}t} + 6 = 0,$$
whence $a = 1$ and $b = 6$.

Rewriting,
$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + \frac{\mathrm{d}y}{\mathrm{d}t} = -6.$$
Integrating with respect to t, we get
$$\frac{\mathrm{d}y}{\mathrm{d}t} + y = -6t + C.$$
Multiplying through by e^t yields
$$\mathrm{e}^t\frac{\mathrm{d}y}{\mathrm{d}t} + \mathrm{e}^ty = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^ty\right) = \mathrm{e}^t\left(-6t + C\right).$$
Integrating with respect to t,
$$\mathrm{e}^ty = \int \mathrm{e}^t\left(-6t + C\right)\,\mathrm{d}t = -6\left(t\mathrm{e}^t - \mathrm{e}^t\right) + C\mathrm{e}^t = \mathrm{e}^t\left(-6t + A\right) + B.$$
Thus,
$$y = -6t + A + B\mathrm{e}^{-t} = -6x^2 + A + B\mathrm{e}^{-x^2}.$$

* * * * *

Problem 10. A damped vibrating spring system is described by the differential equation
$$m\frac{\mathrm{d}^2y}{\mathrm{d}t^2} = -ky - \lambda\frac{\mathrm{d}y}{\mathrm{d}t},$$
where m, k and λ are positive constants. The variable y represents the displacement of the object from equilibrium position in centimetres, and t is time measured in seconds. Given that $m = 1$, $k = 25$ and $\lambda = 10$, and the object was initially released from rest at $y = 1$, find the equation of motion and sketch its graph. Briefly explain if this motion is suitable to be used to close a door.

Solution. We have
$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2}+10\frac{\mathrm{d}y}{\mathrm{d}t}+25y=0.$$
The characteristic equation $r^2+10r+25=(r+5)^2=0$ has a single root $r=-5$. Thus,
$$y=(A+Bt)\,\mathrm{e}^{-5t}.$$
Since $y=1$ when $t=0$, we get $A=1$. Hence,
$$y=(1+Bt)\,\mathrm{e}^{-5t}\implies\frac{\mathrm{d}y}{\mathrm{d}t}=B\mathrm{e}^{-5t}-5\,(1+Bt)\,\mathrm{e}^{-5t}.$$
Since the object is initially at rest, $\frac{\mathrm{d}y}{\mathrm{d}t}=0$ when $t=0$. This gives $B=5$. Thus,
$$y=(1+5t)\,\mathrm{e}^{-5t}.$$



Since the object does not oscillate (y does not change sign) and y approaches $y=0$ quite quickly, the motion is suitable to be used to close a door.

* * * * *

Problem 11. The motion of the tip of a tuning fork can be modelled by the differential equation
$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}+k\frac{\mathrm{d}x}{\mathrm{d}t}+m\omega^2x=0,$$
where x is the displacement of the tip from its equilibrium position at time t and m, k and ω are positive constants. It is known that k is so small that k^2 can be ignored as k models the slight damping due to the resistance of the air. It is given that the tip of the fork is initially in its equilibrium position and moving with speed v in the positive x-direction.

(a) Solve the differential equation.

The amplitude of a vibration is the maximum displacement of the tip from its equilibrium position and one period of a vibration is the time interval between the occurrences of two consecutive amplitudes.

(b) Comment on the period of the vibrations over time and show that the amplitude of successive vibrations follows a geometric progression.

(c) Given that k is no longer small and $k^2>4m^2\omega^2$, describe the behaviour of x as time progresses and sketch a possible graph of x against t. Justify your answer.

Solution.

Part (a). The characteristic equation of the DE is given by $mr^2 + kr + m\omega^2 = 0$. Let the roots be r_1 and r_2. We have

$$r_{1,2} = \frac{-k \pm \sqrt{k^2 - 4m^2\omega^2}}{2m}.$$

Since k^2 can be ignored,

$$r_{1,2} = \frac{-k \pm \sqrt{-4m^2\omega^2}}{2m} = \frac{-k \pm 2m\omega\mathrm{i}}{2m} = -\frac{k}{2m} \pm \omega\mathrm{i}.$$

The general solution is thus given by

$$x = \mathrm{e}^{-\frac{k}{2m}t}\left(A\cos\omega t + B\sin\omega t\right).$$

Since the object is initially at equilibrium, we have $x = 0$ at $t = 0$. There is hence no contribution from the cosine term, i.e. $A = 0$. Thus,

$$x = B\mathrm{e}^{-\frac{k}{2m}t}\sin\omega t.$$

Differentiating with respect to t,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = B\mathrm{e}^{-\frac{k}{2m}t}\left(\omega\cos\omega t - \frac{k}{2m}\sin\omega t\right).$$

Since the object was initially released with speed $v > 0$, we have $\mathrm{d}x/\mathrm{d}t = v$ at $t = 0$. Hence,

$$B\omega = v \implies B = \frac{v}{\omega}.$$

We hence obtain the solution

$$x = \frac{v}{\omega}\mathrm{e}^{-\frac{k}{2m}t}\sin\omega t.$$

Part (b). Let x_n be the (signed) amplitude of the nth vibration, and let t_n be the corresponding time, where $n \in \mathbb{N}$.

To find t_n, we consider the stationary points of x:

$$\left.\frac{\mathrm{d}x}{\mathrm{d}t}\right|_{t=t_n} = \frac{v}{\omega}\mathrm{e}^{-\frac{k}{2m}t_n}\left(\omega\cos\omega t_n - \frac{k}{2m}\sin\omega t_n\right) = 0 \implies \tan\omega t_n = \frac{2m\omega}{k}.$$

Since tangent has period π,

$$t_n = \frac{1}{\omega}\left(\arctan\frac{2m\omega}{k} + \pi n\right).$$

Quite clearly, x has a constant period $2\pi/\omega$.

We now find x_n. Evaluating x at t_n,

$$x_n = \frac{v}{\omega}\exp\left(-\frac{k}{2m\omega}\left[\arctan\frac{2m\omega}{k} + \pi n\right]\right)\sin\left(\arctan\frac{2m\omega}{k} + \pi n\right).$$

Note that $\sin(X + \pi n) = (-1)^n \sin X$. Hence,

$$x_n = \left[-\exp\left(-\frac{k\pi}{2m\omega}\right)\right]^n \underbrace{\left[\frac{v}{\omega}\exp\left(-\frac{k}{2m\omega}\arctan\frac{2m\omega}{k}\right)\sin\left(\arctan\frac{2m\omega}{k}\right)\right]}_{\text{constant}}.$$

Hence,

$$\frac{|x_{n+1}|}{|x_n|} = \mathrm{e}^{-\frac{k\pi}{2m\omega}},$$

whence the amplitudes $\{|x_n|\}$ are in geometric progression with common ratio $\mathrm{e}^{-\frac{k\pi}{2m\omega}}$.

Part (c). Recall that the roots of the characteristic equation are given by

$$r_{1,2} = \frac{-k \pm \sqrt{k^2 - 4m^2\omega^2}}{2m}.$$

If $k^2 > 4m^2\omega^2$, then the roots are real and distinct, whence x has general solution

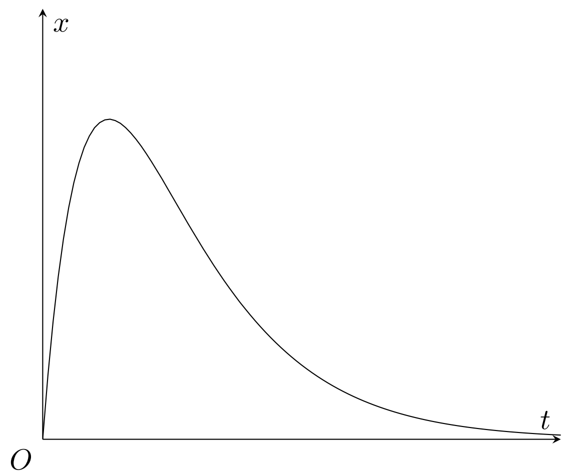
$$x = A\mathrm{e}^{r_1 t} + B\mathrm{e}^{r_2 t}.$$

Since $x = 0$ at $t = 0$, we obtain $A + B = 0$. Thus,

$$x = A\left(\mathrm{e}^{r_1 t} - \mathrm{e}^{r_2 t}\right).$$

Note that both roots are negative (since $\sqrt{k^2 - X^2} < \sqrt{k^2} = k$). Hence, as t tends to infinity, $\mathrm{e}^{r_1 t} - \mathrm{e}^{r_2 t}$ (and by extension x) tends to 0.

A possible graph of x is



Self-Practice B16

Problem 1. Find the general solution of the differential equation

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 4y = 24\mathrm{e}^{-2x}.$$

Show that when $x \to \infty$, the solution can be expressed as a single trigonometric expression.

* * * * *

Problem 2.

(a) By using the substitution $z = xy$, show that the differential equation

$$x\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + (2-4x)\frac{\mathrm{d}y}{\mathrm{d}x} + 4y(x-1) = 0$$

can be simplified into the differential equation

$$\frac{\mathrm{d}^2z}{\mathrm{d}x^2} - 4\frac{\mathrm{d}z}{\mathrm{d}x} + 4z = 0.$$

Hence, find the general solution for y in terms of x.

(b) Using a machine, a particle is accelerated from rest such that at a time t seconds after the machine is turned on, its displacement s from its initial starting point is modelled by the following differential equation:

$$\frac{\mathrm{d}^2s}{\mathrm{d}t^2} - 4\frac{\mathrm{d}s}{\mathrm{d}t} + 4s = \cos t.$$

Find s in terms of t. Hence, find the amount of time required for the particle's speed to exceed the speed of sound ($340\ \mathrm{ms}^{-1}$), giving your answer to the nearest hundredth of a second.

* * * * *

Problem 3. Given that $\frac{\mathrm{d}x}{\mathrm{d}t} = 5x - 9t + y$, $\frac{\mathrm{d}y}{\mathrm{d}t} = y - 4x + 9$, by eliminating the variable y, show that

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} - 6\frac{\mathrm{d}x}{\mathrm{d}t} + 9x = 9t.$$

Find the general solution of x in terms of t and hence obtain the general solution of y in terms of t. Find the ratio $x : y$ when $t \to -\infty$.

* * * * *

Problem 4. A teacher gave his class the following differential equation

$$x\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - \frac{\mathrm{d}y}{\mathrm{d}x} + 16x^3y = 8x^3 \tag{1}$$

and asked them to find the solution. One of the students, Adrian, who had come across (1) before recalled that the solution is $y = \cos^2(x^2)$.

(a) Show that $y = \cos^2(x^2)$ satisfies (1).

Another student, Bobby, decided to use the substitution $t = x^2$ to solve (1).

(b) Show that by using the substitution, (1) can be transformed into the equation

$$\frac{d^2y}{dt^2} + ay = b,$$

where a and b are constants to be determined.

(c) Find the general solution of (1).

(d) Show that Adrian's solution can be obtained from Bobby's solution be choosing suitable values of the arbitrary constants in the solution.

The teacher found out later that there was a typing error in (1). The differential equation should be

$$x\frac{d^2y}{dx^2} - \frac{dy}{dx} - 16x^3y = 8x^3. \tag{2}$$

Deduce the general solution of (2).

* * * * *

Problem 5. Use the substitution $x = \sec\theta$, where $0 < \theta < \pi/2$, to show that the differential equation

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

where k is a positive integer, can be reduced to

$$\frac{d^2y}{d\theta^2} + ky = 2\cos^2\theta.$$

Hence, obtain the general solution for the differential equation in x and y for $k \neq 4$ in the form

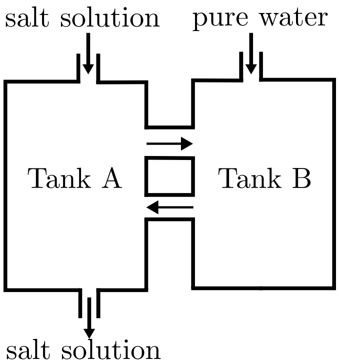
$$y = A\cos\sqrt{k}\,(\operatorname{arcsec} x) + B\sin\sqrt{k}\,(\operatorname{arcsec} k) + f(x),$$

where A and B are arbitrary constants and $f(x)$ is a function of x to be determined.

* * * * *

Problem 6. Two 50-litre tanks, Tank A and Tank B (as shown in the diagram below), containing salt solution are connected by two horizontal pipes. Both tanks have inlets and outlets where salt solution flows in and out of the tanks. The rates of flow in the inlets, outlets and pipes are managed in such a way that both tanks will be full at all times.

Tank A is receiving salt solution at a concentration of 1 gram per litre at a rate of 1 litre per minute and Tank B is receiving pure water at the same rate. Salt solution is flowing out from Tank A from the bottom outlet at a rate of 2 litres per minute. Salt solution flows from Tank A to Tank B through one of the horizontal pipes at a rate of 2 litres per minute and flows in the reverse direction through the other horizontal pipe at a rate of 3 litres per minute.



(a) Suppose that at time t minutes, the amount of salt in Tank A and Tank B are x grams and y grams respectively. Show that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{2}{25}x + \frac{3}{50}y + 1 \quad \text{and} \quad \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{25}x - \frac{3}{50}y.$$

(b) Prove that

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + p\frac{\mathrm{d}y}{\mathrm{d}t} + qy = r,$$

where p, q and r are constants to be determined.

(c) Find the general solution of the differential equation in (b).

(d) By expressing x in terms of t, find the ratio of the amount of salt in Tank A to the amount of salt in Tank B in the long run.

* * * * *

Problem 7 (🌶). A model for the vibrations of a wine glass when struck by an external force is

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \lambda\frac{\mathrm{d}x}{\mathrm{d}t} + \omega^2 x = 0,$$

where λ is a constant due to the external force, ω is a constant of the wine glass, $2\omega > \lambda > 0$, and x is the deformation of the glass.

(a) Find the general solution of the model in the form

$$x = \mathrm{e}^{At}\left(c_1 \cos Bt + c_2 \sin Bt\right),$$

where A and B are expressions to be determined in terms of λ and ω, and c_1 and c_2 are arbitrary constants.

Suppose that the wine glass vibrates at 440 Hz when struck, that is, the period of the oscillation is 1/440 second.

(b) Show that $\sqrt{4\omega^2 - \lambda^2} = 1760\pi$.

If it takes about 2 seconds for the sound to die away, and this happens when the original vibrations have reduced to one hundredth of their initial amplitude,

(c) show that $\lambda = \ln 100$ and hence find ω, correct to three significant figures.

A pure tone at 440 Hz is produced at D decibels and aimed at the glass, forcing it to vibrate at its natural frequency. The glass will shatter if the amplitude of the pure tone is approximately 1. The vibrations are now modelled by

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \lambda\frac{\mathrm{d}x}{\mathrm{d}t} + \omega^2 x = \frac{10^{D/10-8}}{3}\cos(880\pi t).$$

(d) Determine how loud the sound should be, i.e. how large D should be, in order to shatter the glass.

Assignment B16

Problem 1. The charge, Q coulombs, on the capacitor in an electrical circuit is governed by Kirchhoff's Second Law, which satisfies the differential equation

$$L\frac{\mathrm{d}^2Q}{\mathrm{d}t^2}+R\frac{\mathrm{d}Q}{\mathrm{d}t}+\frac{Q}{C}=V(t),$$

where L is the inductance (in henries), R the resistance (in ohms), C the capacitance (in farads) and $V(t)$ is the applied voltage (in volts).

The initial charge Q and initial current $\mathrm{d}Q/\mathrm{d}t$ in a circuit are both zero.

Given that $L=0.5$ henries, $R=10$ ohms, $C=0.02$ farads and $V(t)=50\mathrm{e}^{-10t}$, solve for the charge Q at time t and sketch your solution curve.

Solution. Substituting the given values of L, R, C and $V(t)$, the differential equation becomes

$$0.5\frac{\mathrm{d}^2Q}{\mathrm{d}t^2}+10\frac{\mathrm{d}Q}{\mathrm{d}t}+\frac{Q}{0.02}=50\mathrm{e}^{-10t},$$

which simplifies as

$$\frac{\mathrm{d}^2Q}{\mathrm{d}t^2}+20\frac{\mathrm{d}Q}{\mathrm{d}t}+100Q=100\mathrm{e}^{-10t}.$$

The characteristic equation $r^2+20r+100=(r+10)^2=0$ has the single root $r=-10$. Thus,

$$Q_c=(A+Bt)\,\mathrm{e}^{-10t}.$$

For the particular solution, we try

$$Q_p=Ct^2\mathrm{e}^{-10t} \implies \frac{\mathrm{d}Q_p}{\mathrm{d}t}=C\mathrm{e}^{-10t}\left(2t-10t^2\right) \implies \frac{\mathrm{d}^2Q_p}{\mathrm{d}t^2}=C\mathrm{e}^{-10t}\left(2-40t+100t^2\right).$$

Substituting this into the differential equation, we get

$$C\mathrm{e}^{-10t}\left(2-40t+100t^2\right)+20C\mathrm{e}^{-10t}\left(2t-10t^2\right)+100Ct^2\mathrm{e}^{-10t}=100\mathrm{e}^{-10t},$$

whence $C=50$ upon simplification. Thus,

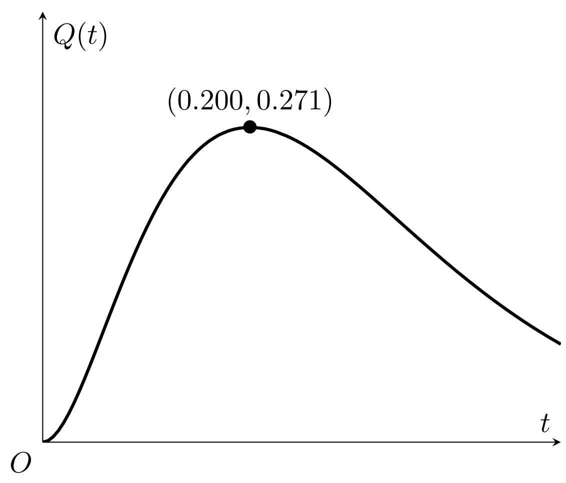
$$Q=Q_c+Q_p=(A+Bt)\,\mathrm{e}^{-10t}+50t^2\mathrm{e}^{-10t}.$$

Note that

$$\frac{\mathrm{d}Q}{\mathrm{d}t}=\mathrm{e}^{-10t}\left(B-10A-10Bt+100t-500t^2\right).$$

Since $Q=0$ and $\mathrm{d}Q/\mathrm{d}t=0$ at $t=0$, we get $A=0$ and $B=10A=0$. Hence,

$$Q=50t^2\mathrm{e}^{-10t}.$$



Problem 2. Given that y is a function of x and $x = \tan\theta$, show that

$$\left(1+x^2\right)\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}\theta}.$$

Hence, show that the differential equation

$$\left(1+x^2\right)^3\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 2\left(1+x^2\right)^2(1+x)\frac{\mathrm{d}y}{\mathrm{d}x} - 3\left(1+x^2\right)y = x^2+6x-1$$

can be expressed as

$$\frac{\mathrm{d}^2y}{\mathrm{d}\theta^2} + a\frac{\mathrm{d}y}{\mathrm{d}\theta} + by = c\sin 2\theta + d\cos 2\theta,$$

where a, b, c and d are constants to be determined.

Hence, find the general solution for y in terms of x.

Solution. Note that

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = \sec^2\theta = \tan^2\theta + 1 = x^2 + 1.$$

Thus,

$$\frac{\mathrm{d}y}{\mathrm{d}\theta} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}\theta} = \left(1+x^2\right)\frac{\mathrm{d}y}{\mathrm{d}x}.$$

Differentiating once more with respect to θ,

$$\frac{\mathrm{d}^2y}{\mathrm{d}\theta^2} = \left(1+x^2\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\frac{\mathrm{d}x}{\mathrm{d}\theta} + 2x\frac{\mathrm{d}x}{\mathrm{d}\theta}\frac{\mathrm{d}y}{\mathrm{d}x} = \left(1+x^2\right)^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 2x\frac{\mathrm{d}y}{\mathrm{d}\theta}.$$

Dividing the given DE by $1+x^2$, we have

$$\left(1+x^2\right)^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 2\left(1+x^2\right)(1+x)\frac{\mathrm{d}y}{\mathrm{d}x} - 3y = \frac{x^2+6x-1}{1+x^2}.$$

We can rewrite this as

$$\left(1+x^2\right)^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 2x\frac{\mathrm{d}y}{\mathrm{d}\theta} + 2\frac{\mathrm{d}y}{\mathrm{d}\theta} - 3y = \frac{x^2+6x-1}{1+x^2},$$

which quickly simplifies as

$$\frac{\mathrm{d}^2y}{\mathrm{d}\theta^2} + 2\frac{\mathrm{d}y}{\mathrm{d}\theta} - 3y = \frac{x^2+6x-1}{1+x^2}.$$

Now, observe that

$$\frac{x^2+6x-1}{1+x^2} = \frac{\tan^2\theta + 6\tan\theta - 1}{\sec^2\theta} = \sin^2\theta + 6\sin\theta\cos\theta - \cos^2\theta = 3\sin 2\theta - \cos 2\theta.$$

Thus,

$$\frac{\mathrm{d}^2y}{\mathrm{d}\theta^2} + 2\frac{\mathrm{d}y}{\mathrm{d}\theta} - 3y = 3\sin 2\theta - \cos 2\theta,$$

whence $a = 2$, $b = -3$, $c = 3$ and $d = -1$.

Note that the characteristic equation $r^2 + 2r - 3 = (r+3)(r-1) = 0$ has roots $r = 1$ and $r = -3$. Thus,

$$y_c = A\mathrm{e}^{\theta} + B\mathrm{e}^{-3\theta}.$$

For the particular solution, we try

$$y_p = C\sin 2\theta + D\cos 2\theta.$$

Note that

$$y_p' = 2C\cos 2\theta - 2D\sin 2\theta \quad \text{and} \quad y_p'' = -4C\sin 2\theta - 4D\cos 2\theta.$$

Substituting this into the DE, we get

$$\begin{aligned}(-4C\sin 2\theta - 4D\cos 2\theta) + 2\,(2C\cos 2\theta - 2D\sin 2\theta) - 3\,(C\sin 2\theta + D\cos 2\theta)\\ = 3\sin 2\theta - \cos 2\theta.\end{aligned}$$

Comparing coefficients, we get

$$\begin{cases}-7C - 4D = 3\\ \ \ 4C - 7D = -1\end{cases},$$

whence $C = -5/13$ and $D = -1/13$. Thus,

$$y = y_c + y_p = A\mathrm{e}^{\theta} + B\mathrm{e}^{-3\theta} - \frac{5}{13}\sin 2\theta - \frac{1}{13}\cos 2\theta.$$

Substituting $\theta = \arctan x$,

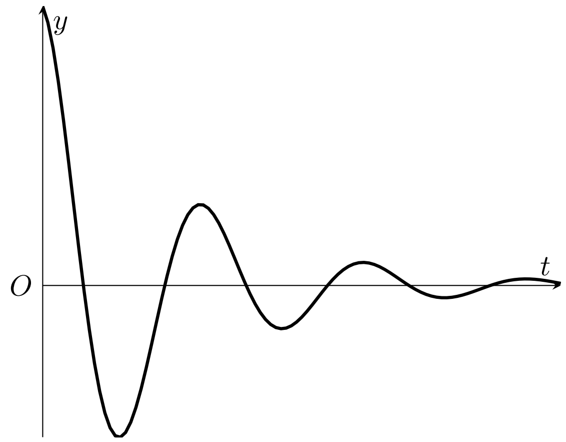
$$y = A\mathrm{e}^{\arctan x} + B\mathrm{e}^{-3\arctan x} - \frac{5}{13}\sin(2\arctan x) - \frac{1}{13}\cos(2\arctan x)\,.$$

* * * * *

Problem 3. An object of mass m, in kilograms, is suspended from one end of a vertical spring of elasticity k, $k > 0$, in a resistive medium with resistivity c, $c > 0$. When the object is pulled down from its equilibrium position and released, the motion of the object can be described by the following differential equation, where y is the displacement, in metres, of the object from the equilibrium position after time t.

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + \frac{c}{m}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{k}{m}y = 0.$$

(a) The diagram below shows how y varies with t for some values of c, m and k.



State the condition(s) for c, m and k for the above scenario.

When an external force, $F(t)$, is applied to the object, the motion of the object can be described by the following modified differential equation:

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + \frac{c}{m}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{k}{m}y = F(t).$$

(b) For the same c, m and k in part (a), sketch a graph of y versus t for

(i) $F(t) = a$,

(ii) $F(t) = b\mathrm{e}^{-t}$,

where a and b are positive constants, showing the long-term behaviour of y.

(c) In another setup, the resistivity is approximately equal to 0, that is $c = 0$. Given that the external force is $F(t) = \sin wt$, where $w^2 = k/m$, the differential equation is now

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + w^2 y = \sin wt.$$

Solve for y, in terms of w and t, if the object is initially at the equilibrium position with zero velocity.

Solution.

Part (a). Note that the characteristic equation of the DE is given by

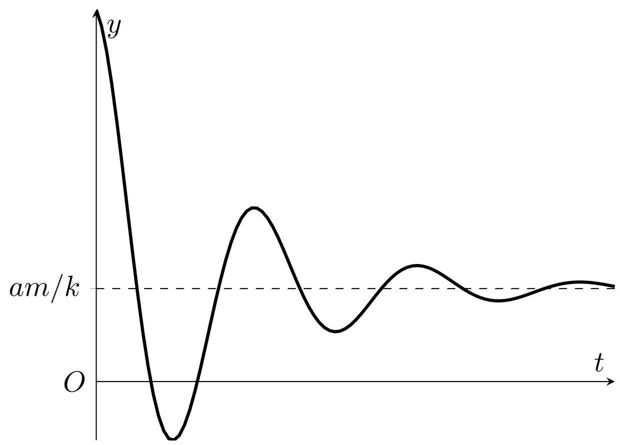
$$r^2 + \frac{c}{m}r + \frac{k}{m} = 0 \implies mr^2 + cr + k = 0.$$

For y to oscillate (i.e. composed of sine and cosine terms), the roots of the characteristic equation must be non-real. We hence obtain the constraint

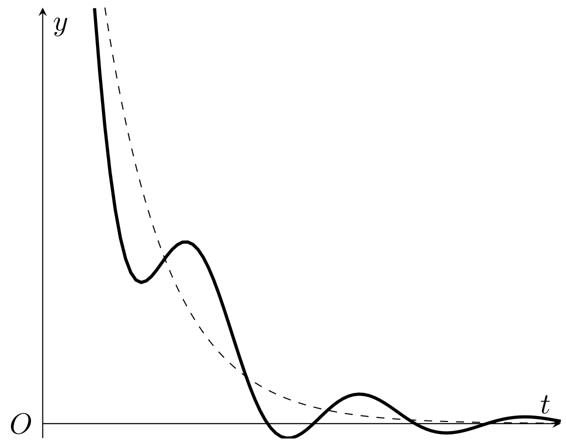
$$\Delta = c^2 - 4mk < 0.$$

Part (b). Note that the given graph represents the complementary solution.

Part (b)(i). Let the particular solution be a constant z. Substituting this into the DE, we get $kz/m = a$, whence $z = am/k$, which is a positive constant. Hence, we simply shift the graph of $y(t)$ in the positive y-axis.



Part (b)(ii). Since $F(t) = b\mathrm{e}^{-t}$, the particular solution is of the form $C\mathrm{e}^{-t}$. The resulting graph hence oscillates around $C\mathrm{e}^{-t}$:



Part (c). Note that the characteristic equation is $r^2 + r^2 = 0$, whence $r = \pm \mathrm{i}w$. Thus,

$$y_c = A\cos wt + B\sin wt.$$

For the particular solution, we try

$$y_p = Ct\sin wt + Dt\cos wt.$$

Note that

$$y_p' = wt\,(C\cos wt - D\sin wt) + C\sin wt + D\cos wt$$

and

$$y_p'' = w^2 t\,(-C\sin wt - D\cos wt) + 2w\,(C\cos wt - D\sin wt)\,.$$

Substituting this into the DE, we get

$$\begin{gathered} w^2 t\,(-C\sin wt - D\cos wt) + 2w\,(C\cos wt - D\sin wt) \\ +w^2\,(Ct\sin wt + Dt\cos wt) = \sin wt. \end{gathered}$$

Comparing coefficients, we obtain $C = 0$ and $D = -1/2w$. Thus,

$$y = y_c + y_p = A\cos wt + B\sin wt - \frac{t}{2w}\cos wt.$$

Note that

$$y' = -Aw\sin wt + Bw\cos wt - \frac{1}{2w}\,(\cos wt - tw\sin wt)\,.$$

Since $y(0) = y'(0) = 0$, we have $A = 0$ and

$$Bw - \frac{1}{2w} = 0 \implies B = \frac{1}{2w^2}.$$

Thus,

$$y = \frac{1}{2w^2}\sin wt - \frac{t}{2w}\cos wt.$$

B17A Linear Algebra - Matrices

Tutorial B17A

Problem 1. Without the use of G.C., find the following matrix products:

(a) $\begin{pmatrix} -1 & 4 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}$

(b) $\begin{pmatrix} 3 \\ 9 \\ 2 \end{pmatrix} \begin{pmatrix} 1 & -6 & 3 \end{pmatrix}$

(c) $\begin{pmatrix} 4 & -1 & 1 \\ 2 & 3 & 0 \end{pmatrix} \begin{pmatrix} 7 & -3 \\ 5 & 4 \\ 1 & 2 \end{pmatrix}$

(d) $\begin{pmatrix} 2 & 2 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} -2 & -3 & 4 \\ 2 & 2 & -3 \\ 1 & 2 & -2 \end{pmatrix}$

Solution.
Part (a).

$$\begin{pmatrix} -1 & 4 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \end{pmatrix}$$

Part (b).

$$\begin{pmatrix} 3 \\ 9 \\ 2 \end{pmatrix} \begin{pmatrix} 1 & -6 & 3 \end{pmatrix} = \begin{pmatrix} 3 & -18 & 9 \\ 9 & -54 & 27 \\ 2 & -12 & 6 \end{pmatrix}$$

Part (c).

$$\begin{pmatrix} 4 & -1 & 1 \\ 2 & 3 & 0 \end{pmatrix} \begin{pmatrix} 7 & -3 \\ 5 & 4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 24 & -14 \\ 29 & 6 \end{pmatrix}$$

Part (d).

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

* * * * *

Problem 2. An orthogonal matrix $\mathbf{M}$ has the property

$$\mathbf{M}\mathbf{M}^{\mathsf{T}} = \mathbf{M}^{\mathsf{T}}\mathbf{M} = \mathbf{I},$$

where $\mathbf{M}^{\mathsf{T}}$ and $\mathbf{I}$ denote the transpose of the matrix $\mathbf{M}$ and the identity matrix respectively.
Given that matrices $\mathbf{A}$ and $\mathbf{B}$ are orthogonal, are the following true or false?

(a) $\mathbf{AB}$ is orthogonal.

(b) $\mathbf{A}+\mathbf{B}$ is orthogonal.

Solution.

Part (a). The statement is true. Since $\mathbf{A}$ and $\mathbf{B}$ are both orthogonal,

$$(\mathbf{AB})^\mathsf{T}(\mathbf{AB}) = \mathbf{B}^\mathsf{T}\left(\mathbf{A}^\mathsf{T}\mathbf{A}\right)\mathbf{B} = \mathbf{B}^\mathsf{T}\mathbf{B} = \mathbf{I},$$

and

$$(\mathbf{AB})(\mathbf{AB})^\mathsf{T} = \mathbf{A}\left(\mathbf{B}\mathbf{B}^\mathsf{T}\right)\mathbf{A}^\mathsf{T} = \mathbf{A}\mathbf{A}^\mathsf{T} = \mathbf{I},$$

thus $\mathbf{AB}$ is orthogonal.

Geometric Approach. A real matrix is orthogonal if and only if it is norm-preserving. Since $\mathbf{A}$ and $\mathbf{B}$ are both orthogonal,

$$\|\mathbf{v}\| = \|\mathbf{Bv}\| = \|\mathbf{ABv}\|,$$

thus $\mathbf{AB}$ is also orthogonal.

Part (b). The statement is false. Let $\mathbf{M}$ be orthogonal. Then $-\mathbf{M}$ is also orthogonal (reflection preserves norm). However, their sum, $\mathbf{0}$, is clearly not orthogonal.

* * * * *

Problem 3. It is given that a matrix $\mathbf{A}$ is symmetric if and only if $\mathbf{A}^\mathsf{T} = \mathbf{A}$. Suppose that $\mathbf{A}$ is a symmetric $m \times m$ matrix and that $\mathbf{P}$ is any $m \times m$ matrix. Prove that $\mathbf{P}^\mathsf{T}\mathbf{AP}$ is symmetric.

Solution.

$$\left(\mathbf{P}^\mathsf{T}\mathbf{AP}\right)^\mathsf{T} = \mathbf{P}^\mathsf{T}\mathbf{A}^\mathsf{T}\left(\mathbf{P}^\mathsf{T}\right)^\mathsf{T} = \mathbf{P}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{P} = \mathbf{P}^\mathsf{T}\mathbf{AP}.$$

* * * * *

Problem 4. For what value(s) of the constant k does the following system of linear equations

$$\begin{cases} x - y = 3 \\ 2x - 2y = k \end{cases}$$

have

(a) no solutions?

(b) exactly one solution?

(c) infinitely many solutions?

Solution. Note that we have $2x - 2y = 6$ and $2x - 2y = k$.

Part (a). If $k \neq 6$, there are no solutions.

Part (b). It is impossible for the system to have a unique solution.

Part (c). If $k = 6$, there are infinitely many solutions.

* * * * *

Problem 5.

(a) Solve the following system of linear equations by using row operations to express the matrix representation of the following system of linear equations in row echelon form.

$$\begin{cases} x_1 + \ \ x_2 + \ \ x_3 = 8 \\ -x_1 - 2x_2 + 3x_3 = 1 \\ 3x_1 - 7x_2 + 4x_3 = 10 \end{cases}.$$

(b) Solve the following system of linear equations by using row operations to express the matrix representation of the following system of linear equations in reduced row echelon form.

$$\begin{cases} x + \ \ y + \ \ z = 0 \\ -2x + 5y + 2z = 0 \\ -7x + 7y + \ \ z = 0 \end{cases}.$$

Solution.

Part (a). Note that the given system of linear equations has matrix representation

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 8 \\ -1 & -2 & 3 & 1 \\ 3 & -7 & 4 & 10 \end{array}\right).$$

Performing row operations on this matrix, we get

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 8 \\ -1 & -2 & 3 & 1 \\ 3 & -7 & 4 & 10 \end{array}\right) \to \begin{array}{r} 2R_1+R_2 \\ -R_2-R_1 \\ -\frac{1}{39}(R_3-13R_1-10R_2) \end{array} \left(\begin{array}{ccc|c} 1 & 0 & 5 & 17 \\ 0 & 1 & -4 & -9 \\ 0 & 0 & 1 & \frac{8}{3} \end{array}\right).$$

We thus recover the system of linear equations

$$\begin{cases} x_1 \qquad + 4x_3 = 17 \\ \qquad x_2 - 4x_3 = -9 \\ \qquad\qquad x_3 = \frac{8}{3} \end{cases},$$

whence we have $x_1 = \frac{11}{3}$, $x_2 = \frac{5}{3}$ and $x_3 = \frac{8}{3}$.

Part (b). Note that the given system of linear equations has matrix representation

$$\begin{pmatrix} 1 & 1 & 1 \\ -2 & 5 & 2 \\ -7 & 7 & 1 \end{pmatrix}.$$

Performing row operations on this matrix, we get

$$\begin{pmatrix} 1 & 1 & 1 \\ -2 & 5 & 2 \\ -7 & 7 & 1 \end{pmatrix} \to \begin{array}{r} \frac{1}{7}(5R_1-R_2) \\ \frac{1}{7}(R_2+2R_1) \\ R_3+3R_1-2R_2 \end{array} \begin{pmatrix} 1 & 0 & \frac{3}{7} \\ 0 & 1 & \frac{4}{7} \\ 0 & 0 & 0 \end{pmatrix}.$$

Since the last row is full of zeroes, we have a free variable. Let $x_3 = t$, where $t \in \mathbb{R}$. Then we recover the system of linear equations

$$\begin{cases} x_1 \qquad + \frac{3}{7}x_3 = 0 \\ \qquad x_2 + \frac{4}{7}x_3 = 0 \\ \qquad\qquad x_3 = t \end{cases},$$

whence $x_1 = -\frac{3}{7}t$, $x_2 = -\frac{4}{7}t$ and $x_3 = t$.

Problem 6. What conditions must the b's satisfy in order for the following system of linear equations to be consistent?

(a) $\begin{cases} x_1 - x_2 + 3x_3 = b_1 \\ 3x_1 - 3x_2 + 4x_3 = b_2 \\ -2x_1 + 2x_2 - 6x_3 = b_3 \end{cases}$

(b) $\begin{cases} 2x_1 + 3x_2 - x_3 + x_4 = b_1 \\ x_1 + 5x_2 + x_3 - 2x_4 = b_2 \\ -x_1 + 2x_2 + 2x_3 - 3x_4 = b_3 \\ 3x_1 + x_2 - 3x_3 + 4x_4 = b_4 \end{cases}$

Solution.

Part (a). We can represent the given system of linear equations with the matrix

$$\left(\begin{array}{ccc|c} 1 & -1 & 3 & b_1 \\ 3 & -3 & 4 & b_2 \\ -2 & 2 & -6 & -b_3 \end{array}\right).$$

Performing row operations, we see that

$$\left(\begin{array}{ccc|c} 1 & -1 & 3 & b_1 \\ 3 & -3 & 4 & b_2 \\ -2 & 2 & -6 & -b_3 \end{array}\right) \to \begin{array}{c} \\ {\scriptstyle R_2-3R_1} \\ {\scriptstyle R_3+2R_1} \end{array} \left(\begin{array}{ccc|c} 1 & -1 & 3 & b_1 \\ 0 & 0 & -5 & b_2 - 3b_1 \\ 0 & 0 & 0 & b_3 + 2b_1 \end{array}\right).$$

For the system to be consistent, we require $b_3 + 2b_1 = 0$.

Part (b). We can represent the given system of linear equations with the matrix

$$\left(\begin{array}{cccc|c} 2 & 3 & -1 & 1 & b_1 \\ 1 & 5 & 1 & -2 & b_2 \\ -1 & 2 & 2 & -3 & b_3 \\ 3 & 1 & -3 & 4 & b_4 \end{array}\right).$$

Performing row operations, we see that

$$\left(\begin{array}{cccc|c} 2 & 3 & -1 & 1 & b_1 \\ 1 & 5 & 1 & -2 & b_2 \\ -1 & 2 & 2 & -3 & b_3 \\ 3 & 1 & -3 & 4 & b_4 \end{array}\right) \to \begin{array}{c} {\scriptstyle R_1-2R_2} \\ \\ {\scriptstyle R_3+R_1-R_2} \\ {\scriptstyle R_4-2R_1+R_2} \end{array} \left(\begin{array}{cccc|c} 0 & -7 & 3 & 5 & b_1 - 2b_2 \\ 1 & 5 & 1 & -2 & b_2 \\ 0 & 0 & 0 & 0 & b_3 + b_1 - b_2 \\ 0 & 0 & 0 & 0 & b_4 - 2b_1 + b_2 \end{array}\right).$$

For the system to be consistent, we require $b_3 + b_1 - b_2 = 0$ and $b_4 - 2b_1 + b_2 = 0$.

* * * * *

Problem 7. Without the use of a graphing calculator, find $\mathbf{A}^{-1}$ for each of the following cases of $\mathbf{A}$.

(a) $\begin{pmatrix} 2 & 3 & 1 \\ 3 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$

(b) $\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$

Solution.

Part (a). Performing row operations on $\left(\mathbf{A} \mid \mathbf{I}\right)$, we have

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 3 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 3 & 0 & 0 & 1 \end{array}\right) \to \begin{array}{r} \frac{1}{18}(R_1+7R_2-5R_3) \\ \frac{1}{18}(7R_1-5R_2+R_3) \\ \frac{1}{18}(-5R_1+R_2+7R_3) \end{array} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{18} & \frac{7}{18} & -\frac{5}{18} \\ 0 & 1 & 0 & \frac{7}{18} & -\frac{5}{18} & \frac{1}{18} \\ 0 & 0 & 1 & -\frac{5}{18} & \frac{1}{18} & \frac{7}{18} \end{array}\right).$$

Thus,

$$\mathbf{A}^{-1} = \frac{1}{18}\begin{pmatrix} 1 & 7 & -5 \\ 7 & -5 & 1 \\ -5 & 1 & 7 \end{pmatrix}.$$

Part (b). By the formula for the inverse of a 2×2 matrix,

$$\mathbf{A}^{-1} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Part (c). Observe that $\mathbf{A}$ represents a rotation of α about the x-axis. Thus, $\mathbf{A}^{-1}$ represents a rotation of $-\alpha$ about the x-axis, i.e.

$$\mathbf{A}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix}.$$

* * * * *

Problem 8. Without the use of a graphing calculator, find the determinants of the following matrices:

(a) $\mathbf{A} = \begin{pmatrix} 0 & 4 \\ -1 & 2 \end{pmatrix}$

(b) $\mathbf{B} = \begin{pmatrix} 2 & -1 & 4 \\ 4 & -3 & 1 \\ 1 & 2 & 1 \end{pmatrix}$

(c) $\mathbf{C} = \begin{pmatrix} 2 & 0 & 0 \\ 4 & -3 & 0 \\ 1 & 2 & 1 \end{pmatrix}$

(d) $\mathbf{D} = \begin{pmatrix} 2 & -1 & 4 \\ 4 & -3 & 1 \\ 3 & 6 & 3 \end{pmatrix}$

Solution.

Part (a).

$$\det\mathbf{A} = (0)(2) - (-1)(4) = 4.$$

Part (b).

$$\det\mathbf{B} = 2\begin{vmatrix} -3 & 1 \\ 2 & 1 \end{vmatrix} - (-1)\begin{vmatrix} 4 & 1 \\ 1 & 1 \end{vmatrix} + 4\begin{vmatrix} 4 & -3 \\ 1 & 2 \end{vmatrix} = 2(-5) - (-1)(3) + 4(11) = 37.$$

Part (c).

$$\det\mathbf{C} = \det\begin{vmatrix} 2 & 0 \\ 4 & -3 \end{vmatrix} = -6.$$

Part (d). Note that $\mathbf{D}$ is simply $\mathbf{B}$ where the third row has been multiplied by 3. Thus,

$$\det \mathbf{D} = 3 \det \mathbf{B} = 111.$$

* * * * *

Problem 9. For the case where $\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & 1 & 3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ -1 & 1 & -1 \end{pmatrix}$, verify the results

(a) $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$,

(b) $\det\left(\mathbf{A}^{-1}\right) = 1/\det(\mathbf{A})$.

Determine also if

(c) $\det(\mathbf{A} + \mathbf{B}) = \det(\mathbf{A}) + \det(\mathbf{B})$,

(d) $\det(\mathbf{A}) = \det\left(\mathbf{A}^\mathsf{T}\right)$.

Solution. Note that $\det A = -3$ and $\det B = 5$.

Part (a). Using G.C.,

$$\mathbf{AB} = \begin{pmatrix} 2 & 1 & 0 \\ 6 & 1 & 3 \\ -1 & 4 & -3 \end{pmatrix}.$$

Hence,

$$\det(\mathbf{AB}) = -15 = (-3)(5) = \det(\mathbf{A}) \det(\mathbf{B}).$$

Part (b). Using G.C.,

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

Hence,

$$\det\left(\mathbf{A}^{-1}\right) = -\frac{1}{3} = \frac{1}{-3} = \frac{1}{\det \mathbf{A}}.$$

Part (c). Note that

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 3 & -1 \\ -1 & 2 & 2 \end{pmatrix}.$$

Hence,

$$\det(\mathbf{A} + \mathbf{B}) = 21 \neq -3 + 5 = \det \mathbf{A} + \det \mathbf{B}.$$

Part (d). Note that

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

Hence,

$$\det \mathbf{A}^\mathsf{T} = -3 = \det \mathbf{A}.$$

Problem 10. It is given that matrices $\mathbf{A} = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 0 & 4 \\ 3 & 1 & 0 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 1 & 12 \\ 3 & 1 & 0 \end{pmatrix}$. Without the use of the G.C., find the inverse of $\mathbf{A}$ and $\mathbf{B}$ if it exists. For each of (a) and (b) below, solve, if possible, the equation, giving your answers in terms of a (where applicable).

(a) $\mathbf{Ax} = (4,\ 1,\ a)^\mathsf{T}$,

(b) $\mathbf{Bx} = (4,\ 1,\ a)^\mathsf{T}$.

Hence, determine whether it is possible for $\mathbf{x}$ to have a unique solution when

$$\mathbf{ABx} = \begin{pmatrix} 4 \\ 1 \\ a \end{pmatrix}.$$

Solution. Performing row operations on $\left(\mathbf{A} \mid \mathbf{I}\right)$, we have

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 3 & 1 & 0 & 0 \\ -1 & 0 & 4 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \to \begin{array}{r} -4R_1+3R_2+4R_3 \\ 12R_1-9R_2-11R_3 \\ -R_1+R_2+R_3 \end{array} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & 3 & 4 \\ 0 & 1 & 0 & 12 & -9 & -11 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right).$$

Thus,

$$\mathbf{A}^{-1} = \begin{pmatrix} -4 & 3 & 4 \\ 12 & -9 & -11 \\ -1 & 1 & 1 \end{pmatrix}.$$

Note that

$$\det \mathbf{B} = 3 \begin{vmatrix} -1 & 1 \\ 3 & 1 \end{vmatrix} - 12 \begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} = 0.$$

Hence, $\mathbf{B}^{-1}$ does not exist.

Part (a). Since $\mathbf{A}$ is invertible, we can pre-multiply $\mathbf{A}^{-1}$ on both sides of the equation $\mathbf{Ax} = (4,\ 1,\ a)^\mathsf{T}$, yielding

$$\mathbf{x} = \begin{pmatrix} -4 & 3 & 4 \\ 12 & -9 & -11 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ a \end{pmatrix} = \begin{pmatrix} -13+4a \\ 39-11a \\ -3+a \end{pmatrix}.$$

Part (b). Note that the equation $\mathbf{Bx} = (4,\ 1,\ a)^\mathsf{T}$ has matrix representation

$$\left(\begin{array}{ccc|c} 2 & 1 & 3 & 4 \\ -1 & 1 & 12 & 1 \\ 3 & 1 & 0 & a \end{array}\right).$$

Performing Gaussian elimination yields

$$\left(\begin{array}{ccc|c} 2 & 1 & 3 & 4 \\ -1 & 1 & 12 & 1 \\ 3 & 1 & 0 & a \end{array}\right) \to \begin{array}{r} \frac{1}{3}R_1+\frac{2}{3}R_2 \\ \frac{1}{3}R_1-\frac{1}{3}R_2 \\ -\frac{4}{3}R_1+\frac{1}{3}R_2+R_3 \end{array} \left(\begin{array}{ccc|c} 0 & 1 & 9 & 2 \\ 1 & 0 & -3 & 1 \\ 0 & 0 & 0 & a-5 \end{array}\right).$$

If $a \neq 5$, the system is inconsistent and there is no solution. If $a = 5$, the system is consistent with one free variable. Let $\mathbf{x} = (x_1,\ x_2,\ x_3)^\mathsf{T}$, with $x_3 = t$, where $t \in \mathbb{R}$. We have

$$\begin{cases} \quad x_2 + 9x_3 = 2 \\ x_1 \qquad\ \ - 3x_3 = 1 \end{cases},$$

whence $x_1 = 1 + 3t$, $x_2 = 2 - 9t$ and $x_3 = t$. Thus,

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ -9 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}.$$

Since $\mathbf{A}$ is invertible,

$$\mathbf{ABx} = \begin{pmatrix} 4 \\ 1 \\ a \end{pmatrix} \implies \mathbf{Bx} = \mathbf{A}^{-1} \begin{pmatrix} 4 \\ 1 \\ a \end{pmatrix}.$$

However, because $\mathbf{B}$ is not invertible, there cannot be a unique solution $\mathbf{x}$ to the above equation (if such a solution exists).

* * * * *

Problem 11. Let $\mathbf{Ax} = \mathbf{0}$ be a homogeneous system of n linear equations in n unknowns that has only the trivial solution. Show that if k is any positive integer, then the system $\mathbf{A}^k\mathbf{x} = \mathbf{0}$ also has only the trivial solution.

Solution. Since $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution, $\det(\mathbf{A}) \neq 0$. Thus, $\det\left(\mathbf{A}^k\right) = \det(\mathbf{A})^k \neq 0$, whence $\mathbf{A}^k\mathbf{x} = \mathbf{0}$ has a unique solution, which must clearly be the trivial solution.

* * * * *

Problem 12.

(a) Let $\mathbf{A}$ be a non-zero square matrix such that $\mathbf{A}^2 = \mathbf{A}$. Determine all possible values of $\det(\mathbf{A})$. Determine if the following statements are true. Justify your answer.

  (i) $\mathbf{I} - \mathbf{A}$ is always invertible.

  (ii) $\mathbf{I} + \mathbf{A}$ is always invertible.

(b) Let $\mathbf{B} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Given that $\mathbf{B}$ is the inverse of a matrix $\mathbf{C}$, and $\mathbf{D}$ is the matrix obtained from $\mathbf{C}$ by adding to the second row of $\mathbf{C}$ twice the first row of $\mathbf{C}$, find $\mathbf{D}^{-1}$ in a similar form to $\mathbf{B}$.

Solution.

Part (a). Taking determinants on both sides of the given equation, we have

$$\det\left(\mathbf{A}^2\right) = \det(\mathbf{A}) \implies \det(\mathbf{A})^2 = \det(\mathbf{A}).$$

Thus, the possible values of $\det(\mathbf{A})$ are 0 and 1. An example of $\mathbf{A}$ with determinant 0 is

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

while an example of $\mathbf{A}$ with determinant 1 is simply $\mathbf{I}$.

In fact, $\mathbf{I}$ is the only such matrix that has determinant 1: if $\mathbf{A}$ is invertible, we can pre-multiply its inverse to the equation $\mathbf{A}^2 = \mathbf{A}$ to get

$$\mathbf{A}^{-1}\mathbf{A}^2 = \mathbf{A}^{-1}\mathbf{A} \implies \mathbf{A} = \mathbf{I}.$$

Geometric Approach. Let $\mathbf{A}$ represent a linear transformation of $\mathbb{R}^n$. $\mathbf{A}$ is idempotent if and only if all vectors $\mathbf{v} \in \operatorname{col} A$ are invariant under $\mathbf{A}$. If $\mathbf{A}$ has non-zero determinant, then its image is $\mathbb{R}^n$, i.e. $\mathbf{Av} = \mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^n$, which is only possible if $\mathbf{A} = \mathbf{I}$.

Part (a)(i). The statement is false. Observe that if $\mathbf{A}$ is idempotent, then $\mathbf{I}-\mathbf{A}$ is also idempotent:

$$(\mathbf{I}-\mathbf{A})^2=\mathbf{I}^2-\mathbf{IA}-\mathbf{AI}+\mathbf{A}^2=\mathbf{I}-\mathbf{A}-\mathbf{A}+\mathbf{A}=\mathbf{I}-\mathbf{A}.$$

Seeking a contradiction, suppose $\mathbf{I}-\mathbf{A}$ is invertible. From the above result, it follows that

$$\mathbf{I}-\mathbf{A}=\mathbf{I} \implies \mathbf{A}=\mathbf{0},$$

which contradicts our assumption that $\mathbf{A}$ is non-zero. Hence, $\mathbf{I}-\mathbf{A}$ is not invertible.
Alternative Approach. Let $\mathbf{v}\in\operatorname{col}\mathbf{A}$. Then

$$(\mathbf{I}-\mathbf{A})\mathbf{v}=\mathbf{v}-\mathbf{A}\mathbf{v}=\mathbf{v}-\mathbf{v}=\mathbf{0} \implies \operatorname{col}(\mathbf{A})\subseteq\ker(\mathbf{I}-\mathbf{A}).$$

If $\mathbf{I}-\mathbf{A}$ is invertible, then its kernel must be trivial. This means that

$$\operatorname{col}(\mathbf{A})=\ker(\mathbf{I}-\mathbf{A})=\{\mathbf{0}\} \implies \mathbf{A}=\mathbf{0},$$

contradicting our assumption that $\mathbf{A}$ is non-zero. Hence, $\mathbf{I}-\mathbf{A}$ is not invertible.
Part (a)(ii). The statement is true. Let $\mathbf{B}=\mathbf{I}+\mathbf{A}$. Since $\mathbf{A}^2=\mathbf{A}$,

$$\mathbf{0}=\mathbf{A}^2-\mathbf{A}=(\mathbf{B}-\mathbf{I})^2-(\mathbf{B}-\mathbf{I})=\mathbf{B}^2-3\mathbf{B}+2\mathbf{I}.$$

Rearranging,

$$\mathbf{I}=\frac{3}{2}\mathbf{B}-\frac{1}{2}\mathbf{B}^2=\mathbf{B}\left(\frac{3}{2}\mathbf{I}-\frac{1}{2}\mathbf{B}\right)=(\mathbf{I}+\mathbf{A})\left(\mathbf{I}-\frac{1}{2}\mathbf{A}\right).$$

Thus, the inverse of $\mathbf{I}+\mathbf{A}$ exists and is given by $\mathbf{I}-\frac{1}{2}\mathbf{A}$.
Alternative Approach. Let $\mathbf{v}\in\ker(\mathbf{I}+\mathbf{A})$. Then

$$(\mathbf{I}+\mathbf{A})\,\mathbf{v}=\mathbf{v}+\mathbf{A}\mathbf{v}=\mathbf{0} \implies \mathbf{A}\mathbf{v}=-\mathbf{v}. \tag{1}$$

Pre-multiplying both sides by $\mathbf{A}$,

$$\mathbf{A}^2\mathbf{v}=-\mathbf{A}\mathbf{v} \implies \mathbf{A}\mathbf{v}=-\mathbf{A}\mathbf{v} \implies \mathbf{A}\mathbf{v}=\mathbf{0}.$$

Substituting this into (1), we have $\mathbf{v}=\mathbf{0}$. Hence, the kernel of $\mathbf{I}+\mathbf{A}$ is trivial, thus $\mathbf{I}+\mathbf{A}$ is invertible.
Part (b). Note that

$$\mathbf{D}=\begin{pmatrix}1&0&0\\2&1&0\\0&0&1\end{pmatrix}\mathbf{C}=\begin{pmatrix}1&0&0\\2&1&0\\0&0&1\end{pmatrix}\mathbf{B}^{-1}.$$

Thus, our goal matrix $\mathbf{D}^{-1}$ is given by

$$\begin{aligned}\mathbf{D}^{-1}&=\left[\begin{pmatrix}1&0&0\\2&1&0\\0&0&1\end{pmatrix}\mathbf{B}^{-1}\right]^{-1}=\mathbf{B}\begin{pmatrix}1&0&0\\2&1&0\\0&0&1\end{pmatrix}^{-1}\\&=\begin{pmatrix}a&b&c\\d&e&f\\g&h&i\end{pmatrix}\begin{pmatrix}1&0&0\\-2&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}a-2b&b&c\\d-2e&e&f\\g-2h&h&i\end{pmatrix}.\end{aligned}$$

Self-Practice B17A

Problem 1. Find the 2×2 matrix $\mathbf{T}$ such that

$$\begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix} = \mathbf{T} \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix}.$$

Solution. We have

$$\mathbf{T} = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & \frac{1}{2} \\ 3 & 0 \end{pmatrix}.$$

* * * * *

Problem 2. If

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

show that the roots of the equation $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$ are $\lambda = -1$ and $\lambda = 4$.

Solution. Note that

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \det \begin{pmatrix} 1 - \lambda & 3 \\ 2 & 2 - \lambda \end{pmatrix} = \lambda^2 - 3\lambda - 4 = (\lambda + 1)(\lambda - 4),$$

so the roots are $\lambda = -1$ and $\lambda = 4$.

* * * * *

Problem 3. Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}.$$

(a) Find the elementary matrices $\mathbf{E}_1$ and $\mathbf{E}_2$ such that $\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}$.

(b) Write $\mathbf{A}^{-1}$ as a product of two elementary matrices.

(c) Write $\mathbf{A}$ as a product of two elementary matrices.

Solution.

Part (a). We can reduce $\mathbf{A}$ to the identity matrix $\mathbf{I}$ using two elementary row operations:

$$\begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} \to_{R_2 - 3R_1} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \to_{\frac{1}{4}R_2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The two elementary matrices representing these operations are given by

$$\mathbf{E}_1 = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{E}_2 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}.$$

Part (b). From $\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}$, we have

$$\mathbf{A}^{-1} = \mathbf{E}_2\mathbf{E}_1 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}.$$

Part (c). Taking the inverse of the previous part, we have

$$\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1} = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

Problem 4. Given that the matrix $\mathbf{A}$ is singular, where

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ a & -1 & -11 \\ -2 & a & 12 \end{pmatrix},$$

find the possible values of a. For each of these values, determine the number of solutions to the equation

$$\mathbf{A}\mathbf{x} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

If there are infinitely many solutions for a particular value of a, give the general solution.

Solution. Note that

$$\det \mathbf{A} = \begin{vmatrix} -1 & -11 \\ a & 12 \end{vmatrix} - 2\begin{vmatrix} 9 & -11 \\ -2 & 12 \end{vmatrix} + \begin{vmatrix} a & -1 \\ -2 & a \end{vmatrix} = a^2 - 13a + 30 = (a-10)(a-3).$$

For $\mathbf{A}$ to be singular, its determinant must be 0, so we have $a = 10$ or $a = 3$.

Case 1: $a = 10$. The equation can be represented by the following augmented matrix, which we reduce to its RREF:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 10 & -1 & -11 & 2 \\ -2 & 10 & 12 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

From the last row, we see that the system is inconsistent, so there are no solutions.

Case 2: $a = 3$. The equation can be represented by the following augmented matrix, which we reduce to its RREF:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 3 & -1 & -11 & 2 \\ -2 & 3 & 12 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & -3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Since the last row is full of zeroes, there are infinitely many solutions.

Let $\mathbf{x} = (x, y, z)^\mathsf{T}$. Then

$$\begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \implies \begin{cases} x \quad - 3z = 1 \\ \quad y + 2z = 1 \end{cases}.$$

Let $z = \lambda \in \mathbb{R}$. Then the general solution is given by

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1+3\lambda \\ 1-2\lambda \\ \lambda \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda\begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix}.$$

* * * * *

Problem 5. Given that

$$\mathbf{Y} = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \quad \text{and} \quad \mathbf{Y}\mathbf{Y}^\mathsf{T} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

show that, if p, q, r, s are real, they all lie in the interval $\left[-\sqrt{2}, \sqrt{2}\right]$.

Solution. Note that

$$\mathbf{YY}^{\mathsf{T}} = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} p & r \\ q & s \end{pmatrix} = \begin{pmatrix} p^2+q^2 & pr+qs \\ pr+qs & r^2+s^2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

so we have $p^2+q^2=2$ and $r^2+s^2=2$. It immediately follows that $p,q,r,s \in \left[-\sqrt{2}, \sqrt{2}\right]$.

* * * * *

Problem 6. Let $\mathbf{Ax}=\mathbf{0}$ be a homogeneous system of n linear equations in n unknowns, and let $\mathbf{Q}$ be an invertible matrix. Show that $\mathbf{Ax}=\mathbf{0}$ has just the trivial solution if and only if $\mathbf{QAx}=\mathbf{0}$ has just the trivial solution.

Solution. Suppose $\mathbf{Ax}=\mathbf{0}$ has just the trivial solution. Then $\mathbf{A}$ is invertible, so $\det \mathbf{A} \neq 0$, whence $\det(\mathbf{QA}) = \det(\mathbf{Q})\det(\mathbf{A}) \neq 0$. Thus, $\mathbf{QA}$ is also invertible, so $\mathbf{QAx}=\mathbf{0}$ has just the trivial solution.

Suppose $\mathbf{QAx}=\mathbf{0}$ has just the trivial solution. Then $\mathbf{QA}$ is invertible, so $\det(\mathbf{QA}) \neq 0$. Since $\mathbf{Q}$ is invertible, we have $\det(\mathbf{Q}) \neq 0$, whence it follows that $\det(\mathbf{A}) \neq 0$, so $\mathbf{A}$ is invertible and $\mathbf{Ax}=\mathbf{0}$ has just the trivial solution.

* * * * *

Problem 7. Let

$$\left(\begin{array}{ccc|c} a & 0 & b & 2 \\ a & a & 4 & 4 \\ 0 & a & 2 & b \end{array}\right)$$

be the augmented matrix for a linear system. For what values of a and b does the system have

- a unique solution?
- a one-parameter solution?
- a two-parameter solution?
- no solution?

Solution. Note that

$$\left(\begin{array}{ccc|c} a & 0 & b & 2 \\ a & a & 4 & 4 \\ 0 & a & 2 & b \end{array}\right) \to \begin{array}{r} \\ {\scriptstyle R_2-R_1} \\ {\scriptstyle R_3+R_1-R_2} \end{array} \left(\begin{array}{ccc|c} a & 0 & b & 2 \\ 0 & a & 4-b & 2 \\ 0 & 0 & b-2 & b-2 \end{array}\right).$$

Case 1: $a \neq 0$, $b \neq 2$. Note that

$$\det \begin{pmatrix} a & 0 & b \\ a & a & 4 \\ 0 & a & 2 \end{pmatrix} = a^2(b-2) \neq 0,$$

so the matrix is invertible and we obtain a unique solution.

Case 2: $a=0$, $b=2$. Our matrix can be further reduced to

$$\left(\begin{array}{ccc|c} a & 0 & b & 2 \\ 0 & a & 4-b & 2 \\ 0 & 0 & b-2 & b-2 \end{array}\right) = \left(\begin{array}{ccc|c} 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Since there are two rows of zeroes, we have a two-parameter solution.

Case 3: $a = 0$, $b \neq 2$. Our matrix can be further reduced to

$$\left(\begin{array}{ccc|c} a & 0 & b & 2 \\ 0 & a & 4-b & 2 \\ 0 & 0 & b-2 & b-2 \end{array}\right) = \left(\begin{array}{ccc|c} 0 & 0 & b & 2 \\ 0 & 0 & 4-b & 2 \\ 0 & 0 & b-2 & b-2 \end{array}\right) \to \left(\begin{array}{ccc|c} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & b-2 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Since $b - 2 \neq 0$, the second row is inconsistent, so there are no solutions in this case.

Case 4: $a \neq 0$, $b = 2$. Substituting $b = 2$ into the reduced augmented matrix, we see that

$$\left(\begin{array}{ccc|c} a & 0 & b & 2 \\ 0 & a & 4-b & 2 \\ 0 & 0 & b-2 & b-2 \end{array}\right) = \left(\begin{array}{ccc|c} a & 0 & 2 & 2 \\ 0 & a & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

There is only one row of zeroes, so we have a one-parameter solution in this case.

To summarize,

Solution	a	b
None	$a = 0$	$b \neq 2$
Unique	$a \neq 0$	$b \neq 2$
One-parameter	$a \neq 0$	$b = 2$
Two-parameter	$a = 0$	$b = 2$

* * * * *

Problem 8. The matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{pmatrix} 6 & 3 & 2 \\ 3 & 2 & 1 \\ 8 & 4 & 3 \end{pmatrix}.$$

By performing row-operations on the matrix $\left(\mathbf{A} \mid \mathbf{I}\right)$, find $\mathbf{A}^{-1}$. Hence, or otherwise, find $\mathbf{B}^{-1}$ where

$$\mathbf{B} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{8} \end{pmatrix}.$$

Given that the real numbers x_1, x_2 and x_3 satisfy the equation

$$\mathbf{B} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix},$$

show that the solution of the equation

$$\mathbf{Bx} = \begin{pmatrix} c_1 + \delta \\ c_2 - \delta \\ c_3 - \delta \end{pmatrix}$$

is

$$\mathbf{x} = \begin{pmatrix} x_1 + 42\delta \\ x_2 - 18\delta \\ x_3 - 96\delta \end{pmatrix}.$$

Solution. We have

$$\left(\begin{array}{ccc|ccc} 6 & 3 & 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 1 & 0 \\ 8 & 4 & 3 & 0 & 0 & 1 \end{array}\right) \to \begin{array}{r} {\scriptstyle 2R_1 - R_2 - R_3} \\ {\scriptstyle 2R_2 - 2R_1} \\ {\scriptstyle 3R_3 - 4R_1} \end{array} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -1 & -1 \\ 0 & 1 & 0 & -1 & 2 & 0 \\ 0 & 0 & 1 & -4 & 0 & 3 \end{array}\right),$$

so

$$\mathbf{A}^{-1} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 0 \\ -4 & 0 & 3 \end{pmatrix}.$$

Note that

$$\mathbf{B} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{8} \end{pmatrix} = \begin{pmatrix} \frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{6} & 0 \\ 0 & 0 & \frac{1}{24} \end{pmatrix} \mathbf{A},$$

so

$$\mathbf{B}^{-1} = \mathbf{A}^{-1} \begin{pmatrix} \frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{6} & 0 \\ 0 & 0 & \frac{1}{24} \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 0 \\ -4 & 0 & 3 \end{pmatrix} \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 24 \end{pmatrix} = \begin{pmatrix} 12 & -6 & -24 \\ -6 & 12 & 0 \\ -24 & 0 & 72 \end{pmatrix}.$$

Note that

$$\mathbf{Bx} = \begin{pmatrix} c_1 + \delta \\ c_2 - \delta \\ c_3 - \delta \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \delta \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \mathbf{B} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \delta \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$

Pre-multiplying $\mathbf{B}^{-1}$ on both sides,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \delta \mathbf{B}^{-1} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \delta \begin{pmatrix} 42 \\ -18 \\ -96 \end{pmatrix} = \begin{pmatrix} x_1 + 42\delta \\ x_2 - 18\delta \\ x_3 - 96\delta \end{pmatrix}.$$

* * * * *

Problem 9. Given that

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 4 & a \\ -5 & a & 11 \end{pmatrix},$$

find the values of a for which the equation $\mathbf{Ax} = \mathbf{b}$ does not have exactly one solution where $\mathbf{x}$ and $\mathbf{b}$ are 3×1 matrices.

Using each of these values of a, find the solutions, if any, of the equation

$$\mathbf{Ax} = \begin{pmatrix} 1 \\ 4 \\ -3 \end{pmatrix}.$$

Solution. Note that

$$\det \mathbf{A} = \begin{vmatrix} 4 & a \\ a & 11 \end{vmatrix} - 2 \begin{vmatrix} 5 & a \\ -5 & 11 \end{vmatrix} + 3 \begin{vmatrix} 5 & 4 \\ -5 & a \end{vmatrix} = -a^2 + 5a - 6 = -(a-3)(a-2).$$

For $\mathbf{Ax} = \mathbf{b}$ to not have a unique solution, we require $\det \mathbf{A} = 0$, so $a = 2$ or $a = 3$.

Case 1: $a = 2$. The equation can be represented by the following augmented matrix, which we reduce to its RREF:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 5 & 4 & 2 & 4 \\ -5 & 2 & 11 & -3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & -\frac{4}{3} & \frac{2}{3} \\ 0 & 1 & \frac{13}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Let $\mathbf{x} = (x,\, y,\, z)^{\mathsf{T}}$. Then the system becomes

$$\begin{cases} x \quad - \frac{4}{3}z \quad = \frac{2}{3} \\ \quad y \quad + \frac{13}{6} = \frac{1}{6} \end{cases}.$$

Let $z = \lambda \in \mathbb{R}$. Then

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2/3 \\ 1/6 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 4/3 \\ -13/6 \\ 1 \end{pmatrix}.$$

Case 2: $a = 3$. The equation can be represented by the following augmented matrix, which we reduce to its RREF:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 5 & 4 & 3 & 4 \\ -5 & 3 & 11 & -3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

From the last row, we see that the system is inconsistent, so it has no solutions.

Assignment B17A

Problem 1. The equations of three planes p, q and r are

$$\begin{cases} 2x + \ \ y + 3z = 4 \\ 8x + 6y + 5z = \mu \\ -4x + 8y + \lambda z = 7 \end{cases}$$

respectively, where λ and μ are constants.

Determine the conditions on λ and μ such that the three planes

(a) intersect at exactly one point,

(b) intersect at a line,

(c) have no point in common.

Solution. The system of equations can be rewritten as the matrix equation

$$\begin{pmatrix} 2 & 1 & 3 \\ 8 & 6 & 5 \\ -4 & 8 & \lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ \mu \\ 7 \end{pmatrix}.$$

The augmented matrix corresponding to this equation is row-equivalent to

$$\left(\begin{array}{ccc|c} 2 & 1 & 3 & 4 \\ 8 & 6 & 5 & \mu \\ -4 & 8 & \lambda & 7 \end{array}\right) \to \begin{array}{r} \\ {\scriptstyle R_2-4R_1} \\ {\scriptstyle R_3+22R_1-5R_2} \end{array} \left(\begin{array}{ccc|c} 2 & 1 & 3 & 4 \\ 0 & 2 & -7 & \mu-16 \\ 0 & 0 & \lambda+41 & 95-5\mu \end{array}\right).$$

Part (a). For the three planes to intersect at exactly one point, the system of equations must have a unique solution. Hence, there must be no row of 0's, whence $\lambda + 41 = 0$. Thus, $\lambda \neq -41$ and $\mu \in \mathbb{R}$.

Part (b). For the three planes to intersect at a line, the system of equations must have infinitely many solutions. Hence, there are must a consistent row of 0's. This gives $\lambda + 41 = 0$ and $95 - 5\mu = 0$, whence $\lambda = -41$ and $\mu = 19$.

Part (c). For the three planes to have no common point, the system of equations must be inconsistent. Thus, $\lambda + 41 = 0$ and $95 - 5\mu \neq 0$, whence $\lambda = -41$ and $\mu \neq 19$.

* * * * *

Problem 2. An $n \times n$ matrix $\mathbf{A}$ is said to be an *involutory matrix* if $\mathbf{A}^2 = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. It is an *idempotent matrix* if $\mathbf{A}^2 = \mathbf{A}$.

(a) Find the possible values of the determinant of an involutory matrix.

(b) State the expression of $\mathbf{A}^{2n+1}$ where $n \in \mathbb{Z}^+$, where $\mathbf{A}$ is an involutory matrix.

(c) Prove that $\mathbf{A}$ is an involutory matrix if and only if $\frac{1}{2}(\mathbf{A} + \mathbf{I})$ is idempotent.

Solution.

Part (a). Taking determinants on both sides,

$$\det \mathbf{A}^2 = \det \mathbf{I} \implies (\det \mathbf{A})^2 = 1 \implies \det \mathbf{A} = \pm 1.$$

Part (b). Clearly,

$$\mathbf{A}^{2n+1} = \mathbf{A}\left[(\mathbf{A})^2\right]^n = \mathbf{A}\mathbf{I}^n = \mathbf{A}.$$

Part (c). Suppose $\mathbf{A}$ is involutory. Then $\mathbf{A}^2 = \mathbf{I}$. Consider $\left[\frac{1}{2}(\mathbf{A}+\mathbf{I})\right]^2$:

$$\left[\frac{1}{2}(\mathbf{A}+\mathbf{I})\right] = \frac{1}{4}\left(\mathbf{A}^2 + \mathbf{AI} + \mathbf{IA} + \mathbf{I}^2\right) = \frac{1}{4}\left(\mathbf{A}^2 + 2\mathbf{A} + \mathbf{I}\right)$$
$$= \frac{1}{4}(\mathbf{I} + 2\mathbf{A} + \mathbf{I}) = \frac{1}{2}(\mathbf{A}+\mathbf{I}).$$

Thus, $\frac{1}{2}(\mathbf{A}+\mathbf{I})$ is idempotent.

Suppose that $\frac{1}{2}(\mathbf{A}+\mathbf{I})$ is idempotent. Then

$$\left[\frac{1}{2}(\mathbf{A}+\mathbf{I})\right]^2 = \frac{1}{2}(\mathbf{A}+\mathbf{I}) \implies (\mathbf{A}+\mathbf{I})^2 = 2(\mathbf{A}+\mathbf{I})$$
$$\implies \mathbf{A}^2 + 2\mathbf{A} + \mathbf{I} = 2\mathbf{A} + 2\mathbf{I} \implies \mathbf{A}^2 = \mathbf{I},$$

hence $\mathbf{A}$ is involutory.

Thus, $\mathbf{A}$ is involutory if and only if $\frac{1}{2}(\mathbf{A}+\mathbf{I})$ is idempotent.

* * * * *

Problem 3. The matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{pmatrix} -5 & 4 & 3 \\ 10 & -7 & -6 \\ -8 & 6 & 5 \end{pmatrix}.$$

(a) By performing row operations on the matrix $\left(\mathbf{A} \mid \mathbf{I}\right)$, find $\mathbf{A}^{-1}$.

(b) Solve the equation $\mathbf{xA} = \begin{pmatrix} -1 & 2 & 3 \end{pmatrix}$, where $\mathbf{x}$ is a 1×3 matrix.

(c) Solve, by multiplying both sides of the equation by $\mathbf{A}^{-1}$, the equation

$$\begin{pmatrix} x & y & z & t \end{pmatrix} \begin{pmatrix} -1 & 2 & 3 \\ -5 & 4 & 3 \\ 10 & -7 & -6 \\ -8 & 6 & 5 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 \end{pmatrix}.$$

Solution.

Part (a). Performing row operations on the augmented matrix $\left(\mathbf{A} \mid \mathbf{I}\right)$, we have

$$\left(\begin{array}{ccc|ccc} -5 & 4 & 3 & 1 & 0 & 0 \\ 10 & -7 & -6 & 0 & 1 & 0 \\ -8 & 6 & 5 & 0 & 0 & 1 \end{array}\right) \to \begin{array}{r} {\scriptstyle -R_1+2R_2+3R_3} \\ {\scriptstyle R_2+2R_1} \\ {\scriptstyle 5R_3-4R_1+2R_2} \end{array} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 2 & 3 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -4 & 2 & 5 \end{array}\right).$$

Thus,

$$\mathbf{A}^{-1} = \begin{pmatrix} -1 & 2 & 3 \\ 2 & 1 & 0 \\ -4 & 2 & 5 \end{pmatrix}.$$

Part (b). Using G.C., we have $\mathbf{x} = \begin{pmatrix} -7 & 6 & 12 \end{pmatrix}$.

Part (c). Post-multiplying the given equation with $\mathbf{A}^{-1}$, we have

$$\begin{pmatrix} x & y & z & t \end{pmatrix} \begin{pmatrix} -1 & 2 & 3 \\ & \mathbf{A} & \end{pmatrix} \mathbf{A}^{-1} = \begin{pmatrix} 2 & -2 & 1 \end{pmatrix} \mathbf{A}^{-1}.$$

Block-multiplying the LHS yields

$$\begin{pmatrix} x & y & z & t \end{pmatrix} \begin{pmatrix} -7 & 6 & 12 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -10 & 4 & 11 \end{pmatrix}.$$

This gives the system of linear equations

$$\begin{cases} -7x + y & & & = -10 \\ 6x & + z & & = 4 \\ 12x & & + t & = 11 \end{cases}.$$

Let $x = \lambda$, where $\lambda \in \mathbb{R}$. Then

$$x = \lambda, \quad y = -10 + 7\lambda, \quad z = 4 - 6\lambda, \quad t = 11 - 12\lambda.$$

B17B Linear Algebra - Linear Spaces

Tutorial B17B

Problem 1. Determine which of the following functions are linear transformations, and if they are, find the matrix representing the linear transformation.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by
$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 2x - 3y \\ 3x + 4y \end{pmatrix}.$$

(b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by
$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 + 1 \\ x_2 - 2 \\ x_3 \end{pmatrix}.$$

(c) $T : \mathbb{R}^2 \to \mathbb{R}$ given by
$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \sqrt{x^2 + y^2}.$$

Solution.
Part (a). Note that
$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = x\begin{pmatrix} 2 \\ 3 \end{pmatrix} + y\begin{pmatrix} -3 \\ 4 \end{pmatrix}.$$

Hence,
$$\begin{aligned} T\left(\begin{pmatrix} u_x \\ u_y \end{pmatrix} + \begin{pmatrix} v_x \\ v_y \end{pmatrix}\right) &= T\left(\begin{pmatrix} u_x + v_x \\ u_y + v_y \end{pmatrix}\right) = (u_x + v_x)\begin{pmatrix} 2 \\ 3 \end{pmatrix} + (u_y + v_y)\begin{pmatrix} -3 \\ 4 \end{pmatrix} \\ &= \left[u_x\begin{pmatrix} 2 \\ 3 \end{pmatrix} + u_y\begin{pmatrix} -3 \\ 4 \end{pmatrix}\right] + \left[v_x\begin{pmatrix} 2 \\ 3 \end{pmatrix} + v_y\begin{pmatrix} -3 \\ 4 \end{pmatrix}\right] = T\left(\begin{pmatrix} u_x \\ u_y \end{pmatrix}\right) + T\left(\begin{pmatrix} v_x \\ v_y \end{pmatrix}\right). \end{aligned}$$

Let $k \in \mathbb{R}$. Then
$$T\left(k\begin{pmatrix} x \\ y \end{pmatrix}\right) = T\left(\begin{pmatrix} kx \\ ky \end{pmatrix}\right) = kx\begin{pmatrix} 2 \\ 3 \end{pmatrix} + ky\begin{pmatrix} -3 \\ 4 \end{pmatrix} = k\left[x\begin{pmatrix} 2 \\ 3 \end{pmatrix} + y\begin{pmatrix} -3 \\ 4 \end{pmatrix}\right] = kT\left(\begin{pmatrix} x \\ y \end{pmatrix}\right).$$

Thus, T is a linear transformation.
Part (b). Note that
$$T\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}.$$

Let $k \in \mathbb{R}$. Then
$$T\left(k\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right) = T\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \neq k\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = kT\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right).$$

Thus, T is not a linear transformation.

Part (c). Note that

$$T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)+T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)=\sqrt{0^2+1^2}+\sqrt{1^2+0^2}=2.$$

However,

$$T\left(\begin{pmatrix}1\\1\end{pmatrix}\right)=\sqrt{1^2+1^2}=\sqrt{2}.$$

Thus,

$$T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)+T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)=2\neq\sqrt{2}=T\left(\begin{pmatrix}1\\1\end{pmatrix}\right)=T\left(\begin{pmatrix}0\\1\end{pmatrix}+\begin{pmatrix}1\\0\end{pmatrix}\right).$$

Thus, T is not a linear transformation.

* * * * *

Problem 2. Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be the linear transformation such that

$$T\left(\begin{pmatrix}1\\1\end{pmatrix}\right)=\begin{pmatrix}0\\2\end{pmatrix}\quad\text{and}\quad T\left(\begin{pmatrix}1\\-1\end{pmatrix}\right)=\begin{pmatrix}2\\0\end{pmatrix}.$$

(a) Compute $T((1,\,4)^{\mathsf{T}})$ and $T((-2,\,1)^{\mathsf{T}})$.

(b) Find the matrix representing the linear transformation T.

Solution.

Part (a). Note that

$$T\left(\begin{pmatrix}1\\0\end{pmatrix}\right)=T\left(\frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}+\frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}\right)=\frac{1}{2}T\left(\begin{pmatrix}1\\1\end{pmatrix}\right)+\frac{1}{2}T\left(\begin{pmatrix}1\\-1\end{pmatrix}\right)=\begin{pmatrix}1\\1\end{pmatrix}.$$

Also note that

$$T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)=T\left(\frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}-\frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}\right)=\frac{1}{2}T\left(\begin{pmatrix}1\\1\end{pmatrix}\right)-\frac{1}{2}T\left(\begin{pmatrix}1\\-1\end{pmatrix}\right)=\begin{pmatrix}-1\\1\end{pmatrix}.$$

Thus,

$$T\left(\begin{pmatrix}1\\4\end{pmatrix}\right)=T\left(\begin{pmatrix}1\\0\end{pmatrix}+4\begin{pmatrix}0\\1\end{pmatrix}\right)=T\left(\begin{pmatrix}1\\0\end{pmatrix}\right)+4T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)=\begin{pmatrix}-3\\5\end{pmatrix}$$

and

$$T\left(\begin{pmatrix}-2\\1\end{pmatrix}\right)=T\left(-2\begin{pmatrix}1\\0\end{pmatrix}+\begin{pmatrix}0\\1\end{pmatrix}\right)=-2T\left(\begin{pmatrix}1\\0\end{pmatrix}\right)+T\left(\begin{pmatrix}0\\1\end{pmatrix}\right)=\begin{pmatrix}-3\\-1\end{pmatrix}.$$

Part (b). The matrix representing T is

$$\begin{pmatrix}1&-1\\1&1\end{pmatrix}.$$

* * * * *

Problem 3. In each part, determine whether the given set of vectors span $\mathbb{R}^3$.

(a) $\left\{\begin{pmatrix}2\\2\\2\end{pmatrix},\begin{pmatrix}0\\0\\3\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix}\right\}$,

(b) $\left\{\begin{pmatrix}1\\2\\6\end{pmatrix}, \begin{pmatrix}3\\4\\1\end{pmatrix}, \begin{pmatrix}4\\3\\1\end{pmatrix}, \begin{pmatrix}3\\3\\1\end{pmatrix}\right\}$.

Solution.

Part (a). Let $(a, b, c)^\mathsf{T} \in \mathbb{R}^3$ be an arbitrary vector. Consider

$$x_1 \begin{pmatrix}2\\2\\2\end{pmatrix} + x_2 \begin{pmatrix}0\\0\\3\end{pmatrix} + x_3 \begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}a\\b\\c\end{pmatrix}. \tag{1}$$

We can rewrite this equation as

$$\begin{pmatrix}2 & 0 & 0\\2 & 0 & 1\\2 & 3 & 1\end{pmatrix} \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}a\\b\\c\end{pmatrix}.$$

However, observe that

$$\det \begin{pmatrix}2 & 0 & 0\\2 & 0 & 1\\2 & 3 & 1\end{pmatrix} = -6 \neq 0.$$

Thus, there will always be solutions x_1, x_2 and x_3 that satisfy (1). Hence, the set of vectors spans $\mathbb{R}^3$.

Part (b). Clearly,

$$\operatorname{span}\left\{\begin{pmatrix}1\\2\\6\end{pmatrix}, \begin{pmatrix}3\\4\\1\end{pmatrix}, \begin{pmatrix}4\\3\\1\end{pmatrix}, \begin{pmatrix}3\\3\\1\end{pmatrix}\right\} \supseteq \operatorname{span}\left\{\begin{pmatrix}1\\2\\6\end{pmatrix}, \begin{pmatrix}3\\4\\1\end{pmatrix}, \begin{pmatrix}4\\3\\1\end{pmatrix}\right\}.$$

Let $(a, b, c)^\mathsf{T} \in \mathbb{R}^3$ be an arbitrary vector. Consider

$$x_1 \begin{pmatrix}1\\2\\6\end{pmatrix} + x_2 \begin{pmatrix}3\\4\\1\end{pmatrix} + x_3 \begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}a\\b\\c\end{pmatrix}, \tag{2}$$

which we can rewrite as

$$\begin{pmatrix}1 & 3 & 4\\2 & 4 & 3\\6 & 1 & 1\end{pmatrix} \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}a\\b\\c\end{pmatrix}.$$

Since

$$\det \begin{pmatrix}1 & 3 & 4\\2 & 4 & 3\\6 & 1 & 1\end{pmatrix} = -39 \neq 0,$$

there will always be solutions x_1, x_2, x_3 that satisfy (2). Hence, the set of vectors span $\mathbb{R}^3$.

* * * * *

Problem 4. Determine whether the set of vectors are linearly independent.

(a) $\left\{\begin{pmatrix}3\\1\\4\end{pmatrix}, \begin{pmatrix}2\\-3\\5\end{pmatrix}, \begin{pmatrix}5\\-2\\9\end{pmatrix}, \begin{pmatrix}1\\4\\-1\end{pmatrix}\right\}$,

(b) $\left\{ \begin{pmatrix} 2 \\ -4 \\ 6 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ -5 \end{pmatrix}, \begin{pmatrix} -2 \\ 5 \\ 8 \end{pmatrix} \right\}$.

Solution.

Part (a). Since the set consists of 4 vectors in $\mathbb{R}^3$ space, it is not linearly independent.

Part (b). Consider

$$x_1 \begin{pmatrix} 2 \\ -4 \\ 6 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 3 \\ -5 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 5 \\ 8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{1}$$

We can rewrite this as

$$\begin{pmatrix} 2 & -1 & -2 \\ -4 & 3 & 5 \\ 6 & -5 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since

$$\det \begin{pmatrix} 2 & -1 & -2 \\ -4 & 3 & 5 \\ 6 & -5 & 8 \end{pmatrix} = 32 \neq 0,$$

the only solution to (1) is the trivial solution $x_1 = x_2 = x_3 = 0$. Thus, the set of vectors is linearly independent.

* * * * *

Problem 5.

(a) Show that the value of

$$\det \begin{pmatrix} -2 & 1 & 2k \\ -1 & 1 & k+1 \\ 2 & k-1 & 1 \end{pmatrix}$$

is independent of k.

(b) State, with a reason, whether the vectors

$$\begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix}$$

are linearly independent.

(c) (i) State, with a reason, whether the system of equations

$$\begin{cases} -2x + y + 6z = 1 \\ -x + y + 4z = 0 \\ 2x + 2y + z = -2 \end{cases}$$

is consistent.

(ii) The three equations given in part (c)(i) are the Cartesian equations of three planes. Describe the geometrical relationship of the three planes.

Solution.

Part (a). Expanding along column 2,

$$\det\begin{pmatrix}-2 & 1 & 2k\\ -1 & 1 & k+1\\ 2 & k-1 & 1\end{pmatrix} = -\begin{vmatrix}-1 & k+1\\ 2 & 1\end{vmatrix} + \begin{vmatrix}-1 & 2k\\ 2 & 1\end{vmatrix} - (k-1)\begin{vmatrix}-2 & 2k\\ -1 & k+1\end{vmatrix}$$
$$= (2k+3) + (-2-4k) + (2k-2) = -1.$$

Part (b). Taking $k = 2$, we see that

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

is invertible (since its determinant is $-1 \neq 0$). Thus,

$$x_1\begin{pmatrix}-2\\ -1\\ 2\end{pmatrix} + x_2\begin{pmatrix}1\\ 1\\ 1\end{pmatrix} + x_3\begin{pmatrix}4\\ 3\\ 1\end{pmatrix} = \begin{pmatrix}-2 & 1 & 4\\ -1 & 1 & 3\\ 2 & 1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} = \begin{pmatrix}0\\ 0\\ 0\end{pmatrix}$$

has only the trivial solution $x_1 = x_2 = x_3 = 0$. Thus, the vectors

$$\begin{pmatrix}-2\\ 1\\ 2\end{pmatrix}, \begin{pmatrix}1\\ 1\\ 1\end{pmatrix}, \begin{pmatrix}4\\ 3\\ 1\end{pmatrix}$$

are linearly independent.

Part (c).

Part (c)(i). The system of equations can be rewritten as

$$\begin{pmatrix}-2 & 1 & 6\\ -1 & 1 & 4\\ 2 & 2 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}1\\ 0\\ -1\end{pmatrix}.$$

Taking $k = 3$, we see that

$$\begin{pmatrix}-2 & 1 & 6\\ -1 & 1 & 4\\ 2 & 2 & 1\end{pmatrix}$$

is invertible (since its determinant is $-1 \neq 0$). Thus, the system is consistent and has a unique solution.

Part (c)(ii). The three planes intersect at a single common point.

* * * * *

Problem 6. Find a basis for the row space and a basis for the column space of the matrix

$$\mathbf{A} = \begin{pmatrix}2 & 1 & 3 & 3\\ 0 & -3 & 1 & -2\\ 4 & 5 & 5 & 8\end{pmatrix}.$$

State the rank of $\mathbf{A}$.

Solution. Observe that the RREF of $\mathbf{A}$ is

$$\begin{pmatrix} 1 & 0 & 5/3 & 7/6 \\ 0 & 1 & -1/3 & 2/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, a basis for the row space of $\mathbf{A}$ is simply

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 5/3 \\ 7/6 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1/3 \\ 2/3 \end{pmatrix} \right\},$$

while a basis for the column space of $\mathbf{A}$ is

$$\left\{ \begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix} \right\}.$$

The rank of $\mathbf{A}$ is given by

$$\operatorname{rank} \mathbf{A} = \dim \operatorname{range} \mathbf{A} = \dim \operatorname{col} \mathbf{A} = 2.$$

* * * * *

Problem 7. In this question, V denotes the set of vectors of the form $(a, b, c, d)^{\mathsf{T}}$, where a, b, c and d are real numbers. You may assume that V forms a linear space under the usual operations of vector addition and multiplication by scalar.

(a) Show that the subset of V for which $a + b + c + d = 0$ forms a linear space.

(b) Show that the subset of V for which $a + b + c + d = 1$ does not form a linear space.

(c) Determine whether the subset for which $a + b = c + d$ and $a + 2b = c + 3d$ forms a linear space.

(d) State the dimension of the linear space defined in part (a) and provide a basis for this linear space.

Solution. Let Π be the null space of a matrix $\mathbf{A}$.

- Π contains the zero vector: $\mathbf{A0} = \mathbf{0}$.
- Π is closed under addition: for any $\mathbf{v}_1, \mathbf{v}_2 \in \Pi$,

$$\mathbf{A}(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{A}\mathbf{v}_1 + \mathbf{A}\mathbf{v}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

- Π is closed scalar multiplication: for any $\mathbf{v} \in \Pi$ and $k \in \mathbb{R}$,

$$\mathbf{A}(k\mathbf{v}) = k\mathbf{A}\mathbf{v} = k\mathbf{0} = \mathbf{0}.$$

Thus, null spaces are linear spaces.

Part (a). Let V_1 be the subset for which $a + b + c + d = 0$. Then

$$V_1 = \left\{ \mathbf{r} \in \mathbb{R}^4 : \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \mathbf{r} = \mathbf{0} \right\},$$

which is a null space and thus forms a linear space.

Part (b). Let V_2 be the subset for which $a+b+c+d=1$. Clearly, V_2 does not contain the zero vector ($a=b=c=d=0$) and is thus not a linear space.

Part (c). Let V_3 be the subset for which $a+b=c+d$ and $a+2b=c+3d$. Then

$$V_3 = \left\{ \mathbf{r} \in \mathbb{R}^4 : \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 2 & -1 & -3 \end{pmatrix} \mathbf{r} = \mathbf{0} \right\},$$

which is clearly a null space. Hence, V_3 is a linear space.

Part (d). The dimension of V_1 is 3. Its basis is

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ -3 \end{pmatrix} \right\}.$$

One can easily see that the three vectors are pairwise orthogonal and are thus linearly independent.

* * * * *

Problem 8. The vector spaces S_1, S_2 and S_3 are given by

$$S_1 = \left\{ x \in \mathbb{R}^4 : x = \lambda \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\},$$

$$S_2 = \left\{ x \in \mathbb{R}^4 : x = \lambda \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right\},$$

$$S_3 = \left\{ x \in \mathbb{R}^4 : x = \lambda \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

(a) Find a basis for the vector space $S_1 \cap S_2$.

(b) Show that $S_1 \cup S_2$ is not a vector space.

(c) Determine whether the set $(S_2 \setminus S_3) \cup \{\mathbf{0}\}$ is a vector space.

Solution.

Part (a). Clearly,

$$S_1 \cap S_2 = \left\{ x \in \mathbb{R}^4 : x = \lambda \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Hence, its basis is simply

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Part (b). Take

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Clearly, $\mathbf{v}_1, \mathbf{v}_2 \in S_1 \cup S_2$. However, their sum is

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix} \notin S_1 \cup S_2.$$

Thus, $S_1 \cup S_2$ is not closed under addition and is thus not a vector space.

Part (c). Take

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Clearly, $\mathbf{v}_1, \mathbf{v}_2 \in (S_2 \setminus S_3) \cup \{\mathbf{0}\}$. However, their sum is

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} \in S_3.$$

Thus, $S_2 \setminus S_3 \cup \{\mathbf{0}\}$ is not closed under addition and is hence not a vector space.

* * * * *

Problem 9.

(a) Three $n \times 1$ column vectors are denoted by $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ and $\mathbf{M}$ is an $n \times n$ matrix. Show that if $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ are linearly dependent then the vectors $\mathbf{M}\mathbf{x}_1$, $\mathbf{M}\mathbf{x}_2$, $\mathbf{M}\mathbf{x}_3$ are also linearly dependent.

(b) The vectors $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$ and the matrix $\mathbf{P}$ are defined as follows:

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ 5 \\ 7 \end{pmatrix}, \mathbf{y}_2 = \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}, \mathbf{y}_3 = \begin{pmatrix} 5 \\ 51 \\ 55 \end{pmatrix}, \mathbf{P} \begin{pmatrix} 1 & -4 & 3 \\ 0 & 2 & 5 \\ 0 & 0 & 7 \end{pmatrix}.$$

Show that $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$ are linearly dependent.

(c) Find a basis for the linear space spanned by the vectors $\mathbf{P}\mathbf{y}_1$, $\mathbf{P}\mathbf{y}_2$, $\mathbf{P}\mathbf{y}_3$.

Solution.

Part (a). Since $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ are linearly dependent, there exists $a, b, c \in \mathbb{R}$ that are not all 0 such that

$$a\mathbf{x}_1 + b\mathbf{x}_2 + c\mathbf{x}_3 = \mathbf{0}.$$

Applying $\mathbf{M}$ to both sides of the equation,

$$a\mathbf{M}\mathbf{x}_1 + b\mathbf{M}\mathbf{x}_2 + c\mathbf{M}\mathbf{x}_3 = \mathbf{0}.$$

Since a, b and c are not all 0, by definition, $\mathbf{M}\mathbf{x}_1$, $\mathbf{M}\mathbf{x}_2$, $\mathbf{M}\mathbf{x}_3$ are also linearly dependent.

Part (b). Observe that

$$8\mathbf{y}_1 - 2\mathbf{y}_2 - \mathbf{y}_3 = 9\begin{pmatrix}1\\5\\7\end{pmatrix} - 2\begin{pmatrix}2\\-3\\4\end{pmatrix} - \begin{pmatrix}5\\51\\55\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Hence, $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$ are linearly dependent.

Part (c). Note that the basis of $\text{span}\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3\}$ is simply $\{\mathbf{y}_1, \mathbf{y}_2\}$. Thus, the basis of $\text{span}\{\mathbf{P}\mathbf{y}_1, \mathbf{P}\mathbf{y}_2, \mathbf{P}\mathbf{y}_3\}$ is $\{\mathbf{P}\mathbf{y}_1, \mathbf{P}\mathbf{y}_2\}$, which works out to be

$$\left\{\begin{pmatrix}2\\45\\49\end{pmatrix}, \begin{pmatrix}26\\14\\28\end{pmatrix}\right\}.$$

* * * * *

Problem 10. In the equation

$$\begin{pmatrix}1 & 1 & -1 & 1\\ 2 & 3 & 0 & 5\\ 1 & 0 & -2 & -6\\ 2 & 5 & 4 & 11-\alpha\end{pmatrix}\mathbf{x} = \begin{pmatrix}0\\0\\0\\\beta\end{pmatrix},$$

α and β are real constants.

(a) Show that if $\alpha = \beta = 0$, then the set of solutions for $\mathbf{x}$ is a vector space V of dimension 1, and find a basis for V with integer elements.

(b) The set of all solutions for which α and β are both non-zero is denoted by S. Show that S is a subset of V, but is not itself a vector space.

Solution.

Part (a). When $\alpha = \beta = 0$, our equation becomes

$$\begin{pmatrix}1 & 1 & -1 & 1\\ 2 & 3 & 0 & 5\\ 1 & 0 & -2 & -6\\ 2 & 5 & 4 & 11\end{pmatrix}\mathbf{x} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}.$$

The RREF of the matrix is

$$\begin{pmatrix}1 & 0 & 0 & -14\\ 0 & 1 & 0 & 11\\ 0 & 0 & 1 & -4\\ 0 & 0 & 0 & 0\end{pmatrix}.$$

We thus have the system of linear equations

$$\begin{cases} x_1 \qquad\qquad - 14x_4 = 0\\ \qquad x_2 \qquad + 11x_4 = 0\,,\\ \qquad\qquad x_3 - \;\; 4x_4 = 0\end{cases}$$

where $x = (x_1,\, x_2,\, x_3,\, x_4)^{\mathsf{T}}$. Taking $x_4 = \lambda$, where $\lambda \in \mathbb{R}$, we get

$$\mathbf{x} = \lambda\begin{pmatrix}14\\-11\\4\\1\end{pmatrix}.$$

This describes a line that passes through the origin. Hence, V is a vector space with dimension 1 and basis $(14,\, -11,\, 4,\, 1)^{\mathsf{T}}$.

Part (b). Observe that any solution to

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ 2 & 3 & 0 & 5 \\ 1 & 0 & -2 & -6 \\ 2 & 5 & 4 & 11-\alpha \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \beta \end{pmatrix}$$

must be of the form

$$\mathbf{x} = \lambda \begin{pmatrix} 14 \\ -11 \\ 4 \\ 1 \end{pmatrix} \tag{1}$$

since the first three rows of its RREF is the same as in part (a). Meanwhile, we can expand the last row as

$$2x_1 + 5x_2 + 4x_3 + (11-\alpha)x_4 = \beta.$$

Substituting (1), we have

$$2(14\lambda) - 5(11\lambda) + 4(4\lambda) + (11-\alpha)\lambda = \beta \implies \lambda = -\frac{\beta}{\alpha}.$$

Thus,

$$S = \left\{ x \in \mathbb{R}^4 : x = -\frac{\beta}{\alpha} \begin{pmatrix} 14 \\ -11 \\ 4 \\ 1 \end{pmatrix}, \alpha, \beta \neq 0 \right\},$$

which is clearly a subset of V. However, because S does not contain the zero vector $(-\beta/\alpha \neq 0)$, it is not a vector space.

* * * * *

Problem 11. The elements of the matrices $\mathbf{A}$ and $\mathbf{B}$ are given by

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

(a) Write down in full the first column of the product $\mathbf{AB}$ and show that this can be put in the form $b_{11}\mathbf{c}_1 + b_{21}\mathbf{c}_2 + b_{31}\mathbf{c}_3$, where

$$\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \mathbf{c}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \mathbf{c}_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}.$$

(b) Write down the corresponding expressions for the second and third columns of $\mathbf{AB}$. Hence, show that the rank of $\mathbf{AB}$ cannot be greater than the rank of $\mathbf{A}$.

(c) For the case where

$$\mathbf{A} = \begin{pmatrix} 1 & \alpha & \beta \\ 2 & 2\alpha+\beta-1 & \alpha+2\beta \\ 5 & 5\alpha+3\beta-3 & 3\alpha+5\beta \end{pmatrix},$$

$\alpha, \beta \in \mathbb{R}$, show that

(i) for all values of α and β, the rank of $\mathbf{A}$ is not greater than 2,

(ii) if $\alpha = 0$ and $\beta = 1$, then, for all 3×3 matrices $\mathbf{B}$, there are at least 2 linearly independent solutions for $\mathbf{x}$ of the equation $\mathbf{ABx} = \mathbf{0}$, where $\mathbf{x} \in \mathbb{R}^3$.

Solution.

Part (a). The first column of $\mathbf{AB}$ is given by

$$\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} \end{pmatrix}$$

$$= b_{11}\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + b_{21}\begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + b_{31}\begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} = b_{11}\mathbf{c}_1 + b_{21}\mathbf{c}_2 + b_{31}\mathbf{c}_3.$$

Part (b). The second column is given by

$$b_{12}\mathbf{c}_1 + b_{22}\mathbf{c}_2 + b_{32}\mathbf{c}_3,$$

while the third column is given by

$$b_{13}\mathbf{c}_1 + b_{23}\mathbf{c}_2 + b_{33}\mathbf{c}_3.$$

Observe that every column of $\mathbf{AB}$ lies in $\operatorname{span}\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$. Thus,

$$\operatorname{col}(\mathbf{AB}) \subseteq \operatorname{col}(\mathbf{A}) \implies \operatorname{rank}(\mathbf{AB}) = \dim\operatorname{col}(\mathbf{AB}) \leq \dim\operatorname{col}(\mathbf{A}) = \operatorname{rank}(\mathbf{A})\,.$$

Part (c).

Part (c)(i). Note that

$$\begin{pmatrix} \alpha \\ 2\alpha + \beta - 1 \\ 5\alpha + 3\beta - 3 \end{pmatrix} = \alpha\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + (\beta - 1)\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$$

and

$$\begin{pmatrix} \beta \\ \alpha + 2\beta \\ 3\alpha + 5\beta \end{pmatrix} = \beta\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + \alpha\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}.$$

Thus, $\operatorname{col}(\mathbf{A})$ is given by

$$\operatorname{span}\left\{\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \alpha\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + (\beta - 1)\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}, \beta\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + \alpha\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}\right\} = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}\right\},$$

whence

$$\operatorname{rank}(\mathbf{A}) = \dim\operatorname{col}(\mathbf{A}) = 2.$$

Part (c)(ii). When $\alpha = 0$ and $\beta = 1$, we have

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 5 & 0 & 5 \end{pmatrix}.$$

Hence,

$$\operatorname{rank}(\mathbf{AB}) \leq \operatorname{rank}(\mathbf{A}) = \dim\operatorname{span}\left\{\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}\right\} = 1.$$

By the rank-nullity theorem,

$$\operatorname{nullity}(\mathbf{AB}) = 3 - \operatorname{rank}(\mathbf{AB}) \geq 3 - 1 = 2.$$

Thus, the kernel of $\mathbf{AB}$ has dimension at least 2, whence there are at least 2 linearly independent solutions to $\mathbf{ABx} = \mathbf{0}$.

* * * * *

Problem 12. The linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^3$ is represented by the matrix $\mathbf{M}$, where

$$\mathbf{M} = \begin{pmatrix} 1 & -2 & 2 & 2 \\ 4 & -7 & \lambda & 5 \\ 3 & \lambda & -7 & 3 \end{pmatrix},$$

where $\lambda \in \mathbb{R}$.

(a) Show that the rank of $\mathbf{M}$ is 2 if $\lambda = -5$ and determine the dimension of the null space, K of T if $\lambda \neq -5$.

(b) If $\lambda = -5$, write down a basis for the range space, R, and the null space, K of T.

(c) If $\lambda = -5$, find the set of vectors $\mathbf{x}$ such that

$$\mathbf{Mx} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

and state whether this set forms a vector space, justifying your answer.

Solution.

Part (a). Performing elementary row operations on $\mathbf{M}$, we get

$$\begin{pmatrix} 1 & -2 & 2 & 2 \\ 4 & -7 & \lambda & 5 \\ 3 & \lambda & -7 & 3 \end{pmatrix} \to \begin{array}{r} \\ {\scriptstyle R_2 - 4R_1} \\ {\scriptstyle (4\lambda+21)R_1 - (\lambda+6)R_2 + R_3} \end{array} \begin{pmatrix} 1 & -2 & 2 & 2 \\ 0 & 1 & \lambda - 8 & -3 \\ 0 & 0 & -(\lambda+5)(\lambda-7) & 3(\lambda+5) \end{pmatrix}.$$

Thus, if $\lambda = -5$, then the last row is entirely 0, whence $\operatorname{rank} \mathbf{M} = 2$. If $\lambda \neq -5$, then the last row is not entirely 0, whence $\operatorname{rank} \mathbf{M} = 3$ and

$$\dim K = 4 - \operatorname{rank} \mathbf{M} = 1.$$

Part (b). When $\lambda = -5$, the RREF of $\mathbf{M}$ is

$$\begin{pmatrix} 1 & 0 & -24 & -4 \\ 0 & 1 & -13 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, the basis of R is simply

$$\left\{ \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ -7 \\ -5 \end{pmatrix} \right\}.$$

Now consider K, the solution set of $\mathbf{Mx} = \mathbf{0}$, which is equivalent to the solution set of

$$\begin{pmatrix} 1 & 0 & -24 & -4 \\ 0 & 1 & -13 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

where $x = (x_1,\, x_2,\, x_3,\, x_4)^{\mathsf{T}}$. This gives the system of linear equations

$$\begin{cases} x_1 \qquad\quad - 24x_3 - 4x_4 = 0 \\ \qquad x_2 - 13x_3 - 3x_4 = 0 \end{cases}.$$

Let $x_3 = s$ and $x_4 = t$ be free variables, with $s, t \in \mathbb{R}$. Then

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 24s + 4t \\ 13s + 3t \\ s \\ t \end{pmatrix} = s \begin{pmatrix} 24 \\ 13 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 4 \\ 3 \\ 0 \\ 1 \end{pmatrix}.$$

Thus, the basis of K is

$$\left\{ \begin{pmatrix} 24 \\ 13 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Part (c). Consider a particular solution to $\mathbf{Mx} = (1,\, 1,\, 0)^{\mathsf{T}}$. Since there are two free variables, we take $\mathbf{x} = (x_1,\, x_2,\, 0,\, 0)^{\mathsf{T}}$. Then

$$\begin{pmatrix} 1 & -2 & 2 & 2 \\ 4 & -7 & -5 & 5 \\ 3 & -5 & -7 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

This gives the system of linear equations

$$\begin{cases} x_1 - 2x_2 = 1 \\ 4x_1 - 7x_2 = 1 \\ 3x_1 - 5x_2 = 0 \end{cases}.$$

Solving, we get $x_1 = -5$ and $x_2 = -3$. Thus, the set of all solutions to $\mathbf{Mx} = (1,\, 1,\, 0)^{\mathsf{T}}$ is

$$\left\{ \mathbf{x} \in \mathbb{R}^4 : \mathbf{x} = \begin{pmatrix} -5 \\ -3 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 24 \\ 13 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 4 \\ 3 \\ 0 \\ 1 \end{pmatrix},\ s, t \in \mathbb{R} \right\}.$$

This set is not a vector space since it does not contain the zero vector.

* * * * *

Problem 13. The linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^4$ is represented by the matrix $\mathbf{A}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -3 & -4 \\ 2 & 5 & 1 & 3 \\ 3 & 7 & -2 & -1 \\ 7 & 16 & -7 & q \end{pmatrix}.$$

The range space of T is denoted by R.

(a) Show that $q = -6$ if the dimension of R is 2, and that $q \neq -6$ if the dimension of R is 3.

(b) For the case where $q = -6$, write down a basis for R, and hence, find a vector $\mathbf{x}$ such that $\mathbf{Ax} = (1,\, 5,\, 6,\, 13)^{\mathsf{T}}$.

(c) The null space of T, for the case where $q \neq -6$, is denoted by K_1. Find a basis for K_1.

(d) The null space of T, for the case where $q = -6$, is denoted by K_2. Without using a calculator, find a basis for K_2. Show that K_1 is a subspace of K_2.

Solution.

Part (a). Performing elementary row operations on $\mathbf{A}$, we get

$$\begin{pmatrix} 1 & 2 & -3 & -4 \\ 2 & 5 & 1 & 3 \\ 3 & 7 & -2 & -1 \\ 7 & 16 & -7 & q \end{pmatrix} \to \begin{array}{r} {\scriptstyle 5R_1-2R_2} \\ {\scriptstyle R_2-2R_1} \\ {\scriptstyle -2R_3-R_1+R_4} \\ {\scriptstyle R_3-R_1-R_2} \end{array} \begin{pmatrix} 1 & 0 & -17 & -26 \\ 0 & 1 & 7 & 11 \\ 0 & 0 & 0 & q+6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Clearly, if $q = -6$, then we have two non-zero rows, whence $\dim R = \operatorname{rank} \mathbf{A} = 2$. If $q \neq -6$, then we have three non-zero rows, whence $\dim R = \operatorname{rank} \mathbf{A} = 3$.

Part (b). When $q = -6$, we have from the above calculation that the RREF of $\mathbf{A}$ is

$$\begin{pmatrix} 1 & 0 & -17 & -26 \\ 0 & 1 & 7 & 11 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, the range space of T has basis

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 6 \\ 16 \end{pmatrix} \right\}.$$

Thus, the solution $\mathbf{x} = \begin{pmatrix} a \\ b \\ 0 \\ 0 \end{pmatrix}$ to $\mathbf{Ax} = (1,\, 5,\, 6,\, 13)^{\mathsf{T}}$ also satisfies the equation

$$a \begin{pmatrix} 1 \\ 2 \\ 3 \\ 7 \end{pmatrix} + b \begin{pmatrix} 2 \\ 5 \\ 7 \\ 16 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 6 \\ 13 \end{pmatrix}.$$

This is equivalent to the system of linear equations

$$\begin{cases} a + \ 2b = 1 \\ 2a + \ 5b = 5 \\ 3a + \ 7b = 6 \\ 7a + 16b = 13 \end{cases},$$

which has the unique solution $x_1 = -5$ and $x_2 = 3$. Thus,

$$\mathbf{x} = \begin{pmatrix} -5 \\ 3 \\ 0 \\ 0 \end{pmatrix}.$$

Part (c). Consider $\mathbf{A}\mathbf{x} = 0$. This is equivalent to solving

$$\begin{pmatrix} 1 & 0 & -17 & -26 \\ 0 & 1 & 7 & 11 \\ 0 & 0 & 0 & q+6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

This yields the system of equations

$$\begin{cases} x_1 \qquad - 17x_3 - \qquad 26x_4 = 0 \\ \qquad x_2 + \ \ 7x_3 + \qquad 11x_4 = 0 \\ \qquad\qquad\qquad (q+6)x_4 = 0 \end{cases}.$$

Note that $x_4 = 0$. Let $x_3 = \lambda$, where $\lambda \in \mathbb{R}$. Then

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \lambda \begin{pmatrix} 17 \\ -7 \\ 1 \\ 0 \end{pmatrix}.$$

Thus, the basis of K_1 is

$$\left\{ \begin{pmatrix} 17 \\ -7 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Part (d). Consider $\mathbf{A}\mathbf{x} = 0$. This is equivalent to solving

$$\begin{pmatrix} 1 & 0 & -17 & -26 \\ 0 & 1 & 7 & 11 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

This yields the system of equations

$$\begin{cases} x_1 \qquad - 17x_3 - 26x_4 = 0 \\ \qquad x_2 + \ \ 7x_3 + 11x_4 = 0 \end{cases}.$$

Let $x_3 = \lambda$ and $x_4 = \mu$. Then

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 17\lambda + 26\mu \\ -7\lambda - 11\mu \\ \lambda \\ \mu \end{pmatrix} = \lambda \begin{pmatrix} 17 \\ -7 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 26 \\ -11 \\ 0 \\ 1 \end{pmatrix}.$$

Since K_1 and K_2 are both null spaces, they must be vector spaces. Since $K_1 \subset K_2$, it follows that K_1 is a subspace of K_2.

* * * * *

Problem 14. Let $\mathbf{u} = (1,\, 1,\, 0)^{\mathsf{T}}$ and $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation

$$T(\mathbf{v}) = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}} \right) \mathbf{u}.$$

(a) Find the null space, $\ker(T)$ and a basis for it. State also its geometrical interpretation and write down its Cartesian equation.

(b) Find the range space of T and its rank. State also a geometrical interpretation of the range space of T.

Solution.

Part (a). Consider $T(\mathbf{v}) = \mathbf{0}$:

$$T(\mathbf{v}) = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right) \mathbf{u} = \mathbf{0} \implies \mathbf{u} \cdot \mathbf{v} = 0 \implies \mathbf{v} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 0.$$

Let $\mathbf{v} = (x,\, y,\, z)^\mathsf{T}$. Expanding the dot product, we see that $\ker(T)$ has Cartesian equation $x + y = 0$, $z \in \mathbb{R}$. Let $y = \lambda$ and $z = \mu$. Then

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\lambda \\ \lambda \\ \mu \end{pmatrix} = \lambda \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Thus, the kernel of T is

$$\ker(T) = \left\{ \mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \lambda \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ \lambda, \mu \in \mathbb{R} \right\},$$

and its basis is

$$\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

$\ker(T)$ is the plane passing through the origin that is normal to $(1,\, 1,\, 0)^\mathsf{T}$.

Part (b). Note that $\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}$ is simply a scalar. Thus,

$$\operatorname{range}(T) = \left\{ \mathbf{x} \in \mathbb{R}^3 : \nu \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},\ \nu \in \mathbb{R} \right\}.$$

The range of T is a line passing through the origin with direction vector $\mathbf{u}$. Thus, the rank of T is 1.

* * * * *

Problem 15. The linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^3$ is represented by the matrix $\mathbf{M}$, where

$$\mathbf{M} = \begin{pmatrix} 1 & 5 & -1 & -2 \\ -1 & -3 & 4 & 3 \\ 1 & 11 & 8 & 1 \end{pmatrix}.$$

(a) (i) Find a basis for $R(T)$, the range space of T. Give a precise geometrical description of $R(T)$.

(ii) Find a basis for $K(T)$, the null space of T.

(iii) Hence, find the general solution of the equation

$$T(\mathbf{x}) = \begin{pmatrix} 6\alpha - 5\beta \\ -4a + 3\beta \\ 12\alpha - 11\beta \end{pmatrix},$$

where $\alpha, \beta \in \mathbb{R}$, leaving your answer in terms of α and β.

(b) Let $V \subseteq \mathbb{R}^3$ be the set that satisfies the following properties:

$$V \cap R(T) = \{\mathbf{0}\}, \quad V \cup R(T) = \mathbb{R}^3.$$

Determine whether V is a subspace of $\mathbb{R}^3$.

Solution.

Part (a). Using G.C., we see that the RREF of $\mathbf{M}$ is

$$\begin{pmatrix} 1 & 0 & -17/2 & -9/2 \\ 0 & 1 & 3/2 & 1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Part (a)(i). $R(T)$ has basis

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ -3 \\ 11 \end{pmatrix} \right\}.$$

$R(T)$ represents the plane passing through the origin that contains the points $(1, -1, 1)$ and $(5, -3, 11)$.

Part (a)(ii). Consider $\mathbf{Mx} = \mathbf{0}$. This is equivalent to

$$\begin{pmatrix} 1 & 0 & -17/2 & -9/2 \\ 0 & 1 & 3/2 & 1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \mathbf{0},$$

which gives the system of linear equations

$$\begin{cases} x_1 \qquad - \frac{17}{2}x_3 - \frac{9}{2}x_4 = 0 \\ \qquad x_2 + \frac{3}{2}x_3 + \frac{1}{2}x_4 = 0 \end{cases}.$$

Let $x_3 = s$ and $x_4 = t$, where $s, t \in \mathbb{R}$. Then $x_1 = \frac{17}{2}s + \frac{9}{2}t$ and $x_2 = -\frac{3}{2}s - \frac{1}{2}t$, whence

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = s \begin{pmatrix} 17/2 \\ -3/2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 9/2 \\ -1/2 \\ 0 \\ 1 \end{pmatrix}.$$

Thus, a basis of $K(T)$ is

$$\left\{ \begin{pmatrix} 17 \\ -3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 9 \\ -1 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Part (a)(iii). Consider

$$\begin{pmatrix} 1 & 5 & -1 & -2 \\ -1 & -3 & 4 & 3 \\ 1 & 11 & 8 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 6\alpha - 5\beta \\ -4a + 3\beta \\ 12\alpha - 11\beta \end{pmatrix}.$$

Expanding the LHS and RHS, we see that

$$x_1 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 5 \\ -3 \\ 11 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 4 \\ 8 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} = \alpha \left[\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 5 \\ -3 \\ 11 \end{pmatrix} \right] - \beta \begin{pmatrix} 5 \\ -3 \\ 11 \end{pmatrix}.$$

It is hence obvious that taking $x_1 = \alpha$, $x_2 = \alpha - \beta$, $x_3 = x_4 = 0$ yields a particular solution to the equation. Thus, the general solution is

$$\mathbf{x} = \begin{pmatrix} \alpha \\ \alpha - \beta \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 17 \\ -3 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} 9 \\ -1 \\ 0 \\ 2 \end{pmatrix}, \quad s, t \in \mathbb{R}.$$

Part (b). From the given equations, it is obvious that $V = \mathbb{R}^3 \setminus R(T) \cup \{\mathbf{0}\}$. Take

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Clearly, both $\mathbf{v}_1$ and $\mathbf{v}_2$ are not in $R(T)$. Thus, $\mathbf{v}_1, \mathbf{v}_2 \in V$. However, their sum

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

is clearly in $R(T)$ and is also not the zero vector, thus it cannot be in V. Hence, V is not closed under addition, thus it is not a subspace of $\mathbb{R}^3$.

* * * * *

Problem 16. Matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ define linear transformations from $\mathbb{R}^4$ to $\mathbb{R}^4$ and are respectively defined as follows:

$$\mathbf{M}_1 = \begin{pmatrix} 1 & 2 & 1 & a \\ 0 & 1 & 1 & b \\ 0 & 0 & 0 & c \\ 0 & 0 & 0 & d \end{pmatrix}, \quad \mathbf{M}_2 = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & c-1 & 1 \\ 0 & 0 & 0 & c+1 \end{pmatrix},$$

where a, b, c and d are real constants.

(a) The null spaces of $\mathbf{M}_1$ and $\mathbf{M}_2$ are denoted by N_1 and N_2 respectively. For the case where $a = b = c = 1$ and $d = 0$, find a basis for N_1 and a basis for N_2. Hence, determine whether $N_1 \cup N_2$ is a vector space.

(b) The range spaces of the linear transformations defined by $\mathbf{M}_1$ and $\mathbf{M}_2$ are denoted by R_1 and R_2 respectively. Given that $R_1 \cup R_2$ is a vector space, find the possible conditions to be satisfied by a, b, c and d.

Solution.

Part (a). When $a = b = c = 1$ and $d = 0$, we have

$$\mathbf{M}_1 = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{M}_2 = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Note also that the RREF of $\mathbf{M}_1$ and $\mathbf{M}_2$ are given by

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & -7 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

respectively.

Consider $\mathbf{M}_1\mathbf{x} = \mathbf{0}$. This is equivalent to solving

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

which yields the system of linear equations

$$\begin{cases} x_1 \quad - x_3 \quad = 0 \\ \quad x_2 + x_3 \quad = 0 \\ \quad\quad x_4 = 0 \end{cases}.$$

Let $x_3 = \lambda$, where $\lambda \in \mathbb{R}$. Then $x_1 = \lambda$ and $x_2 = -\lambda$, whence

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \lambda \\ -\lambda \\ \lambda \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}.$$

Thus, N_1 has basis

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Consider $\mathbf{M}_2\mathbf{x} = \mathbf{0}$. This is equivalent to solving

$$\begin{pmatrix} 1 & 0 & -7 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

which yields the system of linear equations

$$\begin{cases} x_1 \quad - 7x_3 \quad = 0 \\ \quad x_2 + 3x_3 \quad = 0 \\ \quad\quad x_4 = 0 \end{cases}.$$

Let $x_3 = \mu$, where $\mu \in \mathbb{R}$. Then $x_1 = 7\mu$ and $x_2 = -3\mu$, whence

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7\mu \\ -3\mu \\ \mu \\ 0 \end{pmatrix} = \mu \begin{pmatrix} 7 \\ -3 \\ 1 \\ 0 \end{pmatrix}.$$

Thus, N_2 has basis

$$\left\{ \begin{pmatrix} 7 \\ -3 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Clearly,

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} 7 \\ -3 \\ 1 \\ 0 \end{pmatrix}$$

are both in $N_1 \cup N_2$. However, their sum

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{pmatrix} 8 \\ -4 \\ 2 \\ 0 \end{pmatrix}$$

is neither a scalar multiple of $\mathbf{v}_1$ nor $\mathbf{v}_2$, hence it is neither in N_1 nor N_2, so $\mathbf{v}_1 + \mathbf{v}_2 \notin N_1 \cup N_2$. Thus, $N_1 \cup N_2$ is not closed under addition, and it cannot be a vector space.

Part (b). Note that because both R_1 and R_2 are already vector spaces, for their union $R_1 \cup R_2$ to also be a vector space, either $R_1 \subseteq R_2$ or $R_2 \subseteq R_1$.

Performing elementary column operations on $\mathbf{M}_1$ and $\mathbf{M}_2$, we see that the two matrices are column-equivalent to

$$\mathbf{M}_1' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & c \\ 0 & 0 & 0 & d \end{pmatrix} \quad \text{and} \quad \mathbf{M}_2' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c-1 & 1 \\ 0 & 0 & 0 & c+1 \end{pmatrix}.$$

Case 1. Suppose $c \notin \{-1, 1\}$. Then $\mathbf{M}_2'$ has no row or column full of zeroes and thus has full rank. Hence, $\mathbf{M}_2$ also has full rank, i.e. $R_2 = \mathbb{R}^4$. Thus, regardless of what d is, we will always have $R_1 \subseteq R_4 = \mathbb{R}^4$, whence $R_1 \cup R_4$ is a vector space.

Case 2. Suppose $c = 1$. Then

$$\mathbf{M}_1' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & d \end{pmatrix} \quad \text{and} \quad \mathbf{M}_2' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Thus,

$$R_1 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ d \end{pmatrix} \right\} \quad \text{and} \quad R_2 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix} \right\}.$$

Since both R_1 and R_2 have equal dimension (3), we require $R_1 = R_2$, which is only possible if $d = 2$.

Case 3. Suppose $c = -1$. Then

$$\mathbf{M}_1' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & d \end{pmatrix} \quad \text{and} \quad \mathbf{M}_2' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus,

$$R_1 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -d \end{pmatrix} \right\} \quad \text{and} \quad R_2 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Since both R_1 and R_2 have equal dimension (3), we require $R_1 = R_2$, which is only possible if $d = 0$.

Thus, if $R_1 \cup R_2$ is a vector space, then $a, b \in \mathbb{R}$, with

- $c \in \mathbb{R} \setminus \{-1, 1\}$, $d \in \mathbb{R}$, or
- $c = 1$, $d = 2$, or
- $c = -1$, $d = 0$.

Self-Practice B17B

Problem 1. Determine which of the following are subspaces of the space of all real-valued functions f defined on the entire real line.

(a) all f such that $f(x) \leq 0$ for all x.

(b) all f such that $f(0) = 0$.

(c) all f such that $f(0) = 2$.

(d) all constant functions.

Solution. Let V be the space of all real-valued functions defined on $\mathbb{R}$. Note that the zero element of V is the zero function $f(x) \equiv 0$.

Part (a). Let S be the set of all functions $f \in V$ such that $f(x) \leq 0$ for all $x \in \mathbb{R}$. Then $g \in S$, where $g(x) \equiv -1$. Taking $c = -1$, we see that $cg(x) \equiv 1 \not\leq 0$ for all x, so S is not closed under scalar multiplication. Hence, S is not a subspace of V.

Part (b). Let S be the set of all functions $f \in V$ such that $f(0) = 0$. Clearly, the zero function is in S. Let $g, h \in S$ and $c \in \mathbb{R}$. Then

$$g(0) + h(0) = 0 + 0 = 0,$$

so S is closed under addition, and

$$cg(0) = c(0) = 0,$$

so S is closed under scalar multiplication. Thus, S is a subspace of V.

Part (c). Let S be the set of all functions $f \in V$ such that $f(0) = 2$. The zero function is not in S, so S is not a subspace of V.

Part (d). Let S be the set of all constant functions. Clearly, the zero function is in S. Let $g, h \in S$ such that $g(x) \equiv a$ and $h(x) \equiv b$, where a and b are constants. Let also $c \in \mathbb{R}$. Then

$$g(x) + h(x) \equiv a + b,$$

which is a constant, so S is closed under addition. Further,

$$cg(x) \equiv ca,$$

which is also a constant, so S is closed under scalar multiplication. Thus, S is a subspace of V.

* * * * *

Problem 2. Find a basis for the vector space spanned by the vectors $(1,\ 2,\ -2)^{\mathsf{T}}$, $(2,\ 1,\ 3)^{\mathsf{T}}$, $(1,\ -4,\ 12)^{\mathsf{T}}$ and $(3,\ -9,\ 29)^{\mathsf{T}}$. What is the dimension of this space? For what value (or values) of a does $(4,\ 3,\ a)^{\mathsf{T}}$ belong to this space?

Part (a). Consider

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 2 & 1 & -4 & -9 \\ -2 & 3 & 12 & 29 \end{pmatrix}.$$

Its RREF is given by

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

so the column space of $\mathbf{A}$, which is also the span of the given vectors, has basis

$$\left\{\begin{pmatrix}1\\2\\-2\end{pmatrix}, \begin{pmatrix}2\\1\\3\end{pmatrix}\right\}$$

and thus has dimension 2.

Suppose $(4,\,3,\,a)^\mathsf{T}$ belongs to the column space of $\mathbf{A}$. Then there exist $x, y \in \mathbb{R}$ such that

$$x\begin{pmatrix}1\\2\\-2\end{pmatrix} + y\begin{pmatrix}2\\1\\3\end{pmatrix} = \begin{pmatrix}4\\3\\a\end{pmatrix}.$$

From the first two rows, we see that x and y satisfy

$$\begin{cases} x + 2y = 4 \\ 2x + y = 3 \end{cases},$$

so $x = 2/3$ and $y = 5/3$. From the last row, we have

$$a = -2x + 3y = \frac{11}{3}.$$

* * * * *

Problem 3. Given that $(a,\,b,\,c)^\mathsf{T}$ belongs to the row space of the matrix

$$\begin{pmatrix}3 & 2 & 1\\-2 & -2 & 1\\1 & -2 & 7\end{pmatrix},$$

find a linear relation that must be satisfied by a, b and c.

Solution. The given matrix has RREF

$$\begin{pmatrix}1 & 0 & 2\\0 & 1 & -5/2\\0 & 0 & 0\end{pmatrix},$$

so its row space has basis

$$\left\{\begin{pmatrix}1\\0\\2\end{pmatrix}, \begin{pmatrix}0\\1\\-5/2\end{pmatrix}\right\}.$$

Since $(a,\,b,\,c)^\mathsf{T}$ is in the row space, there exist $x, y \in \mathbb{R}$ such that

$$\begin{pmatrix}a\\b\\c\end{pmatrix} = x\begin{pmatrix}1\\0\\2\end{pmatrix} + y\begin{pmatrix}0\\1\\-5/2\end{pmatrix}.$$

From the first two rows, we see that $x = a$ and $y = b$. Thus, from the last row,

$$c = 2a - \frac{5}{2}b.$$

Problem 4. Determine the rank of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 3 & 2 & 3 & 13 \\ 4 & 4 & 9 & 7 \\ 11 & 9 & 17 & 36 \end{pmatrix}.$$

Deduce that if $\mathbf{x}$ is a solution of the equation

$$\mathbf{A}\mathbf{x} = p \begin{pmatrix} 1 \\ 3 \\ 4 \\ 11 \end{pmatrix} + q \begin{pmatrix} 1 \\ 2 \\ 4 \\ 9 \end{pmatrix} + r \begin{pmatrix} 2 \\ 3 \\ 9 \\ 17 \end{pmatrix},$$

where p, q and r are given real numbers, then

$$\mathbf{x} = \begin{pmatrix} p - 2\lambda \\ q - 11\lambda \\ r + 5\lambda \\ \lambda \end{pmatrix},$$

where $\lambda \in \mathbb{R}$.

Hence, or otherwise, for solutions $\mathbf{x} = (\alpha,\ \beta,\ \gamma,\ \delta)^{\mathsf{T}}$ of the equation $\mathbf{A}\mathbf{x} = (4,\ 8,\ 17,\ 37)^{\mathsf{T}}$,

(a) find $\mathbf{x}$ such that $\alpha = 0$,

(b) show that there is no $\mathbf{x}$ for which $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1$.

Solution. The RREF of $\mathbf{A}$ is

$$\begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 11 \\ 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

so $\mathbf{A}$ has rank 3.

Let $\mathbf{x} = (x,\ y,\ z,\ w)^{\mathsf{T}}$. Then the LHS of the equation can be expanded as

$$x \begin{pmatrix} 1 \\ 3 \\ 4 \\ 11 \end{pmatrix} + y \begin{pmatrix} 1 \\ 2 \\ 4 \\ 9 \end{pmatrix} + z \begin{pmatrix} 2 \\ 3 \\ 9 \\ 17 \end{pmatrix} + w \begin{pmatrix} 3 \\ 13 \\ 7 \\ 36 \end{pmatrix} = p \begin{pmatrix} 1 \\ 3 \\ 4 \\ 11 \end{pmatrix} + q \begin{pmatrix} 1 \\ 2 \\ 4 \\ 9 \end{pmatrix} + r \begin{pmatrix} 2 \\ 3 \\ 9 \\ 17 \end{pmatrix},$$

which rearranges as

$$(x-p) \begin{pmatrix} 1 \\ 3 \\ 4 \\ 11 \end{pmatrix} + (y-q) \begin{pmatrix} 1 \\ 2 \\ 4 \\ 9 \end{pmatrix} + (z-r) \begin{pmatrix} 2 \\ 3 \\ 9 \\ 17 \end{pmatrix} + w \begin{pmatrix} 3 \\ 13 \\ 7 \\ 36 \end{pmatrix} = \mathbf{A} \begin{pmatrix} x-p \\ y-q \\ z-r \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \qquad (*)$$

We now find the null space of $\mathbf{A}$. Consider $\mathbf{A}\mathbf{y} = \mathbf{0}$, where $\mathbf{y} = (a,\ b,\ c,\ d)^{\mathsf{T}}$. From the RREF of $\mathbf{A}$, we see that

$$\begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 11 \\ 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{cases} a + 2d = 0 \\ b + 11d = 0 \\ c - 5d = 0 \end{cases}.$$

Let $d = \lambda \in \mathbb{R}$. Then the general solution is

$$\mathbf{y} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \lambda \begin{pmatrix} -2 \\ -11 \\ 5 \\ 1 \end{pmatrix}.$$

Going back to $(*)$, we see that

$$\begin{pmatrix} x-p \\ y-q \\ z-r \\ w \end{pmatrix} = \lambda \begin{pmatrix} -2 \\ -11 \\ 5 \\ 1 \end{pmatrix} \implies \mathbf{x} = \begin{pmatrix} p-2\lambda \\ q-11\lambda \\ r+5\lambda \\ \lambda \end{pmatrix}.$$

Part (a). Take $p = q = r = 1$. Then the solution to

$$\mathbf{Ax} = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 11 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 4 \\ 9 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \\ 9 \\ 17 \end{pmatrix} = (4,\, 8,\, 17,\, 37)^{\mathsf{T}}$$

has the form

$$\mathbf{x} = \begin{pmatrix} 1-2\lambda \\ 1-11\lambda \\ 1+5\lambda \\ \lambda \end{pmatrix}.$$

Since $\alpha = 0$, we take $\lambda = 1/2$ to obtain

$$\mathbf{x} = \begin{pmatrix} 0 \\ -9/2 \\ 7/2 \\ 1/2 \end{pmatrix}.$$

Part (b). Observe that

$$\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = (1-2\lambda)^2 + (1-11\lambda)^2 + (1+5\lambda) + \lambda^2 = 151\lambda^2 - 16\lambda + 3.$$

Suppose now that $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1$ has a solution. Then there exists $\lambda \in \mathbb{R}$ such that $151\lambda^2 - 16\lambda + 2$. The discriminant of this quadratic is $(-16)^2 - 4(151)(2) < 0$, so $\lambda \notin \mathbb{R}$, a contradiction. Thus, there does not exist $\mathbf{x}$ such that $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1$.

* * * * *

Problem 5. The set P of all quadratic polynomials in x is a vector space over $\mathbb{R}$. For each of the following subsets of P, determine whether it is a subspace, and if so, give a basis.

(a) $S_1 = \{f \in P : f(0) = 0\}$.

(b) $S_2 = \{f \in P : f(0) = 1\}$.

(c) $S_3 = \{f \in P : f(1) = 0\}$.

(d) $S_4 = \{f \in P : f(-x) = f(x)\, \forall x \in \mathbb{R}\}$.

Solution. Note that the zero element of P is the zero polynomial $f(x) \equiv 0$.

Part (a). Clearly, the zero polynomial is in S_1. Let $g, h \in S_1$ and $c \in \mathbb{R}$. Observe that

$$g(0) + h(0) = 0 + 0 = 0,$$

so S_1 is closed under addition. Also,

$$cg(0) = c(0) = 0,$$

so S_1 is closed under scalar multiplication. Thus, S_1 is a subspace of P.

Let $f \in S_1$, where $f(x) = ax^2 + bx + c$. Since $f(0) = 0$, we must have $c = 0$, so $f(x) = ax^2 + bx$. Thus, a basis of S_1 is $\{x^2, x\}$.

Part (b). S_2 does not contain the zero polynomial, hence it is not a subspace of P.

Part (c). Clearly, the zero polynomial is in S_3. Let $g, h \in S_3$ and $c \in \mathbb{R}$. Observe that

$$g(1) + h(1) = 0 + 0 = 0,$$

so S_3 is closed under addition. Also,

$$cg(1) = c(1) = 0,$$

so S_3 is closed under scalar multiplication. Thus, S_3 is a subspace of P.

Let $f \in S_3$, where $f(x) = ax^2 + bx + c$. Since $f(1) = 0$, we must have $a + b + c = 0$, so

$$f(x) = ax^2 + bx - a - b = a\left(x^2 - 1\right) + b\left(x - 1\right).$$

Thus, a basis of S_3 is $\{x^2 - 1, x - 1\}$.

Part (d). Clearly, the zero polynomial is in S_4. Let $g, h \in S_4$ and $c \in \mathbb{R}$. Observe that

$$g(x) + h(x) = g(-x) + h(-x),$$

so S_4 is closed under addition. Also,

$$cg(-x) = cg(x),$$

so S_4 is closed under scalar multiplication. Thus, S_4 is a subspace of P.

Let $f \in S_4$, where $f(x) = ax^2 + bx + c$. Since $f(-x) = f(x)$, we must have $b = 0$, so $f(x) = ax^2 + c$. Thus, a basis of S_4 is $\{x^2, 1\}$.

Problem 6. The linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^4$ is represented by the matrix $\mathbf{A}$ where

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 0 & 4 \\ 3 & 1 & 1 & 0 \\ -1 & -5 & -1 & 8 \\ 3 & 8 & 2 & -12 \end{pmatrix}.$$

(a) Find the rank of $\mathbf{A}$ and a basis for the range space of T.

(b) Show that the vector $\mathbf{b} = (4,\ 5,\ 3,\ -2)^{\mathsf{T}}$ belongs to the range space of T.

(c) Find a basis for the null space of T. Hence, find the general solution of $\mathbf{Ax} = \mathbf{b}$.

Solution.

Part (a). The RREF of $\mathbf{A}$ is given by

$$\begin{pmatrix} 1 & 0 & 2/7 & 4/7 \\ 0 & 1 & 1/7 & -12/7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, $\mathbf{A}$ has rank 2, and a basis for the range space of T is given by

$$\left\{ \begin{pmatrix} 1 \\ 3 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ -5 \\ 8 \end{pmatrix} \right\}.$$

Part (b). Consider

$$x \begin{pmatrix} 1 \\ 2 \\ -1 \\ 3 \end{pmatrix} + y \begin{pmatrix} -2 \\ 1 \\ -5 \\ 8 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 3 \\ -2 \end{pmatrix}.$$

Since there exists a solution ($x = 2$, $y = -1$), it follows that $(4,\ 5,\ 3,\ -2)^{\mathsf{T}}$ belongs to the range space of T.

Part (c). Consider $\mathbf{Ax} = \mathbf{0}$, where $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$. From the RREF of $\mathbf{A}$, we have

$$\begin{cases} x \qquad + \frac{2}{7}z + \ \frac{4}{7}w = 0 \\ \qquad y + \frac{1}{7}z - \frac{12}{7}w = 0 \end{cases}.$$

Let $z = 7\lambda$ and $w = 7\mu$, where $\lambda, \mu \in \mathbb{R}$. Then

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \lambda \begin{pmatrix} -2 \\ -1 \\ 7 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -4 \\ 12 \\ 0 \\ 7 \end{pmatrix},$$

so a basis for the null space of T is given by

$$\left\{ \begin{pmatrix} -2 \\ 1 \\ 7 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 12 \\ 0 \\ 7 \end{pmatrix} \right\}.$$

From (b), we know a particular solution to $\mathbf{Ax} = \mathbf{b}$:

$$\begin{pmatrix} 4 \\ 5 \\ 3 \\ -2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 3 \end{pmatrix} + (-1) \begin{pmatrix} -2 \\ 1 \\ -5 \\ 8 \end{pmatrix} + (0) \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix} + (0) \begin{pmatrix} 4 \\ 0 \\ 8 \\ -12 \end{pmatrix} = \mathbf{A} \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Thus, the general solution of $\mathbf{Ax} = \mathbf{b}$ is thus

$$\mathbf{x} = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ -1 \\ 7 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -4 \\ 12 \\ 0 \\ 7 \end{pmatrix}.$$

* * * * *

Problem 7. The linear transformation $L : \mathbb{R}^4 \to \mathbb{R}^3$ is represented by the matrix $\mathbf{A}$, where

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

(a) Find the rank of $\mathbf{A}$ and deduce that the dimension of the null space, N, of L is 2.

(b) Show that there is a basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ for N such that

$$\mathbf{e}_1 = \begin{pmatrix} p \\ 2 \\ q \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{pmatrix} r \\ 0 \\ s \\ 2 \end{pmatrix},$$

where p, q, r and s are integers to be found.

(c) Given that $\mathbf{e}_0 = (1,\ 1,\ 1,\ 1)^\mathsf{T}$, $\mathbf{b} = (4,\ 2,\ -2)^\mathsf{T}$, show that the solution set, W, of the equation $\mathbf{Ax} = \mathbf{b}$ is

$$W = \{\mathbf{e}_0 + \lambda\mathbf{e}_1 + \mu\mathbf{e}_2 \mid \lambda, \mu \in \mathbb{R}\}.$$

The set of elements of $\mathbb{R}^4$ which do not belong to W is denoted by V.

(d) Show that N is a subset of V.

(e) Write down the vectors $\mathbf{e}_3$ and $\mathbf{e}_4$ such that the set $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ forms a basis for $\mathbb{R}^4$, justifying your answer.

Solution.

Part (a). The RREF of $\mathbf{A}$ is

$$\begin{pmatrix} 1 & 0 & 4/5 & -1/5 \\ 0 & 1 & -2/5 & 3/5 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence, the rank of $\mathbf{A}$ is 2, so by the rank-nullity theorem, the dimension of N is $4-2 = 2$.

Part (b). Consider $\mathbf{A}\mathbf{x} = \mathbf{0}$, where $\mathbf{x} = (x,\, y,\, z,\, w)^{\mathsf{T}}$. From the RREF, we have

$$\begin{cases} x \phantom{{}-y} + \frac{4}{5}z - \frac{1}{5}w = 0 \\ y - \frac{2}{5}z + \frac{3}{5}w = 0 \end{cases}.$$

Let $z = 5\lambda$ and $w = 5\mu$, where $\lambda, \mu \in \mathbb{R}$. Then,

$$\mathbf{x} = \lambda \begin{pmatrix} -4 \\ 2 \\ 5 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -3 \\ 0 \\ 5 \end{pmatrix}.$$

Clearly, $\mathbf{e}_1 = (-4,\, 2,\, 5,\, 0)^{\mathsf{T}}$, so $p = -4$ and $q = 5$. Now consider

$$\begin{pmatrix} r \\ 0 \\ 5 \\ 2 \end{pmatrix} = a \begin{pmatrix} -4 \\ 2 \\ 5 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ -3 \\ 0 \\ 5 \end{pmatrix}.$$

From the first and third rows, we have $5b = 2$ and $2a - 3b = 0$, so $a = 3/5$ and $b = 2/5$. Thus,

$$\mathbf{e}_2 = \begin{pmatrix} r \\ 0 \\ 5 \\ 2 \end{pmatrix} = \frac{3}{5} \begin{pmatrix} -4 \\ 2 \\ 5 \\ 0 \end{pmatrix} + \frac{2}{5} \begin{pmatrix} 1 \\ -3 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 3 \\ 2 \end{pmatrix},$$

so $r = -2$ and $s = 3$. The basis for N is

$$\left\{ \begin{pmatrix} -4 \\ 2 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 3 \\ 2 \end{pmatrix} \right\}.$$

Part (c). Note that

$$\mathbf{A}\mathbf{e}_0 = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & -1 & 2 & -1 \\ 1 & -3 & 2 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ -2 \end{pmatrix} = \mathbf{b}.$$

Let $\mathbf{x} \in W$. Then

$$\mathbf{A}\mathbf{x} = \mathbf{b} = \mathbf{A}\mathbf{e}_0 \implies \mathbf{A}\left(\mathbf{x} - \mathbf{e}_0\right) = \mathbf{0},$$

so $\mathbf{x} - \mathbf{e}_0 \in N$, whence $\mathbf{x} = \mathbf{e}_0 + \lambda\mathbf{e}_1 + \mu\mathbf{e}_2$.

Part (d). Suppose $\mathbf{v} \in N$. Then $\mathbf{A}\mathbf{v} = \mathbf{0} \neq \mathbf{b}$, so $\mathbf{v} \notin W$, whence $\mathbf{v} \in V$. Thus, $N \subseteq V$.

Part (e). Note that the RREF of the matrix below is $\mathbf{I}$:

$$\begin{pmatrix} -4 & -2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 5 & 3 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Thus, a basis for $\mathbb{R}^4$ is

$$\left\{ \begin{pmatrix} -4 \\ 2 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\},$$

hence $\mathbf{e}_3 = (1, 0, 0, 0)^\mathsf{T}$ and $\mathbf{e}_4 = (0, 1, 0, 0)^\mathsf{T}$.

* * * * *

Problem 8. The linear transformations $T_1 : \mathbb{R}^4 \to \mathbb{R}^4$ and $T_2 : \mathbb{R}^4 \to \mathbb{R}^4$ are represented by the matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ respectively, where

$$\mathbf{M}_1 = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 0 & \theta & 3 & 3 \\ 3 & 1 & 10 & 4 \\ 0 & 1 & -19 & -10 \end{pmatrix} \quad \text{and} \quad \mathbf{M}_2 = \begin{pmatrix} 2 & 1 & 3 & 2 \\ 1 & 0 & 2 & 3 \\ 1 & 3 & 14 & 16 \\ 1 & 0 & -1 & -2 \end{pmatrix}.$$

The null spaces of T_1 and T_2 are denoted by K_1 and K_2 respectively. The range spaces of T_1 and T_2 are denoted by R_1 and R_2 respectively.

(a) Determine the value of θ given that the dimension of K_1 is 1. Find a basis of R_1.

(b) Write down a basis of K_2 and a basis of R_2.

(c) Prove that $R_1 \cap R_2$ is a vector space. Show that the dimension of $R_1 \cap R_2$ is 2.

(d) Without evaluating $\mathbf{M}_1\mathbf{M}_2$, find a vector in the null space the transformation $T_3 : \mathbb{R}^4 \to \mathbb{R}^4$ represented by $\mathbf{M}_1\mathbf{M}_2$.

Solution.

Part (a). Using G.C., $\mathbf{M}_1$ can be reduced to

$$\mathbf{M}_1 \to \begin{pmatrix} 1 & 0 & 0 & -\frac{23}{22} \\ 0 & 1 & 0 & \frac{27}{22} \\ 0 & 0 & 1 & \frac{13}{22} \\ 0 & \theta & 3 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & -\frac{23}{22} \\ 0 & 1 & 0 & \frac{27}{22} \\ 0 & 0 & 1 & \frac{13}{22} \\ 0 & 0 & 0 & 3 - \theta\left(\frac{27}{22}\right) - 3\left(\frac{13}{22}\right) \end{pmatrix}.$$

Since $\dim K_1 = 1$, we must have exactly one row of zeroes, so

$$3 - \theta\left(\frac{27}{22}\right) - 3\left(\frac{13}{22}\right) \implies \theta = 1.$$

A basis for R_1 is

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 10 \\ -19 \end{pmatrix} \right\}.$$

Part (b). The RREF of $\mathbf{M}_2$ is given by

$$\begin{pmatrix} 1 & 0 & 0 & -\frac{1}{3} \\ 0 & 1 & 0 & -\frac{7}{3} \\ 0 & 0 & 1 & \frac{5}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

A basis for R_2 is

$$\left\{ \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 14 \\ -1 \end{pmatrix} \right\}.$$

Consider $\mathbf{M}_2\mathbf{x} = \mathbf{0}$, where $\mathbf{x} = (x,\, y,\, z,\, w)^\mathsf{T}$. Then

$$\begin{pmatrix} 1 & 0 & 0 & -\frac{1}{3} \\ 0 & 1 & 0 & -\frac{7}{3} \\ 0 & 0 & 1 & \frac{5}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{cases} x \qquad\qquad - \frac{1}{3}w = 0 \\ \qquad y \qquad - \frac{7}{3}w = 0 \\ \qquad\qquad z + \frac{5}{3}w = 0 \end{cases}.$$

Let $w = 3\lambda \in \mathbb{R}$. Then

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 7 \\ -5 \\ 3 \end{pmatrix},$$

so a basis of K_2 is

$$\left\{ \begin{pmatrix} 1 \\ 7 \\ -5 \\ 3 \end{pmatrix} \right\}.$$

Part (c). Let $v \in R_1 \cap R_2$. Then there exist constants $a, b, c, d, e, f \in \mathbb{R}$ such that

$$a \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix} + b \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix} + c \begin{pmatrix} 1 \\ 3 \\ 10 \\ -19 \end{pmatrix} = d \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix} + e \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix} + f \begin{pmatrix} 3 \\ 2 \\ 14 \\ -1 \end{pmatrix}.$$

Rearranging, we have

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 2 \\ 3 & 1 & 10 & 14 \\ 0 & 1 & -19 & -1 \end{pmatrix} \begin{pmatrix} a-d \\ b-e \\ c \\ -f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since the matrix on the LHS is invertible, we must have $a = d$, $b = e$ and $c = f = 0$. Hence, $\mathbf{v}$ is of the form

$$\mathbf{v} = a \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix} + b \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Thus,

$$R_1 \cap R_2 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

and is thus a vector space with dimension 2.

Part (d). Picking $\mathbf{x} = (1,\, 7,\, -5,\, 3)^\mathsf{T}$, which is in K_2, we have

$$\mathbf{M}_1\mathbf{M}_2\mathbf{x} = \mathbf{M}_1\mathbf{0} = \mathbf{0}$$

as desired.

Problem 9.

(a) Let $L : V \to V$ be a linear transformation on the vector space V. A linear transformation is said to be *nilpotent* if there exists $k \in \mathbb{Z}^+$ such that $L^k(x) = 0$ for all $x \in V$. A linear transformation L is said to be *invertible* if there exists a transformation T such that $LT(x) = TL(x) = x$ for all $x \in V$. Show that if L is nilpotent, then the transformation $M = I - L$ is invertible by finding an explicit formula for $(I - L)^{-1}$, where I is the identity transformation.

(b) Let V be the space of quadratic polynomials and L the differential transformation, that is,

$$L\big(ax^2 + bx + c\big) = 2ax + b.$$

Show that L is a linear transformation on V.

(c) Using (a) and (b), find a particular solution $y_p \notin V$ of $y' - y = 5x^2 - 3$.

(d) Hence, find the general solution of the differential equation in (c).

Solution.

Part (a). A simple series expansion gives

$$(I - L)^{-1} = I + L + L^2 + \cdots + L^{k-1} + L^k + \cdots = I + L + L^2 + \cdots + L^{k-1},$$

so M is invertible.

Part (b). Let f, g be quadratic polynomials, where $f(x) = a_1x^2 + b_1x + c_1$ and $g(x) = a_2x^2 + b_2x + c_2$. Then

$$\begin{aligned} L(f+g) &= L\big((a_1 + a_2)\,x^2 + (a_2 + b_2)\,x + (c_1 + c_2)\big) = 2\,(a_1 + a_2)\,x + (b_1 + b_2) \\ &= (2a_1x + b_1) + (2a_2x + b_2) = L(f) + L(g). \end{aligned}$$

Let $k \in \mathbb{R}$. Then

$$L(kf) = L\big(ka_1x^2 + kb_1x + kc_1\big) = 2ka_1x + kb_1 = k\,(2a_1x + b_1) = kL(f).$$

Thus, L is a linear transformation on V.

Part (c). Rewriting the given differential equation, we see that

$$y_p' - y_p = (L - I)y_p = 5x^2 - 3 \implies (I - L)y_p = -5x^2 + 3.$$

Since $L^3 = 0$, by (a), we have the particular solution

$$\begin{aligned} y_p &= (I - L)^{-1}\left(-5x^2 + 3\right) = \left(I + L + L^2\right)\left(-5x^2 + 3\right) \\ &= \left(-5x^2 + 3\right) + (-10x) + (-10) = -5x^2 - 10x - 7. \end{aligned}$$

Part (d). The associated homogeneous differential equation is $y' = y$, so the complementary solution is $y_c = Ae^x$. Thus, the general solution is given by

$$y = y_c + y_p = Ae^x - 5x^2 - 10x - 7.$$

Assignment B17B

Problem 1. In $\mathbb{R}^2$, a horizontal shear is a mapping that takes a generic point with position vector $(x,\, y)^{\mathsf{T}}$ to the point with position vector $(x+my,\, y)^{\mathsf{T}}$, where m is a fixed parameter called the *shear factor*.

Show that every horizontal shear mapping in $\mathbb{R}^2$ is a linear transformation. State the matrix that represents the horizontal shear with shear factor m.

Solution. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+my \\ y \end{pmatrix}.$$

Let $\alpha, \beta \in \mathbb{R}$. Observe that

$$\begin{aligned} T\left(\alpha \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \beta \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) &= T\left(\begin{pmatrix} \alpha x_1 + \beta x_2 \\ \alpha y_1 + \beta y_2 \end{pmatrix}\right) = \begin{pmatrix} \alpha x_1 + \beta x_2 + m\alpha y_1 + m\beta y_2 \\ \alpha y_1 + \beta y_2 \end{pmatrix} \\ &= \alpha \begin{pmatrix} x_1 + m y_1 \\ y_1 \end{pmatrix} + \beta \begin{pmatrix} x_2 + m y_2 \\ y_2 \end{pmatrix} = \alpha T\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) + \beta T\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right). \end{aligned}$$

Hence, T preserves both addition and scalar multiplication, whence it is a linear transformation.

Note that

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+my \\ y \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} m \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Thus,

$$\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$$

is the matrix representation of T.

* * * * *

Problem 2. Given that

$$\left\{ \begin{pmatrix} 1 \\ a \\ b \end{pmatrix}, \begin{pmatrix} b \\ 1 \\ a \end{pmatrix}, \begin{pmatrix} a \\ b \\ 1 \end{pmatrix} \right\}$$

is not a basis for $\mathbb{R}^3$, prove that $a^3 - 3ab + b^3 + 1 = 0$.

Solution. Since the vectors do not form a basis of $\mathbb{R}^3$, they must be linearly dependent. Thus, there exist $x_1, x_2, x_3 \in \mathbb{R}$ that are not all 0 such that

$$x_1 \begin{pmatrix} 1 \\ a \\ b \end{pmatrix} + x_2 \begin{pmatrix} b \\ 1 \\ a \end{pmatrix} + x_3 \begin{pmatrix} a \\ b \\ 1 \end{pmatrix} = \mathbf{0},$$

which is equivalent to the matrix equation

$$\begin{pmatrix} 1 & b & a \\ a & 1 & b \\ b & a & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0}.$$

Thus, the above matrix has a non-trivial kernel, whence its determinant is 0. Thus,

$$0 = \det \begin{pmatrix} 1 & b & a \\ a & 1 & b \\ b & a & 1 \end{pmatrix} = 1\begin{vmatrix} 1 & b \\ a & 1 \end{vmatrix} - a\begin{vmatrix} b & a \\ a & 1 \end{vmatrix} + b\begin{vmatrix} b & a \\ 1 & b \end{vmatrix} = a^3 - 3ab + b^3 + 1.$$

Problem 3. Let $\mathbf{A}$ be an $n \times n$ matrix and W to be the subset $\{\mathbf{u} \in \mathbb{R}^n \mid \mathbf{Au} = \mathbf{u}\}$ of $\mathbb{R}^n$.

(a) Show that W is a subspace of $\mathbb{R}^n$.

(b) Let

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

Find a basis of W.

Solution.

Part (a). Note that

$$W = \{\mathbf{u} \in \mathbb{R}^n \mid \mathbf{Au} = \mathbf{u}\} = \{\mathbf{u} \in \mathbb{R}^n \mid (\mathbf{A} - \mathbf{I})\mathbf{u} = \mathbf{0}\},$$

which is a null space in $\mathbb{R}^n$. Hence, W is a subspace of $\mathbb{R}^n$.

Part (b). Consider the solutions to $(\mathbf{A} - \mathbf{I})\mathbf{u} = \mathbf{0}$:

$$(\mathbf{A} - \mathbf{I})\mathbf{u} = \mathbf{0} \implies \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

We immediately see that $x, y \in \mathbb{R}$ and $z = 0$. Letting $x = \lambda$ and $y = \mu$, where $\lambda, \mu \in \mathbb{R}$, we have

$$\mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \lambda \\ \mu \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Thus, a basis of W is

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

* * * * *

Problem 4. The linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^3$ is represented by the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & -2 & a \\ 2 & -1 & 3 & -5 \\ -3 & -3 & 0 & 3 \end{pmatrix}$$

where a is a real constant.

(a) It is given that the dimension of the null space of T is 2. Find the value of a. Hence, find a basis for the null space of T.

(b) Show that R, the range space of T, is a plane, and find the Cartesian equation of R.

(c) Let V be a vector space spanned by $\mathbf{v}$ where $\mathbf{v} = (0, b, c)^{\mathsf{T}}$, $b, c \in \mathbb{R}$. If $R \cup V$ is a vector space, find a relationship between b and c.

Solution.

Part (a). Since $\dim\ker T = 2$, we have $\dim\operatorname{range} T = 4 - 2 = 2$. Since

$$\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix}$$

are linearly independent, they form a basis for $\operatorname{range} T$. Hence, there exist $\lambda, \mu \in \mathbb{R}$ such that

$$\begin{pmatrix} a \\ -5 \\ 3 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ -1 \\ -3 \end{pmatrix},$$

which is equivalent to the system of linear equations

$$\begin{cases} \lambda + 3\mu = a \\ 2\lambda - \mu = -5 \\ 3\lambda - 3\mu = 3 \end{cases}.$$

Solving the last two equations simultaneously yields $\lambda = -2$ and $\mu = 1$, whence $a = 1$. Thus,

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & -2 & 1 \\ 2 & -1 & 3 & -5 \\ -3 & -3 & 0 & 3 \end{pmatrix},$$

and its RREF is

$$\begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Consider the equation $\mathbf{Ax} = \mathbf{0}$. Then $\mathbf{x}$ also satisfies

$$\begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

which is equivalent to the system of linear equations

$$\begin{cases} x_1 \phantom{{}+x_2} + x_3 - 2x_4 = 0 \\ \phantom{x_1 +{}} x_2 - x_3 + x_4 = 0 \end{cases}.$$

Let $x_3 = \alpha$ and $x_4 = \beta$, where $\alpha, \beta \in \mathbb{R}$. Then

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -\alpha + 2\beta \\ \alpha - \beta \\ \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

Thus, $\ker T$ has basis

$$\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Part (b). We have $\dim R = \dim \operatorname{range} T = 2$, hence R is a plane. From the RREF of $\mathbf{A}$, the basis of R is

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix} \right\}.$$

Since

$$\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} \times \begin{pmatrix} 3 \\ -1 \\ -3 \end{pmatrix} = -\begin{pmatrix} 9 \\ 6 \\ 7 \end{pmatrix},$$

R has scalar product form

$$R : \mathbf{r} \cdot \begin{pmatrix} 9 \\ 6 \\ 7 \end{pmatrix} = 0,$$

which translates into the Cartesian equation $9x + 6y + 7z = 0$.

Part (c). Since R and V are both vector spaces, for their union $R \cup V$ to also be a vector space, we require either $R \subseteq V$ or $V \subseteq R$. However, since $\dim R = 2 > 1 \geq \dim V$, we can only have $V \subset R$. Thus, there exist $s, t \in \mathbb{R}$ such that

$$\begin{pmatrix} 0 \\ b \\ c \end{pmatrix} = s \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + t \begin{pmatrix} 3 \\ -1 \\ -3 \end{pmatrix}.$$

We immediately have $s = -3t$. Thus, $b = -7t$ and $c = 6t$, whence $6b = -7c$.

* * * * *

Problem 5. The matrix $\mathbf{A}$ and the vectors $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$ are defined as follows:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -3 & 4 & 11 & -10 \\ 4 & 5 & 6 & 3 \\ 6 & -2 & -10 & 14 \end{pmatrix}, \mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{x}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

The vector space V is the set of all vectors of the form $\lambda_1 \mathbf{A}\mathbf{x}_1 + \lambda_2 \mathbf{A}\mathbf{x}_2 + \lambda_3 \mathbf{A}\mathbf{x}_3 + \lambda_4 \mathbf{A}\mathbf{x}_4$, where $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$.

(a) Show that $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$ form a basis of $\mathbb{R}^4$.

(b) Find the rank of $\mathbf{A}$. Deduce the dimension of the null space of $\mathbf{A}$.

(c) Explain why the dimension of V is 2 and state a basis of V.

(d) The vector $(p,\, q,\, 23,\, 6)^{\mathsf{T}}$ belongs to V. Find p and q.

Solution.

Part (a). Let $\mathbf{B} = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 \end{pmatrix}$. In full,

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Observe that $\det \mathbf{B} = 1 \neq 0$. Thus, the column space of $\mathbf{B}$ (i.e. the span of $\mathbf{x}_1, \ldots, \mathbf{x}_4$) is $\mathbb{R}^4$. Also, the columns of $\mathbf{B}$ are linearly independent. Thus, $\mathbf{x}_1, \ldots, \mathbf{x}_4$ form a basis of $\mathbb{R}^4$.

Part (b). Note that $\mathbf{A}$ has RREF

$$\begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, the dimension of the column space of $\mathbf{A}$ is $\operatorname{rank} \mathbf{A} = 2$. By the rank-nullity theorem, the dimension of the null space of $\mathbf{A}$ is $4 - 2 = 2$.

Part (c). Note that V is the range (or column space) of $\mathbf{AB} = \begin{pmatrix} \mathbf{Ax}_1 & \mathbf{Ax}_2 & \mathbf{Ax}_3 & \mathbf{Ax}_4 \end{pmatrix}$. Thus, $\dim V = \operatorname{rank}(\mathbf{AB}) = \operatorname{rank} \mathbf{A} = 2$. Note that in the second-last step, we used the fact that $\mathbf{B}$ has full rank and thus does not affect the rank of $\mathbf{AB}$. A basis of V is

$$\{\mathbf{Ax}_1, \mathbf{Ax}_2\} = \left\{ \begin{pmatrix} 8 \\ 7 \\ 10 \\ 6 \end{pmatrix}, \begin{pmatrix} 8 \\ 7 \\ 3 \\ -2 \end{pmatrix} \right\}.$$

Part (d). There exists $\lambda, \mu \in \mathbb{R}$ for which

$$\begin{pmatrix} p \\ q \\ 23 \\ 6 \end{pmatrix} = \lambda \begin{pmatrix} 8 \\ 7 \\ 10 \\ 6 \end{pmatrix} + \mu \begin{pmatrix} 8 \\ 7 \\ 3 \\ -2 \end{pmatrix}.$$

The last two rows give the system of linear equations

$$\begin{cases} 10\lambda + 3\mu = 23 \\ 6\lambda - 2\mu = 6 \end{cases},$$

whence $\lambda = 32/19$ and $\mu = 39/19$. Thus,

$$p = 8(\lambda + \mu) = \frac{568}{19}, \quad q = 7(\lambda + \mu) = \frac{497}{19}.$$

B17C Linear Algebra - Eigenvalues and Eigenvectors

Tutorial B17C

Problem 1. For each of the following matrices $\mathbf{A}$, determine the eigenvalue(s) and corresponding eigenvector(s). Where $\mathbf{A}$ is diagonalizable, write down the matrix $\mathbf{Q}$ and $\mathbf{D}$ where $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^{-1}$.

(a) $\begin{pmatrix} 0 & -1 \\ -2 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 0 \\ -3 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} -3 & 0 \\ 0 & -3 \end{pmatrix}$

(d) $\begin{pmatrix} 19 & -9 & -6 \\ 25 & -11 & -9 \\ 17 & -9 & -4 \end{pmatrix}$ (e) $\begin{pmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{pmatrix}$ (f) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{pmatrix}$

Solution.

Part (a). Consider $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$:

$$0 = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & -1 \\ -2 & -\lambda \end{vmatrix} = \lambda^2 - 2 \implies \lambda = \pm\sqrt{2}.$$

Let $\mathbf{x} = (x,\, y)^\mathsf{T} \neq \mathbf{0}$ be an eigenvector.

Case 1: $\lambda = \sqrt{2}$. Consider $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$:

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \begin{pmatrix} -\sqrt{2} & -1 \\ -2 & -\sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solving, we get $x + \frac{1}{\sqrt{2}}y = 0$. Taking $y = -\sqrt{2}$, we have $\mathbf{x} = \big(1,\, -\sqrt{2}\big)^\mathsf{T}$.

Case 2: $\lambda = -\sqrt{2}$. Consider $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$:

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \begin{pmatrix} \sqrt{2} & -1 \\ -2 & \sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solving, we get $x - \frac{1}{\sqrt{2}}y = 0$. Taking $y = \sqrt{2}$, we have $\mathbf{x} = \big(1,\, \sqrt{2}\big)^\mathsf{T}$.

Thus,

$$\mathbf{Q} = \begin{pmatrix} 1 & 1 \\ -\sqrt{2} & \sqrt{2} \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{pmatrix}.$$

Part (b). Consider $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$:

$$0 = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 2-\lambda & 0 \\ -3 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 \implies \lambda = 2.$$

Let $\mathbf{x} = (x,\, y)^\mathsf{T} \neq \mathbf{0}$ be an eigenvector. Consider $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$:

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \begin{pmatrix} 0 & 0 \\ -3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solving, we get $x = 0$ and $y \in \mathbb{R}$. Taking $y = 1$, we have $(0,\, 1)^\mathsf{T}$.

Since there are fewer eigenvectors (1) than dimensions (2), $\mathbf{A}$ is not diagonalizable.

Part (c).

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

so $\lambda = -3$ with eigenvectors $(1,\, 0)^{\mathsf{T}}$ and $(0,\, 1)^{\mathsf{T}}$.

Part (d). Note that

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= |19| + |-11| + |-4| = 4, \\ \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 &= \begin{vmatrix} 19 & -9 \\ 25 & -11 \end{vmatrix} + \begin{vmatrix} -11 & -9 \\ -9 & -4 \end{vmatrix} + \begin{vmatrix} 19 & -6 \\ 17 & -4 \end{vmatrix} = 5, \\ \lambda_1\lambda_2\lambda_3 &= \begin{vmatrix} 19 & -9 & -6 \\ 25 & -11 & -9 \\ 17 & -9 & -4 \end{vmatrix} = 2. \end{aligned}$$

Thus, the characteristic polynomial of $\mathbf{A}$ is $-\lambda^3 + 4\lambda^2 - 5\lambda + 2$. Solving, we get $\lambda = 1, 2$.

Let $\mathbf{x} = (x,\, y,\, z)^{\mathsf{T}} \neq \mathbf{0}$ be an eigenvector.

Case 1: $\lambda = 1$. Consider $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$:

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \begin{pmatrix} 18 & -9 & -6 \\ 25 & -12 & -9 \\ 17 & -9 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Reducing to RREF, we have

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -4/3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Letting $z = t \in \mathbb{R}$, we have

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 1 \\ 4/3 \\ 1 \end{pmatrix}.$$

Taking $t = 3$, our eigenvector is $(3,\, 4,\, 3)^{\mathsf{T}}$.

Case 2: $\lambda = 2$. Consider $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$:

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \begin{pmatrix} 17 & -9 & -6 \\ 25 & -13 & -9 \\ 17 & -9 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Reducing to RREF, we have

$$\begin{pmatrix} 1 & 0 & -3/4 \\ 0 & 1 & -3/4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Letting $z = t \in \mathbb{R}$, we have

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 3/4 \\ 3/4 \\ 1 \end{pmatrix}.$$

Taking $t = 4$, our eigenvector is $(3,\, 3,\, 4)^{\mathsf{T}}$.

Since there are fewer eigenvectors (2) than dimensions (3), $\mathbf{A}$ is not diagonalizable.

Part (e). Note that the characteristic polynomial of $\mathbf{A}$ is simply $(5-\lambda)^3$. Hence, the only eigenvalue of $\mathbf{A}$ is $\lambda = 5$. Let $\mathbf{x} = (x,\, y,\, z)^{\mathsf{T}} \neq \mathbf{0}$ be an eigenvector. Then

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Hence, $x = y = 0$ and $z = t \in \mathbb{R}$. Thus, the only eigenvector is $(0,\, 0,\, 1)^{\mathsf{T}}$. Since there are fewer eigenvectors (1) than dimensions (3), $\mathbf{A}$ is not diagonalizable.

Part (f). Consider $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$:

$$0 = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 3 & 0 & 1-\lambda \end{vmatrix} = \lambda^2(1-\lambda) \implies \lambda = 0, 1.$$

Let $\mathbf{x} = (x,\, y,\, z)^{\mathsf{T}} \neq \mathbf{0}$ be an eigenvector.

Case 1: $\lambda = 0$. Consider $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$:

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Let $y = s \in \mathbb{R}$ and $z = t \in \mathbb{R}$. Then $x = -1/3t$, so

$$\mathbf{x} = \begin{pmatrix} -t/3 \\ s \\ t \end{pmatrix} = s \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{3}t \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}.$$

Thus, the corresponding eigenvectors are $(0,\, 1,\, 0)^{\mathsf{T}}$ and $(-1,\, 0,\, 3)^{\mathsf{T}}$.

Case 2: $\lambda = 1$. Consider $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$:

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 3 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Thus, $x = y = 0$ while $z = t \in \mathbb{R}$. Hence, the corresponding eigenvector is $(0,\, 0,\, 1)^{\mathsf{T}}$.

Thus,

$$\mathbf{Q} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 3 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

* * * * *

Problem 2. Find the eigenvalues and corresponding eigenvectors of the matrix $\mathbf{A}$, where

$$\mathbf{A} = \begin{pmatrix} -3 & 5 & 5 \\ -4 & 6 & 5 \\ 4 & -4 & -3 \end{pmatrix}.$$

Hence, find a matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{P}^{-1}\mathbf{D}\mathbf{P}$. Find also a diagonal matrix $\mathbf{E}$ such that $\mathbf{A}^3 = \mathbf{P}^{-1}\mathbf{E}\mathbf{P}$.

Solution. Note that

$$\begin{aligned}
\lambda_1+\lambda_2+\lambda_3 &= |-3|+|6|+|-3| = 0\\
\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_3\lambda_1 &= \begin{vmatrix}-3 & 5\\ -4 & 6\end{vmatrix} + \begin{vmatrix}6 & 5\\ -4 & -3\end{vmatrix} + \begin{vmatrix}-3 & 5\\ 4 & -3\end{vmatrix} = -7\\
\lambda_1\lambda_2\lambda_3 &= \begin{vmatrix}-3 & 5 & 5\\ -4 & 6 & 5\\ 4 & -4 & -3\end{vmatrix} = -6.
\end{aligned}$$

Hence, the characteristic polynomial of $\mathbf{A}$ is $-\lambda^3+7\lambda-6$, whence the roots are $\lambda = -3, 1, 2$.

Note that

$$\mathbf{A}-\lambda\mathbf{I} = \begin{pmatrix}-3-\lambda & 5 & 5\\ -4 & 6-\lambda & 5\\ 4 & -4 & -3-\lambda\end{pmatrix}.$$

The eigenvectors are thus

$$\begin{aligned}
\mathbf{e}_1 &= \frac{1}{20}\begin{pmatrix}-3-(-3)\\ 5\\ 5\end{pmatrix} \times \begin{pmatrix}4\\ -4\\ -3-(-3)\end{pmatrix} = \begin{pmatrix}1\\ 1\\ -1\end{pmatrix}\\
\mathbf{e}_2 &= \frac{1}{4}\begin{pmatrix}-3-1\\ 5\\ 5\end{pmatrix} \times \begin{pmatrix}4\\ -4\\ -3-1\end{pmatrix} = \begin{pmatrix}0\\ 1\\ -1\end{pmatrix}\\
\mathbf{e}_3 &= -\frac{1}{5}\begin{pmatrix}-3-2\\ 5\\ 5\end{pmatrix} \times \begin{pmatrix}4\\ -4\\ -3-2\end{pmatrix} = \begin{pmatrix}1\\ 1\\ 0\end{pmatrix}.
\end{aligned}$$

Thus,

$$\mathbf{P} = \begin{pmatrix}1 & 0 & 1\\ 1 & 1 & 1\\ -1 & -1 & 0\end{pmatrix}^{-1} = \begin{pmatrix}1 & -1 & -1\\ -1 & 1 & 0\\ 0 & 1 & 1\end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix}-3 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 2\end{pmatrix}.$$

Note that $\mathbf{A}^3 = \left(\mathbf{P}^{-1}\mathbf{D}\mathbf{P}\right)^3 = \mathbf{P}^{-1}\mathbf{D}^3\mathbf{P}$. Hence,

$$\mathbf{E} = \mathbf{D}^3 = \begin{pmatrix}-27 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 8\end{pmatrix}.$$

* * * * *

Problem 3. Find the eigenvalues and corresponding eigenvectors of the matrix $\mathbf{A}$, where

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

Hence, or otherwise,

(a) find the eigenvalues and corresponding eigenvectors of the matrix $\mathbf{A}+10\mathbf{I}$, where $\mathbf{I}$ is the unit matrix of order 3,

(b) find a matrix $\mathbf{P}$ such that $\mathbf{PAP}^{-1}$ is a diagonal matrix.

Solution. Note that

$$\lambda_1 + \lambda_2 + \lambda_3 = |2| + |-1| + |1| = 2$$
$$\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 = \begin{vmatrix} 2 & -3 \\ 1 & -2 \end{vmatrix} + \begin{vmatrix} -1 & 1 \\ 3 & 1 \end{vmatrix} + \begin{vmatrix} 2 & 0 \\ -1 & 1 \end{vmatrix} = -1$$
$$\lambda_1\lambda_2\lambda_3 = \begin{vmatrix} 2 & -3 & 0 \\ 1 & -1 & 1 \\ -1 & 3 & 1 \end{vmatrix} = -2.$$

Hence, the characteristic equation of $\mathbf{A}$ is $-\lambda^3 + 2\lambda^2 + \lambda - 2$, whence its roots are $\lambda = -1, 1, 2$.

Note that

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

The eigenvectors are thus

$$\mathbf{e}_1 = -\frac{1}{3}\begin{pmatrix} 2-(-1) \\ -3 \\ 0 \end{pmatrix} \times \begin{pmatrix} 1 \\ -1-(-1) \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix},$$
$$\mathbf{e}_2 = -\begin{pmatrix} 2-1 \\ -3 \\ 0 \end{pmatrix} \times \begin{pmatrix} 1 \\ -1-1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix},$$
$$\mathbf{e}_3 = \frac{1}{3}\begin{pmatrix} 2-2 \\ -3 \\ 0 \end{pmatrix} \times \begin{pmatrix} 1 \\ -1-2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Part (a). Note that

$$(\mathbf{A} + 10\mathbf{I})\,\mathbf{e} = \mathbf{A}\mathbf{e} + 10\mathbf{e} = \lambda\mathbf{e} + 10\mathbf{e} = (\lambda + 10)\mathbf{e}.$$

Hence, the eigenvalues of $\mathbf{A} + 10\mathbf{I}$ are 9, 11 and 12. Their corresponding eigenvectors are $(1,\ 1,\ -1)^\mathsf{T}$, $(3,\ 1,\ -1)^\mathsf{T}$ and $(-1,\ 0,\ 1)^\mathsf{T}$.

Part (b). Note that $\mathbf{A} = \mathbf{QDQ}^{-1}$, where

$$\mathbf{Q} = \begin{pmatrix} 1 & 3 & -1 \\ 1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Rearranging, we have $\mathbf{Q}^{-1}\mathbf{AQ} = \mathbf{D}$. Hence,

$$\mathbf{P} = \mathbf{Q}^{-1} = \frac{1}{2}\begin{pmatrix} -1 & 2 & -1 \\ 1 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix}.$$

* * * * *

Problem 4.

(a) Given two square matrices $\mathbf{A}$ and $\mathbf{B}$ of the same order, show that if there exists a non-singular matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{PBP}^{-1}$, then, $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvalues.

(b) Hence, by considering the product

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 1 & 5 & 0 \\ 2 & 3 & 7 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

or otherwise, find the eigenvalues and corresponding eigenvectors of the matrix $\mathbf{M}$, where

$$\mathbf{M} = \begin{pmatrix} 2 & -3 & -5 \\ 1 & 5 & 1 \\ 2 & 3 & 9 \end{pmatrix}.$$

(c) Find a matrix $\mathbf{Q}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{M} = \mathbf{QDQ}^{-1}$, and deduce that if n is a positive integer, then

$$\mathbf{M}^n = \mathbf{Q} \begin{pmatrix} 4^n & 0 & 0 \\ 0 & 5^n & 0 \\ 0 & 0 & 7^n \end{pmatrix} \mathbf{Q}^{-1}.$$

Solution.

Part (a). Let $\mathbf{B} = \mathbf{QDQ}^{-1}$, where $\mathbf{D}$ is a diagonal matrix. Then the principal diagonal of $\mathbf{D}$ contains the eigenvalues of $\mathbf{B}$. Further,

$$\mathbf{A} = \mathbf{PBP}^{-1} = \mathbf{PQDQ}^{-1}\mathbf{P}^{-1} = (\mathbf{PQ})\,\mathbf{D}\,(\mathbf{PQ})\,,$$

so the principal diagonal of $\mathbf{D}$ also contains the eigenvalues of $\mathbf{A}$. Thus, $\mathbf{A}$ and $\mathbf{B}$ must have the same eigenvalues.

Part (b). Since

$$\begin{pmatrix} 4 & 0 & 0 \\ 1 & 5 & 0 \\ 2 & 3 & 7 \end{pmatrix}$$

is triangular, its eigenvalues are simply the entries of its principal diagonal, i.e. $\lambda = 4, 5, 7$. Since

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 1 & 5 & 0 \\ 2 & 3 & 7 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 1 & 5 & 0 \\ 2 & 3 & 7 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1},$$

it follows from (a) that $\mathbf{M}$ also has eigenvalues $\lambda = 4, 5, 7$.

Part (c). Note that

$$\begin{pmatrix} 4 & 0 & 0 \\ 1 & 5 & 0 \\ 2 & 3 & 7 \end{pmatrix} - \lambda\mathbf{I} = \begin{pmatrix} 4-\lambda & 0 & 0 \\ 1 & 5-\lambda & 0 \\ 2 & 3 & 7-\lambda \end{pmatrix}.$$

The eigenvectors are hence

$$\mathbf{e}_1 = \begin{pmatrix} 2 \\ 3 \\ 7-4 \end{pmatrix} \times \begin{pmatrix} 1 \\ 5-4 \\ 0 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \\ -1 \end{pmatrix},$$

$$\mathbf{e}_2 = \begin{pmatrix} 4-5 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 2 \\ 3 \\ 7-5 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ -3 \end{pmatrix},$$

$$\mathbf{e}_3 = \begin{pmatrix} 4-7 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 1 \\ 5-7 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 6 \end{pmatrix}.$$

Thus,

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -3 & 0 & 0 \\ 3 & 2 & 0 \\ -1 & -3 & 6 \end{pmatrix} = \begin{pmatrix} -2 & 3 & -6 \\ 3 & 2 & 0 \\ -1 & -3 & -6 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{pmatrix}.$$

Note also that

$$\mathbf{M} = \underbrace{\left(\mathbf{QDQ}^{-1}\right)\left(\mathbf{QDQ}^{-1}\right)\ldots\left(\mathbf{QDQ}^{-1}\right)}_{n \text{ times}} = \mathbf{QD}^n\mathbf{Q}^{-1} = \mathbf{Q}\begin{pmatrix} 4^n & 0 & 0 \\ 0 & 5^n & 0 \\ 0 & 0 & 7^n \end{pmatrix}\mathbf{Q}^{-1}.$$

* * * * *

Problem 5. The matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{pmatrix} -1 & 5 & 0 \\ 2 & 3 & b \\ 0 & a & -1 \end{pmatrix}.$$

(a) Given that 5 is an eigenvalue of $\mathbf{A}$ with eigenvector $(5,\, 6,\, 2)^\mathsf{T}$, find the values of a and b.

(b) Find the other eigenvalues of $\mathbf{A}$ and their corresponding eigenvectors.

(c) Hence, state the matrices P and D such that $\mathbf{A} = \mathbf{PDP}^{-1}$

The matrix $\mathbf{B}$ is such that $\mathbf{B} = \mathbf{A}^2 - 2\mathbf{A} + 3\mathbf{I}$.

(d) Find a matrix $\mathbf{Q}$ and a diagonal matrix $\mathbf{E}$ such that $\mathbf{E} = \mathbf{QBQ}^{-1}$.

Solution.

Part (a). We have

$$\begin{pmatrix} -1 & 5 & 0 \\ 2 & 3 & b \\ 0 & a & -1 \end{pmatrix}\begin{pmatrix} 5 \\ 6 \\ 2 \end{pmatrix} = 5\begin{pmatrix} 5 \\ 6 \\ 2 \end{pmatrix} \implies \begin{pmatrix} 25 \\ 28+2b \\ 6a-2 \end{pmatrix} = \begin{pmatrix} 25 \\ 30 \\ 10 \end{pmatrix}.$$

Thus, $a = 2$ and $b = 1$.

Part (b). Note that

$$5 + \lambda_2 + \lambda_3 = |-1| + |3| + |-1| = 1 \quad \text{and} \quad 5\lambda_2\lambda_3 = \begin{vmatrix} -1 & 5 & 0 \\ 2 & 3 & 1 \\ 0 & 2 & -1 \end{vmatrix} = 15.$$

By inspection, we have $\lambda_2 = -3$ and $\lambda_3 = -1$.

Note that

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} -1-\lambda & 5 & 0 \\ 2 & 3-\lambda & 1 \\ 0 & 2 & -1-\lambda \end{pmatrix}.$$

The eigenvectors are thus

$$\mathbf{e}_2 = \begin{pmatrix} -1-(-3) \\ 5 \\ 0 \end{pmatrix} \times \begin{pmatrix} 2 \\ 3-(-3) \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 2 \end{pmatrix}$$

$$\mathbf{e}_3 = \frac{1}{5}\begin{pmatrix} -1-(-1) \\ 5 \\ 0 \end{pmatrix} \times \begin{pmatrix} 2 \\ 3-(-1) \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}.$$

Part (c). We have

$$\mathbf{P} = \begin{pmatrix} 5 & 5 & 1 \\ 6 & -2 & 0 \\ 2 & 2 & -2 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Part (d). Note that the eigenvalues of $\mathbf{B}$ are $\lambda^2 - 2\lambda + 3 = 18, 18, 6$. Hence,

$$\mathbf{B} = \mathbf{A}^2 - 2\mathbf{A} + 3\mathbf{I} = \mathbf{P}\left(\mathbf{D}^2 - 2\mathbf{D} + 3\mathbf{I}\right)\mathbf{P}^{-1} = \mathbf{P}\begin{pmatrix} 18 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & 6 \end{pmatrix}\mathbf{P}^{-1}.$$

Thus,

$$\mathbf{Q} = \mathbf{P}^{-1} = \frac{1}{48}\begin{pmatrix} 2 & 6 & 1 \\ 6 & -6 & 3 \\ 8 & 0 & -20 \end{pmatrix} \quad \text{and} \quad \mathbf{E} = \begin{pmatrix} 18 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

* * * * *

Problem 6.

(a) Determine the eigenvalues of a square matrix, $\mathbf{A}$, if

(i) $\mathbf{A}^n = \mathbf{0}$ for some positive integer n,

(ii) $\mathbf{A}^3 = \mathbf{A}$.

(b) The matrices $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. The corresponding eigenvalues of $\mathbf{A}$ are λ_1, λ_2 and λ_3 while the corresponding eigenvalues of $\mathbf{B}$ are μ_1, μ_2 and μ_3.

(i) Show that the matrix $\mathbf{A} + \mathbf{B}$ has eigenvalues $\lambda_1 + \mu_1$, $\lambda_2 + \mu_2$ and $\lambda_3 + \mu_3$ with corresponding common eigenvectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

It is given that

$$\mathbf{A} = \begin{pmatrix} 0 & -1 & 0 \\ -4 & -9 & -6 \\ 5 & 11 & 7 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} -4 & -16 & -11 \\ -9 & -27 & -19 \\ 14 & 44 & 31 \end{pmatrix}$$

and $\mu_1 = -3$, $\mu_2 = 2$, $\mu_3 = 1$.

(ii) Find λ_1, λ_2, λ_3, where $\lambda_1 < \lambda_2 < \lambda_3$, and the corresponding $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

(iii) Find matrices $\mathbf{R}$ and $\mathbf{S}$ and a diagonal matrix $\mathbf{D}$ such that $(\mathbf{A} + \mathbf{B})^5 = \mathbf{RDS}$.

Solution.

Part (a). Let λ be an eigenvalue of $\mathbf{A}$.

Part (a)(i). Since $\mathbf{A}^5 = \mathbf{0}$, we have $\lambda^5 = 0$. Hence, the eigenvalues of $\mathbf{A}$ are all 0.

Part (a)(ii). Since $\mathbf{A}^3 = \mathbf{A}$, we have $\lambda^3 = \lambda$. Hence, $\lambda = -1, 0, 1$.

Part (b).

Part (b)(i). Since

$$(\mathbf{A} + \mathbf{B})\,\mathbf{e}_i = \mathbf{A}\mathbf{e}_i + \mathbf{B}\mathbf{e}_i = \lambda_i\mathbf{e}_i + \mu_i\mathbf{e}_i = (\lambda_i + \mu_i)\mathbf{e}_i,$$

it follows that $\mathbf{A} + \mathbf{B}$ has eigenvalues $\lambda_i + \mu_i$ with corresponding eigenvectors $\mathbf{e}_i$ for $i = 1, 2, 3$.

Part (b)(ii). We have the system of equations

$$\begin{aligned}
\lambda_1+\lambda_2+\lambda_3 &= \left|0\right| + \left|-9\right| + \left|7\right| = -2,\\
\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_3\lambda_1 &= \begin{vmatrix} 0 & -1 \\ -4 & -9 \end{vmatrix} + \begin{vmatrix} -9 & -6 \\ 11 & 7 \end{vmatrix} + \begin{vmatrix} 0 & 0 \\ 5 & 7 \end{vmatrix} = -1\\
\lambda_1\lambda_2\lambda_3 &= \begin{vmatrix} 0 & -1 & 0 \\ -4 & -9 & -6 \\ 5 & 11 & 7 \end{vmatrix} = 2.
\end{aligned}$$

Hence, the characteristic polynomial of $\mathbf{A}$ is $-\lambda^3-2\lambda^2+\lambda+2$. Solving, we have $\lambda_1=-2$, $\lambda_2=-1$ and $\lambda_3=1$.

Note that

$$\mathbf{A}-\lambda_i\mathbf{I} = \begin{pmatrix} -\lambda_i & -1 & 0 \\ -4 & -9-\lambda_i & -6 \\ 5 & 11 & 7-\lambda_i \end{pmatrix}.$$

Hence, we take

$$\begin{aligned}
\mathbf{e}_1 &= \frac{1}{6}\begin{pmatrix} -(-2) \\ -1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -4 \\ -9-(-2) \\ -6 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix},\\
\mathbf{e}_2 &= \frac{1}{6}\begin{pmatrix} -(-1) \\ -1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -4 \\ -9-(-1) \\ -6 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}\\
\mathbf{e}_3 &= \frac{1}{6}\begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -4 \\ -9-1 \\ -6 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.
\end{aligned}$$

Part (b)(iii). Note that $\mathbf{A}+\mathbf{B} = \mathbf{P}\mathbf{E}\mathbf{P}^{-1}$, where

$$\mathbf{P} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ -3 & -2 & 1 \end{pmatrix}$$

and

$$\mathbf{E} = \begin{pmatrix} \lambda_1+\mu_1 & 0 & 0 \\ 0 & \lambda_2+\mu_2 & 0 \\ 0 & 0 & \lambda_3+\mu_3 \end{pmatrix} = \begin{pmatrix} -5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Hence, $(\mathbf{A}+\mathbf{B})^5 = \mathbf{P}\mathbf{E}^5\mathbf{P}^{-1} = \mathbf{R}\mathbf{D}\mathbf{S}$. We thus take

$$\begin{aligned}
\mathbf{R} = \mathbf{P} &= \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ -3 & -2 & 1 \end{pmatrix},\\
\mathbf{S} = \mathbf{P}^{-1} &= \begin{pmatrix} 1 & 3 & 2 \\ -1 & -4 & -3 \\ 1 & 1 & 1 \end{pmatrix},\\
\mathbf{D} = \mathbf{E}^5 &= \begin{pmatrix} -3125 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 32 \end{pmatrix}.
\end{aligned}$$

Problem 7. A square matrix $\mathbf{A}$ has eigenvector $\mathbf{x}$ with corresponding eigenvalue λ.

(a) Show that $\mathbf{x}$ is also an eigenvector of $\mathbf{A}+\mathbf{A}^2+\mathbf{A}^3+\mathbf{A}^4$ with corresponding eigenvalue $\lambda+\lambda^2+\lambda^3+\lambda^4$.

Let $\mathbf{B} = \mathbf{A}+\mathbf{A}^2+\mathbf{A}^3+\mathbf{A}^4$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 4 \\ 0 & 2 & 8 \\ 0 & 0 & -3 \end{pmatrix}.$$

(b) Using the result from part (a), find the eigenvectors and corresponding eigenvalues of $\mathbf{B}$.

(c) Hence, write down a non-singular matrix $\mathbf{Q}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{B} = \mathbf{Q}\mathbf{D}\mathbf{Q}^{-1}$.

Solution.

Part (a). Observe that

$$\begin{aligned}\left(\mathbf{A}+\mathbf{A}^2+\mathbf{A}^3+\mathbf{A}^4\right)\mathbf{x} &= \mathbf{A}\mathbf{x}+\mathbf{A}^2\mathbf{x}+\mathbf{A}^3\mathbf{x}+\mathbf{A}^4\mathbf{x} \\ &= \lambda\mathbf{x}+\lambda^2\mathbf{x}+\lambda^3\mathbf{x}+\lambda^4\mathbf{x} = \left(\lambda+\lambda^2+\lambda^3+\lambda^4\right)\mathbf{x}.\end{aligned}$$

Hence, $\mathbf{x}$ is also an eigenvector of $\mathbf{A}+\mathbf{A}^2+\mathbf{A}^3+\mathbf{A}^4$ with corresponding eigenvalue $\lambda+\lambda^2+\lambda^3+\lambda^4$.

Part (b). Since $\mathbf{A}$ is triangular, the entries on its principal diagonal are precisely its eigenvectors. Hence, $\lambda = 1, 2, -3$. Thus, $\mathbf{B}$ has eigenvalues $\lambda+\lambda^2+\lambda^3+\lambda^4 = 4, 30, 60$.

Note that

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} 1-\lambda & 3 & 4 \\ 0 & 2-\lambda & 8 \\ 0 & 0 & -3-\lambda \end{pmatrix}.$$

The eigenvectors of $\mathbf{A}$ and $\mathbf{B}$ are hence

$$\begin{aligned}\mathbf{e}_1 &= -\frac{1}{12}\begin{pmatrix} 1-1 \\ 3 \\ 4 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ -3-1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \\ \mathbf{e}_2 &= -\frac{1}{5}\begin{pmatrix} 1-2 \\ 3 \\ 4 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ -3-2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \\ \mathbf{e}_3 &= \frac{1}{4}\begin{pmatrix} 1-(-3) \\ 3 \\ 4 \end{pmatrix} \times \begin{pmatrix} 0 \\ 2-(-3) \\ 8 \end{pmatrix} = \begin{pmatrix} 1 \\ -8 \\ 5 \end{pmatrix}.\end{aligned}$$

Part (c). We have

$$\mathbf{Q} = \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & -8 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 30 & 0 \\ 0 & 0 & 60 \end{pmatrix}.$$

Problem 8. Two $n \times n$ square matrices $\mathbf{A}$ and $\mathbf{B}$ are said to commute if $\mathbf{AB} = \mathbf{BA}$.

(a) Show that any two diagonal $n \times n$ square matrices commute.

(b) Let $\mathbf{A}$ and $\mathbf{B}$ be two $n \times n$ square matrices with the same eigenvectors. By diagonalizing the square matrices, show that $\mathbf{A}$ and $\mathbf{B}$ commute.

(c) The matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{pmatrix} -5 & -5 & -2 \\ 8 & 4 & 4 \\ 16 & 10 & 7 \end{pmatrix}.$$

Find the eigenvalues of $\mathbf{A}$ and corresponding eigenvectors with integer entries.

(d) Write down a matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$.

(e) The matrix $\mathbf{B}$ is given by

$$\mathbf{B} = \begin{pmatrix} -11 & 3 & -6 \\ 8 & -2 & 4 \\ 16 & -6 & 9 \end{pmatrix}.$$

Determine whether $\mathbf{B}$ is diagonalizable with your matrix $\mathbf{P}$ from (d), and hence deduce whether $\mathbf{A}$ and $\mathbf{B}$ commutes.

Solution.

Part (a). Let $\mathbf{x}_i$ be the ith standard basis vector, where $1 \leq i \leq n$. Note that $\mathbf{x}_i$ is an eigenvector of a diagonal matrix $\mathbf{D}$ with corresponding eigenvalue d_{ii}. Hence,

$$\mathbf{ABx}_i = \mathbf{A}\left(b_{ii}\mathbf{x}_i\right) = a_{ii}b_{ii}\mathbf{x}_i \quad \text{and} \quad \mathbf{BAx}_i = \mathbf{B}\left(a_{ii}\mathbf{x}_i\right) = b_{ii}a_{ii}\mathbf{x}_i.$$

Hence, $\mathbf{x}_i$ has the same image under $\mathbf{AB}$ and $\mathbf{BA}$, whence $\mathbf{AB} = \mathbf{BA}$.

Part (b). Since $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvectors, we can write $\mathbf{A} = \mathbf{PDP}^{-1}$ and $\mathbf{B} = \mathbf{PEP}^{-1}$, where $\mathbf{D}$ and $\mathbf{E}$ are diagonal matrices. Then

$$\mathbf{AB} = \left(\mathbf{PDP}^{-1}\right)\left(\mathbf{PEP}^{-1}\right) = \mathbf{PDEP}^{-1}$$

and

$$\mathbf{BA} = \left(\mathbf{PEP}^{-1}\right)\left(\mathbf{PDP}^{-1}\right) = \mathbf{PEDP}^{-1}.$$

Since $\mathbf{D}$ and $\mathbf{E}$ are diagonal, by part (a), they commute, so $\mathbf{DE} = \mathbf{ED}$. Consequently, $\mathbf{AB} = \mathbf{BA}$, so $\mathbf{A}$ and $\mathbf{B}$ also commute.

Part (c). We have the system of equations

$$\begin{aligned}
\lambda_1 + \lambda_2 + \lambda_3 &= \left|-5\right| + \left|4\right| + \left|7\right| = 6, \\
\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 &= \begin{vmatrix} -5 & -5 \\ 8 & 4 \end{vmatrix} + \begin{vmatrix} 4 & 4 \\ 10 & 7 \end{vmatrix} + \begin{vmatrix} -5 & -2 \\ 16 & 7 \end{vmatrix} = 5 \\
\lambda_1\lambda_2\lambda_3 &= \begin{vmatrix} -5 & -5 & -2 \\ 8 & 4 & 4 \\ 16 & 10 & 7 \end{vmatrix} = -12.
\end{aligned}$$

Hence, the characteristic equation of $\mathbf{A}$ is $-\lambda^3 + 6\lambda^2 - 5\lambda - 12$. Solving, we get $\lambda = -1, 3, 4$.

Note that

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} -5-\lambda & -5 & -2 \\ 8 & 4-\lambda & 4 \\ 16 & 10 & 7-\lambda \end{pmatrix}.$$

The corresponding eigenvectors are thus

$$\mathbf{e}_1 = \frac{1}{10}\begin{pmatrix} -5-(-1) \\ -5 \\ -2 \end{pmatrix} \times \begin{pmatrix} 8 \\ 4-(-1) \\ 4 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix},$$
$$\mathbf{e}_2 = \frac{1}{2}\begin{pmatrix} -5-3 \\ -5 \\ -2 \end{pmatrix} \times \begin{pmatrix} 8 \\ 4-3 \\ 4 \end{pmatrix} = \begin{pmatrix} -9 \\ 8 \\ 16 \end{pmatrix},$$
$$\mathbf{e}_3 = \frac{1}{20}\begin{pmatrix} -5-4 \\ -5 \\ -2 \end{pmatrix} \times \begin{pmatrix} 8 \\ 4-4 \\ 4 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}.$$

Part (d). Rearranging, we have $\mathbf{A} = \mathbf{PDP}^{-1}$. Thus,

$$\mathbf{P} = \begin{pmatrix} -1 & -9 & -1 \\ 0 & 8 & 1 \\ 2 & 16 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Part (e). Observe that

$$\mathbf{B}\mathbf{e}_1 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} = \mathbf{e}_1, \quad \mathbf{B}\mathbf{e}_2 = \begin{pmatrix} 27 \\ -24 \\ -48 \end{pmatrix} = -3\mathbf{e}_1, \quad \mathbf{B}\mathbf{e}_3 = \begin{pmatrix} 2 \\ -2 \\ -4 \end{pmatrix} = -2\begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} = -2\mathbf{e}_3.$$

Hence, $\mathbf{A}$ and $\mathbf{B}$ share the share eigenvectors. Thus, by part (b), $\mathbf{A}$ and $\mathbf{B}$ commute.

* * * * *

Problem 9. The linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ is represented by the matrix $\mathbf{A}$, where

$$\mathbf{A} = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}.$$

(a) Show that T transforms any point on the line $y = x$ to a point on the same line.

(b) Explain what happens to the points on the line $4y+x = 0$ when they are transformed by T.

(c) State the two eigenvalues of $\mathbf{A}$ and state two eigenvectors corresponding to the two eigenvalues.

Solution.

Part (a). Let $\mathbf{x}$ be a point on the line $y = x$. Then $\mathbf{x}$ has the form $t\,(1,\, 1)^{\mathsf{T}}$, where $t \in \mathbb{R}$. Now observe that

$$T\mathbf{x} = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}\left[t\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right] = t\begin{pmatrix} -2 \\ -2 \end{pmatrix} = -2\mathbf{x}.$$

Hence, the image of $\mathbf{x}$ under T is a point on the same line $y = x$.

Part (b). Let $\mathbf{x}$ be a point on the line $4y + x = 0$. Then $\mathbf{x}$ has the form $t\,(-4,\, 1)^{\mathsf{T}}$, where $t \in \mathbb{R}$. Now observe that

$$T\mathbf{x} = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}\left[t\begin{pmatrix} -4 \\ 1 \end{pmatrix}\right] = t\begin{pmatrix} -12 \\ 3 \end{pmatrix} = 3\mathbf{x}.$$

Hence, the image of $\mathbf{x}$ under T is a point on the same line $4y + x = 0$.

Part (c). The eigenvalues of $\mathbf{A}$ are $\lambda = -2, 3$, and their corresponding eigenvectors are $(1,\ 1)^{\mathsf{T}}$ and $(-4,\ 1)^{\mathsf{T}}$.

* * * * *

Problem 10.

(a) (i) Suppose that an invertible 3×3 matrix $\mathbf{A}$ has non-zero eigenvalues λ_1, λ_2, λ_3 with corresponding eigenvectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Show that the matrix $\mathbf{A}^{-1}$ has the same eigenvectors and find the corresponding eigenvalues of $\mathbf{A}^{-1}$.

(ii) Another 3×3 matrix $\mathbf{B}$ has eigenvalues μ_1, μ_2, μ_3 with corresponding eigenvectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Determine the eigenvalues and the corresponding eigenvectors of the matrix $\mathbf{A}^{-1}\mathbf{B}$. Hence, find the eigenvalue of the matrix $\mathbf{C}$ corresponding to the eigenvector $\mathbf{e}_1$, where

$$\mathbf{C} = \mathbf{I} + \mathbf{A}^{-1}\mathbf{B} + \left(\mathbf{A}^{-1}\mathbf{B}\right)^2 + \cdots + \left(\mathbf{A}^{-1}\mathbf{B}\right)^n,$$

where $n \in \mathbb{Z}^+$ and $\mu_1 \neq \lambda_1$.

(b) The matrix

$$\mathbf{E} = \begin{pmatrix} a & b \\ -1 & 0 \end{pmatrix},$$

where $a, b \in \mathbb{R}$, has real eigenvalues β_1, β_2.

(i) If $\beta_1 \neq \beta_2$, show that $a^2 > 4b$.

(ii) Assume that $a \neq 0$. State the value of b when $\mathbf{E}$ is singular. Find a matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{E} = \mathbf{PEP}^{-1}$ when $\mathbf{E}$ is singular.

Solution.

Part (a).

Part (a)(i). We have $\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$. Pre-multiplying by $\frac{1}{\lambda_i}\mathbf{A}^{-1}$ yields $\frac{1}{\lambda_i}\mathbf{e}_i = \mathbf{A}^{-1}\mathbf{e}_i$. Hence, $\mathbf{A}^{-1}$ has the same eigenvectors $\mathbf{e}_i$ with corresponding eigenvalues $1/\lambda_i$.

Part (a)(ii). Note that

$$\mathbf{A}^{-1}\mathbf{B}\mathbf{e}_i = \mathbf{A}^{-1}\mu_i\mathbf{e}_i = \frac{\mu_i}{\lambda_i}\mathbf{e}_i.$$

Thus,

$$\begin{aligned} \mathbf{C}\mathbf{e}_1 &= \left[\mathbf{I} + \mathbf{A}^{-1}\mathbf{B} + \left(\mathbf{A}^{-1}\mathbf{B}\right)^2 + \cdots + \left(\mathbf{A}^{-1}\mathbf{B}\right)^n\right]\mathbf{e}_1 \\ &= \left[1 + \frac{\mu_1}{\lambda_1} + \left(\frac{\mu_1}{\lambda_1}\right)^2 + \cdots + \left(\frac{\mu_1}{\lambda_1}\right)^n\right]\mathbf{e} = \frac{(\mu_1/\lambda_1)^{n+1} - 1}{\mu_1/\lambda_1 - 1}\mathbf{e}_1. \end{aligned}$$

Thus, the eigenvalue corresponding to $\mathbf{e}_1$ is $\frac{(\mu_1/\lambda_1)^{n+1}-1}{\mu_1/\lambda_1 - 1}$.

Part (b).

Part (b)(i). The characteristic polynomial of $\mathbf{E}$ is $\lambda^2 - a\lambda + b$. Since there are two real and distinct eigenvalues, the discriminant of the characteristic polynomial must be strictly greater than 0. Hence, $a^2 - 4b > 0 \implies a^2 > 4b$.

Part (b)(ii). For $\mathbf{E}$ to be singular, we require $b = 0$. In this case, the characteristic polynomial of $\mathbf{E}$ is $\lambda^2 - a\lambda$. Its roots are $\lambda = 0, a$. Let $\mathbf{x} = (x,\ y)^{\mathsf{T}} \neq 0$ be an eigenvector.

Case 1: $\lambda = 0$. Consider $(\mathbf{E} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$:

$$\begin{pmatrix} a & 0 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Hence, $x = 0$, while y is free. Hence, $\mathbf{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Case 2: $\lambda = a$. Consider $(\mathbf{E} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$:

$$\begin{pmatrix} 0 & 0 \\ -1 & -a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We get the equation $x + ay = 0$. Taking $y = 1$, we have $\mathbf{x} = (-a,\, 1)^{\mathsf{T}}$.

Thus,

$$\mathbf{P} = \begin{pmatrix} 0 & -a \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}.$$

* * * * *

Problem 11. A microbiologist wants to investigate the growth of three different species of microorganisms in a controlled habitat. She sets up the following model

$$\begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix} = \mathbf{Q} \begin{pmatrix} a_{n-1} \\ b_{n-1} \\ c_{n-1} \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} 0.42 & 0.076 & 1.16 \\ 0.6 & 0.68 & -1.2 \\ -0.1 & 0.02 & 1.2 \end{pmatrix},$$

where a_n, b_n and c_n represents the number of microorganisms A, B and C (in billions) respectively, n hours after the start of the experiment.

(a) Give an observation that the microbiologist may expect to see pertaining to the population of each of the microorganisms A, B and C, justifying your answer by drawing references to the entries in the matrix $\mathbf{Q}$.

(b) Explain numerically why the model will fail in predicting the growth of the microorganisms for the initial population of $a_0 = 10$, $b_0 = 20$ and $c_0 = 30$.

The microbiologist wants to predict the long-term growth of the microorganisms for the case where the scenario in (b) does not occur. She turns to a mathematician for help. The mathematician advises her to diagonalize $\mathbf{Q}$ into the form $\mathbf{PDP}^{-1}$, such that

$$\mathbf{P} = \begin{pmatrix} 5 & 1 & 2 \\ -10 & 5 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

(c) Find the matrices $\mathbf{D}$ and $\mathbf{P}^{-1}$ corresponding to the given matrix $\mathbf{P}$.

(d) Show how this diagonalization process can be used to help the microbiologist predict that the population of the microorganisms stabilize in the long run. Determine the equilibrium population of the microorganisms in the long run for the case where $a_0 = 40$, $b_0 = 20$ and $c_0 = 6$.

(e) Comment on one possible drawback of the model.

Solution.

Part (a). Microorganism C will likely have a large population. Further, $q_{13} = 1.16$ means that a high population of C will result in a large population of A, while $q_{23} = -1.2$ means that a large population of C will result in a small population of C.

Part (b). Note that

$$\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} = \mathbf{Q} \begin{pmatrix} a_0 \\ b_0 \\ c_0 \end{pmatrix} = \begin{pmatrix} 0.42 & 0.076 & 1.16 \\ 0.6 & 0.68 & -1.2 \\ -0.1 & 0.02 & 1.2 \end{pmatrix} \begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix} = \begin{pmatrix} 40.52 \\ -16.4 \\ 35.4 \end{pmatrix}.$$

This is absurd since populations cannot be negative. Hence, the model fails when $a_0 = 10$, $b_0 = 20$ and $c_0 = 30$.

Part (c). Clearly,

$$\mathbf{P}^{-1} = \frac{1}{25} \begin{pmatrix} 5 & -1 & -10 \\ 10 & 3 & -20 \\ -5 & 1 & 35 \end{pmatrix}.$$

Rearranging the given equation, we obtain

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{Q}\mathbf{P} = \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Part (d). Note that

$$\lim_{n\to\infty} \mathbf{D}^n = \lim_{n\to\infty} \begin{pmatrix} 0.5^n & 0 & 0 \\ 0 & 0.8^n & 0 \\ 0 & 0 & 1^n \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence,

$$\lim_{n\to\infty} \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix} = \lim_{n\to\infty} \mathbf{Q}^n \begin{pmatrix} a_0 \\ b_0 \\ c_0 \end{pmatrix} = \mathbf{P} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{P}^{-1} \begin{pmatrix} 40 \\ 20 \\ 6 \end{pmatrix} = \begin{pmatrix} 2.4 \\ 0 \\ 1.2 \end{pmatrix}.$$

Thus, in the long run, microorganisms A and C will have a population of 2.4 and 1.2 billion respectively, while microorganism B will die out.

Part (e). The model does not take depletion of resources (e.g. space and food) into account.

Self-Practice B17C

Problem 1. The vector $\mathbf{x}$ is an eigenvector of the matrices $\mathbf{A}$ and $\mathbf{B}$ with corresponding eigenvalues λ and μ respectively. Show that $\mathbf{x}$ is an eigenvector of $\mathbf{AB}$ with corresponding eigenvalues $\lambda\mu$.

Find the eigenvalues and corresponding eigenvectors of $\mathbf{A}$, where

$$\mathbf{A} = \begin{pmatrix} -11 & 3 & -6 \\ 8 & -2 & 4 \\ 16 & -6 & 9 \end{pmatrix}.$$

The matrix $\mathbf{B}$ has eigenvectors

$$\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}, \begin{pmatrix} 9 \\ -8 \\ -16 \end{pmatrix},$$

with corresponding eigenvalues -1, 3, 4 respectively.

(a) Without evaluating $\mathbf{AB}$ and $\mathbf{BA}$, determine whether $\mathbf{AB} = \mathbf{BA}$. Justify your conclusion.

(b) Find a matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $(\mathbf{AB})^2 = \mathbf{PDP}^{-1}$.

Solution. By the definition of an eigenvector, one has $\mathbf{Ax} = \lambda\mathbf{x}$ and $\mathbf{Bx} = \mu\mathbf{x}$. Thus,

$$\mathbf{ABx} = \mathbf{A}\mu\mathbf{x} = \mu\mathbf{Ax} = \mu\lambda\mathbf{x} = \lambda\mu\mathbf{x},$$

so $\mathbf{x}$ is an eigenvector of $\mathbf{AB}$ with corresponding eigenvalue $\lambda\mu$.

Let the characteristic polynomial of $\mathbf{A}$ be $\chi(\lambda) = \lambda^3 - c_2\lambda^2 + c_1\lambda - c_0$. Then

$$\begin{aligned} c_0 &= |\mathbf{A}| = 6, \\ c_1 &= \begin{vmatrix} -11 & 3 \\ 8 & -2 \end{vmatrix} + \begin{vmatrix} -2 & 4 \\ -6 & 9 \end{vmatrix} + \begin{vmatrix} -11 & -6 \\ 16 & 9 \end{vmatrix} = 1, \\ c_2 &= |-11| + |-2| + |9| = -4, \end{aligned}$$

Thus, the characteristic polynomial of $\mathbf{A}$ is

$$\chi(\lambda) = \lambda^3 + 4\lambda^2 + \lambda - 6 = (\lambda - 1)(\lambda + 2)(\lambda + 3),$$

so $\mathbf{A}$ has eigenvalues $\lambda = 1, -2, -3$.

Note that

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} -11-\lambda & 3 & -6 \\ 8 & -2-\lambda & 4 \\ 16 & -6 & 9-\lambda \end{pmatrix}.$$

Case 1: $\lambda = 1$. Observe that

$$\begin{pmatrix} -12 \\ 3 \\ -6 \end{pmatrix} \times \begin{pmatrix} 8 \\ -3 \\ 4 \end{pmatrix} = -6 \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}.$$

We hence take $\mathbf{x} = (1,\ 0,\ -2)^\mathsf{T}$ to be the corresponding eigenvector of $\lambda = 1$.

Case 2: $\lambda = -2$. Observe that

$$\begin{pmatrix} -9 \\ 3 \\ -6 \end{pmatrix} \times \begin{pmatrix} 8 \\ 0 \\ 4 \end{pmatrix} = 12 \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}.$$

We hence take $\mathbf{x} = (1,\, -1,\, -2)^\mathsf{T}$ to be the corresponding eigenvector of $\lambda = -2$.

Case 3: $\lambda = -3$. Observe that

$$\begin{pmatrix} -8 \\ 3 \\ -6 \end{pmatrix} \times \begin{pmatrix} 8 \\ 1 \\ 4 \end{pmatrix} = 2 \begin{pmatrix} 9 \\ -8 \\ -16 \end{pmatrix}.$$

We hence take $\mathbf{x} = (9,\, -8,\, -16)^\mathsf{T}$ to be the corresponding eigenvector of $\lambda = -3$.

Part (a). Since $\mathbf{A}$ and $\mathbf{B}$ share the same linearly independent eigenvectors, we can write $\mathbf{A} = \mathbf{P}\mathbf{D}_A\mathbf{P}^{-1}$ and $\mathbf{B} = \mathbf{P}\mathbf{D}_B\mathbf{P}^{-1}$ for some diagonal matrices $\mathbf{D}_A$ and $\mathbf{D}_B$. Then

$$\mathbf{AB} = \left(\mathbf{P}\mathbf{D}_A\mathbf{P}^{-1}\right)\left(\mathbf{P}\mathbf{D}_B\mathbf{P}^{-1}\right) = \mathbf{P}\mathbf{D}_A\mathbf{D}_B\mathbf{P}^{-1}.$$

Since diagonal matrices commute, we see that

$$\mathbf{AB} = \mathbf{P}\mathbf{D}_B\mathbf{D}_A\mathbf{P}^{-1} = \left(\mathbf{P}\mathbf{D}_B\mathbf{P}^{-1}\right)\left(\mathbf{P}\mathbf{D}_A\mathbf{P}^{-1}\right) = \mathbf{BA}.$$

Part (b). Observe that

$$(\mathbf{AB})^2 = \mathbf{P}\left(\mathbf{D}_A\mathbf{D}_B\right)^2\mathbf{P}^{-1},$$

so the desired matrices are

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 9 \\ 0 & -1 & -8 \\ -2 & -2 & -16 \end{pmatrix}$$

and

$$\mathbf{D} = \left(\mathbf{D}_A\mathbf{D}_B\right)^2 = \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}\right]^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 36 & 0 \\ 0 & 0 & 144 \end{pmatrix}.$$

* * * * *

Problem 2. The eigenvector $\mathbf{x}$ is an eigenvector of the matrix $\mathbf{A}$, with corresponding eigenvalue λ, and $\mathbf{x}$ is also an eigenvector of the matrix $\mathbf{B}$, with corresponding eigenvalue μ. Prove that $\mathbf{x}$ is an eigenvector of the matrix $p\mathbf{A} + q\mathbf{B}$, with corresponding eigenvalue $p\lambda + q\mu$, where $p, q \in \mathbb{R}$. Find the eigenvalues and corresponding eigenvectors of the matrix

$$\mathbf{L} = \begin{pmatrix} -11 & 3 & -6 \\ 8 & -2 & 4 \\ 16 & -6 & 9 \end{pmatrix}.$$

The matrix $\mathbf{M}$ has eigenvectors

$$\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}, \begin{pmatrix} 9 \\ -8 \\ -16 \end{pmatrix},$$

with corresponding eigenvalues -1, 3, 4 respectively. Find a matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $(2\mathbf{L} + 3\mathbf{M})^4 = \mathbf{PDP}^{-1}$.

Solution. By the definition of an eigenvector, one has $\mathbf{Ax} = \lambda\mathbf{x}$ and $\mathbf{Bx} = \mu\mathbf{x}$. Thus,

$$(p\mathbf{A} + q\mathbf{B})\,\mathbf{x} = p\mathbf{Ax} + q\mathbf{Bx} = p\lambda\mathbf{x} + q\mu\mathbf{x} = (p\lambda + q\mu)\,\mathbf{x},$$

so $\mathbf{x}$ is an eigenvector of $p\mathbf{A} + q\mathbf{B}$ with corresponding eigenvalue $p\lambda + q\mu$.

Let the characteristic polynomial of $\mathbf{L}$ be $\chi(\lambda) = \lambda^3 - c_2\lambda^2 + c_1\lambda - c_0$. Then

$$\begin{aligned}
c_0 &= |\mathbf{L}| = 6, \\
c_1 &= \begin{vmatrix} -11 & 3 \\ 8 & -2 \end{vmatrix} + \begin{vmatrix} -2 & 4 \\ -6 & 9 \end{vmatrix} + \begin{vmatrix} -11 & -6 \\ 16 & 9 \end{vmatrix} = 1, \\
c_2 &= |-11| + |-2| + |9| = -4,
\end{aligned}$$

Thus, the characteristic polynomial of $\mathbf{L}$ is

$$\chi(\lambda) = \lambda^3 + 4\lambda^2 + \lambda - 6 = (\lambda - 1)(\lambda + 2)(\lambda + 3),$$

so $\mathbf{A}$ has eigenvalues $\lambda = 1, -2, -3$.

Note that

$$\mathbf{L} - \lambda\mathbf{I} = \begin{pmatrix} -11-\lambda & 3 & -6 \\ 8 & -2-\lambda & 4 \\ 16 & -6 & 9-\lambda \end{pmatrix}.$$

Case 1: $\lambda = 1$. Observe that

$$\begin{pmatrix} -12 \\ 3 \\ -6 \end{pmatrix} \times \begin{pmatrix} 8 \\ -3 \\ 4 \end{pmatrix} = -6 \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}.$$

We hence take $\mathbf{x} = (1,\ 0,\ -2)^\mathsf{T}$ to be the corresponding eigenvector of $\lambda = 1$.

Case 2: $\lambda = -2$. Observe that

$$\begin{pmatrix} -9 \\ 3 \\ -6 \end{pmatrix} \times \begin{pmatrix} 8 \\ 0 \\ 4 \end{pmatrix} = 12 \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}.$$

We hence take $\mathbf{x} = (1,\ -1,\ -2)^\mathsf{T}$ to be the corresponding eigenvector of $\lambda = -2$.

Case 3: $\lambda = -3$. Observe that

$$\begin{pmatrix} -8 \\ 3 \\ -6 \end{pmatrix} \times \begin{pmatrix} 8 \\ 1 \\ 4 \end{pmatrix} = 2 \begin{pmatrix} 9 \\ -8 \\ -16 \end{pmatrix}.$$

We hence take $\mathbf{x} = (9,\ -8,\ -16)^\mathsf{T}$ to be the corresponding eigenvector of $\lambda = -3$.

Note that $2\mathbf{L} + 3\mathbf{M}$ has eigenvalues $2\lambda_i + 3\mu_i = -1, 5, 6$ with corresponding eigenvectors

$$\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}, \begin{pmatrix} 9 \\ -8 \\ -16 \end{pmatrix}.$$

Thus, the desired matrices are

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 9 \\ 0 & -1 & -8 \\ -2 & -2 & -16 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} (-1)^4 & 0 & 0 \\ 0 & 5^4 & 0 \\ 0 & 0 & 6^4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 625 & 0 \\ 0 & 0 & 1296 \end{pmatrix}.$$

* * * * *

Problem 3. Given

$$\mathbf{a} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

find a 3×3 matrix $\mathbf{U}$ having eigenvectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ corresponding to the eigenvalues 1, -1 and 2 respectively.

For the transformation $\mathbf{y} = \mathbf{U}\mathbf{x}$ from vector $\mathbf{x}$ to vector $\mathbf{y}$,

(a) find all the invariant points and all the invariant lines,

(b) given that $\mathbf{x} = (3,\, 2,\, -3)^{\mathsf{T}}$ is in the plane spanned by $\mathbf{b}$ and $\mathbf{c}$, find $\mathbf{U}^{10}\mathbf{x}$.

Solution. We have

$$\mathbf{U} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 1 & 0 \\ -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 \\ 1 & 1 & 0 \\ -1 & -1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -3 & 0 \\ 1 & -1 & 1 \\ -1 & 3 & 1 \end{pmatrix}.$$

Part (a). The invariant points correspond to eigenvectors with eigenvalue 1, so they have the form $\lambda\,(3,\, 1,\, -1)^{\mathsf{T}}$ for some $\lambda \in \mathbb{R}$. The invariant lines correspond to all other eigenvectors, so they have the form $\mu\,(1,\, 1,\, -1)^{\mathsf{T}}$ or $\nu\,(1,\, 0,\, -1)^{\mathsf{T}}$ for some $\mu, \nu \in \mathbb{R}$.

Part (b). Note that $\mathbf{x} = 2\mathbf{b} + \mathbf{c}$. Thus,

$$\mathbf{U}^{10}\mathbf{x} = \mathbf{U}^{10}\,(2\mathbf{b} + \mathbf{c}) = 2\mathbf{U}^{10}\mathbf{b} + \mathbf{U}^{10}\mathbf{c} = 2(-1)^{10}\mathbf{b} + 2^{20}\mathbf{c} = \begin{pmatrix} 1026 \\ 0 \\ -1026 \end{pmatrix}.$$

* * * * *

Problem 4. A Leslie matrix is often used to model population dynamics for different stages of a life cycle of certain species of animals of interest. A general form of a Leslie matrix is given by

$$\begin{pmatrix} f_0 & f_1 & f_2 & \cdots & f_{n-2} & f_{n-1} \\ s_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{n-2} & 0 \end{pmatrix},$$

where f_i and s_i are non-negative real numbers.

(a) Consider a 3×3 Leslie matrix

$$\mathbf{L} = \begin{pmatrix} f_0 & f_1 & f_2 \\ s_0 & 0 & 0 \\ 0 & s_1 & 0 \end{pmatrix}.$$

If $\mathbf{L}$ has complex eigenvalues, prove that $\mathbf{L}$ has exactly one positive eigenvalue.

(b) In the study of the population of locusts in a particular region, it is of important to track the number of locusts in various stages of their life cycle. In particular, the tracking of the number of eggs, nymphs (young locust) and adults are of interest. Let $x_1(t)$, $x_2(t)$ and $x_3(t)$ denote the number of eggs, nymphs and adults at time t, where time in measured in years. From years of studies, the relations between x_1, x_2 and x_3 are given as follows:

$$\begin{aligned} x_1(t+1) &= 1000x_3(t), \\ x_2(t+1) &= 0.02x_1(t), \\ x_3(t+1) &= 0.05x_2(t). \end{aligned}$$

Let $\mathbf{x}_t = (x_1(t),\, x_2(t),\, x_3(t))^{\mathsf{T}}$. The relations can be represented by a Leslie matrix $\mathbf{M}$, where $\mathbf{x}_{t+1} = \mathbf{M}\mathbf{x}_t$.

(i) Identify the matrix $\mathbf{M}$ that represents the growth.

(ii) Given at time $t = 0$, there are only 50 adults and no eggs or nymphs, compute the values of x_1, x_2 and x_3 for the first 6 years.

(iii) Using your answer to (b)(ii), predict the behaviour of the population of locusts.

(c) Consider a population in which both juveniles and adults can reproduce. Denote $v_1(t)$ and $v_2(t)$ as the number of juveniles and adults at time t respectively, and let $\mathbf{v}_t = (v_1(t), v_2(t))^\mathsf{T}$. The recurrence relation for $\mathbf{v}_t$ is given by $\mathbf{v}_{t+1} = \mathbf{A}\mathbf{v}_t$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 0.5 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_0 = \begin{pmatrix} a \\ b \end{pmatrix}.$$

(i) Find the eigenvalues and eigenvectors of the matrix $\mathbf{A}$.

(ii) By using a suitable diagonalization of the matrix $\mathbf{A}$, express $\mathbf{v}_t$ in terms of a, b and t.

(iii) Hence, find the long term proportion of juveniles and adults in this population.

Solution.

Part (a). Let α be a complex eigenvalue. Then α is a root to the characteristic polynomial $\chi(\lambda)$ of $\mathbf{L}$. Since $\chi(\lambda)$ has real coefficients, by the conjugate root theorem, α^* is also a solution to $\chi(\lambda)$ and is hence an eigenvalue too. Let β be the remaining eigenvalue. Since the product of eigenvalues is equal to the determinant of $\mathbf{L}$, we have

$$\alpha\alpha^*\beta = \det \mathbf{L} = f_2 s_0 s_1.$$

But $\alpha\alpha^* = |\alpha|^2$, so

$$\beta = \frac{f_2 s_0 s_1}{|\alpha|^2},$$

which is clearly a positive number (since we are given $f_2, s_0, s_1 > 0$).

Part (b).

Part (b)(i). Note that

$$\mathbf{x}_{t+1} = \begin{pmatrix} x_1(t+1) \\ x_2(t+1) \\ x_3(t+1) \end{pmatrix} = \begin{pmatrix} 1000x_3(t) \\ 0.02x_1(t) \\ 0.05x_2(t) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1000 \\ 0.02 & 0 & 0 \\ 0 & 0.05 & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix},$$

so the desired matrix is

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & 1000 \\ 0.02 & 0 & 0 \\ 0 & 0.05 & 0 \end{pmatrix}.$$

Part (b)(ii).

t	$x_1(t)$	$x_2(t)$	$x_3(t)$
1	50000	0	0
2	0	1000	0
3	0	0	50
4	50000	0	0
5	0	1000	0
6	0	0	50

Part (b)(iii). The population will remain in this cycle, where there are only 50 adults every third year.

Part (c).

Part (c)(i). Note that

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 1-\lambda & 4 \\ 0.5 & -\lambda \end{vmatrix} = \lambda^2 - \lambda - 2 = (\lambda + 1)(\lambda - 2),$$

so the eigenvalues of $\mathbf{A}$ are $\lambda = -1, 2$.

Case 1: $\lambda = -1$. Let $\mathbf{x} = (x,\, y)^{\mathsf{T}}$ be an eigenvector. Then

$$\begin{pmatrix} 1-(-1) & 4 \\ 0.5 & -(-1) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Using G.C., $\mathbf{x} = -2\lambda$ and $\mathbf{y} = \lambda$, where $\lambda \in \mathbb{R}$, so we take $\mathbf{x} = (-2,\, 1)^{\mathsf{T}}$ to be the corresponding eigenvector.

Case 2: $\lambda = 2$. Let $\mathbf{x} = (x,\, y)^{\mathsf{T}}$ be an eigenvector. Then

$$\begin{pmatrix} 1-(2) & 4 \\ 0.5 & -(2) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Using G.C., $\mathbf{x} = 4\mu$ and $\mathbf{y} = \mu$, where $\mu \in \mathbb{R}$, so we take $\mathbf{x} = (4,\, 1)^{\mathsf{T}}$ to be the corresponding eigenvector.

Part (c)(ii). Write $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, where

$$\mathbf{P} = \begin{pmatrix} -2 & 4 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Then

$$\mathbf{v}_t = \mathbf{A}^t \mathbf{v}_0 = \mathbf{P}\mathbf{D}^t\mathbf{P}^{-1}\mathbf{v}_0.$$

Substituting, we get

$$\mathbf{v}_t = \begin{pmatrix} -2 & 4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (-1)^t & 0 \\ 0 & 2^t \end{pmatrix} \left[\frac{1}{6} \begin{pmatrix} -1 & 4 \\ 1 & 2 \end{pmatrix} \right] \begin{pmatrix} a \\ b \end{pmatrix},$$

which simplifies to

$$\mathbf{v}_t = \frac{1}{6} \begin{pmatrix} -2(-1)^t(-a+4b) + 4 \cdot 2^t(a+2b) \\ (-1)^t(-a+4b) + 2^t(a+2b) \end{pmatrix}.$$

Part (c)(iii). The proportion of juveniles to adults in the long term is given by

$$\lim_{t \to \infty} \frac{\frac{1}{6}\left[-2(-1)^t(-a+4b) + 4 \cdot 2^t(a+2b)\right]}{\frac{1}{6}\left[(-1)^t(-a+4b) + 2^t(a+2b)\right]} = \lim_{t \to \infty} \frac{4 \cdot 2^t(a+2b)}{2^t(a+2b)} = 4.$$

Assignment B17C

Problem 1.

(a) The matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{pmatrix} 1 & c & 3 \\ 4 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

It is given that $\mathbf{A}$ has an eigenvalue of 6. Find the value of c and the remaining eigenvalues.

(b) Hence, find matrices $\mathbf{P}$ and $\mathbf{D}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix.

(c) It is given that three functions y_1, y_2, y_3 are the solutions of the following system of differential equations:

$$\begin{aligned} \frac{\mathrm{d}y_1}{\mathrm{d}x} &= y_1 + cy_2 + 3y_3, \\ \frac{\mathrm{d}y_2}{\mathrm{d}x} &= 4y_1 + y_2, \\ \frac{\mathrm{d}y_3}{\mathrm{d}x} &= 3y_1 + y_3, \end{aligned}$$

where c is the value found in (a).

By considering $\mathbf{U} = \mathbf{P}^{-1}\mathbf{Y}$, where

$$\mathbf{U} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \quad \text{and} \quad \mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

show that the above system can be rewritten as $\mathbf{U}' = \mathbf{DU}$, where

$$\mathbf{U}' = \begin{pmatrix} \mathrm{d}u_1/\mathrm{d}x \\ \mathrm{d}u_2/\mathrm{d}x \\ \mathrm{d}u_3/\mathrm{d}x \end{pmatrix}.$$

You may assume that $\mathbf{U}' = \mathbf{P}^{-1}\mathbf{Y}'$, where

$$\mathbf{Y}' = \begin{pmatrix} \mathrm{d}y_1/\mathrm{d}x \\ \mathrm{d}y_2/\mathrm{d}x \\ \mathrm{d}y_3/\mathrm{d}x \end{pmatrix}.$$

(d) Hence, or otherwise, find the general solution of the functions y_1, y_2, y_3 in terms of x.

Solution.

Part (a). Let the characteristic polynomial of $\mathbf{A}$ be $\chi(\lambda) = -\lambda^3 + E_1\lambda^2 - E_2\lambda^1 + E_3$. We have

$$\begin{aligned} E_1 &= |1| + |1| + |1| = 3, \\ E_2 &= \begin{vmatrix} 1 & c \\ 4 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix} = -6 - 4c, \\ E_3 &= \begin{vmatrix} 1 & c & 3 \\ 4 & 1 & 0 \\ 3 & 0 & 1 \end{vmatrix} = -8 - 4c. \end{aligned}$$

Thus,

$$\chi(\lambda) = -\lambda^3 + 3\lambda^2 + (6+4c)\,\lambda - (8+4c)\,.$$

Since 6 is an eigenvalue, it is a root to the characteristic polynomial. Thus,

$$\chi(6) = -6^3 + 3\left(6^2\right) + (6+4c)\,6 - (8+4c) = 0 \implies c = 4.$$

Hence,

$$\chi(\lambda) = -\lambda^3 + 3\lambda^2 + 22\lambda - 24 = -\left(\lambda-6\right)\left(\lambda-1\right)\left(\lambda+4\right).$$

The other eigenvalues are thus 1 and -4.

Part (b). Note that

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} 1-\lambda & 4 & 3 \\ 4 & 1-\lambda & 0 \\ 3 & 0 & 1-\lambda \end{pmatrix}.$$

Case 1: $\lambda = 6$. Note that

$$\begin{pmatrix} 4 \\ 1-6 \\ 0 \end{pmatrix} \times \begin{pmatrix} 3 \\ 0 \\ 1-6 \end{pmatrix} = 5\begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix}.$$

We thus take $(5,\,4,\,3)^{\mathsf{T}}$ to be our eigenvector corresponding to $\lambda = 6$.

Case 2: $\lambda = 1$. Note that

$$\begin{pmatrix} 1-1 \\ 4 \\ 3 \end{pmatrix} \times \begin{pmatrix} 3 \\ 0 \\ 1-1 \end{pmatrix} = 3\begin{pmatrix} 0 \\ 3 \\ -4 \end{pmatrix}.$$

We thus take $(0,\,3,\,-4)^{\mathsf{T}}$ to be our eigenvector corresponding to $\lambda = 1$.

Case 1: $\lambda = -4$. Note that

$$\begin{pmatrix} 4 \\ 1-(-4) \\ 0 \end{pmatrix} \times \begin{pmatrix} 3 \\ 0 \\ 1-(-4) \end{pmatrix} = -5\begin{pmatrix} -5 \\ 4 \\ 3 \end{pmatrix}.$$

We thus take $(-5,\,4,\,3)^{\mathsf{T}}$ to be our eigenvector corresponding to $\lambda = -4$.

Thus,

$$\mathbf{P} = \begin{pmatrix} 5 & 0 & -5 \\ 4 & 3 & 4 \\ 3 & -4 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{pmatrix}.$$

Part (c). Rewriting the given system of equations, we have

$$\mathbf{Y}' = \mathbf{AY} = \mathbf{PDP}^{-1}\mathbf{Y} \implies \mathbf{P}^{-1}\mathbf{Y}' = \mathbf{DP}^{-1}\mathbf{Y}.$$

Since $\mathbf{U} = \mathbf{P}^{-1}\mathbf{Y}$ and $\mathbf{U}' = \mathbf{P}^{-1}\mathbf{Y}'$, we have $\mathbf{U}' = \mathbf{DU}$ as desired.

Part (d). $\mathbf{U}' = \mathbf{DU}$ expands as

$$\mathbf{U}' = \begin{pmatrix} \mathrm{d}u_1/\mathrm{d}x \\ \mathrm{d}u_2/\mathrm{d}x \\ \mathrm{d}u_3/\mathrm{d}x \end{pmatrix} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

This gives us the system of differential equations

$$\frac{\mathrm{d}u_1}{\mathrm{d}x} = 6u_1, \quad \frac{\mathrm{d}u_2}{\mathrm{d}x} = u_2, \quad \frac{\mathrm{d}u_3}{\mathrm{d}x} = -4u_3,$$

which we can easily solve:

$$u_1 = c_1 e^{6x}, \quad u_2 = c_2 e^{x}, \quad u_3 = c_3 e^{-4x}.$$

Since $\mathbf{U} = \mathbf{P}^{-1}\mathbf{Y}$, we have

$$\mathbf{Y} = \mathbf{PU} \implies \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 5 & 0 & -5 \\ 4 & 3 & 4 \\ 3 & -4 & 3 \end{pmatrix} \begin{pmatrix} c_1 e^{6x} \\ c_2 e^{x} \\ c_3 e^{-4x} \end{pmatrix},$$

so

$$\begin{aligned} y_1 &= 5c_1 e^{6x} - 5c_3 e^{-4x}, \\ y_2 &= 4c_1 e^{5x} + 3c_2 e^{x} + 4c_3 e^{-4x}, \\ y_3 &= 3c_1 e^{6x} - 4c_2 e^{x} + 3c_3 e^{-4x}. \end{aligned}$$

* * * * *

Problem 2. The 3×3 non-singular matrix $\mathbf{A}$ has eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ with corresponding eigenvalues α, β and γ respectively. The three eigenvectors are linearly independent, and the eigenvalues are all non-zero real numbers. The eigenvectors of the 3×3 matrix $\mathbf{B}$ are also $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and the corresponding eigenvalues are $\alpha - \beta\gamma$, $\beta - \gamma\alpha$ and $\gamma - \alpha\beta$ respectively.

(a) The characteristic equation of $\mathbf{A}$ is $x^3 - x^2 + kx + 4 = 0$, where k is a real constant. Find an expression for the matrix $\mathbf{B}$ but in terms of the matrix $\mathbf{A}$.

The transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is represented by the matrix $\mathbf{B}$.

(b) Show that $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ forms a basis for the range of T.

Solution.

Part (a). Note that

$$x^3 - x^2 + kx + 4 = (x - \alpha)(x - \beta)(x - \gamma).$$

By Vieta's formula, we see that $\alpha\beta\gamma = -4$. Thus, $\mathbf{B}$ has eigenvalues $\alpha + 4/\alpha$, $\beta + 4/\beta$ and $\gamma + 4/\gamma$, so $\mathbf{B} = \mathbf{A} + 4\mathbf{A}^{-1}$.

Part (b). Let λ be an eigenvalue of $\mathbf{A}$. The corresponding eigenvalue of $\mathbf{B}$ is $\lambda + 4/\lambda$, which is non-zero, since

$$\lambda + \frac{4}{\lambda} = \frac{1}{\lambda}\left(\lambda^2 + 4\right)$$

has no real roots. Hence, $\mathbf{B}$ has full rank, so range$(T) = \mathbb{R}^3$. But because the eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are linearly independent, they span $\mathbb{R}^3$, so they form a basis for range(T).

* * * * *

Problem 3. The matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{pmatrix} p & 5 & 0 \\ -1 & -1 & 1 - p \\ -1 & 2 & 1 \end{pmatrix}.$$

(a) Find the possible number(s) of real eigenvalues of $\mathbf{A}$, and the corresponding range of values of p.

Joe makes the following assertion: "If $p = -1$, then $\mathbf{A}$ is not diagonalizable."

(b) Explain, with working, whether Joe's assertion is true.

Solution.

Part (a). Let $\chi(\lambda) = -\lambda^3 + E_1\lambda^2 - E_2\lambda + E_3$. Note that

$$E_1 = |p| + |-1| + |1| = p,$$
$$E_2 = \begin{vmatrix} p & 5 \\ -1 & -1 \end{vmatrix} + \begin{vmatrix} -1 & 1-p \\ 2 & 1 \end{vmatrix} + \begin{vmatrix} p & 0 \\ -1 & 1 \end{vmatrix} = 2p + 2,$$
$$E_3 = \begin{vmatrix} p & 5 & 0 \\ -1 & -1 & 1-p \\ -1 & 2 & 1 \end{vmatrix} = p\,(2p+2)\,.$$

Thus,

$$\chi(\lambda) = -\lambda^3 + p\lambda^2 - (2p+2)\,\lambda + p\,(2p+2) = (-\lambda + p)\left(\lambda^2 + 2p + 2\right).$$

Case 1. If $2p + 2 < 0$, i.e. $p < -1$, then $\mathbf{A}$ has three real and distinct eigenvalues, namely p, $\sqrt{-(2p-2)}$ and $\sqrt{-(2p-2)}$.

Case 2. If $2p+2 = 0$, i.e. $p = -1$, then $\mathbf{A}$ has two real and distinct eigenvalues, namely $p = -1$ and 0.

Case 3. If $2p + 2 > 0$, i.e. $p > -1$, then $\mathbf{A}$ has one real eigenvalue, namely p.

Part (b). When $p = -1$, $\mathbf{A}$ has two eigenvalues, namely -1 and 0. Note that

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} -1-\lambda & 5 & 0 \\ -1 & -1-\lambda & 2 \\ -1 & 2 & 1-\lambda \end{pmatrix}.$$

Case 1: $\lambda = -1$. Note that

$$\begin{pmatrix} 0 \\ 5 \\ 0 \end{pmatrix} \times \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} = 5 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

The eigenvector corresponding to $\lambda = -1$ is thus $(2,\ 0,\ 1)^{\mathsf{T}}$.

Case 2: $\lambda = 0$. Note that

$$\begin{pmatrix} -1 \\ 5 \\ 0 \end{pmatrix} \times \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}.$$

The eigenvector corresponding to $\lambda = 0$ is thus $(5,\ 1,\ 3)^{\mathsf{T}}$.

Thus, $\mathbf{A}$ has two independent eigenvectors when $p = -1$. Since there are less than 3 independent eigenvectors, $\mathbf{A}$ is not diagonalizable, so Joe is correct.

B18 Correlation and Regression

Tutorial B18

Problem 1. The product moment correlation coefficient is denoted by r. Comment on the validity of following:

(a) $r = 0$ for a set of data (x, y) implies x and y are unrelated.

(b) If x is the number of cigarettes smoked per day by lung cancer patients and y is the age of the lung cancer patients at death, then $r = -0.9$ implies that smoking more cigarettes per day causes lung cancer patients to die at a younger age.

(c) The value of r for a sample (x, y) being 1 means that a linear relation holds for x and y.

Solution.

Part (a). False. $r = 0$ implies that x and y are not linearly correlated; x and y could be related by another model (e.g. quadratic).

Part (b). False. Though $r = -0.9$ implies that x and y have a strong negative linear correlation, it does not imply a causal relationship between x and y.

Part (c). False. If $r = 1$, we can say that a linear relation holds for x and y within the range provided by the sample. However, outside this range, we cannot say that x and y still share a linear relationship.

* * * * *

Problem 2. For a random sample of 12 observations of pairs of values $(x,\, y)$, the equation of the regression line of y on x is $y = 4.82 - 2.25x$. The sum of the 12 values of x is 20.64 and the product moment correlation coefficient for the sample is -0.3.

(a) Find the sum of the 12 values of y.

(b) Find the estimated value of y when $x = 2.8$ and comment on the reliability of this estimate.

Part (a).

Part (b). Note that $\overline{x} = 20.64/12 = 1.72$. Since $(\overline{x}, \overline{y})$ lies on the regression line, we have

$$\overline{y} = 4.82 - 2.25\,(1.72) = 0.95,$$

so the sum of the 12 values of y is $12\,(0.95) = 11.4$.

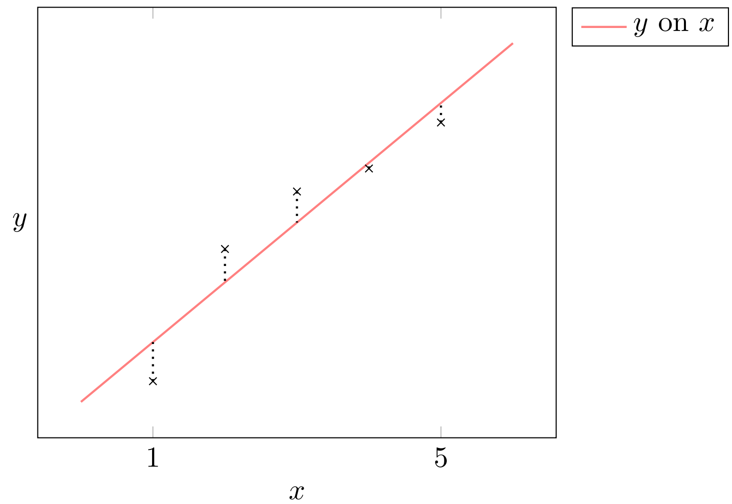
Part (c). When $x = 2.8$, we have

$$y = 4.82 - 2.25\,(2.8) = -1.48.$$

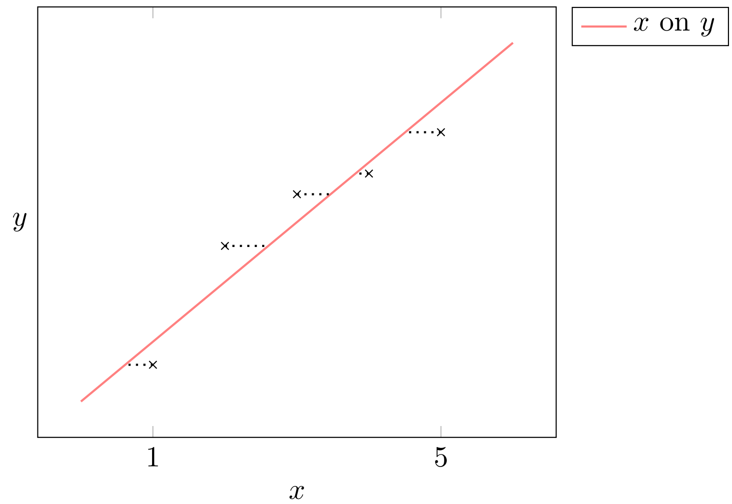
Since $|r| = 0.3$, there is a weak linear relationship between x and y, so the estimate is unreliable.

Problem 3. With the aid of suitable scatter diagrams, describe the differences between the least squares linear regression line of y on x and that of x on y. Show clearly on these diagrams, the distances which are used to draw the least squares linear regression lines from 5 data points. Explain why these distances are squared.

Solution.



The least squares linear regression line of y on x is the line that minimizes the squares of the vertical distances (dotted lines) between the data points to the line.



The least squares linear regression line of x on y is the line that minimizes the squares of the horizontal distances (dotted lines) between the data points to the line.

The regression lines aim to minimize the "distance" between the line and the data points. This is equivalent to minimizing the squared deviations. Hence, the distances (deviations) are squared.

* * * * *

Problem 4. An engineering company makes cranes. The numbers, x, sold in each three-month period for two years, together with the profits, y thousand dollars, on the sale of these cranes are given in the following table.

x	15	17	13	21	16	22	14	18
y	290	350	270	430	340	410	300	360

(a) Give a sketch of the scatter diagram for the data as shown on your calculator.

(b) Find $\overline{x}$ and $\overline{y}$ and mark the point $(\overline{x}, \overline{y})$ on your scatter diagram.

(c) Calculate the equation of the regression line of y on x, and draw this line on your scatter diagram. Interpret the gradient of this line in the context of question.

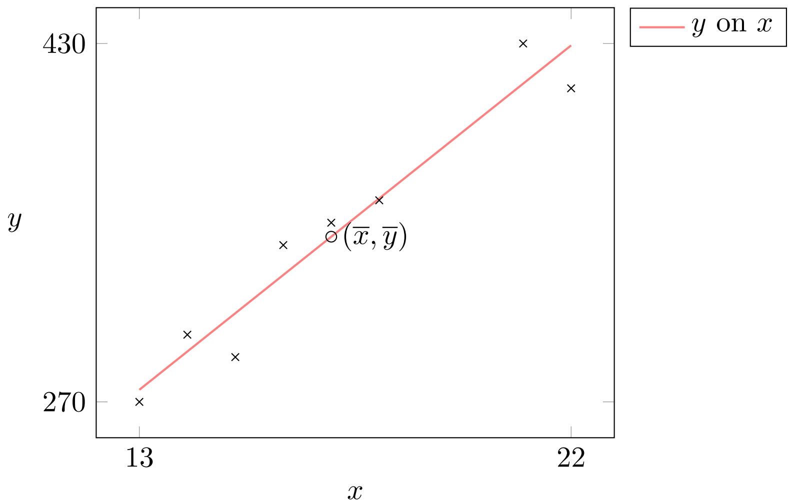
(d) Calculate the product moment correlation coefficient, and comment on its value in relation to your scatter diagram.

(e) For the next three-month period, the sales target is 20 cranes. Estimate the corresponding profit.

(f) The company's sales director uses the regression line in part (c) to predict the profit if 40 cranes were to be sold in a three-month period. Comment on the validity of this prediction.

Solution.

Part (a).



Part (b). Using G.C., $\overline{x} = 17$ and $\overline{y} = 343.75$.

Part (c). Using G.C., the equation of the regression line of y on x is $y = 17.083x + 53.333$. Each additional crane yields a profit of \$17 083.

Part (d). Using G.C., $r = 0.969$, indicating that x and y have a strong positive linear correlation.

Part (e). Using the regression line of y on x at $x = 20$, we have

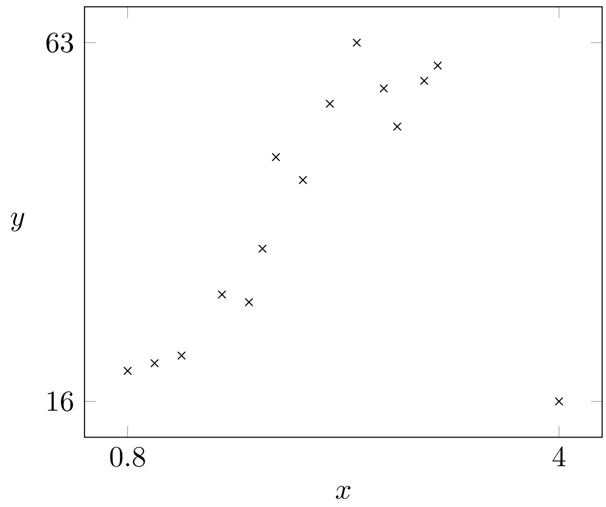
$$y = 17.083\,(20) + 53.333 = 395,$$

so the corresponding profit is \$395 000.

Part (f). $x = 40$ lies outside the given range of values $(13 \leq x \leq 22)$, so the estimate is an extrapolation and is hence unreliable.

* * * * *

Problem 5. A study was carried out to investigate possible links between the weights of hens (x kg) and their eggs (y g). A sample of 15 hens was chosen at random and the weights of these hens and their eggs were noted. The scatter diagram and the summarized information for the sample are shown below. The linear product moment coefficient was also computed and found to be 0.200.



n	$\sum x$	$\sum x^2$	$\sum y$	$\sum y^2$	$\sum xy$
15	33.9	85.99	690	34432	1591.2

By referring to the scatter diagram and the given value of the linear product moment correlation coefficient, comment on the appropriateness of a linear model.

One of the points, $(4, 16)$, was identified as an outlier and removed.

(a) For the remaining sample of size 14, recalculate the values in the table above and determine the value of the linear product moment correlation coefficient. Show your workings clearly.

(b) Use a suitable regression line to estimate the weight of an egg laid by a hen weighing 4 kg, giving your answer to the nearest grams.

(c) Comment on the reliability of your answer.

Solution.

Part (a). We have

$$\begin{aligned}\sum x &= 33.9 - 4 = 29.9,\\ \sum x^2 &= 85.99 - 4^2 = 69.99,\\ \sum y &= 690 - 16 = 674,\\ \sum y^2 &= 34432 - 16^2 = 34176,\\ \sum xy &= 1591.2 - (4)(16) = 1527.2,\end{aligned}$$

so

$$r = \frac{\sum xy - \frac{1}{n}\sum x \sum y}{\sqrt{\left[\sum x^2 - \frac{1}{n}\left(\sum x\right)^2\right]\left[\sum y^2 - \frac{1}{n}\left(\sum y\right)^2\right]}} = 0.852.$$

Part (b). The equation of the regression line y on x is given by

$$y - \overline{y} = b\left(x - \overline{x}\right),$$

where

$$b = \frac{\sum xy - \frac{1}{n}\sum x \sum y}{\sum x^2 - \frac{1}{n}\left(\sum x\right)^2} = 14.3063.$$

Thus,

$$y = 14.306x + 17.589.$$

At $x = 4$, $y = 75$, so the weight of an egg laid by a hen weighing 4 kg is approximately 75 g.

Part (c). $x = 4$ is outside the given range of values (since we removed the point $(4, 16)$). Thus, the estimate is an extrapolation and is hence unreliable.

* * * * *

Problem 6. The table gives the world record time, in seconds above 3 minutes 30 seconds, for running 1 mile as at 1st January in various years.

Year, x	1930	1940	1950	1960	1970	1980	1990	2000
Time, t	40.4	36.4	31.3	24.5	21.1	19.0	16.3	13.1

(a) Draw a scatter diagram to illustrate the data.

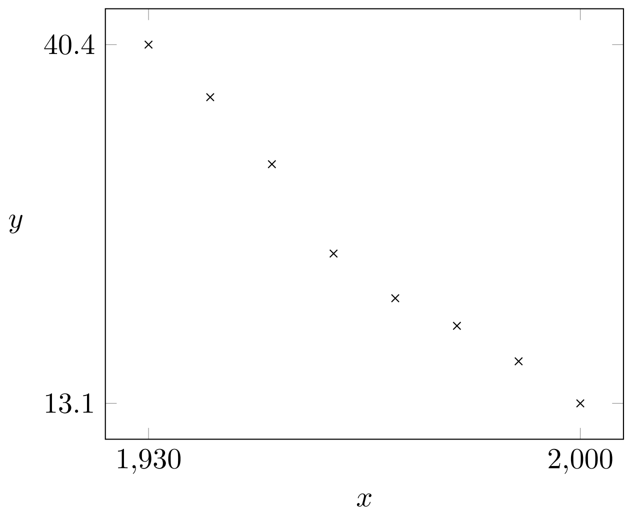
(b) Comment on whether a linear model would be appropriate, referring to both the scatter diagram and the context of the question.

(c) Explain why in this context a quadratic model would probably not be appropriate for long-term predictions.

(d) Fit a model of the form $\ln t = a + bx$ to the data, and use it to predict the world record time as at 1st January 2010. Comment on the reliability of your prediction.

Solution.

Part (a).



Part (b). Based on the scatter plot, a linear model is appropriate. However, it is unlikely for the world record time to keep on decreasing at its current rate; one would expect it to taper off (approach a limiting value). Thus, a linear model is not appropriate in the given context.

Part (c). The rate of decrease in time will likely decrease, not increase, as a quadratic model would predict.

Part (d). The equation of the regression line $\ln t$ on x is

$$\ln t = -0.39512x + 801.67.$$

At $x = 2010$, we have $t = e^{2.43589} = 11.426$. Hence, the model predicts the world recorded time, at 1st January 2010, to be 3 minutes and 41 seconds.

* * * * *

Problem 7.

(a) Sketch a scatter diagram that might be expected for the case when x and y are related approximately by $y = a + bx^2$, where a is positive and b is negative. Your diagram should include 5 points, approximately equally spaced with respect to x, and with all x- and y-values positive.

The table gives the values of seven observations of bivariate data, x and y.

x	2.0	2.5	3.0	3.5	4.0	4.5	5.0
y	18.8	16.9	14.5	11.7	8.6	4.9	0.8

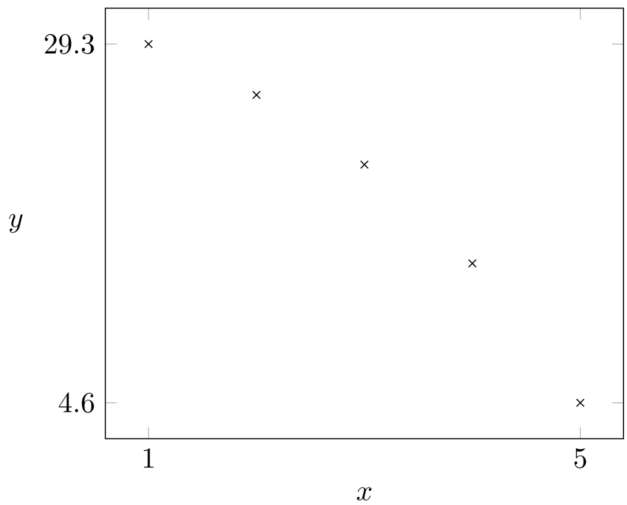
(b) Calculate the value of the product moment correlation coefficient, and explain why its value does not necessarily mean that the best model for the relationship between x and y is $y = c + dx$.

(c) Explain how to use the values obtained by calculating product moment correlation coefficients to decide, for this data, whether $y = a + bx^2$ or $y = c + dx$ is the better model.

(d) It is desired to use the data in the table to estimate the value of y for which $x = 3.2$. Find the equation of the least-squares regression line of y on x^2. Use your equation to calculate the desired estimate.

Solution.

Part (a).



Part (b). Using G.C., $r = -0.992$. The rate of decrease of y is not constant; it seems to be decreasing at an increasing rate. Hence, a linear model may not be best.

Part (c). The product moment correlation coefficient for $y = a + bx^2$ is $r = -0.99998$, which is much closer to -1 than the coefficient for $y = cx + d$, indicating that $y = a + bx^2$ is the better model.

Part (d). The equation for the regression line of y on x^2 is

$$y = -0.85621x^2 + 22.230.$$

At $x = 3.2$, $y = 13.5$.

* * * * *

Problem 8. A certain metal discolours when exposed to air. To protect the metal against discolouring, it is treated with a chemical. In an experiment, different quantities, x ml, of the chemical were applied to standard samples of the metal, and the times, t hours, for the metal to discolour were measured. The results are given in the table.

x	1.2	2.0	2.7	3.8	4.8	5.6	6.9
t	2.2	4.5	5.8	7.3	7.6	9.0	9.9

(a) Calculate the product moment correlation coefficient between x and t, and explain whether your answer suggests that a linear model is appropriate.

(b) Draw a scatter diagram for the data.

One of the values of t appears to be incorrect.

(c) Indicate the corresponding point on your diagram by labelling it P, and explain why the scatter diagram for the remaining points may be consistent with a model of the form $t = a + b \ln x$.

(d) Omitting P, calculate the least squares estimate of a and b for the model $t = a + b \ln x$.

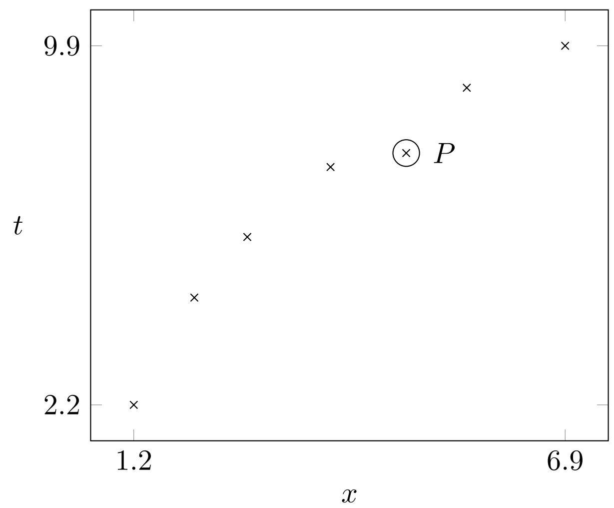
(e) Assume that the value of x at P is correct. Estimate the value of t for this value of x.

(f) Comment on the use of the model in part (d) in predicting the value of t when $x = 8.0$.

Solution.

Part (a). Using G.C., $r = 0.970$. Since $|r|$ is close to 1, there is a strong linear correlation between x and t, so a linear model is appropriate.

Part (b).



Part (c). Removing P, the product moment correlation coefficient of t on $\ln x$ is $r = 0.99998$, indicating a near perfect linear correlation between t and $\ln x$, suggesting that $t = a + b \ln x$ is a suitable model.

Part (d). Using G.C., $a = 1.4247$ and $b = 4.3966$.

Part (e). At $x = 4.8$, we have

$$t = 1.4247 + 4.3966 \ln 4.8 = 8.3.$$

Part (f). $x = 8.0$ is outside the given range of values ($1.2 \leq x \leq 6.9$), hence the estimate is an extrapolation and will be unreliable.

* * * * *

Problem 9. Amy is revising for a mathematics examination and takes a different practice paper each week. Her marks, $y\%$ in week x, are as follows:

Week x	1	2	3	4	5	6
Percentage mark y	38	63	67	75	71	82

(a) Draw a scatter diagram showing these marks.

(b) Suggest a possible reason why one of the marks does not seem to follow the trend.

(c) It is desired to predict Amy's marks on future papers. Explain why, in this context, neither a linear nor a quadratic model is likely to be appropriate.

It is desired to fit a model of the form $\ln(L - y) = a + bx$, where L is a suitable constant. The product moment correlation coefficient between x and $\ln(L - y)$ is denoted by r. The following table gives values of r for some positive values of L.

L	91	92	93
r		-0.929944	-0.929918

(d) Calculate the value of r for $L = 91$, giving your answer correct to 6 decimal places.

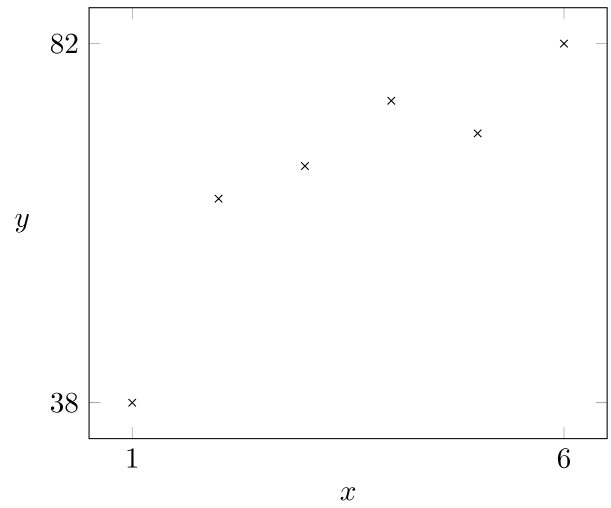
(e) Use the table and your answer to part (d) to suggest with a reason which of 91, 92 or 93 is the most appropriate value of L.

(f) Using the value for L, calculate the values of a and b, and use them to predict the week in which Amy will obtain her first mark of at least 90%.

(g) Give an interpretation, in context, of the value of L.

Solution.

Part (a).



Part (b). The paper might have been much more difficult than usual, so she scored lower than usual.

Part (c). There is a maximum score for the papers. Since linear and quadratic models grow without bound, they are not appropriate.

Part (d). Using G.C., $\ln(91 - y)$ and x have a product moment correlation coefficient of $r = -0.929744$.

Part (e). $L = 92$ is the most suitable, as its value of r is the closest to -1.

Part (f). The regression line of $\ln(92 - y)$ on x is

$$\ln(92 - y) = -0.27960x + 4.1045,$$

so $a = 4.1045$ and $b = -0.27960$. If $y \geq 90$, then

$$\ln(92 - 90) \geq -0.27960 + 4.1045,$$

so the least integral x is 13. Hence, Amy will obtain her first mark of at least 90% in week 13.

Part (g). L is the highest mark obtainable by Amy.

* * * * *

Problem 10. In an experiment, the following information was gathered about air pressure P, measured in inches of mercury, at different heights above sea level h, measured in feet.

h	2000	5000	10000	15000	20000	25000	30000	35000	40000	45000
P	27.8	24.9	20.6	16.9	13.8	11.1	8.89	7.04	5.52	4.28

(a) Draw a scatter diagram for these values, labelling the axes.

(b) Find, correct to 4 decimal places, the product moment correlation coefficient between

  (i) h and P,

  (ii) $\ln h$ and P,

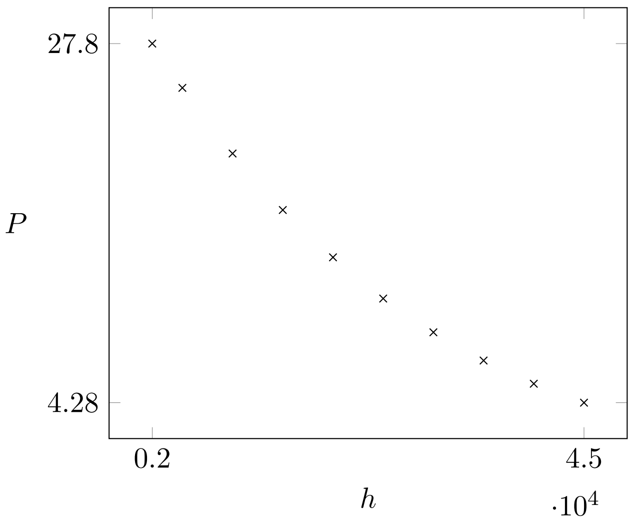
  (iii) $\sqrt{h}$ and P.

(c) Using the most appropriate case from part (b), find the equation which best models air pressure at different heights.

(d) Given that 1 metre = 3.28 feet, re-write your equation from part (c) so that it can be used to estimate the air pressure when the height is measured in metres.

Solution.

Part (a).



Part (b).

Part (b)(i). Using G.C., $r = -0.98073$.

Part (b)(ii). Using G.C., $r = -0.97480$.

Part (b)(iii). Using G.C., $r = -0.99864$.

Part (c). The most appropriate case is $\sqrt{h}$ and P, since its value of r is the closest to -1. Its regression line is given by

$$P = -0.14687\sqrt{h} + 34.789.$$

Part (d). The equation becomes

$$P = -0.14687\sqrt{3.28h} + 34.789 = -0.26599\sqrt{h} + 34.789.$$

* * * * *

Problem 11. A website about electric motors gives information about the percentage efficiency y of motors depending on their power x, measured in horsepower. Xian has copied the following table for a particular type of electric motor, but he has copied one of the efficiency values wrongly.

x	1	1.5	2	3	5	7.5	10	20	30	40	50
y	72.5	82.5	84.0	87.4	87.5	88.5	89.5	90.2	91.0	91.7	92.4

(a) Plot a scatter diagram for these values. On your diagram, circle the point that Xian has copied wrongly.

For parts (b), (c) and (d) of this question you should exclude the point for which Xian has copied the efficiency value wrongly.

(b) Explain from your scatter diagram why the relationship between x and y should not be modelled by an equation of the form $y = ax + b$.

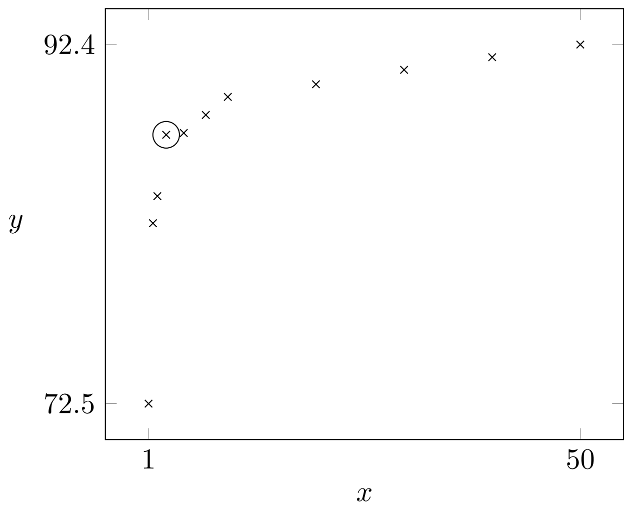
(c) Suppose that the relationship between x and y is modelled by an equation of the form $y = c/x + d$, where c and d are constants. State with reason whether each of c and d is positive or negative.

(d) Find the product moment correlation coefficient and the constants c and d for the model in part (c).

(e) Use the model $y = c/x + d$, with the values of c and d found in part (d), to estimate the efficiency value (y) that Xian copied wrongly. Give two reasons why you would expect this estimate to be reliable.

Solution.

Part (a).



Part (b). y is increasing at a decreasing rate, not at a constant rate as a linear model would suggest.

Part (c). c is negative since the rate of increase is decreasing. d is positive since the y-values are positive.

Part (d). Using G.C., $r = -0.97955$. The regression line of y on $1/x$ is

$$y = \frac{-17.484}{x} + 91.750.$$

Part (e). At $x = 3$, $y = 85.9$. $x = 3$ is within the given range of values ($1 \leq x \leq 50$), so the estimate is an interpolation. Further, $|r|$ is close to 1. Thus, the estimate is reliable.

* * * * *

Problem 12. An athletic coach believes that athletes with longer legs can run faster. He selected 10 of his athletes and recorded their leg lengths, x metres and their times, t seconds, in a 100 m race. The results are given in the table.

x	0.70	0.76	0.80	0.84	0.85	0.89	0.92	0.95	0.98	1.00
t	13.90	12.73	12.12	11.89	11.80	11.42	11.29	10.94	11.00	10.80

It is given that the value of the product moment correlation coefficient for this data is -0.963, correct to 3 decimal places.

(a) State, with a reason, whether the value of the product moment correlation coefficient would be different if the leg lengths had been measured in centimetres instead.

(b) One of the athletes, Aaron, had missed the race. Assuming a linear model, the coach decides to use a regression line to estimate Aaron's 100 m race timing by measuring his leg length. Explain which of the least squares regression lines, x on t or t on x, should be used.

(c) Draw a scatter diagram to illustrate the data.

(d) Aaron disagreed with the coach and claimed that x and t do not have a linear correlation. Comment on Aaron's statement with reference to the scatter diagram.

(e) To be fair to Aaron, the coach considered another possible model for the relationship between x and t: $t = a + b/x^2$, where a and b are constants.

(i) Find the value of the product moment correlation coefficient between t and $1/x^2$, and hence explain why this new model is better than the linear model.

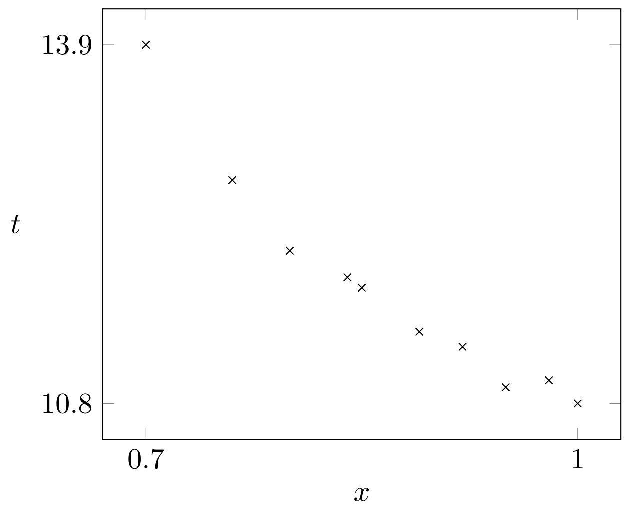
(ii) The coach wants to train an athlete to run the 100 m race in 10 seconds. Calculate the equation of the regression line based on the new model, and use it to estimate the minimum leg length required for the potential athlete. Comment on the reliability of the estimate.

Solution.

Part (a). r is invariant under scaling, hence r will remain the same even if the leg lengths were measured in cm.

Part (b). Since the leg length (x) is given, he should use the regression line of t on x.

Part (c).



Part (d). t seems to approach a limiting value as x increases. Hence, Aaron's statement is correct.

Part (e).

Part (e)(i). Using G.C., $r = 0.9951$. Since $|r|$ is now closer to 1 ($0.9951 > 0.963$), the new model is better than the linear model.

Part (e)(ii). The regression line t on $1/x^2$ is given by

$$t = \frac{2.8616}{x^2} + 7.8603.$$

When $t = 10$, we have $x = 1.16$. Since t is outside the given range of values ($10.80 \leq t \leq 13.90$), the estimate is an extrapolation and thus unreliable.

Assignment B18

Problem 1. The table below gives the observed values of bivariate x and y.

x	20	30	34	35	36	40	42
y	32	25	a	22	26	18	19

It is given that the equation of the regression line y on x is $y = 43.5 - 0.602x$.

(a) Find the value of a correct to the nearest integer.

(b) Using the result in part (a), write down the equation of the regression line x on y and the value of the product moment correlation coefficient between x and y.

Solution.

Part (a). Using G.C., $\overline{x} = 33.857$. Since $(\overline{x}, \overline{y})$ lies on the regression line y on x, we have

$$\frac{32 + 25 + a + 22 + 26 + 18 + 19}{7} = \overline{y} = 43.5 - 0.602\,(33.857)\,,$$

whence $a = 20$ (to the nearest integer).

Part (b). Using G.C., the regression line x on y is

$$x = -1.3176y + 64.349$$

and the product moment correlation coefficient is $r = -0.891$.

* * * * *

Problem 2. A medical statistician wishes to carry out a test to see if there is any correlation between the head circumference and body length of newborn babies. A random sample of ten newborn babies have their head circumference, c cm and body length, l cm measured. This bivariate data is illustrated in the table below.

c	31.0	32.0	33.5	34.0	34.0	51.0	35.0	36.0	36.5	37.5
l	45.0	49.0	49.0	47.0	50.0	34.0	50.0	53.0	51.0	51.0

One particular data has been recorded incorrect with its values of c and l interchanged. Identify the point.

(a) Make the necessary correction and use a suitable regression line to estimate the length of a baby whose head has the circumference of

   (i) 34.5 cm,

   (ii) 45.0 cm.

(b) Give a reason why the estimation found in (a)(ii) may not be a good one.

Solution. The point is $(c, l) = (51.0, 34.0)$.

Part (a). Since c is given, we use the regression line l on c. Using G.C., this is given by

$$l = 0.86981c + 19.722.$$

Part (a)(i). When $c = 34.5$, $l = 49.7$ cm.

Part (a)(ii). When $c = 45.0$, $l = 58.9$ cm.

Part (b). $c = 45.0$ is out of the given range of values ($31.0 \leq c \leq 37.5$). Hence, the estimate found is an extrapolation, so it is unreliable.

* * * * *

Problem 3. A car is placed in a wind tunnel and the drag force F for different wind speeds v, in appropriate units, is recorded. The results are shown in the table.

v	0	4	8	12	16	20	24	28	32	36
F	0	2.5	5.1	8.8	11.2	13.6	17.6	22.0	27.8	33.9

(a) Draw the scatter diagram for these values, labelling the axes correctly.

It is thought that the drag force F can be modelled by one of the formulae

$$F = a + bc \quad \text{or} \quad F = c + dv^2,$$

where a, b, c and d are constants.

(b) Find, correct to 4 decimal places, the value of the product moment correlation coefficient between

(i) v and F,

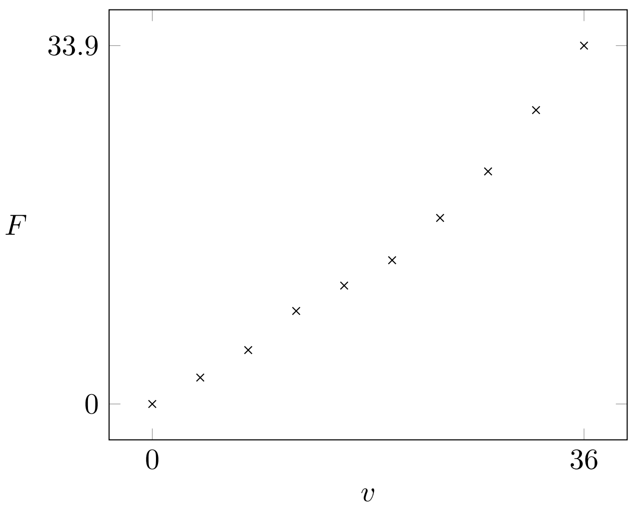
(ii) v^2 and F.

(c) Use your answers to parts (a) and (b) to explain why of $F = a + bv$ or $F = c + dv^2$ is the better model.

(d) It is required to estimate the value of v for which $F = 26.0$. Find the equation of a suitable regression line, and use it to find the required estimate. Explain why neither the regression line of v on F nor the regression line of v^2 on F should be used.

Solution.

Part (a).



Part (b).

Part (b)(i). Using G.C., $r = 0.9860$.

Part (b)(ii). Using G.C., $r = 0.9906$.

Part (c). Since $r = 0.9907$ is closer to 1 than $r = 0.9860$, there is a stronger linear correlation between v^2 and F than v and F. Further, from the scatter diagram of v and F, there is a slight curvature present, so a linear model may not be suitable for v and F.

Part (d). Using G.C., the regression line F on v^2 is given by

$$F = 0.024242v^2 + 3.1957.$$

When $F = 26.0$, $v = 30.7$. Note that we reject $v = -30.7$ since $v \geq 0$.

v is the independent variable while F is the dependent variable, so the regression lines v on F and v^2 on F should not be used.

* * * * *

Problem 4. The number of employees, y, who stay back and continue in the office t minutes after 5 pm on a particular day in a company is recorded. The results are shown in the table.

t	15	30	45	60	75	90	105
y	30	19	15	13	12	11	10

(a) Draw a scatter diagram for these values, labelling the axes clearly.

(b) Find, correct to 4 decimal places, the product moment correlation coefficient between

(i) t and y,

(ii) $\sqrt{t}$ and t,

(iii) $1/t$ and y.

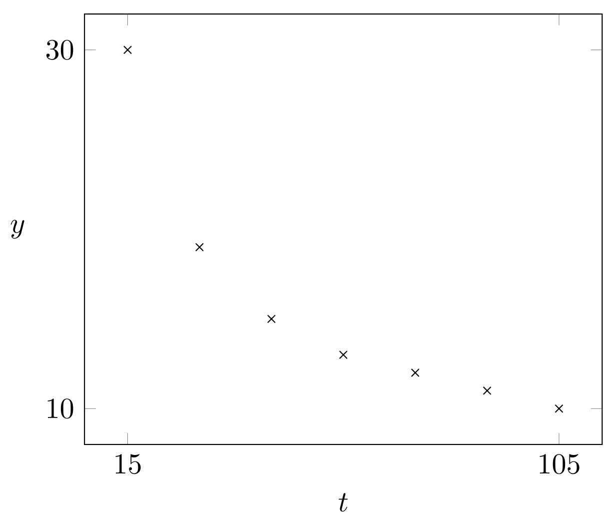
Hence, state with a valid reason, which of the above models is the most appropriate model of the relationship between t and y.

(c) Using the model you chose in part (b), find the equation for the relationship between t and y.

(d) Predict, to the nearest whole number, the number of employees who stay back and continue to work in the office at 7 pm on that particular day. Comment on the reliability of your prediction.

Solution.

Part (a).



Part (b).

Part (b)(i). Using G.C., $r = -0.8745$.

Part (b)(ii). Using G.C., $r = -0.9288$.

Part (b)(iii). Using G.C., $r = 0.9993$.

The model between $1/t$ and y is the most appropriate, since its $|r|$ is the closest to 1 among the three, thus it has the strongest linear correlation among the three models.

Part (c). Since t is independent, we use the regression line $1/t$ on y. Using G.C., this is given by

$$y = \frac{344.60}{t} + 7.2048.$$

Part (d). Note that 7 pm corresponds to $t = 120$, which gives $y = 10$ (to the nearest integer). Thus, the number of employees staying back until 7 pm is 10. However, because $t = 120$ is outside the given range of values ($15 \leq t \leq 105$), the estimate is an extrapolation and hence unreliable.

B19 Non-Parametric Tests

Tutorial B19

Problem 1. In your own words, explain the rejection criterion for the Sign Test using the test statistic K_-, the number of negative signs, for the left-tail, right-tail and two-tail test.

Solution. Let m be the population median and let m_0 be a fixed value.

For the left-tail test, H_0: $m = m_0$ and H_1: $m < m_0$. In this case, H_0 will be rejected if the observed number of negatives, k_-, is too large (i.e. more than some critical value corresponding to the level of significance).

$$\mathbb{P}[K_- \geq k_-] \leq \frac{\alpha}{100}.$$

For the right-tail test, H_0: $m = m_0$ and H_1: $m > m_0$. In this case, H_0 will be rejected if k_- is too small (i.e. less than some critical value corresponding to the level of significance).

$$\mathbb{P}[K_- \leq k_-] \leq \frac{\alpha}{100}.$$

For the two-tail test, H_0: $m = m_0$ and H_1: $m \neq m_0$. In this case, H_0 will be rejected if k_- is too small or too large (i.e. less/greater than some critical value corresponding to the level of significance).

$$2 \min\{\mathbb{P}[K_- \geq k_-]\,, \mathbb{P}[K_- \leq k_-]\} \leq \frac{\alpha}{100}.$$

* * * * *

Problem 2. Show that if the sign test is applied to $n = 5$ pairs, the null hypothesis will never be rejected in favour of a two-sided alternative hypothesis at the 5% level of significance, no matter how extreme the sample results are. Why does this imply there is no point in carrying out this test when $n = 5$ (or less)?

What is the corresponding value of n in the case where the null hypothesis will never be rejected in favour of a one-sided alternative hypothesis at the 5% level of significance.

Solution. Let m be the population median, and let m_0 be a fixed value. Our hypotheses are H_0: $m = m_0$, H_1: $m \neq m_0$. Let K_+ be the number of observed data greater than m_0. Under a sign test,

$$K_+ \sim \mathrm{B}\left(5, \frac{1}{2}\right).$$

In the extreme case where all observed data aligns with our alternative hypothesis, say (without loss of generality) $k_+ = 0$, then the p-value is $2\,\mathbb{P}[K_+ \leq 0] = 2/2^5$, which is greater than the 5% significance level we took. Hence, we will never reject H_0. Since there is only one possible outcome, there is no point in carrying out this test when $n = 5$. A similar situation occurs when $n \leq 5$. The corresponding lowest p-value is $2/2^n$, which is always greater than 5% for all $n \leq 5$.

For a one-sided H_1, the lowest p-value is now $1/2^n$. For H_0 to never be rejected, we require

$$\frac{1}{2^n} \geq 0.05 \implies n \leq 4.$$

Problem 3. In the Wilcoxon matched-pair signed rank test, why should $P+Q = \frac{1}{2}n(n+1)$?

Solution. P corresponds to the sum of the positive ranks while Q corresponds to the sum of the negative ranks. Altogether, $P+Q$ corresponds to the sum of all n ranks, so

$$P + Q = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

* * * * *

Problem 4. The manufacturers of an electric water heater claim that their heaters will heat 500 litres of water from a temperature of 10°C to a temperature of 55°C in, on average, no longer than 12 minutes. In order to test this claim, 14 randomly chosen heaters are bought and the times (x minutes) to heat 500 litres of water from 10°C to 55°C are measured. Correct to 1 decimal place, the results are as follows.

13.2, 12.2, 11.4, 14.5, 11.6, 12.9, 12.4, 10.3, 12.3, 11.8, 11.0, 13.0, 12.1, 12.6.

Stating, in each case, any assumption necessary for validity, test the manufacturer's claim at the 10% significance level using

(a) a t-test,

(b) a sign test.

Solution.

Part (a). Let m be the median time taken, in minutes. Our hypotheses are H_0: $m = 12$, H_1: $m > 12$. We perform a sign test at the 10% significance level. Let K_+ be the number of values larger than 12. Under H_0, $K_+ \sim \mathrm{B}(14, 1/2)$. From the sample, $k_+ = 9$, so the p-value is 0.21, which is greater than the 10% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence at the 10% significance level to reject the manufacturer's claim.

Part (b). Let X be the time taken in minutes. Our hypotheses are H_0: $\mu = 12$, H_1: $\mu > 12$. Assuming that X is normally distributed, we perform a t-test at the 10% significance level. From the sample, $\overline{x} = 12.236$ and $s = 1.0315$. Under H_0,

$$\frac{\overline{X} - 12}{s/\sqrt{14}} \sim \mathrm{t}(13).$$

The p-value is 0.204, which is greater than our significance level of 10%, hence we do not reject H_0 and conclude there is insufficient evidence at the 10% significance level to reject the manufacturer's claim.

* * * * *

Problem 5. In order to compare the effectiveness of two mail delivery services, A and B, two samples of 12 identical deliveries were arranged. The number of hours taken for each delivery was recorded, with the following results, to the nearest half hour.

A	26.0	21.0	35.0	24.5	26.0	31.0	28.5	18.5	25.0	27.5	15.5	29.5
B	26.5	20.0	27.0	27.0	24.5	34.0	33.5	20.5	28.5	32.0	19.5	37.0

(a) It is required to test, at the 5% significance level, whether the data indicate that, on average, service A takes a shorter time for its deliveries than service B. Without assuming that the data are samples taken from normal distributions, perform a suitable test, clearly stating your hypothesis.

(b) Service A claims that its average delivery time is 25 hours. Use a non-parametric test, at the 10% significance level, to test this claim against the alternative hypothesis that the average delivery time exceeds 25 hours.

Solution.

Part (a). We perform a Wilcoxon matched-pair signed rank test. Let m be the median of $B-A$, in hours. Our hypotheses are H_0: $m=0$ and H_1: $m>0$. We take a 5% significance level.

From the sample, the ranks are

$B-A$	0.5	−1.0	−8.0	2.5	−1.5	3.0	5.0	2.0	3.5	4.5	4.0	7.5
Rank	1	2	12	5	3	6	10	4	7	9	8	11

Let P and Q be the sum of the ranks corresponding to the positive and negative differences respectively. Let T be the smaller of the two. From the above table, we see that $p=61$ and $q=17$, so $t=17$. From the formula list, with $n=12$, we reject H_0 if $t\leq 17$. Since $t=17\leq 17$, we reject H_0 and conclude there is sufficient evidence at a 5% significance level that service A takes a shorter time for its deliveries than service B.

Part (b). We perform a sign test. Let m' be the median time taken by service A's deliveries, in hours. Our hypotheses are H_0: $m'=25$, H_1: $m'>25$. We take a 10% significance level. Let K_+ be the number of values larger than 25. From the sample, the signs are

$$+,\ -,\ +,\ -,\ +,\ +,\ +,\ -,\ 0,\ +,\ -,\ +,$$

so $k_+=7$. We discard the 0 and reduce our sample size to $n=11$. Under H_0, $K_+\sim\mathrm{B}(11,1/2)$. Our p-value is thus $\mathbb{P}[K_+\geq 7]=0.274$, which is greater than our 10% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence at a 10% significance level that the median time taken by service A is greater than 25 hours.

* * * * *

Problem 6. The weights of fish in two populations were compared by analysing the differences in the weights of a sample of pairs of fish (one from each population) matched by length. The weight differences, in g, for 10 pairs of fish were as follows:

$$11,\ -13,\ -125,\ -210,\ -73,\ 2,\ 3,\ -147,\ -12,\ -4.$$

Give a reason why a parametric test is unsuitable in the context of this question.

Perform two tests, each at 5% significance level, to test for a difference in average weights of fish in the two populations, where one test

(a) ignores the magnitudes of the differences; while the other

(b) uses both the signs and magnitudes of the difference.

Draw an overall conclusion from (a) and (b).

Solution. The differences in the weights of the fish vary wildly, and hence do not seem to follow a familiar distribution. Thus, we cannot assume any underlying distribution, so a parametric test is unsuitable in the context of this question.

Let m be the median difference in weights, in grams. Our hypotheses are H_0: $m=0$, H_1: $m\neq 0$. We take a 5% significance level.

Part (a). We perform a sign test. Let K_+ be the number of positive differences. From the sample, we see that $k_+ = 3$. Under H_0, $K_+ \sim \mathrm{B}(10, 1/2)$, so the p-value is $2\,\mathbb{P}[K_+ \leq 3] = 0.344$, which is greater than our significance level of 5%. Thus, we do not reject H_0 and conclude there is insufficient evidence at the 5% significance level that there is a difference in the average weights of fish in the two populations.

Part (b). We perform a Wilcoxon matched-pair signed rank test. From the given data, the ranks are

Differences	11	−13	−125	−210	−73	2	3	−147	−12	−4
Rank	4	6	8	10	7	1	2	9	5	3

Let P and Q be the sum of ranks corresponding to the positive and negative differences respectively. Let T be the smaller of the two. From the above table, we see that $p = 7$ and $q = 48$, so $t = 7$. From the formula list, we reject H_0 if $t \leq 8$. Since $t = 7 \leq 8$, we reject H_0 and conclude there is sufficient evidence at the 5% significance level that there is a difference in the average weights of fish in the two populations.

Overall, we use the result of the Wilcoxon matched-pair signed rank test and reject H_0. This is because the positive values are relatively small (11, 2, 3) while the negatives are relatively large (−125, −210, −147), thus considering the magnitude is very important.

* * * * *

Problem 7. Briefly describe circumstances in which each of the following are used:

(a) parametric tests of significance,

(b) non-parametric tests of significance.

It is believed that the material from which running tracks are made has a significant effect on the times taken for athletes to run specified distances. In order to test this, 12 athletes ran on two tracks over a distance of 200 m. One track was made from synthetic material and the other from cinders. The times, in seconds, are given in the table.

Athlete	A	B	C	D	E	F	G	H	I	J	K	L
Synthetic	26.5	26.3	24.9	25.7	26.5	24.8	26.1	27.0	24.4	24.7	24.6	24.5
Cinder	27.3	26.4	26.6	25.1	26.0	27.0	26.8	27.0	25.4	26.7	25.0	24.7

(c) Use a Wilcoxon matched-pair signed rank test to show that, at the 1% significance level, there is insufficient evidence that the median time on the synthetic track is lower than that on the cinder track. State the lowest significance level at which it can be concluded that the median time on the synthetic track is lower.

(d) By using a sign test, show that, at the 10% significance level, the median times on each of the tracks could be 26 seconds.

Solution.

Part (a). Parametric tests are used when we can make assumptions about the underlying distribution of the parameter we wish to test, e.g. when the data is normally distributed.

Part (b). Non-parametric tests are used when we cannot make assumptions about the underlying distribution of the parameter we wish to test, e.g. when the data is normally distributed and there is a small sample size.

Part (c). Consider the difference "cinder" − "synthetic". Let m be the median time of these differences, measured in seconds. We perform a Wilcoxon matched-pair signed rank test. Our hypotheses are H_0: $m = 0$ and H_1: $m > 0$. We take a 1% significance level.

From the sample, the ranks are given by

$C-S$	0.8	0.1	1.7	-0.6	-0.5	2.2	0.7	0	1	2	0.4	0.2
Rank	7	1	9	5	4	11	6	–	8	10	3	2

We discard the 0 and reduce our sample size to $n = 11$. Let P and Q be the sum of the ranks corresponding to the positive and negative differences. Let T be the smaller of the two. From the above table, $p = 57$ and $q = 9$, so $t = 9$. From the formula list, we reject H_0 if $t \leq 7$. Since $t = 9 > 7$, we do not reject H_0 and conclude there is insufficient evidence at the 1% significance level that the median time on the synthetic track is lower than that on the cinder track.

The lowest significance level at which it can be concluded that the median time on the synthetic track is lower is 2.5%.

Part (d). Let m' be the median time taken on a track. We perform two sign tests. Our hypotheses are H_0: $m' = 26$ and H_1: $m' \neq 26$. We take a 10% level of significance. Let K_+ be the number of values larger than 26.

Case 1: *Synthetic Track.* From the data, the signs are

$$+, +, -, -, +, -, +, +, -, -, -, -,$$

so $k_+ = 5$. Under H_0, $K_+ \sim \mathrm{B}(12, 1/2)$, so the p-value is $2\,\mathbb{P}[K_+ \leq 5] = 0.774$, which is greater than our 10% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence at the 10% significance level that the median time on the synthetic track differs from 26 seconds.

Case 2: *Cinder Track.* From the data, the signs are

$$+, +, +, -, +, +, +, +, -, +, -, -,$$

so $k_+ = 8$. Under H_0, $K_+ \sim \mathrm{B}(12, 1/2)$, so the p-value is $2\,\mathbb{P}[K_+ \geq 8] = 0.388$, which is greater than our 10% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence at the 10% significance level that the median time on the cinder track differs from 26 seconds.

* * * * *

Problem 8. For the case of paired samples, explain briefly

(a) the circumstances under which the t-test would be appropriate; and

(b) the relative advantages and disadvantages of the sign test and of the Wilcoxon matched-pair signed rank test.

A teacher in charge of bowling wants to find out if switching to a Class III coach affected the bowling team's A division results. To do that, he randomly selected 8 students and compared their A division results in 2016 (taught by the Class II coach) and in 2017 (taught by the Class III coach). The total pin falls of their results in 2016 and 2017 are recorded in the table below.

Student	1	2	3	4	5	6	7	8
Total pin fall (2016)	1758	1961	1787	1626	1600	1859	1764	1680
Total pin fall (2017)	1757	1964	2023	1984	1610	1857	1990	1744

(c) State explicitly suitable null and alternative hypothesis.

(d) Using the sign test, carry out a test of the null hypothesis at the 5% significance level, and state your conclusions.

(e) Using the Wilcoxon matched-pair signed rank test, carry out a test of the null hypothesis at the 5% significance level, and state your conclusions.

(f) Comment on the conclusions of the 2 tests.

Solution.

Part (a). A t-test is appropriate if the differences can be assumed to be normally distributed.

Part (b). The sign test is easier to compute, while the Wilcoxon matched-pair signed rank test is more powerful, as it considers both sign and magnitude.

Part (c). Consider the difference "pinfall in 2017" − "pinfall in 2016". Let m be the median of these differences. Our hypotheses are H_0: $m = 0$ and H_1: $m \neq 0$.

Part (d). We perform a sign test at the 5% significance level. Let K_+ be the number of positive differences. From the data, the signs are

$$-, +, +, +, +, -, +, +,$$

so $k_+ = 6$. Under H_0, $K_+ \sim \mathrm{B}(8, 1/2)$, so the p-value is $2\,\mathbb{P}[K_+ \geq 6] = 0.289$, which is greater than our 5% significance level. This, we do not reject H_0 and conclude there is insufficient evidence at the 5% significance level that the average pinfall differs between 2016 and 2017.

Part (e). We perform a Wilcoxon matched-pair signed rank test at the 5% significance level. From the data, the ranks are

Difference	−1	3	236	358	10	−2	226	64
Rank	1	3	7	8	4	2	6	5

Let P and Q be the sum of ranks corresponding to the positive and negative differences respectively. Let T be the smaller of the two. From the above table, $p = 33$ and $q = 3$, so $t = 3$. From the formula list, we reject H_0 if $t \leq 3$. Sicne $t = 3 \leq 3$, we reject H_0 and conclude there is sufficient evidence at the 5% significance level that the average pinfall differs between 2016 and 2017.

Part (f). The Wilcoxon matched-pair signed rank test is more powerful than the sign test as it takes both sign and magnitude into account. Thus, we use the result of the Wilcoxon matched-pair signed rank test, so we reject H_0.

* * * * *

Problem 9. A device for reducing air conditioning costs has been produced, and in order to test its effectiveness, 11 households were selected at random and the device was fitted. The annual costs for the year before fitting the device and for the year after fitting the device are shown in teh table. It may be assumed that the price of electricity had not risen over the two-year period and that the weather patterns in the two years were similar.

Cost before ($)	756	650	855	533	796	1128	591	656	976	844	681
Cost after ($)	711	608	833	551	776	1096	608	648	942	859	644

Stating your null and alternative hypotheses, perform two non-parametric tests, each at the 5% significance level, to determine whether the average cost had decreased after fitting the device.

State how the conclusion in the Wilcoxon matched-pairs signed rank test would be affected if the price of electricity had risen in the year after fitting the device.

The manufacturer claims that the average reduction in the annual bill would be at least $35. Test the manufacturer's claim using a 5% significance level.

Solution. Consider the difference "cost after" − "cost before". Let m be the median difference in cost. Our hypotheses are H_0: $m = 0$, H_1: $m < 0$. We take a 5% significance level.

Case 1: *Sign Test.* Let K_- be the number of negative differences. From the data, the signs are

$$-, -, -, +, -, -, +, -, -, +, -,$$

so $k_- = 8$. Under H_0, $K_- \sim \mathrm{B}(11, 1/2)$, so the p-value is $\mathbb{P}[K_- \geq 8] = 0.113$, which is greater than our 5% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence at the 5% significance that the average cost decreased after fitting the device.

Case 2: *Wilcoxon Matched-Pair Signed Rank Test.* From the data, the ranks are

Difference	−45	−42	−22	18	−20	−32	17	−8	−34	13	−37
Rank	11	10	6	4	5	7	3	1	8	2	9

Let P and Q be the sum of ranks corresponding to the positive and negative difference respectively. Let T be the smaller of the two. From the above table, $p = 9$ and $q = 57$, so $t = 9$. From the formula list, we reject H_0if $t \leq 13$. Since $t = 9 \leq 13$, we reject H_0 and conclude there is sufficient evidence at the 5% significance that the average cost decreased after fitting the device.

If the price of electricity had risen in the year after fitting the device, the difference between the cost before and the cost after will generally decrease throughout. This results in more positive ranks, therefore the value of t associated with the Wilcoxon matched-pair signed rank test will likely increase, so the conclusion may change.

Let m be the median cost reduction. Our hypotheses are H_0: $m = 35$, H_1: $m < 35$. We perform a sign test at a 5% significance level. Let K_- be the number of cost reductions less than \$35. From the data, the signs are

$$+, +, -, -, -, -, -, -, -, -, +,$$

so $k_- = 8$. Under H_0, $K_- \sim \mathrm{B}(11, 1/2)$, so the p-value is $\mathbb{P}[K_- \geq 8] = 0.113$, which is greater than our 5% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence at the 5% significance level that the average cost reduction is less than \$35.

Assignment B19

Problem 1. Explain why it is better to use a Wilcoxon matched-pair signed rank test, rather than a sign test, to test for a difference between two populations.

The task completion times, in minutes, for a random sample of 12 operatives using two different methods are given in the table below.

Operative	1	2	3	4	5	6	7	8	9	10	11	12
Method A	9.1	8.6	8.2	9.0	8.7	9.1	9.5	8.9	10.0	9.6	9.5	8.3
Method B	8.4	8.8	7.6	9.4	9.2	8.2	9.4	7.9	8.7	8.4	8.7	8.0

(a) Use both tests mentioned above to test, at the 5% significance level, whether Method B results in a smaller median completion time than Method A. Comment on the results.

(b) Test, at the 5% significance level, whether the median time for Method A is 8.3 minutes.

Solution. A Wilcoxon matched-pair signed rank test accounts for the sign and magnitude of the differences between samples of two populations, while a sign test only accounts for the sign. Hence, a Wilcoxon matched-pair signed rank test is more powerful than a sign test.

Part (a). Consider the differences "Method A" − "Method B". Let m be the median of these differences. Our hypotheses are H_0: $m = 0$, H_1: $m > 0$. We take a 5% significance level.

A − B	0.7	−0.2	0.6	−0.4	−0.5	0.9	0.1	1	1.3	1.2	0.8	0.3
Rank	7	2	6	4	5	9	1	10	12	11	8	3

Case 1: *Sign Test.* Let K_+ be the number of positive differences. From the above table, $k_+ = 9$. Under H_0, $K_+ \sim \mathrm{B}(12, 1/2)$. Hence, the p-value is $\mathbb{P}[K_+ \geq 9] = 0.0730$, which is greater than our 5% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence to claim at a 5% significance level that Method B results in a smaller median completion time than Method A.

Case 2: *Wilcoxon Matched-Pair Signed Rank Test.* Let P and Q be the sum of ranks corresponding to the positive and negative differences respectively. Let T be the smaller of the two. From the above table, $p = 67$ and $q = 11$, so $t = 11$. From the formula list, we reject H_0 if $t \leq 17$. Since $t = 11 \leq 17$, we reject H_0 and conclude there is sufficient evidence to claim at a 5% significance level that Method B results in a smaller median completion time than Method A.

Since the Wilcoxon test is more powerful than the sign test, we should reject H_0.

Part (b). Let m' be the median of completion times for Method A. We perform a sign test at a 5% significance level. Our hypothesis are H_0: $m' = 8.3$, H_1: $m' \neq 8.3$. Let K_+ be the number of completion times for Method A that takes longer than 8.3 minutes.

From the data, the signs are

$$+, +, -, +, +, +, +, +, +, +, +, 0,$$

so $k_+ = 10$. We discard the 0 and reduce our sample size to $n = 11$. Under H_0, $K_+ \sim \mathrm{B}(11, 1/2)$, so the p-value is $2\,\mathbb{P}[K_+ \geq 10] = 0.0117$, which is less than our 5% significance level. Thus, we reject H_0 and conclude there is sufficient evidence to claim at a 5% significance level that the median time for Method A differs from 8.3 minutes.

Problem 2. A group of 14 students from the 19/20 DHS FM course participated in an experiment on writing speeds using dominant and non-dominant hands. Each student was to write with each of their 2 hands in 30 seconds the alphabets A to Z, and repeating if time permits. Out of the 14, 7 were randomly assigned to use the dominant hand, followed by the non-dominant hand whereas the rest were to use the non-dominant followed by the dominant.

The number of alphabets each student wrote were recorded in the table below.

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Dominant	57	52	60	46	26	63	54	65	67	48	78	70	45	40
Non-dominant	35	21	20	16	23	31	17	28	36	41	33	22	13	25

Taking the 14 students as a random sample of the DHS 19/20 cohort, test at the 1% significance level, the hypothesis that DHS students are able to write, on average, at least 3 times as fast using their dominant hand as compared to their non-dominant hand.

Solution. Consider the difference "Dominant" $-$ 3×"Non-dominant", and let m be its median. Our hypotheses are H_0: $m = 0$ and H_1: $m < 0$. We perform a sign test at a 1% significance level. Let K_+ be the number of positive differences.

From the data, the signs of the differences are

$$-, -, 0, -, -, -, +, -, -, -, -, +, +, -,$$

so $k_+ = 3$. We also discard the 0 and reduce our sample size to $n = 13$. Under H_0, $K_+ \sim \mathrm{B}(13, 1/2)$, so the p-value is $\mathbb{P}[K_+ \leq 3] = 0.0461$, which is greater than our 1% significance level. Thus, we do not reject H_0 and conclude there is insufficient evidence to claim at the 1% significance level that the students are able to write, on vaerage, at least 3 times as fast using their dominant hand as compared to their non-dominant hand.

Part XI

H3 Mathematics

Mathematical Proofs and Reasoning

An Introduction to the Mathematical Vernacular

Problem 1. Prove that the sum of even and even is even.

Proof. Let a and b be even. By definition, there exists $a', b' \in \mathbb{Z}$ such that $a = 2a'$ and $b = 2b'$. Thus,

$$a + b = 2a' + 2b' = 2\left(a' + b'\right) = 2c,$$

where $c = a' + b'$. Since c is an integer, by definition, $a + b$ is even. Hence, the sum of even and even is even. $\square$

* * * * *

Problem 2. Prove that the sum of even and odd is odd.

Proof. Let a be even and let b be odd. By definition, there exists $a', b' \in \mathbb{Z}$ such that $a = 2a'$ and $b = 2b' + 1$. Thus,

$$a + b = 2a' + \left(2b' + 1\right) = 2\left(a' + b'\right) + 1 = 2c + 1,$$

where $c = a' + b'$. Since c is an integer, by definition, $a + b$ is odd. Hence, the sum of even and odd is odd. $\square$

* * * * *

Problem 3. Prove that the sum of odd and odd is even.

Proof. Let a and b be odd. By definition, there exists $a', b' \in \mathbb{Z}$ such that $a = 2a' + 1$ and $b = 2b' + 1$. Thus,

$$a + b = 2\left(a' + 1\right) + 2\left(b' + 1\right) = 2\left(a' + b' + 2\right) = 2c,$$

where $c = a' + b' + 2$. Since c is an integer, by definition, $a + b$ is even. Hence, the sum of odd and odd is even. $\square$

An Introduction to Proofs

Problem 1. Let m and N be positive integers. Prove that $\sqrt[m]{N}$ is either an integer or an irrational.

Proof. Let $x = \sqrt[m]{N}$. Let A be the nearest integer to x.

Consider $(x - A)^n$. By the binomial theorem,

$$(x - A)^n = \sum_{k=0}^{n} \binom{n}{k} x^k (-A)^{n-k}.$$

Since $x^m = N \in \mathbb{Z}$, the above n-degree polynomial reduces to an $m-1$ degree polynomial with integer coefficients, i.e.

$$(x - A)^n = \sum_{k=0}^{m-1} c_k x^k, \tag{1}$$

where $\{c_k\}$ are integers.

Now, suppose $x \in \mathbb{Q}$. Then we can write $x = p/q$, where $p, q \in \mathbb{Z}$ and $q \neq 0$. Substituting this into (1), we get

$$(x - A)^n = \sum_{k=0}^{m-1} c_k \left(\frac{p}{q}\right)^k = \sum_{k=0}^{m-1} \frac{c_k p^k}{q^k}.$$

By combining all terms into a single fraction, we can write

$$(x - A)^n = \frac{l}{p^{m-1}},$$

where l is an integer. Thus, the only possible values that $(x - A)^n$ can take on are

$$\ldots, \frac{-2}{p^{m-1}}, \frac{-1}{p^{m-1}}, 0, \frac{1}{p^{m-1}}, \frac{2}{p^{m-1}}, \ldots.$$

Observe that $1/p^{m-1}$ is constant with respect to n, i.e. p and m do not depend on n. Since $|x - A| < 1$, for arbitrarily large n, we can make $(x - A)^n$ as close to 0 as we wish. In other words, we can always find an n large enough such that

$$|(x - A)^n| < \frac{1}{p^{m-1}}.$$

Thus, $(x - A)^n$ must be 0, whence $x = A \in \mathbb{Z}$. Hence, if x is rational, it must necessarily be an integer. This completes the proof. □

* * * * *

Problem 2. Prove that π is irrational.

Proof. Seeking a contradiction, suppose $\pi = p/q$, where $p, q \in \mathbb{Z}$ with $q \neq 0$.

Consider the function $\sin x$. It is well known that $\sin x$

- is non-negative for all $x \in [0, \pi]$, with equality only when $x = 0$ or $x = \pi$; and
- attains a maximum at $\pi/2$.

Now consider the $2n$th degree polynomial $f(x) = x^n(p - qx)^n$. Clearly, $f(x)$ is non-negative on $[0, \pi]$ and has roots only at $x = 0$ and $x = \pi$. Additionally, $f(x)$ attains a maximum of $(\pi/2)^{2n}$ at $x = \pi/2$. Thus, $f(x)$ also satisfies the above two properties.

Consider now the integral

$$I = \int_0^\pi f(x) \sin x \,\mathrm{d}x.$$

Since both $f(x)$ and $\sin x$ are non-negative on $[0, \pi]$, it follows that I must also be non-negative on $[0, \pi]$. Additionally, since $f(x) \sin x \not\equiv 0$ on $[0, \pi]$, we have the strict lower bound

$$0 < I.$$

We can also bound I from above:

$$I = \int_0^\pi f(x) \sin x \,\mathrm{d}x \le \int_0^\pi \left(\frac{\pi}{2}\right)^{2n} \mathrm{d}x = \frac{\pi^{2n+1}}{2^{2n}} \le p^{2n+1}.$$

Putting both inequalities together,

$$0 < I \le p^{2n+1}. \tag{1}$$

We now evaluate I. Repeatedly integrating by parts, we get

$$I = \sum_{k=0}^{2n+1} \left[f^{(k)}(x) \sin^{(-k-1)}(x)\right]_0^\pi = \sum_{k=0}^{2n+1} \left[f^{(k)}(\pi) \sin^{(-k-1)}(\pi) - f^{(k)}(0) \sin^{(-k-1)}(0)\right].$$

Note that the sum ends at $k = 2n + 1$ since $f^{(k)} = 0$ for $k \ge 2n + 2$. Also observe that

$$\sin^{(-k-1)}(x) = \begin{cases} -\cos x, & k \equiv 0 \pmod 4 \\ -\sin x, & k \equiv 1 \pmod 4 \\ \cos x, & k \equiv 2 \pmod 4 \\ \sin x, & k \equiv 3 \pmod 4 \end{cases}.$$

The odd k terms hence vanish. We are thus left with

$$I = \sum_{k=0}^{n} (-1)^{k+1} \left[f^{(2k)}(\pi) + f^{(2k)}(0)\right].$$

We now consider $f^{(2k)}(x)$. Firstly, notice that $f(x) = f(\pi - x)$. Hence, by differentiating this repeatedly, we get $f^{(2k)}(0) = f^{(2k)}(\pi)$, so

$$I = 2 \sum_{k=0}^{n} (-1)^{k+1} f^{(2k)}(0).$$

Now, observe that when expanded, $f(x)$ is of the form

$$f(x) = \sum_{i=n}^{2n} a_i x^i,$$

where $\{a_i\}$ are integers. Repeatedly differentiating this yields

$$f^{(k)}(x) = x^n (px - q)^n = \sum_{i=n}^{2n} a_i (i)(i-1) \dots (i-k+1) x^{i-k}.$$

Thus,

$$f^{(k)}(0) = \begin{cases} 0, & 0 \le k < n \\ a_k k!, & n \le k \le 2n \end{cases}.$$

Thus, $f^{(k)}(0)$ is divisible by $k!$ and by extension $n!$ too (since $n \le k$). Hence, I is divisible by $n!$, i.e. $I = Cn!$ for some integer C. From Inequality (1), we have

$$0 < Cn! < p^{2n+1}.$$

However, $n!$ grows much faster than p^{2n+1}. Thus, for sufficiently large n, the inequality does not hold, a contradiction. Hence, π must be irrational. □

* * * * *

Problem 3. Prove that e is irrational.

Proof. By definition,

$$\mathrm{e} = \sum_{n=0}^{\infty} \frac{1}{n!}.$$

Seeking a contradiction, suppose e is rational, i.e. $\mathrm{e} = a/b$, where $a, b \in \mathbb{Z}$ and $b \neq 0$. Define x as

$$x = b!\left(\mathrm{e} - \sum_{n=0}^{b} \frac{1}{n!}\right). \tag{1}$$

Replacing e with a/b, we get

$$x = b!\left(\frac{a}{b} - \sum_{n=0}^{b} \frac{1}{n!}\right) = a(b-1)! - \sum_{n=0}^{b} \frac{b!}{n!}.$$

Since $b!/n!$ is an integer for $0 \le n \le b$, it follows that x is also an integer.

Using the definition of e, we can rewrite (1) as

$$x = b!\left(\sum_{n=0}^{\infty} \frac{1}{n!} - \sum_{n=0}^{b} \frac{1}{n!}\right) = \sum_{n=b+1}^{\infty} \frac{b!}{n!}.$$

It follows that $x > 0$. Now, observe that

$$\begin{aligned} x &= \sum_{n=b+1}^{\infty} \frac{b!}{n!} \\ &= \frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \frac{1}{(b+1)(b+2)(b+3)} + \dots \\ &< \frac{1}{b+1} + \frac{1}{(b+1)^2} + \frac{1}{(b+1)^3} + \dots \\ &= \frac{1}{b+1}\left(\frac{1}{1-\frac{1}{b+1}}\right) = \frac{1}{b} \le 1. \end{aligned}$$

Hence, $0 < x < 1$ but $x \in \mathbb{Z}$, a contradiction. Thus, e must be irrational. □

Problem 4. Let a and b be positive integers such that b is not a perfect power of a. Prove that $\log_a b$ is irrational.

Proof. Seeking a contradiction, suppose $\log_a b$ is rational. Then $\log_a b = m/n$, where $m, n \in \mathbb{Z}$ and $n \neq 0$. Note that $m, n > 0$ since $a, b > 1$. Clearly, we have

$$b = a^{m/n} \implies b^n = a^m.$$

Since b is not a perfect power of a, this implies that the integer $k = b^n = a^m$ has two distinct prime factorizations, which is a clear contradiction of the Fundamental Theorem of Algebra. Hence, $\log_a b$ must be irrational. $\square$

Problem-Solving

Problem 1. A warden wishes to give a group of 10 prisoners a change to be released early. He tells the group that the following day, he will line all 10 of them on a flight of stairs such that each prisoner faces downwards and can only see the heads of the prisoners in front of him. The warden will then put a hat, either black or white, on each of the prisoners' heads. The prisoners, starting from the highest stair, are then to 'guess' the colour of the hat on their own head, calling out only "'Black"' or "White"; and they can hear all preceding guesses. If at least 9 of them match what they say with the colour of that hat on their head, the whole group is released early. The warden gives the group time to discuss a strategy beforehand.

(a) Is there a strategy to guarantee that at least nine of them will get the colour of the hat correct?

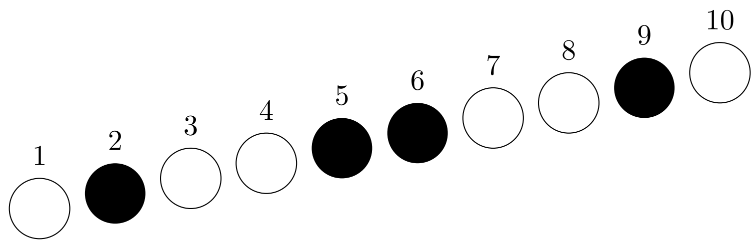
(b) What if the warden used black, white, and red hats?

(c) What if there are now m prisoners and n different hats, where $m \geq n$.

Solution.

Part (a). Label the prisoners $n = 1, 2, \dots, 10$, starting from the prisoner on the lowest stair. Let u_n be the colour of Prisoner n's hat, where 0 represents a white hat, and 1 represents a black hat. Let $S_n = u_1 + u_2 + \cdots + u_{n-1}$ be the sum of 'hats' visible by the nth prisoner. For instance, if Prisoner 10 sees 7 black hats and 2 white hats, then $S_{10} = 7$.

The prisoners can guarantee an early release. The strategy is to let Prisoner 10 say "Black" if S_{10} is even, and "White" if S_{10} is odd. This enables the 9 other prisoners to logically deduce the colour of their own hat by comparing the parity of black hats with what they see.



To illustrate this, consider the above illustration. Here, Prisoner 10 sees 4 black hats ($S_{10} = 4$), thus Prisoner 10 says "Black". This tells the other prisoners that $S_{10} \equiv 0 \pmod{2}$. Since Prisoner 9 only sees 3 black hats ($S_9 = 3$), he deduces that

$$u_9 = S_{10} - S_9 \equiv 1 \pmod{2} \implies u_9 = 1,$$

i.e. his hat is black. He thus says "Black". It is now Prisoner 8's turn. From the preceding answers, he deduces that S_9 is odd, whence

$$u_8 = S_9 - S_8 \equiv 0 \pmod{2} \implies u_8 = 0,$$

i.e. his hat is white. A chain of similar reasoning continues all the way to Prisoner 1, at which point Prisoners 1 – 9 have correctly guessed the colours of their hats.

Part (b). Following a similar argument, the prisoners can simply number the colours as $1, 2, \dots, n$ and consider $S_k - S_{k-1}$ modulo n, from which Prisoners $1 - (m-1)$ will be able to deduce their own hat colour.

Problem 2. Two players are playing a coin game. An even number of coins of non-unique integer values are placed in random order in a row. The players take turns collecting a coin from either end of the row. The player with the highest total value of his collected coins wins. Does either player have a winning strategy, and what is the strategy if there is one?

Solution. Player 1 can guarantee a win or a draw. Index the coins from 1 to $2n$. Observe that Player 1 will always have the choice of an odd and even index, e.g. coins 1 (odd) and $2n$ (even). Thus, the indices of the coins available to Player 2 will have the same parity, and it will be opposite that of the index taken by Player 1 on the previous turn. For instance, if Player 1 takes coin 1 at the start (odd parity), then Player 2 must choose between coins 2 and $2n$ (both even parity). This means that Player 1 can take all coins whose index is of the same parity, e.g. all coins with odd index, or all coins with even index. Player 1 can thus take the parity which results in a higher sum and win the game.

Problem Set 1

Problem 1. Determine whether each of the following statements is true or false. Give a direct proof if it is true, and give a counter-example if it is false.

(a) The set of prime numbers is closed under addition.

(b) The set of positive rational numbers is closed under division.

Solution. (a) is false (since 3 and 5 are prime but their sum, 8, is not), while (b) is true.

Proof of (b). Let a/b and c/d be positive rational numbers, i.e. a, b, c and d are positive integers. Then

$$\frac{a/b}{c/d} = \frac{ad}{bc}.$$

Since both ad and bc are positive integers, it follows that ad/bc is a positive rational number. Hence, the set of positive rational numbers is closed under division. □

* * * * *

Problem 2. Let a, b and c be non-zero integers. Use the definition of divisibility and write down a direct proof for each of the following statements. (Indicate every step clearly).

(a) If a divides b, then ac divides bc.

(b) If a divides b and b divides a, then $a = \pm b$.

Solution.

Proof of (a). Since a divides b, we have

$$b = ka$$

for some integer k. Multiplying this equation through by c,

$$bc = k(ac).$$

Hence, by the definition of divisibility, ac divides bc. □

Proof of (b). Since a divides b, we have

$$b = k_1 a$$

for some integer k_1. Similarly, since b divides a, we have

$$a = k_2 b$$

for some integer k_2. Substituting this into the first equation,

$$b = k_1 k_2 b \implies k_1 k_2 = 1.$$

Since k_1 and k_2 are integers, we either have $k_1 = k_2 = 1$ or $k_1 = k_2 = -1$. Thus, $a = b$ or $a = -b$, i.e. $a = \pm b$. □

Problem 3. Show that 3 divides $n(n+1)(2n+1)$ for any integer n.
Extension. Let $k \geq 3$. Let S and P be the sum and product of $k-1$ consecutive integers, starting from n. Prove that $k \mid PS$ for all $n \in \mathbb{Z}$. (The original problem is the $k=3$ case).

Proof 1. Observe that

$$n(n+1)(2n+1) = 6\sum_{k=1}^{n} k^2 = 3\left(2\sum_{k=1}^{n} k^2\right).$$

Since $2\sum_{k=1}^{n} k^2$ is an integer, $n(n+1)(2n+1)$ is a multiple of 3. □

Proof 2. Observe that

$$n(n+1)(2n+1) = 2(n-1)(n)(n+1) + 3n(n+1).$$

This must be divisible by 3, since $(n-1)(n)(n+1)$ (three consecutive integers) and $3n(n+1)$ are both divisible by 3. □

Proof of Extension. Note that

$$P = n(n+1)(n+2)\dots(n+k-2).$$

Trivially, $k \mid P$ for all $n \equiv 0, 2, 3, \dots, k-1 \pmod{k}$. We hence consider only $n \equiv 1 \pmod{k}$. Note that

$$S = \sum_{i=0}^{k-1}(n+i) = (k-1)n + \frac{(k-1)(k-2)}{2}.$$

Hence,

$$\begin{aligned} SP &= n(n+1)\dots(n+k-2)\left[(k-1)n + \frac{(k-1)(k-2)}{2}\right] \\ &\equiv (1)(2)\dots(k-1)\left[-1 + \frac{(k-1)(k-2)}{2}\right] = (k-1)!\left(\frac{k(k-3)}{2}\right) \pmod{k}. \end{aligned}$$

For all $k \geq 3$, we have $2 \mid (k-1)!$, whence $\frac{1}{2}(k-3)(k-1)!$ is an integer, thus $SP \equiv 0 \pmod{k}$, i.e. $k \mid SP$. □

* * * * *

Problem 4. Prove that for all integers a, if the remainder is NOT 2 when a is divided by 4, then $4 \mid a^3 + 23a$.

Solution. Observe that

$$a^3 + 23a = (a-1)a(a+1) + 24a \equiv (a-1)a(a+1) \pmod{4}.$$

Case 1. If $a \equiv 0 \pmod{4}$, i.e. a is a multiple of 4, then $a^3 + 23a$ is trivially a multiple of 4.

Case 2. If $a \equiv 1, 3 \pmod{4}$, i.e. a is odd, then both $a-1$ and $a+1$ are even and contribute at least one factor of 2 each to $(a-1)a(a+1)$. Hence, $a^3 + 23a$ is divisible by $2^2 = 4$.

Case 3. If $a \equiv 2 \pmod{4}$, then a contributes only one factor of two. Additionally, both $a-1$ and $a+1$ are odd and do not contribute any factors of two. Thus, $(a-1)a(a+1)$ has only one factor of 2 and is not divisible by 4.

Problem 5. For any integer $n > 1$, let the standard factored form of n be given by

$$n = p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}.$$

Prove that n is a perfect square if and only if $k_1, k_2, \dots, k_r$ are all even integers.

Solution.

Proof. We begin by proving the backwards case. Suppose $k_1, k_2, \dots, k_r$ are all even integers. We can write $k_i = 2k_i'$ for all $1 \leq i \leq r$. Then

$$n = p_1^{k_1} p_2^{k_2} \dots p_r^{k_r} = p_1^{2k_1'} p_2^{2k_2'} \dots p_r^{2k_r'} = \left(p_1^{k_1'} p_2^{k_2'} \dots p_r^{k_r'} \right)^2.$$

Since $p_1^{k_1'} p_2^{k_2'} \dots p_r^{k_r'}$ is an integer, n is a perfect square.

We now prove the forwards case. Since n is a perfect square, we have $n = m^2$ for some positive integer m. Let the prime factorization of m be given by

$$m = q_1^{k_1'} q_2^{k_2'} \dots q_r^{k_r'},$$

where q_i are primes and k_1' are non-negative integers. Then

$$n = \left(q_1^{k_1'} q_2^{k_2'} \dots q_r^{k_r'} \right)^2 = q_1^{2k_1'} q_2^{2k_2'} \dots q_r^{2k_r'}.$$

Note that this is exactly the prime factorization of n. Also notice that all the exponents are multiples of 2 and are hence even. ☐

* * * * *

Problem 6. For all integers a and b, prove that $3 \mid ab$ if and only if $3 \mid a$ or $3 \mid b$.

Solution.

Proof. The backwards case is trivial. We hence only consider the forwards case. We prove this claim using the contrapositive. Suppose $3 \nmid a$ and $3 \nmid b$. Then

$$a \equiv n_1 \pmod 3, \quad b \equiv n_2 \pmod 3,$$

where n_1 and n_2 are integers with $0 < n_1, n_2 < 2$. Without loss of generality, suppose $n_1 \leq n_2$.

Applying standard properties of modular arithmetic, we obtain

$$ab \equiv n_1 n_2 \pmod 3.$$

Case 1. Suppose $n_1 = n_2 = 1$. Then $ab \equiv 1 \not\equiv 0 \pmod 3$.

Case 2. Suppose $n_1 = 1$, $n_2 = 2$. Then $ab \equiv 2 \not\equiv 0 \pmod 3$.

Case 3. Suppose $n_1 = n_2 = 2$. Then $ab \equiv 4 \equiv 1 \not\equiv 0 \pmod 3$.

In any case, $ab \not\equiv 0 \pmod 3$, i.e. 3 does not divide ab. By the contrapositive, it follows that 3 divides ab if 3 divides a or b. ☐

* * * * *

Problem 7. Let a, b and n be integers with $n > 1$. Suppose $a \equiv b \pmod n$. Prove the following:

- $ka \equiv kb \pmod{kn}$ for any positive integer k.

- If m is a common divisor of a, b and n, and $1 < m < n$, then

$$\frac{a}{m} \equiv \frac{b}{m} \pmod{\frac{n}{m}}.$$

Solution.

Proof of (a). Since $a \equiv b \pmod{n}$, we have $a = cn + b$ for some integer c. Multiplying this through by k, we have $ak = c(nk) + bk$. Hence, $ak \equiv bk \mod nk$. □

Proof of (b). Since $a \equiv b \pmod{n}$, we have $a = cn + b$ for some integer c. Since m is a common divisor of a, b and n, we have $a = ma'$, $b = mb'$ and $n = mn'$ for integers a', b' and m'. Dividing through by m, we get

$$\frac{a}{m} = \frac{cn}{m} + \frac{b}{n} \implies a' = cn' + b'.$$

Hence, $a' \equiv b' \pmod{n'}$, i.e.

$$\frac{a}{m} \equiv \frac{b}{m} \pmod{\frac{n}{m}}.$$

□

Problem Set 2

Problem 1. Is each of the following statements true or false? Give a proof if it is true, and give a counter-example if it is false.

(a) For each pair of real numbers x and y, if $x + y$ is irrational, then x is irrational and y is irrational.

(b) For each pair of real numbers x and y, if $x + y$ is irrational, then x is irrational or y is irrational.

Solution. The first statement is false: Take $x = 0$ and $y = \sqrt{2}$. Then $x + y = \sqrt{2}$ is irrational, but $x = 0$ is rational.

The second statement is true.

Proof of (b). Suppose x and y are rational. Then $x + y$ is also rational ($\mathbb{Q}$ is closed under addition). Thus, by the contrapositive, if $x + y$ is irrational, either x or y must be irrational. $\square$

* * * * *

Problem 2. Determine whether each of the following real numbers is rational or irrational. Justify your answers.

(a) $\sqrt{3} + \sqrt{5}$;

(b) $\sqrt{2} + \sqrt{8}$;

(c) $\left(1 + \sqrt{2}\right) / \left(1 + \sqrt{3}\right)$.

Solution.

Part (a). Seeking a contradiction, suppose $\sqrt{3} + \sqrt{5}$ is rational. Then

$$\sqrt{3} - \sqrt{5} = \frac{3^2 - 5^2}{\sqrt{3} + \sqrt{5}}$$

is also rational. Thus, both $\sqrt{3}$ and $\sqrt{5}$ are rational. But 3 and 5 are not perfect squares, so by P6, they must be irrational, a contradiction. Thus, $\sqrt{3} + \sqrt{5}$ must be irrational.

Part (b). Note that

$$\sqrt{2} + \sqrt{8} = \sqrt{2} + 2\sqrt{2} = 3\sqrt{2}.$$

Seeking a contradiction, suppose $3\sqrt{2}$ is rational. Then $\sqrt{2}$ must also be rational. But 2 is not a perfect square, so by P6, $\sqrt{2}$ is irrational, a contradiction. Thus, $\sqrt{2} + \sqrt{8}$ must be irrational.

Part (c). Rationalizing the fraction,

$$\frac{1 + \sqrt{2}}{1 + \sqrt{3}} = \frac{\left(1 + \sqrt{2}\right)\left(1 - \sqrt{3}\right)}{\left(1 + \sqrt{3}\right)\left(1 - \sqrt{3}\right)} = \frac{1 + \sqrt{2} - \sqrt{3} - \sqrt{6}}{-8}.$$

Using an identical argument as part (a), one can show that the numerator is irrational, whence the original fraction must also be irrational.

* * * * *

Problem 3. Use proof by contradiction to show that the sum of squares of two odd integers is not divisible by 4.

Proof. Seeking a contradiction, suppose there exist two odd integers $k_1 = 2n_1 + 1$ and $k_2 = 2n_2 + 1$ such that $k_1^2 + k_2^2 \equiv 0 \pmod 4$. However,

$$k_1^2 + k_2^2 = (2n_1+1)^2 + (2n_2+1)^2 = \left(4n_1^2 + 4n_1 + 1\right) + \left(4n_2^2 + 4n_2 + 1\right) \equiv 2 \pmod 4.$$

Thus, $0 \equiv 2 \pmod 4$, a contradiction. Hence, the sum of squares of two odd integers is not divisible by 4. □

* * * * *

Problem 4. Prove that there are no integers a and n with $n \geq 2$ and $a^2 + 1 = 2^n$.

Proof. Note that the only possible remainders of $a^2+1 \pmod 4$ are 1 and 2. However, for $n \geq 2$, we have $2^n \equiv 0 \pmod 4$. Since $0 \not\equiv 1, 2 \pmod 4$, the desired statement holds. □

* * * * *

Problem 5.

(a) Let p be a prime number greater than 2. Write down the possible remainders of p when divided by 4.

Fermat's Little Theorem states that if p is prime and a is an integer, which is not divisible by p, then $a^{p-1} \equiv 1 \pmod p$.

(b) Use Fermat's Little Theorem to prove that if p is a prime number greater than 2, and there exists an integer z such that $z^2 \equiv -1 \pmod p$, then p is not congruent to $3 \pmod 4$.

(c) Write down the possible remainders of w^2 when divided by 8, where w is an integer.

Solution.

Part (a). All primes greater than 2 are odd. Thus, the only possible remainders when p is divided by 4 are 1 and 3.

Proof of (b). Since $z^2 \equiv -1 \not\equiv 0 \pmod p$, we know that $p \nmid z$. Thus, by Fermat's Little Theorem, we have

$$z^{p-1} \equiv 1 \pmod p.$$

Squaring the given congruence, we also have

$$z^4 \equiv 1 \pmod p.$$

Seeking a contradiction, suppose p is congruent to $3 \pmod 4$. Then $p - 1 = 4k + 2$ for some integer k. Thus,

$$z^{p-1} = z^{4k+2} = \left(z^4\right)^k z^2 \equiv 1^k(-1) = -1 \pmod p,$$

which is only possible for $p = 2$, a contradiction. Thus, p cannot be congruent to 3 (mod 4). □

Part (c). The possible remainders of w^2 when divided by 8 are 0, 1, and 4.

* * * * *

Problem 6. The Unique Factorization Theorem states that every integer $n > 1$ has a unique standard factored form, i.e. there is exactly one way to express

$$n = p_1^{k_1} p_2^{k_2} \dots p_t^{k_t}$$

where $p_1 < p_2 < \cdots < p_t$ are distinct primes and $k_1, k_2, \ldots, k_t$ are some positive integers.

Use the Unique Factorization Theorem to prove that, if a positive n is not a perfect square, then $\sqrt{n}$ is irrational.

Proof. We prove the claim via the contrapositive. Suppose $\sqrt{n}$ is rational, where n is an integer. Write $\sqrt{n} = a/b$ for integers a, b with $b \neq 0$. Squaring, we get

$$n = \frac{a^2}{b^2} \implies b^2 n = a^2. \tag{1}$$

Let $\nu_p(z)$ represent the power of p in the factorization of an integer z. From (1), we have

$$\nu_p\left(b^2\right) + \nu_p(n) = \nu_p\left(a^2\right) \implies 2\nu_p(b) + \nu_p(n) = 2\nu_p(a) \implies \nu_p(n) = 2\left[\nu_p(a) - \nu_p(b)\right],$$

which is even. Hence, all prime factors of n have an even power, thus n is a perfect square.

Hence, by the contrapositive, if n is not a perfect square, then $\sqrt{n}$ is irrational. $\square$

Problem Set 3

Problem 1. Use Mathematical Induction to prove the following:

(a) For each positive integer n, $1(1!) + 2(2!) + 3(3!) + \cdots + n(n!) = (n+1)! - 1$.

(b) For each positive integer n, $(-7)^n - 9^n$ is divisible by 16.

(c) For each positive integer n with $n \geq 3$, $\left(1 + \frac{1}{n}\right)^n < n$.

(d) For each positive integer $n \geq 6$, $n^3 < n!$.

Proof of (a). Let $n \in \mathbb{N}$, and let $P(n)$ be the statement

$$P(n) : \sum_{i=1}^{n} i(i!) = (n+1)! - 1.$$

The base case $P(1)$ is trivial:

$$\sum_{i=1}^{1} i(i!) = 1 = (1+1)! - 1.$$

Suppose that $P(k)$ is true for some $k \in \mathbb{N}$. Then

$$\sum_{i=1}^{k+1} i(i!) = \sum_{i=1}^{k} i(i!) + (k+1)(k+1)!$$
$$= [(k+1)! - 1] + (k+1)(k+1)! = (k+2)(k+1)! - 1 = (k+2)! - 1.$$

Thus, $P(k) \implies P(k+1)$. Hence, by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. $\square$

Proof of (b). Let $n \in \mathbb{N}$ and let $P(n)$ be the statement

$$P(n) : 16 \mid (-7)^n - 9^n.$$

The base case $P(1)$ clearly holds, since $(-7)^1 - 9^1 = -16$, which is divisible by 16. Suppose that $P(k)$ is true for some $k \in \mathbb{N}$. Then

$$(-7)^{k+1} - 9^{k+1} = (-7)(-7)^k - 9(9)^k = (-7)\left[(-7)^k - 9^k\right] - 16(9)^k$$
$$\equiv (-7)(0) - (0)(9^n) = 0 \pmod{16}.$$

Thus, $P(k) \implies P(k+1)$. Hence, by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. $\square$

Alternative proof of (b). Let $u_n = (-7)^n - 9^n$. Then u_n is the general solution of a second-order homogenous linear recurrence relation with characteristic equation $(x+7)(x-9) = x^2 - 2x - 63$. Thus,

$$u_n = 2u_{n-1} - 63u_{n-2} \tag{1}$$

for $n \geq 2$ with initial conditions $u_0 = 0$ and $u_1 = -16$. Since both u_0 and u_1 are divisible by 16, and (1) is homogenous, it follows that $16 \mid u_n = (-7)^n - 9^n$ for all $n \in \mathbb{Z}_{\geq 0}$. $\square$

Proof of (c). Let $n \in \mathbb{Z}_{\geq 3}$, and let $P(n)$ be the statement

$$P(n) : \left(1 + \frac{1}{n}\right)^n < n.$$

The base case $P(3)$ is trivial:

$$\left(1 + \frac{1}{3}\right)^3 = \frac{64}{27} < 3.$$

Suppose that $P(k)$ is true for some $k \in \mathbb{Z}_{\geq 3}$. Then

$$\left(1 + \frac{1}{k+1}\right)^{k+1} < \left(1 + \frac{1}{k}\right)^{k+1} < k\left(1 + \frac{1}{k}\right) = k + 1.$$

Thus, $P(k) \implies P(k+1)$. Hence, by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{Z}_{\geq 3}$. $\square$

Proof of (d). We begin by proving that $n^3 - (n+1)^2 > 0$ for all integers $n \geq 6$. Firstly, observe that

$$n^3 - (n+1)^2 = n^3 - n^2 - 2n - 1 = \left(n^2 - 2\right)(n-1) - 3,$$

which is increasing for $n \geq 6$. Thus,

$$n^3 - (n+1)^2 \geq \left(6^2 - 2\right)(6-1) - 3 > 0.$$

Let $P(n)$ be the statement that $n^3 < n!$. We now prove that $P(n)$ is true for all integers $n \geq 6$. The base case is trivial:

$$6^3 = 216 < 720 = 6!.$$

Suppose that $P(k)$ is true for some integer $k \geq 6$. Then

$$(k+1)^3 = (k+1)(k+1)^2 < (k+1)k^3 < (k+1)k! = (k+1)!.$$

Thus, $P(k) \implies P(k+1)$. Hence, by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{Z}_{\geq 6}$. $\square$

* * * * *

Problem 2. Let $f_1, f_2, \ldots, f_n, \ldots$ be the Fibonacci sequence. That is, the sequence is defined recursively by

$$f_n = f_{n-1} + f_{n-2}$$

for all $n \geq 3$, with initial conditions $f_1 = 1$ and $f_2 = 1$.

Prove each of the following:

(a) For each positive integer n, f_{5n} is a multiple of 5.

(b) For each positive integer n, $f_1 + f_3 + \cdots + f_{2n-1} = f_{2n}$.

(c) For each positive integer n, $2f_n + 3f_{n+1} = f_{n+4}$.

(d) For each positive integer n, f_n is even if and only if $3 \mid n$.

Solution. We first prove two identities involving the Fibonacci sequence:

Lemma 35.0.1 (Honsberger's Identity). For all $m, n \in \mathbb{N}_0$, $f_{m+n} = f_{m+1}f_n + f_m f_{n-1}$.

Proof. f_k counts the number of ways to climb $k-1$ steps, taking either 1 or 2 steps each time. Now consider climbing $m+n-1$ steps.

Case 1: We step on the mth step. We effectively climb m steps before climbing another $n-1$ steps. This gives a total of $f_{m+1}f_n$ possibilities.

Case 2: We do not step on the $(m-1)$th step. We effectively climb the first $m-1$ steps, are forced to jump 2 steps to the $(m+1)$th step, and climb the remaining $n-2$ steps. This gives a total of $f_m f_{n-1}$ possibilities.

Thus, the total number of ways to climb $m+n-1$ is

$$f_{m+n} = f_{m+1}f_n + f_m f_{n-1}.$$

□

Lemma 35.0.2. For all $m, n \in \mathbb{Z}_{>2}$, $m \mid n \iff f_m \mid f_n$.

Proof. Let $n = km - r$, where $k, r \in \mathbb{Z}$ with $k \in \mathbb{N}$ and $0 \le r < m$. Note that $m \mid n \iff r = 0$. Let $P(k)$ be the statement

$$P(k) : r = 0 \iff f_m \mid f_{km-r}.$$

Base case: Note that $f_m \ge f_{m-r}$ with equality only when $r = 0$. Thus, $r = 0 \iff f_m \mid f_{m-r}$, hence $P(1)$ holds.

Suppose $P(k)$ is true for some $k \in \mathbb{N}$. Let $f_{km-r} = af_m + b$, where $a, b \in \mathbb{Z}$ and $0 \le b < f_m$. Consider $f_{(k+1)m-r}$. Using Honsberger's identity

$$\begin{aligned} f_{(k+1)m-r} = f_{km-r}f_{m-1} + f_m f_{km-r+1} &= (af_m + b) f_{m-1} + f_m f_{km-r+1} \\ &= f_m (af_{m-1} + f_{km-r+1}) + bf_{m-1}. \end{aligned}$$

It is well-known fact that f_m and f_{m-1} are always coprime. Hence,

$$f_m \mid f_{(k+1)m-r} \iff f_m \mid bf_{m-1} \iff f_m \mid b \iff b = 0.$$

However, by our induction hypothesis,

$$r = 0 \iff f_m \mid f_{km-r} \iff b = 0.$$

Thus,

$$r = 0 \iff f_m \mid f_{(k+1)m-r}.$$

Hence, $P(k) \implies P(k+1)$, and by the principle of mathematical induction, $P(k)$ holds for all $k \in \mathbb{N}$. □

Part (a). By Lemma 35.0.2, $5 = f_5 \mid f_{5n}$ for all positive integers n.

Part (b). Let $P(n)$ be the statement that

$$P(n) : f_1 + f_3 + \cdots + f_{2n-1} = f_{2n}.$$

The base case $P(1)$ clearly holds, since $f_1 = 1 = f_2$. Now suppose $P(k)$ holds for some $k \in \mathbb{N}$. Then

$$f_1 + f_3 + \cdots + f_{2(k+1)-1} = f_1 + f_3 + \cdots + f_{2k-1} + f_{2k+1} = f_{2k} + f_{2k+1} = f_{2k+2} = f_{2(k+1)}.$$

Hence, $P(k) \implies P(k+1)$. Thus, by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$.

Part (c). Using Honsberger's identity, $f_{n+4} = f_3 f_n + f_4 f_{n+1} = 2f_n + 3f_{n+1}$.

Part (d). By Lemma 35.0.2, $2 = f_3 \mid f_n$ if and only if $3 \mid n$.

* * * * *

Problem 3. Euclid's lemma states that for any prime p and any integers a and b, if $p \mid ab$ and $p \nmid a$, then $p \mid b$.

Use mathematical induction together with Euclid's lemma to prove that for any prime p and any integers $q_1, q_2, \ldots, q_n$, if $p \mid q_1 q_2 \ldots q_n$, then $p \mid q_i$ for some i.

Proof. Let $P(n)$ be the statement that for any prime p and any integers $q_1, q_2, \ldots, q_n$, if $p \mid q_1 q_2 \ldots q_n$, then $p \mid q_i$ for some i.

The base case $P(1)$ is trivial: since $p \mid q_1$, we are done. The case $P(2)$ is also trivial: suppose $p \mid q_1 q_2$. If $p \mid q_1$, we are done. If not, by Euclid's lemma, $p \mid q_2$ and we are done.

Now suppose that $P(k)$ holds for some $k \in \mathbb{N}$. Suppose $p \mid q_1 q_2 \ldots q_{k+1}$. If $p \mid q_{k+1}$, we are done. Else, $p \mid q_1 q_2 \ldots q_k$, then by our inductive hypothesis, there exists some $1 \leq i \leq k$ such that $p \mid q_i$ and we are done. Thus, $P(k) \implies P(k+1)$ and by the principle of mathematical induction, $P(n)$ holds for all $n \in \mathbb{N}$. $\square$

* * * * *

Problem 4. Suppose you want to prove $P(n)$ is true (only) for all integers $n \geq 7$ that are not divisible by 4 using a version of mathematical induction as follows:

i. Basis step: $P(a)$, $P(b)$, $P(c)$ are true; and

ii. Inductive step: $(\forall k \in \mathbb{Z}^+) P(k) \implies P(k+d)$ is true.

What should the values for a, b, c, d?

Solution. Clearly, $a = 7$, $b = 9$, $c = 10$ and $d = 4$.

* * * * *

Problem 5. Find the mistake in the following "proof" that purports to show that every non-negative integer power of every non-zero real number is 1.

"Proof". Let r be any non-zero real number and let the predicate $P(n)$ be

$$P(n) : r^n = 1.$$

Basis step: $P(0)$ is true because $r^0 = 1$ by definition of 0-th power.

Inductive step: $P(k-1)$ and $P(k) \implies P(k+1)$.

Suppose that $r^{k-1} = 1$ and $r^k = 1$. This is the induction hypothesis. We must show that $r^{k+1} = 1$. Now

$$r^{k+1} = r^{k+k-(k-1)} = \frac{r^k \cdot r^k}{r^{k-1}} = \frac{1 \cdot 1}{1} = 1.$$

Thus, $r^{k+1} = 1$. Hence, the inductive step is proven. $\square$

Solution. Since the inductive step requires $P(k-1)$ and $P(k)$, the basis step should also show that $P(1)$ holds. However, $P(1)$ is clearly false, since $r^1 \neq 1$ in general.

Problem Set 4

Problem 1. Prove that for every pair of irrational numbers p and q such that $p < q$, there is an irrational x such that $p < x < q$.

Proof. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, it follows that $\mathbb{Q} + \sqrt{2}$ is also dense in $\mathbb{R}$. Hence, there must exist some irrational x (of the form $q + \sqrt{2}$, where $q \in \mathbb{Q}$) such that $p < x < q$. □

* * * * *

Problem 2. Show that there is one and only one integer t such that t, $t+2$, $t+4$ are all prime numbers.

Solution. Let $T = \{t, t+2, t+4\}$. Observe that T forms a complete residue system modulo 3. Hence, $3 \in T$.

Case 1. If $t = 3$, then $T = \{3, 5, 7\}$, which are all prime.

Case 2. If $t + 2 = 3$, then $t = 1$ which is not prime.

Case 3. If $t + 4 = 3$, then $t = -1$ which is not prime.

Thus, t, $t+2$ and $t+4$ are all prime only when $t = 3$.

* * * * *

Problem 3. Show that there exist integers x and y that satisfy

$$(2n+1)x + (9n+4)y = 1$$

for every integer n.

Proof 1. Observe that

$$\gcd(2n+1, 9n+4) = \gcd(2n+1, n) = \gcd(1, n) = 1.$$

Thus, by Bézout's identity, there exist integers x and y such that

$$(2n+1)x + (9n+4)y = 1$$

for all integers n. □

Proof 2. Rearranging, we get

$$x = \frac{1 - (9n+4)y}{2n+1}.$$

Let $y = 2k$ for some integer k. Then

$$x = \frac{1 - 18nk - 8k}{2n+1} = \frac{k+1}{2n+1} - 9k.$$

Taking $k = 2n$, we have $x = 1 - 9k = 1 - 18n$ and $y = 2k = 4n$. Indeed, one can verify that

$$(2n+1)x + (9n+4)y = (2n+1)(1-18n) + (9n+4)(4n) \equiv 1.$$

□

* * * * *

Problem 4. Given n real numbers $a_1, a_2, \ldots, a_n$, show that there exists an a_i $(1 \leq i \leq n)$ such that a_i is greater than or equal to the mean value of the n numbers.

Solution. Let m be the mean of the n real numbers. Seeking a contradiction, suppose there does not exist an a_i such that $a_i \geq m$. Then $a_k < m$ for all $1 \leq k \leq n$, from which it follows that

$$m = \frac{a_1 + a_2 + \cdots + a_n}{n} < \frac{m + m + \cdots + m}{n} = m,$$

a contradiction. Thus, there must exist some a_i greater than or equal to m.

* * * * *

Problem 5. Determine whether the two statements below are true or false. Justify your answers.

(a) There is an irrational number a such that for all irrational numbers b, ab is rational.

(b) For every irrational number a, there is an irrational number b such that ab is rational.

Solution.

Part (a). The statement is false. There are uncountably many irrationals, so there are uncountably many products ab, where b is irrational. However, there are countably many rationals. Hence, there must exist some irrational b such that ab is irrational.

Part (b). The statement is true. Take $b = 1/a$, which is irrational if a is irrational. Then $ab = 1$ which is rational.

* * * * *

Problem 6. Prove that there are infinitely many prime numbers that are congruent to 3 modulo 4.

Solution. Seeking a contradiction, suppose there are finitely many prime numbers congruent to 3 modulo 4. Label them $p_1, p_2, \ldots, p_n$. Now consider

$$P = 2\left(p_1 p_2 \ldots p_n\right) + 1 \equiv 2\left(3^n\right) + 1 \equiv 3 \pmod{4}.$$

By construction, $p_i \nmid P$ for all $1 \leq i \leq n$. Thus, P must also be a prime with residue 3 modulo 4, a contradiction. Thus, there are infinitely many primes congruent to 3 modulo 4.

* * * * *

Problem 7. Prove that, for any positive integer n, there is a perfect square m^2 such that $n \leq m^2 \leq 2n$.

Solution. Seeking a contradiction, suppose there exists an $n \in \mathbb{Z}^+$ such that there does not exist a perfect square in $[n, 2n]$. Then there exists some $m \in \mathbb{Z}$ such that $m^2 < n$ and $2n < (m+1)^2$. Putting the two inequalities together,

$$2m^2 < 2n < (m+1)^2 \implies 2m^2 + 2 \leq (m+1)^2 \implies (m-1)^2 \leq 0,$$

which immediately implies $m = 1$. However, we can rule this possibility out, since $1 \leq 1^2 \leq 2$. Thus, such an m cannot exist, a contradiction. Thus, for all $n \in \mathbb{Z}^+$, there must exist a perfect square m^2 such that $n \leq m^2 \leq 2n$.

* * * * *

Problem 8.

(a) For any positive integer a and real number t, it is given that t can be written as $an + p$ where n is an integer and $a > p \geq 0$. Prove that

$$\int_0^a \left\lfloor \frac{x+t}{a} \right\rfloor \mathrm{d}x = t.$$

(b) For any positive integers a and b and real number x,

(i) prove that

$$\left\lfloor \frac{\lfloor x/a \rfloor}{b} \right\rfloor = \left\lfloor \frac{x}{ab} \right\rfloor,$$

and

(ii) find

$$\int_0^{ab} (fg(x) - gf(x))\,\mathrm{d}x,$$

where $f(x) = \lfloor (x+a)/b \rfloor$ and $g(x) = \lfloor (x+b)/a \rfloor$.

Solution.

Part (a). Writing $t = an + p$, we have

$$\int_0^a \left\lfloor \frac{x+t}{a} \right\rfloor \mathrm{d}x = \int_0^a \left\lfloor \frac{x+an+p}{a} \right\rfloor \mathrm{d}x = \int_0^a \left(n + \left\lfloor \frac{x+p}{a} \right\rfloor \right) \mathrm{d}x = an + \int_0^a \left\lfloor \frac{x+p}{a} \right\rfloor \mathrm{d}x.$$

Now observe that

$$\left\lfloor \frac{x+p}{a} \right\rfloor = \begin{cases} 0, & x \in [0, a-p], \\ 1, & x \in [a-p, a]. \end{cases}$$

Thus,

$$\int_0^a \left\lfloor \frac{x+t}{a} \right\rfloor \mathrm{d}x = an + \int_0^a \left\lfloor \frac{x+p}{a} \right\rfloor \mathrm{d}x = an + \left(\int_0^{a-p} 0\,\mathrm{d}x + \int_{a-p}^a 1\,\mathrm{d}x \right) = an + p = t.$$

Part (b).

Part (b)(i). Write $x = a(bd + r_2) + r_1$, where $d, r_1, r_2 \in \mathbb{Z}$ with $0 \le r_1 < a$ and $0 \le r_2 < b$. Then

$$\left\lfloor \frac{\lfloor x/a \rfloor}{b} \right\rfloor = \left\lfloor \frac{bd + r_2}{b} \right\rfloor = d,$$

and

$$\left\lfloor \frac{x}{ab} \right\rfloor = \left\lfloor \frac{abd + ar_2 + r_1}{ab} \right\rfloor = d + \left\lfloor \frac{ar_2 + r_1}{ab} \right\rfloor.$$

Since

$$ar_2 + r_1 \le a(b-1) + (a-1) = ab - 1 < ab,$$

it follows that

$$\left\lfloor \frac{x}{ab} \right\rfloor = d.$$

Thus,

$$\left\lfloor \frac{\lfloor x/a \rfloor}{b} \right\rfloor = d = \left\lfloor \frac{x}{ab} \right\rfloor,$$

as desired.

Part (b)(ii). Note that

$$fg(x) = \left\lfloor \frac{\lfloor (x+b)/a \rfloor + a}{b} \right\rfloor = \left\lfloor \frac{\lfloor (x+b+a^2)/a \rfloor}{b} \right\rfloor = \left\lfloor \frac{x + (b+a^2)}{ab} \right\rfloor.$$

Using the result from part (a), we have

$$\int_0^{ab} fg(x)\,\mathrm{d}x = \int_0^{ab} \left\lfloor \frac{x + (b+a^2)}{ab} \right\rfloor \mathrm{d}x = b + a^2.$$

Similarly,

$$gf(x) = \left\lfloor \frac{\lfloor (x+a)/b \rfloor + b}{a} \right\rfloor = \left\lfloor \frac{\lfloor (x+a+b^2)/b \rfloor}{a} \right\rfloor = \left\lfloor \frac{x+\left(a+b^2\right)}{ab} \right\rfloor.$$

Using the result from part (a),

$$\int_0^{ab} gf(x)\,\mathrm{d}x = \int_0^{ab} \left\lfloor \frac{x+\left(a+b^2\right)}{ab} \right\rfloor \mathrm{d}x = a+b^2.$$

Thus,

$$\int_0^{ab} (fg(x) - gf(x))\,\mathrm{d}x = \left(b+a^2\right) - \left(a+b^2\right).$$

* * * * *

Problem 9. Let $S = \{1, 2, \ldots, 50\}$ and let D be a subset of S of size 27.

(a) Show that there are 25 subsets of S of the form $\{a, a+5\}$ whose union is S. Apply the pigeonhole principle to prove that D must contain two numbers that differ by exactly 5.

(b) Prove that D must contain two numbers that differ by exactly 6. Show that D does not necessarily contain two numbers that differ by exactly 7.

(c) Determine the maximum possible size of a subset of S that contains no four consecutive numbers.

(d) Determine the maximum possible size of a subset of S that contains no two numbers whose sum is a multiple of 10.

Solution.

Part (a). Let

$$L = \{1, \ldots, 5\} \cup \{11, \ldots, 15\} \cup \cdots \cup \{41, \ldots 45\}.$$

For each $a \in L$, define $l_a = \{a, a+5\}$. Then

$$S = \bigcup_{a \in L} l_a.$$

Since $|L| = 25$, there are 25 subsets of S of the form $\{a, a+5\}$ whose union is S.

All 27 elements of D must be placed into the 25 subsets l_a. Since $27 > 25$, by the pigeonhole principle, at least one subset must have both its elements in D. That is, D contains two numbers that differ by exactly 5.

Part (b). Let

$$L = \{1, \ldots, 6\} \cup \{13, \ldots, 18\} \cup \cdots \cup \{37, \ldots, 42\} \cup \{43, 44\}.$$

For each $a \in L$, define $l_a = \{a, a+6\}$. Note that $|L| = 26$, and

$$S = \bigcup_{a \in L} l_a.$$

All 27 elements of D must be placed into the 26 subsets l_a. Since $27 > 26$, by the pigeonhole principle, at least one subset must have both its elements in D. That is, D contains two numbers that differ by exactly 6.

However, D does not necessarily contain two numbers that differ by exactly 7. For instance, the set

$$\{1, \ldots, 7\} \cup \{15, \ldots, 21\} \cup \{29, \ldots, 35\} \cup \{43, \ldots, 48\}$$

contains 27 elements such that no two differ by exactly 7.

Part (c). Let D be a subset of S that contains no four consecutive numbers. At best, every group of four consecutive numbers have 3 elements in D, i.e. $|D| \leq \frac{3}{4} \times |S| = 37.5$, i.e. $|D| \leq 37$. Indeed, we can construct such a set with 37 elements:

$$D = \{1, 2, 3\} \cup \{5, 6, 7\} \cup \{9, 10, 11\} \cup \cdots \cup \{45, 46, 47\} \cup \{49, 50\}.$$

Part (d). Let D be a subset of S that contains no two numbers whose sum is a multiple of 10. Minimally, D contains all integers that have units digit $\{1, 2, 3, 4\}$ or $\{6, 7, 8, 9\}$ (both have the same number of elements). We can then add on a number with units digit 0 and another with units digit 5; we cannot have multiple numbers with units digit 0 or 5 since we can sum them to get a multiple of 10. Thus, the maximum size of D is

$$\max |D| = 4 \times \frac{50}{10} + 2 = 22.$$

Problem Set 5

Problem 1. Let a, b and $n > 1$ be integers. Prove that if $m > 1$ is a divisor of n and $a \equiv b \pmod{n}$, then $a \equiv b \pmod{m}$.

Proof. Since m is a divisor of n, there exists some integer k such that $n = mk$. Since $a \equiv b \pmod{n}$, there exists some integer l such that

$$a - b = ln = (lk)m.$$

Since lm is an integer, by definition, $a \equiv b \pmod{m}$. $\square$

* * * * *

Problem 2. Prove that if $3 \nmid a$, then $3 \mid a^2 + 5$.

Proof. Since $3 \nmid a$, by Fermat's Little Theorem, $a^2 + 5 \equiv 1 + 5 \equiv 0 \pmod{3}$. Hence, $3 \mid a^2 + 5$. $\square$

* * * * *

Problem 3. If m and n are any two integers with the same parity, then $4 \mid m^2 - n^2$.

Proof. Since m and n have the same parity, both $m + n$ and $m - n$ are even. Hence, $m^2 - n^2 = (m + n)(m - n)$ have two factors of 2, thus it is divisible by 4. $\square$

* * * * *

Problem 4. For any integers n, if 3 divides $1 + n + n^2$, then $n \equiv 1 \pmod{3}$.
Extension. If p is a prime and $p \mid 1 + n + n^2 + \cdots + n^{p-1}$, then $n \equiv 1 \pmod{p}$.

Proof. Since $3 \mid 1 + n + n^2$, we have $3 \nmid n$. Hence, by Fermat's Little Theorem,

$$0 = 1 + n + n^2 \equiv 2 + n \pmod{3}.$$

Thus, $n \equiv 1 \pmod{3}$. $\square$

Proof of Extension. We have

$$0 \equiv 1 + n + n^2 + \cdots + n^{p-1} = \frac{n^p - 1}{n - 1} \pmod{p}.$$

Since p is prime, $\mathbb{Z}/p\mathbb{Z}$ is an integral domain, so $n^p - 1 \equiv 0 \pmod{p}$. Invoking Fermat's Little Theorem, $n - 1 \equiv 0 \pmod{p}$ and the desired result immediately follows. $\square$

* * * * *

Problem 5. Prove that there is no smallest positive rational number.

Proof. Seeking a contradiction, suppose there exists a smallest positive rational number, say $x = m/n$, where $m, n \in \mathbb{Z}^+$ (WLOG) and $n \neq 0$. Consider $x/2 = m/2n$. Clearly, $x/2$ is positive and rational. Thus, $0 < x/2 < x$, contradicting the minimality of x. Hence, there is no smallest positive rational number. $\square$

* * * * *

Problem 6. A sequence is defined by $a_1 = 2$, $a_2 = 4$, $a_{n+2} = 5a_{n+1} - 6a_n$ for all $n \geq 1$. Prove that $a_2 = 2^n$ for all $n \in \mathbb{Z}^+$.

Proof. The characteristic equation of a_n is $m^2 - 5m + 6 = (m-3)(m-2)$. Hence,

$$a_n = A \cdot 2^n + B \cdot 3^n$$

for some constants A and B. From the initial conditions,

$$a_1 = 2 = 2A + 3B, \quad a_2 = 4 = 4A + 9B.$$

Solving, we get $A = 1$ and $B = 0$. Thus, $a_n = 2^n$. $\square$

* * * * *

Problem 7. Let $f_1, f_2, \ldots$ denote the Fibonacci sequence. Show that, for all $n \in \mathbb{Z}^+$, $f_2 + f_4 + \cdots + f_{2n} = f_{2n+1} - 1$.

Proof. Let $P(n)$ be the statement $f_2 + f_4 + \cdots + f_{2n} = f_{2n+1} - 1$. Since $f_2 = 1 = 2 - 1 = f_3 - 1$, $P(1)$ is true. Likewise, since $f_2 + f_4 = 1 + 3 = 5 - 1 = f_5 - 1$, $P(2)$ is also true.

Suppose $P(k-1)$ and $P(k)$ are true for some positive integer k. Then

$$f_2 + f_4 + \cdots + f_{2k} + f_{2k+2} = f_{2k+1} - 1 + f_{2k+2} = f_{2k+3} - 1,$$

whence $P(k+1)$ is true. Thus, by strong induction, the desired result holds. $\square$

* * * * *

Problem 8. Show that if m and $m+2$ are both primes with $m > 3$, then $m+1$ is divisible by 6.

Proof. For primes greater than 3, the only possible residues modulo 6 are 1 and 5. This immediately gives $m \equiv 5 \pmod 6$, whence $m + 1 \equiv 0 \pmod 6$. $\square$

* * * * *

Problem 9. Let a, b be integers, not both zero. If $\gcd(a, b) = 1$, show that the possible values for $\gcd(a+b, a-b)$ can be 1 and 2.

Proof. Let $b = a - k$ for some integer k. Then

$$1 = \gcd(a, b) = \gcd(a, a-k) = \gcd(a, -k) = \gcd(a, k)\,.$$

Note also that

$$\gcd(a+b, a-b) = \gcd(2a, a-b) = \gcd(2a, k)\,.$$

If k is odd, then $\gcd(2, k) = 1$ so

$$\gcd(2a, k) = \gcd(2, k) \times \gcd(a, k) = 1.$$

If k is even, then $\gcd(2, k) = 2$ so

$$\gcd(2a, k) = \gcd(2, k) \times \gcd(a, k) = 2.$$

$\square$

Analysis 1.1 Functions and Graphs

Tutorial A1.1 Set 1

Problem 1. The function f is defined for all $x \in \mathbb{R}$ by

$$f(x) = \begin{cases} k, & |x| \leq l, \\ 0, & |x| > l, \end{cases}$$

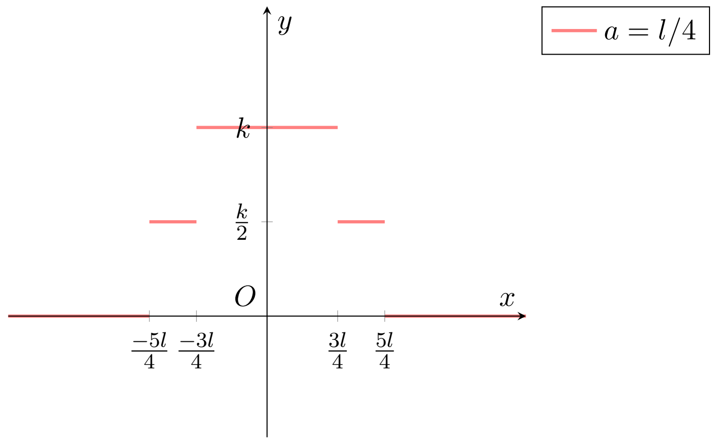
where k and l are positive constants.

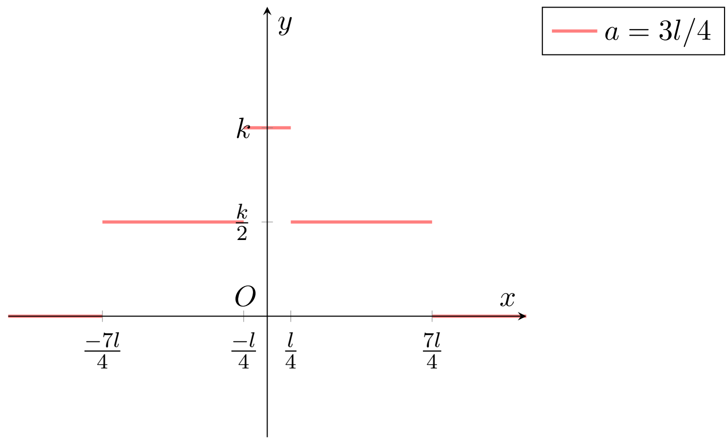
Sketch on three separate diagrams, using the same scales for each, the graph of the function g defined by

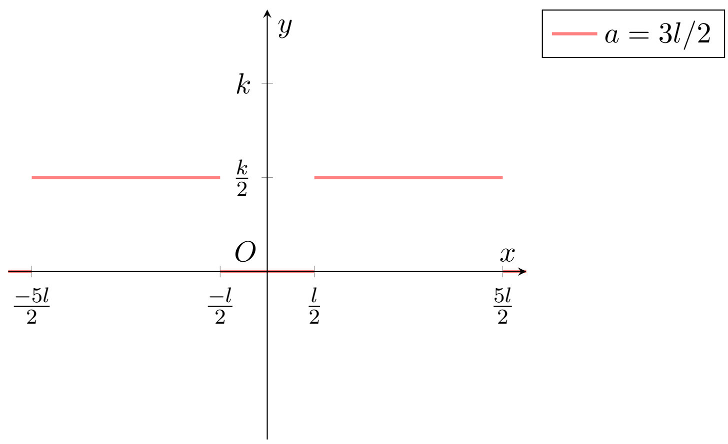
$$g(x) = \frac{f(x+a) + f(x-a)}{2}$$

in the cases $a = l/4$, $a = 3l/4$ and $a = 3l/2$.

Solution.







* * * * *

Problem 2. Prove that

(a) $\lfloor\sqrt{x}\rfloor = \left\lfloor\sqrt{\lfloor x\rfloor}\right\rfloor$,

(b) $\lceil\sqrt{x}\rceil = \left\lceil\sqrt{\lceil x\rceil}\right\rceil$.

Is $\lceil\sqrt{x}\rceil = \left\lceil\sqrt{\lfloor x\rfloor}\right\rceil$? If so, give a proof. If not, provide a counterexample.

Proof of (a). Let $n = \lfloor\sqrt{x}\rfloor$, where $n \in \mathbb{Z}$. By the definition of the floor function,

$$n \leq \sqrt{x} < n + 1.$$

Squaring, we get

$$n^2 \leq x < (n+1)^2.$$

Taking the floor, we have

$$n^2 \leq \lfloor x\rfloor \leq x < (n+1)^2.$$

Rooting, we have

$$n \leq \sqrt{\lfloor x\rfloor} < n + 1.$$

Once again, by the definition of the floor function, $n = \left\lfloor\sqrt{\lfloor x\rfloor}\right\rfloor$. Thus,

$$\lfloor\sqrt{x}\rfloor = n = \left\lfloor\sqrt{\lfloor x\rfloor}\right\rfloor.$$

□

Proof of (b). Let $n = \lceil\sqrt{x}\rceil$, where $n \in \mathbb{Z}$. By the definition of the ceiling function,

$$n - 1 < \sqrt{x} \leq n.$$

Squaring, we get

$$(n-1)^2 < x \leq n^2.$$

Taking the ceiling, we have

$$(n-1)^2 < x \leq \lceil x\rceil \leq n^2.$$

Rooting, we have

$$n - 1 < \sqrt{\lceil x\rceil} \leq n.$$

Once again, by the definition of the ceiling function, $n = \left\lceil \sqrt{\lceil x \rceil} \right\rceil$. Thus,

$$\lceil \sqrt{x} \rceil = n = \left\lceil \sqrt{\lceil x \rceil} \right\rceil .$$

□

Part (c). It is not true that $\lceil \sqrt{x} \rceil = \left\lceil \sqrt{\lfloor x \rfloor} \right\rceil$. Take $x = 1.21$. Then

$$\lceil \sqrt{x} \rceil = \left\lceil \sqrt{1.21} \right\rceil = \lceil 1.1 \rceil = 2,$$

but

$$\left\lceil \sqrt{\lfloor x \rfloor} \right\rceil = \left\lceil \sqrt{\lfloor 1.1 \rfloor} \right\rceil = \left\lceil \sqrt{1} \right\rceil = \lceil 1 \rceil = 1.$$

* * * * *

Problem 3. Given that $f : \mathbb{R} \to \mathbb{R}$ such that for all $x, y \in \mathbb{R} \setminus \{0\}$, $f(xy) = f(x) + f(y)$, find the values of $f(1)$ and $f(-1)$. Hence, show that f is even. Give an example of a function that satisfies the above properties and sketch its graph.

Solution. Taking $x = 1$, we see that

$$f(y) = f(1) + f(y) \implies f(1) = 0.$$

Taking $x = y = -1$,

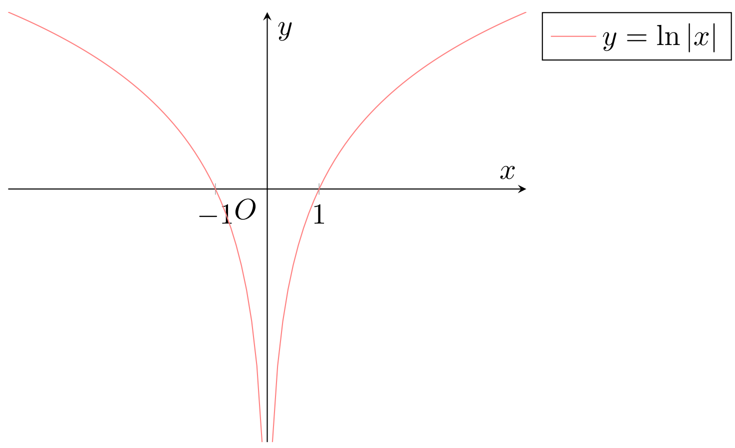
$$f(1) = f(-1) + f(-1) \implies f(-1) = 0.$$

Taking $y = -1$,

$$f(-x) = f(x) + f(-1) = f(x).$$

Thus, by definition, f is even.

An example of f is $f(x) = \ln |x|$.



* * * * *

Problem 4. If f and g are convex functions, show that the function h given by $h(x) = f(x)g(x)$ is not necessarily a convex function with a suitable example.

Solution. Let $f(x) = x^2$ and $g(x) = x^2 - 1$. Then $h(x) = x^2 \left(x^2 - 1\right)$ is clearly not convex.

Problem 5. For a triangle ABC with corresponding angles a, b and c, show that

$$\sin a + \sin b + \sin c \leq \frac{3\sqrt{3}}{2}$$

and determine when equality holds.

Proof. Since $y = \sin x$ is concave, by Jensen's inequality,

$$\frac{\sqrt{3}}{2} = \sin \frac{\pi}{3} = \sin \frac{a+b+c}{3} \geq \frac{\sin a + \sin b + \sin c}{3}.$$

The desired inequality follows immediately. $\square$

Tutorial A1.1 Set 2

Problem 1. The functions f, g are defined on $\mathbb{R}$ such that for any $x, y \in \mathbb{R}$,

$$f(x-y) = f(x)g(y) - f(y)g(x) \quad \text{and} \quad f(1) \neq 0.$$

(a) Prove that f is an odd function.

(b) If $f(1) = f(2)$, find the value of $g(1) + g(-1)$.

Solution.
Part (a). Let $x = y$. Then

$$f(0) = f(x)g(x) - f(x)g(x) = 0.$$

Let $y = 0$. Then

$$f(x) = f(x)g(0) - f(0)g(x) = f(x)g(0) \implies g(0) = 1.$$

Let $x = 0$. Then

$$f(-y) = f(0)g(y) - f(y)g(0) = -f(y).$$

Thus, f is odd.
Part (b). Let $x = 1$ and $y = -1$. Then

$$f(2) = f(1)g(-1) - f(-1)g(1).$$

Since f is odd, we have $f(-1) = -f(1)$. Further, we are given $f(2) = f(1)$. Hence,

$$f(1) = f(1)g(-1) + f(1)g(1) \implies g(-1) + g(1) = 1.$$

* * * * *

Problem 2. The function h is defined for $x \in \mathbb{R}$ by

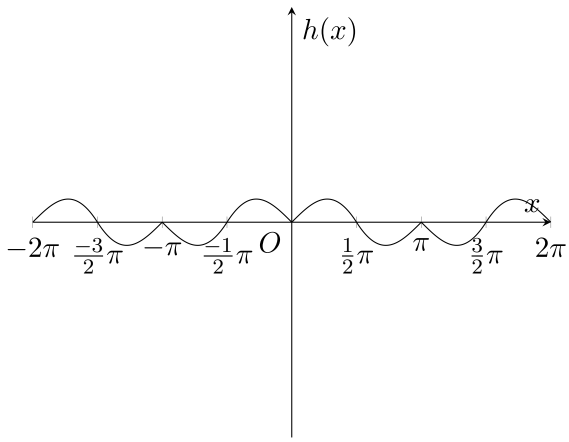
$$h(x) = x\cos x, \quad 0 \leq x \leq \frac{\pi}{2}$$

with the additional properties

$$h(-x) = h(x) \quad \text{and} \quad h(\pi + x) = -h(x).$$

Sketch the graph of h for $-2\pi \leq x \leq 2\pi$.

Solution.



Problem 3. Functions f and g are defined for $x \in \mathbb{R}$ by

$$f(x) = ax + b, \quad g(x) = cx + d,$$

where a, b, c and d are constants with $a \neq 0$. Given that $gf = f^{-1}g$, show that at least one of the following statements is true:

- g is a constant function,
- f^2 is the identity function,
- g^2 is the identity function.

Proof. Note that $f^{-1} = (x - b)/a$. Hence, the condition $gf = f^{-1}g$ implies that

$$c(ax + b) + d = \frac{(cx + d) - b}{a}.$$

Comparing coefficients of x and constant terms, we see that

$$\frac{c}{a} = ac \quad \text{and} \quad \frac{d - b}{a} = bc + d.$$

Case 1: $c = 0$. Then $g(x) = d$ is a constant function.
Case 2: $c \neq 0$. Then $1/a = a \implies a = \pm 1$.
Case 2a: $a = -1$. Then $f(x) = -x + b$, whence $f^2(x) = -(-x + b) + b = x$ is the identity function.
Case 2b: $a = 1$. Then $d - b = bc + d \implies b(c + 1) = 0$. If $b = 0$, then $f(x) = x$ is the identity function. If $c = -1$, then $g(x) = -x + d$, whence $g^2(x) = -(-x + d) + d = x$ is the identity function. ☐

* * * * *

Problem 4. Let $f : \mathbb{Q} \to \mathbb{Z}$ be a function satisfying the following:

- For any $x, y \in \mathbb{Q}$, we have $f(x + y) = f(x) + f(y)$,
- For any $r \in \mathbb{Z}$ and any $x \in \mathbb{Q}$, we have $f(rx) = rf(x)$.

(a) (i) Explain why for any $n \in \mathbb{Z}$, we have $f(1/n) \in \mathbb{Z}$.
(ii) Suppose that $0 \neq f(1)$. Let $p \in \mathbb{Z}$ be a prime such that $p \nmid f(1)$. By writing $f(1)$ as $f(p \cdot 1/p)$, explain why $f(1/p)$ cannot be an integer. Hence, prove that $f(1) = 0$.

(b) Show that for any $a \in \mathbb{Q}$, $f(a) = 0$.

Solution.

Part (a).

Part (a)(i). $f(1/n) \in \operatorname{Im} f = \mathbb{Z}$.

Part (a)(ii). Suppose $f(1) \neq 0$. Then there exists a prime p such that $p \nmid f(1)$. Then

$$f(1) = f\left(p \cdot \frac{1}{p}\right) = pf\left(\frac{1}{p}\right) \implies f\left(\frac{1}{p}\right) = \frac{f(1)}{p}.$$

Since $p \nmid f(1)$, it follows that $f(1/p)$ is not an integer. However, this contradicts (a)(i). Hence, our assumption that $f(1) \neq 0$ is false, i.e. $f(1) = 0$.

Part (b). From the second property, it follows that for any $r \in \mathbb{Z}$,

$$f(r) = f(r \cdot 1) = rf(1) = 0.$$

Let $a \in \mathbb{Q}$. Write $a = \alpha/\beta$, where $\alpha, \beta \in \mathbb{Z}$ with $\beta \neq 0$. Then

$$f(a) = \frac{1}{\beta}\beta f(a) = \frac{1}{\beta}f(\beta a) = \frac{1}{\beta}f(\alpha) = 0.$$

* * * * *

Problem 5. Let $f : \mathbb{R} \to \mathbb{R}$ be an additive function, i.e. for any $x, y \in \mathbb{R}$, we have $f(x+y) = f(x) + f(y)$.

(a) (i) Show that $f(0) = 0$.

(ii) Prove that for any $n \in \mathbb{Z}$ and any $x \in \mathbb{R}$, we have $f(nx) = nf(x)$.

(iii) Prove that for any $r \in \mathbb{Q}$ and any $x \in \mathbb{R}$, we have $f(rx) = rf(x)$.

(iv) Deduce that for any $r, x \in \mathbb{R}$, we have $f(rx) = rf(x)$.

(b) Use (a)(iv) to show that if there exists $M \in \mathbb{R}^+$ such that $|f(x)| \leq M$ for all $x \in \mathbb{R}$, then f is identically zero.

(c) Let $g : \mathbb{R} \to \mathbb{R}$ be a periodic function with period $T > 0$, and suppose that there exists $a \in \mathbb{R}$ such that $|g(x)| \leq N$ for all $x \in [a, a+T]$. Let n be the largest integer such that $a + nT \leq y$. Prove that $|g(x)| \leq N$ for all $x \in \mathbb{R}$.

(d) Suppose further that g is an additive function. Show that there exists $\alpha \in \mathbb{R}$ such that $g(x) = \alpha x$ for all $x \in \mathbb{R}$.

Solution.

Part (a).

Part (a)(i). Taking $x = y = 0$,

$$f(0) = f(0) + f(0) \implies f(0) = 0.$$

Part (a)(ii). For $n \in \mathbb{N}_0$, we have

$$f(nx) = f(\underbrace{x + x + \cdots + x}_{n \text{ times}}) = \underbrace{f(x) + f(x) + \cdots + f(x)}_{n \text{ times}} = nf(x).$$

Now, observe that

$$0 = f(0) = f(nx + (-nx)) = f(nx) + f(-nx) \implies f(-nx) = -f(nx) = -nf(x).$$

Thus, $f(nx) = nf(x)$ for all $n \in \mathbb{Z}$.

Part (a)(iii). Let b be a non-zero integer. Then

$$f(x) = f\left(\underbrace{\frac{x}{b} + \cdots + \frac{x}{b}}_{b \text{ times}}\right) = \underbrace{f\left(\frac{x}{b}\right) + \cdots + f\left(\frac{x}{b}\right)}_{b \text{ times}} = bf\left(\frac{x}{b}\right).$$

Thus,

$$f\left(\frac{x}{b}\right) = \frac{1}{b}f(x).$$

Let $r \in \mathbb{Q}$. Without loss of generality, write $r = a/b$, where $a, b \in \mathbb{Z}$ and $b \geq 1$. Then

$$f(rx) = f\left(\frac{ax}{b}\right) = af\left(\frac{x}{b}\right) = \frac{a}{b}f(x) = rf(x).$$

Part (a)(iv). Since $\mathbb{Q}$ is dense in $\mathbb{R}$, it follows that we have $f(rx) = rf(x)$ for all $r, x \in \mathbb{R}$.

Part (b). Seeking a contradiction, suppose $f(x) \neq 0$ for some $x \in \mathbb{R}$. Then $f(kx) = kf(x)$ is unbounded, a contradiction. Thus, $f(x)$ must be identically zero.

Part (c). For all $x \in \mathbb{R}$, we have

$$a \leq x - kT < a + T$$

where $k = \lfloor (x-a)/T \rfloor$. Since g has period T, it follows that

$$|g(x)| = |g(x - kT)| \leq N.$$

Part (d). By part (b), it must be that $g(x) \equiv 0$. Thus, $\alpha = 0$.

Analysis 1.2 Differentiation

Tutorial A1.2 Set 1

Problem 1. Let $f(x) = x^{\mathrm{e}}/\mathrm{e}^x$, where $x > 0$. Find the maximum value of $f(x)$ and hence prove that $\mathrm{e}^\pi > \pi^{\mathrm{e}}$.

Solution. Note that $f'(x) = x^{\mathrm{e}}\mathrm{e}^{-x}\left(\mathrm{e}/x - 1\right)$. For stationary points, $f'(x) = 0$. Since $x > 0$, this only occurs when $\mathrm{e}/x - 1$, whence $x = \mathrm{e}$. By the first derivative test, we see that this is a maximum. Thus, the maximum value of $f(x)$ is $f(\mathrm{e}) = 1$.

Note that $f(x)$ is decreasing for $x > \mathrm{e}$. Since $\pi > \mathrm{e}$, it follows that

$$\frac{\pi^{\mathrm{e}}}{\mathrm{e}^\pi} = f(\pi) < f(\mathrm{e}) = 1 \implies \pi^{\mathrm{e}} < \mathrm{e}^\pi.$$

* * * * *

Problem 2. By applying Rolle's Theorem on the function $f(x) = \mathrm{e}^{-x} - \sin x$, show that there is at least one real root of $\mathrm{e}^x \cos x = -1$ between any two real roots of $\mathrm{e}^x \sin x = 1$.

Proof. Let α be a root of $\mathrm{e}^x \sin x = 1$. Then

$$\mathrm{e}^\alpha \sin\alpha = 1 \implies 1 - \mathrm{e}^\alpha \sin\alpha = 0 \implies \mathrm{e}^{-\alpha} - \sin\alpha = 0.$$

Hence, α is also a root of $f(x) = 0$.

Let α_1 and α_2 be two distinct roots of $\mathrm{e}^x \sin x = 1$. Then α_1 and α_2 are also roots of $f(x) = 0$, i.e. $f(\alpha_1) = f(\alpha_2) = 0$. By Rolle's Theorem, it follows that there exists some $\beta \in (\alpha_1, \alpha_2)$ such that

$$f'(\beta) = -\mathrm{e}^{-\beta} - \cos\beta = 0 \implies \cos\beta = -\mathrm{e}^{-b} \implies \mathrm{e}^{-b}\cos\beta = -1.$$

Hence, β is a root of $\mathrm{e}^x \cos x = -1$. Thus, there is at least one real root of $\mathrm{e}^x \cos x = -1$ (given by β) between any two real roots of $\mathrm{e}^x \sin x = 1$ (given by α_1 and α_2). $\square$

* * * * *

Problem 3. By using the Theorem of the Mean, show that

$$\frac{\pi}{6} + \frac{\sqrt{3}}{15} < \arcsin 0.6 < \frac{\pi}{6} + \frac{1}{8}.$$

Proof. Let $f(x) = \arcsin x$. By the Theorem of the Mean, there exists some $c \in (0.5, 0.6)$ such that

$$f'(c) = \frac{f(0.6) - f(0.5)}{0.6 - 0.5} \implies \frac{1}{\sqrt{1 - c^2}} = \frac{\arcsin 0.6 - \pi/6}{0.1}.$$

Observe that $1/\sqrt{1 - x^2}$ is increasing on $(0.5, 0.6)$. Hence,

$$\frac{2\sqrt{3}}{3} = \frac{1}{\sqrt{1 - 0.5^2}} < \frac{1}{\sqrt{1 - c^2}} < \frac{1}{\sqrt{1 - 0.6^2}} = \frac{5}{4}.$$

Thus,

$$\frac{2\sqrt{3}}{3} < \frac{\arcsin 0.6 - \pi/6}{0.1} < \frac{5}{4} \implies \frac{\pi}{6} + \frac{\sqrt{3}}{15} < \arcsin 0.6 < \frac{\pi}{6} + \frac{1}{8}.$$

$\square$

Problem 4. Show that the line with gradient m through the point $(t, 2t)$ is a tangent to the curve $xy = k^2$, where $k \neq 0$, if $t^2(2-m)^2 + 4mk^2 = 0$. Hence, or otherwise, investigate how the number of tangents from the point $(t, 2t)$ to the curve $xy = k^2$ varies as t varies.

Solution. Note that $y = k^2/x$. Hence, $y' = -k^2/x^2$. The tangent to $y = k^2/x$ at $(x_0, k^2/x_0)$ is thus given by

$$y - \frac{k^2}{x_0} = -\frac{k^2}{x_0^2}(x - x_0) \implies y = -\frac{k^2}{x_0^2}x + \frac{2k^2}{x_0}.$$

Meanwhile, the line l with gradient m passing through $(t, 2t)$ has equation $y = mx + (2-m)t$. Thus, for l to be tangent to $y = k^2/x$, we require

$$-\frac{k^2}{x_0^2} = m \quad \text{and} \quad \frac{2k^2}{x_0} = (2-m)t.$$

Eliminating x_0, we see that

$$\frac{(2-m)^2t^2}{m} = \frac{4k^4/x_0^2}{-k^2/x_0^2} = -4k^2 \implies t^2(2-m)^2 + 4mk^2 = 0.$$

Expanding this quadratic in m, we have

$$t^2m^2 + \left(4k^2 - 4t^2\right)m + 4t^2 = 0. \tag{1}$$

Notice that the number of tangents from the point $(t, 2t)$ to the curve $xy = k^2$ is exactly the number of possible m's, which is determined by the discriminant of the quadratic in (1). One can easily calculate this discriminant as

$$\Delta = \left(4k^2 - 4t^2\right) = 4\left(t^2\right)\left(4t^2\right) = 16k^2\left(k^2 - 2t^2\right).$$

We now examine the number of tangents case by case. Without loss of generality, suppose $k > 0$.

Case 1. Suppose there are two tangents. Then $D > 0 \implies k^2 - 2t^2 > 0$. Thus, $-\frac{1}{\sqrt{2}}k < t < \frac{1}{\sqrt{2}}k$.

Case 2. Suppose there is one tangent. Then $D = 0 \implies k^2 - 2t^2 = 0$. Thus, $t = \pm\frac{1}{\sqrt{2}}k$.

Case 3. Suppose there are no tangents. Then $D < 0 \implies k^2 - 2t^2 < 0$. Thus, $t < -\frac{1}{\sqrt{2}}k$ or $t > \frac{1}{\sqrt{2}}k$.

* * * * *

Problem 5.

(a) Let $f(t) = \mathrm{e}^t/t$. Show that the minimum value of $f(t)$ occurs when $t = 1$.

(b) With the aid of a diagram, or otherwise, prove that if $\mathrm{e}^x/x = \mathrm{e}^y/y$, where $y > x > 0$, then $xy < 1$.

Solution.

Part (a). Note that $f'(t) = (t-1)\mathrm{e}^t/t^2$. For stationary points, $f'(t) = 0$, which can only occur when $t = 1$. By the first derivative test, we see that this is a minimum. Hence, the minimum value of $f(t)$ is $f(1) = \mathrm{e}$.

Part (b). Taking logarithms on both sides, we get

$$x - \ln x = y - \ln y \implies x - y = \ln\frac{x}{y}.$$

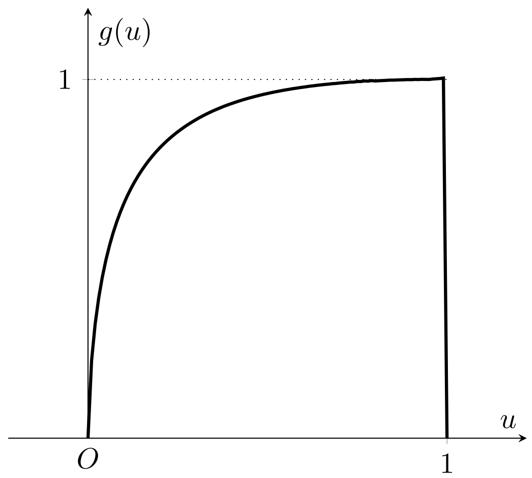
Let $u = x/y$. Since $0 < x < y$, it follows that $0 < u < 1$. Substituting this into our equation yields

$$uy - y = \ln u \implies y = \frac{\ln u}{u-1} \implies x = \frac{u \ln u}{u-1}.$$

Thus,

$$xy = \frac{u \ln^2 u}{(u-1)^2}.$$

Let the RHS be $g(u)$.



From the above graph, it is clear that the maximum value of $g(u)$ is 1, whence $xy < 1$.

* * * * *

Problem 6. By differentiating the series $(1+x)^n$ with respect to x, where n is an integer greater than 1, show that

$$\sum_{r=1}^{n} r \binom{n}{r} = n2^{n-1}.$$

Find a similar expression for the sum

$$\sum_{r=1}^{n} r(r-1) \binom{n}{r}.$$

Hence, or otherwise, show that

$$\sum_{r=1}^{n} r^2 \binom{n}{r} = n(n+1)2^{n-2}.$$

Solution. Note that

$$(1+x)^n = \sum_{r=1}^{n} \binom{n}{r} x^r.$$

Differentiating with respect to x, we have

$$n(1+x)^{n-1} = \sum_{r=1}^{n} r \binom{n}{r} x^{r-1}. \tag{1}$$

Taking $x = 1$, we see that

$$\sum_{r=1}^{n} r \binom{n}{r} = n2^{n-1}.$$

Differentiating (1) once more, we have

$$n(n-1)(1+x)^{n-2} = \sum_{r=1}^{n} r(r-1)\binom{n}{r}x^{r-2}.$$

Taking $x = 1$,

$$\sum_{r=1}^{n} r(r-1)\binom{n}{r} = n(n-1)2^{n-2}.$$

Thus,

$$\begin{aligned}\sum_{r=1}^{n} r^2\binom{n}{r} &= \sum_{r=1}^{n} [r(r-1)+r]\binom{n}{r} = n2^{n-1} = n(n-1)2^{n-2} + n2^{n-1}\\ &= n(n-1)2^{n-2} + n(2)2^{n-2} = n(n+1)2^{n-2}.\end{aligned}$$

Tutorial A1.2 Set 2

Problem 1. Let f be a differentiable function on $(0, \infty)$ and suppose that

$$\lim_{x\to\infty} \big(f(x) + f'(x)\big) = L.$$

By considering $f(x) = \frac{\mathrm{e}^x f(x)}{\mathrm{e}^x}$, show that $\lim_{x\to\infty} f(x) = L$ and $\lim_{x\to\infty} f'(x) = 0$.

Proof. By L'Hôpital's Rule,

$$\lim_{x\to\infty} \frac{\mathrm{e}^x f(x)}{\mathrm{e}^x} = \lim_{x\to\infty} \frac{\mathrm{e}^x f'(x) + \mathrm{e}^x f(x)}{\mathrm{e}^x} = \lim_{x\to\infty} \big(f(x) + f'(x)\big) = L.$$

Hence,

$$\lim_{x\to\infty} f'(x) = \lim_{x\to\infty} \big(f(x) + f'(x)\big) - \lim_{x\to\infty} f(x) = L - L = 0.$$

□

* * * * *

Problem 2. It is given that the functions $f(x)$ and $g(x)$ are non-constant, differentiable functions that are defined on $\mathbb{R}$, and $f(0) = f(2) = g(0) = g(2) = 0$. Suppose that $g''(x) \neq 0$ for all $x \in (0, 2)$.

(a) Show that $g(x) \neq 0$ for all $x \in (0, 2)$.

(b) Using part (a), or otherwise, prove that there exists $d \in (0, 2)$ such that

$$\frac{f(d)}{g(d)} = \frac{f''(d)}{g''(d)}.$$

Solution.

Part (a). By way of contradiction, suppose $g(x_0) = 0$ for some $x \in (0, 2)$. Then $g(0) = g(x_0) = g(2) = 0$. By Rolle's Theorem, there exists some $x_1 \in (0, x_0)$ and $x_2 \in (x_0, 2)$ such that

$$g'(x_1) = 0 \quad \text{and} \quad g'(x_2) = 0.$$

Applying Rolle's Theorem once more, we see that there exists some $x_3 \in (x_1, x_2)$ such that $g''(x_3) = 0$, a contradiction. Thus, $g(x) \neq 0$ for all $x \in (0, 2)$.

Part (b). Let

$$h(x) = f(x)g''(x) - g(x)f''(x).$$

Clearly, $h(x)$ is continuous on $[0, 2]$. Note also that $h(0) = h(2) = 0$. We now integrate $h(x)$ over $(0, 2)$:

$$\int_0^2 h(x)\,\mathrm{d}x = \int_0^2 f(x)g''(x)\,\mathrm{d}x - \int_0^2 g(x)f''(x)\,\mathrm{d}x.$$

Integrating by parts, the first integral reduces to

$$\int_0^2 f(x)g''(x)\,\mathrm{d}x = \big[f(x)g'(x)\big]_0^2 - \int_0^2 f'(x)g'(x)\,\mathrm{d}x = -\int_0^2 f'(x)g'(x)\,\mathrm{d}x.$$

Similarly, the second integral reduces to

$$\int_0^2 g(x)f''(x)\,\mathrm{d}x = \big[g(x)f'(x)\big]_0^2 - \int_0^2 f'(x)g'(x)\,\mathrm{d}x = \int_0^2 f'(x)g'(x)\,\mathrm{d}x.$$

Thus,

$$\int_0^2 h(x)\,\mathrm{d}x = -\int_0^2 f'(x)g'(x)\,\mathrm{d}x + \int_0^2 f'(x)g'(x)\,\mathrm{d}x = 0. \tag{1}$$

If $h(x)$ is not identically zero, then from (1), it follows that $h(x)$ attains positive and negative values in $(0,2)$. Thus, by the Intermediate Value Theorem, there exists some $d \in (0,2)$ such that $h(d) = 0$. If $h(x)$ is identically zero, then $h(d) = 0$ for all $d \in (0,2)$. In any case, we get

$$\frac{f(d)}{g(d)} = \frac{f''(d)}{g''(d)}$$

upon rearrangement.

* * * * *

Problem 3. Let $f : [0,1] \to \mathbb{R}$ be a differentiable function defined on $[0,1]$. Suppose that

$$f'(c) = f''(c) = f'''(c) = 0$$

and $f^{(4)}(c) > 0$ for some $c \in (0,1)$.

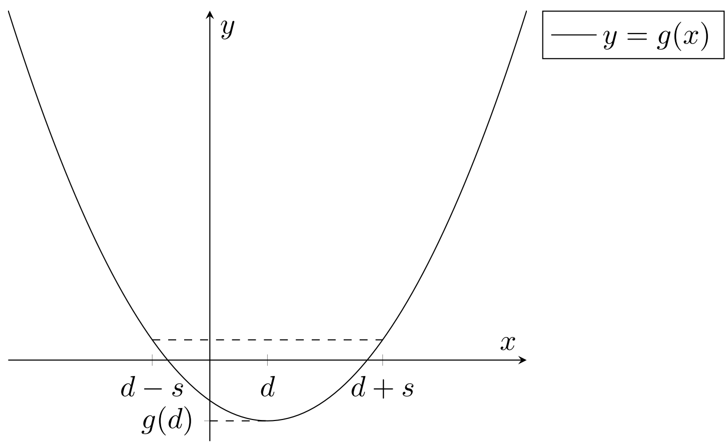
(a) Let $g : [a,b] \to \mathbb{R}$ be a function that is twice differentiable on (a,b), where $a < b$. Suppose that there exists $d \in (a,b)$ such that $g'(d) = 0$ and $g''(d) > 0$. With the aid of a diagram, explain why there exists $s > 0$ such that for all $x \in [d-s, d+s] \setminus \{d\} \subseteq [a,b]$, $g(x) > g(d)$.

(b) Hence, show that f attains a minimum point at $x = c$.

(c) Write down a similar result for f to attain a maximum point when $x = c$.

Solution.

Part (a).



From the above diagram, it is immediately clear that there exists some $s > 0$ such that for all $x \in [d-s, d+s] \setminus \{d\} \subseteq [a,b]$, $g(x) > g(d)$.

Part (b). Take $g(x) = f''(x)$. Since $g'(c) = f'''(c) = 0$ and $g''(c) = f^{(4)}(c) > 0$, by (a), there exists $s > 0$ such that $f''(x) > f''(c)$ for all $x \in [c-s, c+s] \setminus \{c\} \subseteq [0,1]$. This means that $f(c)$ is a minimum.

Part (c). f attains a maximum point when $x = c$ if $f^{(4)}(c) < 0$.

Problem 4. Let $f : [0,1] \to \mathbb{R}$ be a differentiable function such that $f(0) = 0$, $f(1) = 1$. Prove that there exists $c, d \in (0,1)$ such that $c \neq d$ and

$$\frac{f'(c)}{c^2} + \frac{f'(d)}{d} = 5.$$

Proof. Consider the function $g(x) = f(x) - x^3$. Since $g(0) = 0$ and $g(1) = 0$, by Rolle's Theorem, there exists a $c \in (0,1)$ such that

$$g'(c) = 0 \implies f'(c) - 3c^2 = 0 \implies \frac{f'(c)}{c^2} = 3.$$

Similarly, consider the function $h(x) = f(x) - x^2$. Since $h(0) = h(1) = 0$, by Rolle's Theorem, there exists a $d \in (0,1)$ such that

$$h'(d) = 0 \implies f'(d) - 2d = 0 \implies \frac{f'(d)}{d} = 2.$$

Thus, there exists distinct $c, d \in (0,1)$ such that

$$\frac{f'(c)}{c^2} + \frac{f'(d)}{d} = 5.$$

$\square$

Analysis 1.3 Integration

Tutorial A1.3 Set 1

Problem 1.

(a) Let

$$I = \int_0^a \frac{f(x)}{f(x)+f(a-x)}\,\mathrm{d}x.$$

Use a substitution to show that

$$I = \int_0^a \frac{f(a-x)}{f(x)+f(a-x)}\,\mathrm{d}x$$

and hence evaluate I in terms of a. Use this result to evaluate the integrals

$$\int_0^1 \frac{\ln(x+1)}{\ln(2+x-x^2)}\,\mathrm{d}x \quad \text{and} \quad \int_0^{\pi/2} \frac{\sin x}{\sin(x+\pi/4)}\,\mathrm{d}x.$$

(b) Using a suitable substitution, evaluate

$$\int_{1/2}^2 \frac{\sin x}{x\,(\sin x + \sin 1/x)}\,\mathrm{d}x.$$

Solution.

Part (a). Under the substitution $u = a - x$, we have

$$I = \int_0^a \frac{f(x)}{f(x)+f(a-x)}\,\mathrm{d}x = -\int_a^0 \frac{f(a-u)}{f(a-u)+f(u)}\,\mathrm{d}u = \int_0^a \frac{f(a-u)}{f(a-u)+f(u)}\,\mathrm{d}u.$$

Renaming the dummy variable back to x, we have

$$I = \int_0^a \frac{f(a-x)}{f(x)+f(a-x)}\,\mathrm{d}x$$

as desired.

Observe that

$$2I = \int_0^a \frac{f(x)}{f(x)+f(a-x)}\,\mathrm{d}x + \int_0^a \frac{f(a-x)}{f(x)+f(a-x)}\,\mathrm{d}x = \int_0^a \mathrm{d}x = a,$$

so $I = a/2$.

We have

$$\int_0^1 \frac{\ln(x+1)}{\ln(2+x-x^2)}\,\mathrm{d}x = \int_0^1 \frac{\ln(x+1)}{\ln((x+1)(2-x))}\,\mathrm{d}x = \int_0^1 \frac{\ln(x+1)}{\ln(x+1)+\ln(2-x)}\,\mathrm{d}x.$$

Let $f(x) = \ln(1+x)$. Then

$$\int_0^1 \frac{\ln(x+1)}{\ln(2+x-x^2)}\,\mathrm{d}x = \int_0^1 \frac{f(x)}{f(x)+f(1-x)}\,\mathrm{d}x = \frac{1}{2}.$$

We have

$$\int_0^{\pi/2} \frac{\sin x}{\sin(x+\pi/4)}\,\mathrm{d}x = \frac{2}{\sqrt{2}}\int_0^{\pi/2} \frac{\sin x}{\sin x + \cos x} = \frac{2}{\sqrt{2}}\int_0^{\pi/2} \frac{\sin x}{\sin x + \sin(\pi/2 - x)}.$$

Let $f(x) = \sin x$. Then

$$\int_0^{\pi/2} \frac{\sin x}{\sin(x+\pi/4)}\,\mathrm{d}x = \frac{2}{\sqrt{2}}\frac{f(x)}{f(x)+f(\pi/2-x)} = \frac{2}{\sqrt{2}}\cdot\frac{\pi/2}{2} = \frac{\pi}{2\sqrt{2}}.$$

Part (b). Let

$$I = \int_{1/2}^{2} \frac{\sin x}{x\,(\sin x + \sin 1/x)}\,\mathrm{d}x. \tag{1}$$

Under the substitution $x \mapsto 1/x$, we have

$$I = \int_2^{1/2} \frac{\sin 1/x}{(1/x)\,(\sin 1/x + \sin x)}\left(-\frac{1}{x^2}\right)\mathrm{d}x = \int_{1/2}^{2} \frac{\sin 1/x}{x\,(\sin x + \sin 1/x)}\,\mathrm{d}x. \tag{2}$$

Adding (1) and (2) together,

$$\begin{aligned} 2I &= \int_{1/2}^{2} \frac{\sin x}{x\,(\sin x + \sin 1/x)}\,\mathrm{d}x + \int_{1/2}^{2} \frac{\sin 1/x}{x\,(\sin x + \sin 1/x)}\,\mathrm{d}x \\ &= \int_{1/2}^{2} \frac{1}{x}\,\mathrm{d}x = [\ln x]_{1/2}^{2} = 2\ln 2. \end{aligned}$$

Thus, $I = \ln 2$.

* * * * *

Problem 2.

(a) Show that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\sqrt{2}}\arctan\frac{\tan x}{\sqrt{2}}\right) = \frac{1}{1+\cos^2 x}.$$

(b) Use (a) to show that

$$\int_0^{\pi} \frac{x}{1+\cos^2 x}\,\mathrm{d}x = \frac{\pi^2}{2\sqrt{2}}.$$

Solution.

Part (a). We have

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\sqrt{2}}\arctan\frac{\tan x}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}\left(\frac{1}{1+(\tan x/\sqrt{2})^2}\right)\left(\frac{1}{\sqrt{2}}\sec^2 x\right) \\ &= \frac{1}{2}\left(\frac{1}{1+\tan^2(x)/2}\right)\left(\frac{1}{\cos^2 x}\right) = \frac{1}{2+\tan^2 x}\frac{1}{\cos^2 x} = \frac{1}{2\cos^2 x + \sin^2 x} = \frac{1}{1+\cos^2 x}. \end{aligned}$$

Part (b). Let the target integral be I. Note that

$$I = \int_0^{\pi} \frac{x}{1+\cos^2 x}\,\mathrm{d}x = \int_0^{\pi/2} \frac{x}{1+\cos^2 x}\,\mathrm{d}x + \int_{\pi/2}^{\pi} \frac{x}{1+\cos^2 x}\,\mathrm{d}x.$$

Applying the transformation $x \mapsto \pi - x$ to the second integral, we get

$$\begin{aligned} I &= \int_0^{\pi/2} \frac{x}{1+\cos^2 x}\,\mathrm{d}x + \left(\pi\int_0^{\pi/2} \frac{1}{1+\cos^2 x}\,\mathrm{d}x - \int_0^{\pi/2} \frac{x}{1+\cos^2 x}\,\mathrm{d}x\right) \\ &= \pi\int_0^{\pi/2} \frac{1}{1+\cos^2 x}\,\mathrm{d}x = \pi\left[\frac{1}{\sqrt{2}}\arctan\frac{\tan x}{\sqrt{2}}\right]_0^{\pi/2}. \end{aligned}$$

The $x = 0$ term vanishes, so

$$I = \frac{\pi}{\sqrt{2}} \lim_{x \to \frac{\pi}{2}^-} \arctan \frac{\tan x}{\sqrt{2}}.$$

As $x \to (\pi/2)^-$, $\tan x \to \infty$. Thus, $\arctan \frac{\tan x}{\sqrt{2}} \to \frac{\pi}{2}$, whence

$$I = \frac{\pi}{\sqrt{2}} \cdot \frac{\pi}{2} = \frac{\pi^2}{2\sqrt{2}}.$$

* * * * *

Problem 3. (In this question all indices n are integers)
Let

$$I_n = \int_0^{\pi/2} \sin^n x \, \mathrm{d}x, \quad n \geq 0.$$

(a) Show that

$$I_n = \frac{n-1}{n} I_{n-2}, \quad n \geq 2.$$

(b) Use (a) to show that

$$I_{2n+1} = \frac{2 \cdot 4 \cdot 6 \cdots (2n)}{3 \cdot 5 \cdot 7 \cdots (2n+1)}, \quad n \geq 1.$$

(c) Use (a) to show that

$$I_{2n} = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \frac{\pi}{2}, \quad n \geq 1.$$

(d) Use (a) or (c) to show that

$$\frac{I_{2n+2}}{I_{2n}} = \frac{2n+1}{2n+2}, \quad n \geq 1.$$

(e) By considering the integral and comparing $\sin^{k+1} x$ with $\sin^k x$, show that $I_{2n+2} \leq I_{2n+1} \leq I_{2n}$, $n \geq 1$.

(f) Use (d) and (e) to show that

$$\frac{2n+1}{2n+2} \leq \frac{I_{2n+1}}{I_{2n}}, \quad n \geq 1.$$

Hence, deduce that

$$\lim_{n \to \infty} \frac{I_{2n+1}}{I_{2n}} = 1.$$

(g) Use (b), (c) and (f) to show that

$$\frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdots = \frac{\pi}{2}.$$

Solution.

Part (a). Note that

$$I_n = \int_0^{\pi/2} \sin^2 x \sin^{n-2} x \, \mathrm{d}x = \int_0^{\pi/2} \left(1-\cos^2 x\right) \sin^{n-2}\, \mathrm{d}x$$
$$= I_{n-2} - \int_0^{\pi/2} \cos x \left(\cos x \sin^{n-2} x\right)\, \mathrm{d}x.$$

Integrating by parts, we obtain

$$I_n = I_{n-2} - \left(\left[\frac{\cos x \sin^{n-1} x}{n-1} \right]_0^{\pi/2} + \frac{1}{n-1} \int_0^{\pi/2} \sin^n x \, \mathrm{d}x \right).$$

We thus have

$$I_n = I_{n-2} - \frac{1}{n-1} I_n \implies I_n = \frac{n-1}{n} I_{n-2}.$$

Part (b). Observe that

$$I_{2n+1} = \frac{2n}{2n+1} I_{2n-1} = \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} I_{2n-3} = \cdots = \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} \cdots \frac{2}{3} I_1.$$

Since

$$I_1 = \int_0^{\pi/2} \sin x \, \mathrm{d}x = [-\cos x]_0^{\pi/2} = 1,$$

we conclude that

$$I_{2n+1} = \frac{2 \cdot 4 \cdot 6 \cdots (2n)}{3 \cdot 5 \cdot 7 \cdots (2n+1)}.$$

Part (c). Observe that

$$I_{2n} = \frac{2n-1}{2n} I_{2n-2} = \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} I_{2n-4} = \cdots = \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} \cdots \frac{1}{2} I_0.$$

Since

$$I_0 = \int_0^{\pi/2} \sin^0 x \, \mathrm{d}x = \frac{\pi}{2},$$

we conclude that

$$I_{2n} = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \frac{\pi}{2}.$$

Part (d). We have

$$I_{2n+2} = \frac{2n+1}{2n+2} I_{2n} \implies \frac{I_{2n+2}}{I_n} = \frac{2n+1}{2n+2}.$$

Part (e). Since $|\sin x| \leq 1$, it follows that $\sin^n x \geq \sin^{n+1} x$ for all real x. This in turn implies that $I_{n+1} \leq I_n$. Thus,

$$I_{2n+2} \leq I_{2n+1} \leq I_{2n}.$$

Part (f). Dividing the inequality in (e) throughout by I_{2n}, we have

$$\frac{2n+1}{2n+2} \leq \frac{I_{2n+2}}{I_{2n}} = \frac{I_{2n+1}}{I_{2n}} \leq \frac{I_{2n}}{I_{2n}} = 1.$$

Taking the limit as $n \to \infty$, we see that

$$1 = \lim_{n\to\infty} \frac{2n+1}{2n+2} \leq \lim_{n\to\infty} \frac{I_{2n+1}}{I_{2n}} \leq 1.$$

Thus, by the Squeeze Theorem, it follows that

$$\lim_{n\to\infty} \frac{I_{2n+1}}{I_{2n}} = 1.$$

Part (g). The above limit implies that

$$\lim_{n\to\infty} I_{2n+1} = \lim_{n\to\infty} I_{2n}.$$

Thus,

$$\frac{2\cdot 4\cdot 6\cdots}{3\cdot 5\cdot 7\cdots} = \frac{1\cdot 3\cdot 5\cdots}{2\cdot 4\cdot 6\cdots}\frac{\pi}{2}.$$

Rearranging, we recover the Wallis product:

$$\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots.$$

* * * * *

Problem 4. It is given that the following integrals converge. Evaluate the following integrals.

(a) $\displaystyle\int_1^\infty \frac{3x-1}{4x^3-x^2}\,\mathrm{d}x.$

(b) $\displaystyle\int_0^\infty x^2\mathrm{e}^{-x}\,\mathrm{d}x.$

(c) $\displaystyle\int_{-\infty}^\infty \frac{1}{4x^2+9}\,\mathrm{d}x.$

Solution.

Part (a). Let I be the target integral. We have

$$\begin{aligned}I &= \int_1^\infty \frac{3x-1}{4x^3-x^2}\,\mathrm{d}x = \int_1^\infty \frac{3x-1}{x^2(4x-1)}\,\mathrm{d}x = \int_1^\infty\left(\frac{4x-1}{x^2(4x-1)} - \frac{x}{x^2(4x-1)}\right)\mathrm{d}x\\ &= \int_1^\infty \frac{1}{x^2}\,\mathrm{d}x - \int_1^\infty \frac{1}{x(4x-1)}\,\mathrm{d}x = \left[-\frac{1}{x}\right]_1^\infty - \int_1^\infty \frac{1}{x(4x-1)}\,\mathrm{d}x\\ &= 1 - \int_1^\infty \frac{1}{x(4x-1)}\,\mathrm{d}x = 1 - \int_1^\infty \frac{1}{(2x-1/4)^2-(1/4)^2}\,\mathrm{d}x.\end{aligned}$$

Under the substitution $u = 2x - 1/4$, the integral evaluates to

$$I = 1 - \frac{1}{2}\int_{7/4}^\infty \frac{1}{u^2-(1/4)^2}\,\mathrm{d}u = 1 - \frac{1}{2}\left[\frac{1}{2(1/4)}\ln\left(\frac{u-1/4}{u+1/4}\right)\right]_{7/4}^\infty = 1 - \ln\frac{4}{3}.$$

Part (b). Integrating by parts, we have

$$\int_0^\infty x^2\mathrm{e}^{-x}\,\mathrm{d}x = \left[-\mathrm{e}^{-x}\left(x^2+2x+2\right)\right]_0^\infty = 2.$$

Part (c). Let I be the target integral. We have

$$I = \int_{-\infty}^\infty \frac{1}{4x^2+9}\,\mathrm{d}x = \int_0^\infty \frac{2}{(2x)^2+3^2}\,\mathrm{d}x.$$

Under the substitution $u = 2x$, the integral evaluates to

$$I = \int_0^\infty \frac{1}{u^2+3^2}\,\mathrm{d}u = \left[\frac{1}{3}\arctan\frac{u}{3}\right]_0^\infty = \frac{\pi}{6}.$$

Problem 5. Determine which of the following integrals converge.

(a) $\displaystyle\int_0^\infty \frac{\mathrm{d}x}{\sqrt{x^2+1}}$, (b) $\displaystyle\int_{-\infty}^\infty \frac{2\,\mathrm{d}x}{\sqrt{\mathrm{e}^x+\mathrm{e}^{-x}}}$.

Solution.

Part (a). Clearly, the integral $\int_0^\infty \frac{1}{x}\,\mathrm{d}x$ diverges. Since

$$\lim_{x\to\infty} \frac{1/\sqrt{x^2+1}}{1/x} = 1,$$

by the limit comparison test, it follows that $\int_0^\infty \frac{\mathrm{d}x}{\sqrt{x^2+1}}$ also diverges.

Part (b). Note that the integral

$$\int_0^\infty \mathrm{e}^{-x/2}\,\mathrm{d}x = -2\left[\mathrm{e}^{-x/2}\right]_0^\infty = 2$$

converges. Since

$$\lim_{x\to\infty} \frac{4/\sqrt{e^x+e^{-x}}}{\mathrm{e}^{-x/2}} = \lim_{x\to\infty} \frac{4}{\sqrt{1+\mathrm{e}^{-2x}}} = 4,$$

by the limit comparison test, it follows that

$$\int_{-\infty}^\infty \frac{2\,\mathrm{d}x}{\sqrt{\mathrm{e}^x+\mathrm{e}^{-x}}} = \int_0^\infty \frac{4\,\mathrm{d}x}{\sqrt{\mathrm{e}^x+\mathrm{e}^{-x}}}$$

converges.

Tutorial A1.3 Set 2

Problem 1. Let $I = \int_0^{2\pi} \frac{1}{2-\cos x} \,\mathrm{d}x$. Explain the error in the following argument:

> *Since $|\cos x| \leq 1$, it follows that $1/(2-\cos x) > 0$, and, interpreting the integral as an area, it follows that I is positive. However, putting $t = \tan(x/2)$,*
>
> $$I = \int_{\tan 0}^{\tan \pi} \frac{\frac{2}{1+t^2}}{2 - \frac{1-t^2}{1+t^2}} \,\mathrm{d}t = 2\int_0^0 \frac{1}{1+3t^2} \,\mathrm{d}t = 0.$$
>
> *Thus, the positive number I is equal to 0.*

Prove that $I = \int_0^{\pi} \frac{1}{2-\cos x} \,\mathrm{d}x$, and deduce that $I = \frac{2\pi\sqrt{3}}{3}$.

Solution. $t = \tan(x/2)$ is discontinuous at $x = \pi$. Hence, direct substitution is not allowed.

Splitting I, we have

$$I = \int_0^{2\pi} \frac{1}{2-\cos x} \,\mathrm{d}x = \int_0^{\pi} \frac{1}{2-\cos x} \,\mathrm{d}x + \int_{\pi}^{2\pi} \frac{1}{2-\cos x} \,\mathrm{d}x.$$

Applying the substitution $x \mapsto 2\pi - x$ on the latter integral,

$$\begin{aligned} I &= \int_0^{\pi} \frac{1}{2-\cos x} \,\mathrm{d}x + \int_{\pi}^{0} \frac{1}{2-\cos(2\pi - x)} (-\,\mathrm{d}x) \\ &= \int_0^{\pi} \frac{1}{2-\cos x} \,\mathrm{d}x + \int_0^{\pi} \frac{1}{2-\cos x} \,\mathrm{d}x = 2\int_0^{\pi} \frac{1}{2-\cos x} \,\mathrm{d}x. \end{aligned}$$

Using the substitution $t = \tan(x/2)$, we have

$$\begin{aligned} I &= 2\int_0^{\pi} \frac{1}{2-\cos x} \,\mathrm{d}x = 2\int_0^{\infty} \frac{\frac{2}{1+t^2}}{2 - \frac{1-t^2}{1+t^2}} \,\mathrm{d}t = 4\int_0^{\infty} \frac{1}{1+3t^2} \,\mathrm{d}t \\ &= 4\left[\frac{1}{\sqrt{3}} \arctan\left(\sqrt{3}x\right)\right]_0^{\infty} = \frac{4}{\sqrt{3}} \cdot \frac{\pi}{2} = \frac{2\pi\sqrt{3}}{3}. \end{aligned}$$

* * * * *

Problem 2. Without using G.C., evaluate

(a) $\displaystyle\int_0^{\pi/2} \sin x \cos 2x \sin 3x \,\mathrm{d}x$,

(b) $\displaystyle\int_1^2 \frac{1}{x+\sqrt{x}} \,\mathrm{d}x$.

Solution.

Part (a). Note that

$$\begin{aligned} &\sin x \cos 2x \sin 3x = \cos 2x \left(\frac{\cos(3x - x) - \cos(3x + x)}{2}\right) = \frac{1}{2}\cos^2 2x - \frac{1}{2}\cos 2x \cos 4x \\ &= \frac{1}{2}\left(\frac{1+\cos 4x}{2}\right) - \frac{1}{2}\cos 2x \left(1 - 2\sin^2 2x\right) = \frac{1}{4} + \frac{\cos 4x}{4} - \frac{\cos 2x}{2} + \cos 2x \sin^2 2x. \end{aligned}$$

Hence,

$$\begin{aligned} I &= \int_0^{\pi/2} \left(\frac{1}{4} + \frac{\cos 4x}{4} - \frac{\cos 2x}{2} + \cos 2x \sin^2 2x \right) \mathrm{d}x \\ &= \left[\frac{x}{4} + \frac{\sin 4x}{16} - \frac{\sin 2x}{4} + \frac{\sin^3 2x}{6} \right]_0^{\pi/2} = \frac{\pi}{8}. \end{aligned}$$

Part (b). Consider the substitution $u = 1 + \sqrt{x}$.

$$I = \int_1^2 \frac{1}{\sqrt{x}\,(1+\sqrt{x})} \,\mathrm{d}x = \int_1^{1+\sqrt{2}} \frac{2\,\mathrm{d}u}{u} = 2\,[\ln u]_2^{1+\sqrt{2}} = 2 \ln \frac{1+\sqrt{2}}{2}.$$

* * * * *

Problem 3. Let $f : [0,1] \to \mathbb{R}$ be a continuous and twice differentiable function on (a,b). Suppose that

$$\int_0^1 f(x)\,\mathrm{d}x = f(0) = f(1).$$

(a) Let $G(x) = \int_0^x f(t)\,\mathrm{d}t$. Explain why $G'(x) = f(x)$.

(b) Let $F(x) = \int_0^x [f(t) - f(1)]\,\mathrm{d}t$. Show that there exists $c \in (0,1)$ such that $f(c) = f(1)$.

(c) Hence, show that there exists $d \in (0,1)$ such that $f''(d) = 0$.

Solution.

Part (a). By the Fundamental Theorem of Calculus,

$$G'(x) = \frac{\mathrm{d}}{\mathrm{d}x} \int_0^x f(t)\,\mathrm{d}t = f(x).$$

Part (b). Note that $F(x) = G(x) - xf(1)$. Notice further that

$$F(0) = \int_0^0 [f(t) - f(1)]\,\mathrm{d}t = 0 \quad \text{and} \quad F(1) = \int_0^1 f(t)\,\mathrm{d}t - f(1) = f(1) - f(1) = 0.$$

Since

$$F'(x) = G'(x) - f(1) = f(x) - f(1),$$

by Rolle's Theorem, there exists some $c \in (0,1)$ such that

$$F'(c) = 0 \implies f(c) - f(1) = 0 \implies f(c) = f(1).$$

Part (c). By the Mean Value Theorem, there exists some $c_1 \in (0,1)$ such that

$$f'(c_1) = \frac{f(1) - f(0)}{1 - 0} = 0.$$

Likewise, there exists some $c_2 \in (0,1)$ such that

$$f'(c_2) = \frac{f(1) - f(c)}{1 - c} = 0.$$

Invoking Rolle's Theorem on $f'(x)$, it follows that there exists some $d \in (c_1, c_2) \subseteq (0,1)$ such that

$$f''(d) = 0.$$

Problem 4.

(a) Let

$$I_n = \int \frac{1}{x^n\left(ax^2+bx+c\right)} \,\mathrm{d}x$$

for $x > 0$. Show that

$$aI_{n-2} + bI_{n-1} + cI_n = \frac{1}{(1-n)\,x^{n-1}} + k$$

for $n \geq 2$, where k is an arbitrary constant.

(b) Hence, find

$$\int \frac{x^3+2}{x^2\left(x^2+2\right)} \,\mathrm{d}x.$$

(c) Let $f(x)$, $g(x)$ and $h(x)$ be functions of x such that $f'(x) = x^2 g'(x)$ and $h'(x) = \left(1+x^4\right) g'(x)g(x)$. Use integration by parts to find the following:

(a) $\displaystyle\int xg(x)\,\mathrm{d}x$,

(b) $\displaystyle\int xf(x)g(x)\,\mathrm{d}x$.

Solution.

Part (a). We have

$$\begin{aligned}
&aI_{n-2} + bI_{n-1} + cI_n \\
&\quad= \int \left(\frac{a}{x^{n-2}\left(ax^2+bx+c\right)} + \frac{b}{x^{n-1}\left(ax^2+bx+c\right)} + \frac{c}{x^n\left(ax^2+bx+c\right)}\right) \mathrm{d}x \\
&\quad= \int \left(\frac{ax^2}{x^n\left(ax^2+bx+c\right)} + \frac{bx}{x^n\left(ax^2+bx+c\right)} + \frac{c}{x^n\left(ax^2+bx+c\right)}\right) \mathrm{d}x \\
&\quad= \int \frac{ax^2+bx+c}{x^n\left(ax^2+bx+c\right)} = \int \frac{1}{x^n}\,\mathrm{d}x = \frac{1}{(1-n)\,x^{n-1}} + k.
\end{aligned}$$

Part (b). Let $n=2$, $a=1$, $b=0$ and $c=2$. Using the above result,

$$I_0 + 2I_2 = -\frac{1}{x} + k.$$

Rearranging,

$$\int \frac{2}{x^2\left(x^2+2\right)}\,\mathrm{d}x = 2I_2 = -\frac{1}{x} + k - \underbrace{\int \frac{1}{x^2+2}\,\mathrm{d}x}_{I_0} = -\frac{1}{x} + k - \frac{1}{\sqrt{2}}\arctan\frac{x}{\sqrt{2}}.$$

Thus, the target integral is given by

$$\begin{aligned}
\int \frac{x^3+2}{x^2\left(x^2+2\right)}\,\mathrm{d}x &= \int \frac{x}{x^2+2}\,\mathrm{d}x + \int \frac{2}{x^2\left(x^2+2\right)}\,\mathrm{d}x \\
&= \frac{1}{2}\ln\!\left(x^2+2\right) - \frac{1}{x} + k - \frac{1}{\sqrt{2}}\arctan\frac{x}{\sqrt{2}}.
\end{aligned}$$

Part (c). For brevity, we write $f(x)$ as f, $g(x)$ as g, etc.

Part (c)(i). Integrating by parts, we see that

$$\int xg \, \mathrm{d}x = \frac{1}{2}gx^2 - \frac{1}{2}\int x^2 g' \, \mathrm{d}x = \frac{1}{2}x^2 g - \frac{1}{2}\int f' \, \mathrm{d}x = \frac{x^2 g - f}{2} + C.$$

Part (c)(ii). Using the above result to integrate by parts, we see that

$$\int xfg \, \mathrm{d}x = \frac{x^2 fg - f^2}{2} - \frac{1}{2}\int \left(x^2 f' g - ff'\right) \, \mathrm{d}x.$$

Clearly,

$$\int ff' \, \mathrm{d}x = \frac{1}{2}f^2 + C.$$

Also,

$$\int x^2 f' g \, \mathrm{d}x = \int x^4 g' g \, \mathrm{d}x = \int \left(h' - g'g\right) \, \mathrm{d}x = h - \frac{1}{2}g^2 + C.$$

Putting everything together,

$$\int xfg \, \mathrm{d}x = \frac{x^2 fg - f^2}{2} - \frac{1}{2}\left(h - \frac{1}{2}g^2 - \frac{1}{2}f^2\right) + C = \frac{2x^2 fg - f^2 - 2h + g^2}{4} + C.$$

* * * * *

Problem 5.

(a) Show that if $n > 0$ then

$$\int_n^\infty \frac{1}{x^2 + n^2} \, \mathrm{d}x = \frac{\pi}{4n}.$$

(b) Show that if $0 < a < b$, then

$$\int_b^\infty \frac{1}{x^2 + n^2} \, \mathrm{d}x \leq \int_a^\infty \frac{1}{x^2 + n^2} \, \mathrm{d}x.$$

(c) Hence, deduce that

$$\sum_{n=1}^{N} \frac{1}{n} > \frac{4}{\pi}\int_n^\infty \frac{N}{x^2 + N^2} \, \mathrm{d}x,$$

where N is an integer, $N > 1$.

Solution.

Part (a). We have

$$\int_n^\infty \frac{1}{x^2 + n^2} \, \mathrm{d}x = \left[\frac{1}{n}\arctan\frac{x}{n}\right]_n^\infty = \frac{1}{n}\left(\frac{\pi}{2} - \frac{\pi}{4}\right) = \frac{\pi}{4n}.$$

Part (b). Clearly,

$$\int_b^\infty \frac{1}{x^2 + n^2} \, \mathrm{d}x = \frac{\pi}{4b} < \frac{\pi}{4a} = \int_a^\infty \frac{1}{x^2 + n^2} \, \mathrm{d}x.$$

Part (c). We have

$$\sum_{n=1}^{N} \frac{1}{n} > 1 = \frac{4}{\pi}\left(N \cdot \frac{\pi}{4N}\right) = \frac{4}{\pi}\int_n^\infty \frac{N}{x^2 + N^2} \, \mathrm{d}x.$$

Problem 6. The functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to [-1, \infty)$ are defined by

$$f(x) = \begin{cases} -x, & x \leq 0, \\ -x^2, & x > 0, \end{cases} \quad \text{and} \quad g(x) = \begin{cases} 2014\mathrm{e}^x, & x < 0, \\ x - 1, & x \geq 0. \end{cases}$$

(a) Given that the function f is bijective, find f^{-1} in a similar form.

(b) By expressing $fg(x)$ and $gf(x)$ in a similar form, solve $fg(x) > gf(x)$.

Given the greatest integer function $h : \mathbb{R} \to \mathbb{Z}$ such that $h(x) = \lfloor x \rfloor$,

(c) Show that for $n \in \mathbb{Z}^+$,

$$\int_{-\infty}^{n} hg(x) \, \mathrm{d}x = a \ln b - \ln(b!) + \frac{n(n-c)}{2},$$

where a, b and c are integers to be determined.

Solution.

Part (a). For $x \leq 0$, $f(x) = -x \geq 0$, so $f^{-1} = -x$ for $x \geq 0$. For $x > 0$, $f(x) = -x^2 < 0$, so $f^{-1} = \sqrt{-x}$ for $x < 0$. Thus,

$$f^{-1}(x) = \begin{cases} -x, & x \geq 0, \\ \sqrt{-x}, & x < 0. \end{cases}$$

Part (b). We have

$$fg(x) = f \begin{cases} 2014\mathrm{e}^x, & x < 0 \\ x - 1, & 0 \leq x \leq 1 \\ x - 1, & x > 1, \end{cases} = \begin{cases} -\left(2014\mathrm{e}^x\right)^2, & x < 0, \\ -(x-1), & 0 \leq x \leq 1, \\ -(x-1)^2, & x > 1. \end{cases}$$

Similarly, we have

$$gf(x) = \begin{cases} -x, & x \leq 0, \\ -x^2, & x > 0, \end{cases} = \begin{cases} -x - 1, & x \leq 0, \\ 2014\mathrm{e}^{-x^2}, & x > 0. \end{cases}$$

To solve the inequality $fg(x) > gf(x)$, we consider the following cases:

Case 1: $x < 0$. The inequality simplifies down to

$$-\left(2014\mathrm{e}^x\right)^2 > -x - 1 \implies 2014^2\mathrm{e}^{2x} < x + 1,$$

which is never satisfied for all $x < 0$.

Case 2: $0 \leq x \leq 1$. The inequality simplifies down to

$$-(x-1) > 2014\mathrm{e}^{-x^2},$$

which is clearly never satisfied since $-(x-1) < 0$ for all $x \in [0, 1]$, but $2014\mathrm{e}^{-x^2} > 0$.

Case 3: $x > 1$. The inequality simplifies down to

$$-(x-1)^2 > 2014\mathrm{e}^{-x^2}.$$

For identical reasons (the LHS is non-positive but the RHS is positive), this is never satisfied.

Thus, we conclude that $fg(x) > gf(x)$ has no solutions.

Part (c). We have

$$\int_{-\infty}^{n} hg(x)\,\mathrm{d}x = \int_{-\infty}^{0} hg(x)\,\mathrm{d}x + \int_{0}^{n} hg(x)\,\mathrm{d}x = \underbrace{\int_{-\infty}^{0} \lfloor 2014\mathrm{e}^{x} \rfloor \,\mathrm{d}x}_{I_1} + \underbrace{\int_{0}^{n} \lfloor x-1 \rfloor \,\mathrm{d}x}_{I_2}.$$

We first consider I_2. Observe that

$$\lfloor x-1 \rfloor = k-1, \quad k \le x < k-1.$$

Thus,

$$I_2 = \sum_{k=0}^{n-1}(k-1) = \frac{n(n-1)}{2} - n = \frac{n(n-3)}{2}.$$

We now consider I_1. Under the transformation $x \mapsto -x$, we obtain

$$I_1 = \int_{0}^{\infty} \lfloor 2014\mathrm{e}^{-x} \rfloor \,\mathrm{d}x.$$

Observe that

$$\lfloor 2014\mathrm{e}^{-x} \rfloor = k, \quad \ln\frac{2014}{k+1} \le x < \ln\frac{2014}{k}.$$

Thus,

$$I_1 = \sum_{k=1}^{2013} \int_{\ln(2014/k+1)}^{\ln(2014/k)} k\,\mathrm{d}x = \sum_{k=1}^{2013} k\left(\ln\frac{2014}{k} - \ln\frac{2014}{k+1}\right) = \sum_{k=1}^{2013} k\left[\ln(k+1) - \ln k\right].$$

Expanding this sum, we quickly see that it telescopes:

$$\begin{aligned} I_1 &= (\ln 2 - \ln 1) + 2(\ln 3 - \ln 2) + 3(\ln 4 - \ln 3) + \cdots + 2013(\ln 2014 - \ln 2013) \\ &= 2013\ln 2014 - (\ln 1 + \ln 2 + \ln 3 + \cdots + \ln 2013) = 2013\ln 2014 - \ln 2013!. \end{aligned}$$

Note that we can also write I_1 as $2014\ln 2014 - \ln 2014!$.

Putting everything together, we see that

$$\int_{-\infty}^{n} hg(x)\,\mathrm{d}x = 2014\ln 2014 - \ln 2014! + \frac{n(n-3)}{2},$$

whence $a = 2014$, $b = 2014$ and $c = 3$.

Analysis 1.4 Differential Equations

Tutorial A1.4 Set 1

Problem 1. The notation $(1+D)^{[n]}(y)$ is used to denote

$$y+\binom{n}{1}\frac{\mathrm{d}y}{\mathrm{d}x}+\binom{n}{2}\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+\cdots+\frac{\mathrm{d}^ny}{\mathrm{d}x^n}.$$

Show that

$$(1+D)^{[1]}\left[(1+D)^{[1]}(y)\right]=(1+D)^{[2]}(y).$$

Use an integrating factor to solve the differential equation

$$(1+D)^{[1]}(y)=x.$$

Use your solution of this equation to solve the differential equation

$$(1+D)^{[2]}(y)=x.$$

Solution.

$$\begin{aligned}(1+D)^{[1]}\left[(1+D)^{[1]}(y)\right]&=(1+D)^{[1]}\left(y+\frac{\mathrm{d}y}{\mathrm{d}x}\right)=\left(y+\frac{\mathrm{d}y}{\mathrm{d}x}\right)+\left(\frac{\mathrm{d}y}{\mathrm{d}x}+\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right)\\&=y+\binom{2}{1}\frac{\mathrm{d}y}{\mathrm{d}x}+\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=(1+D)^{[2]}(y).\end{aligned}$$

The differential equation $(1+D)^{[1]}(y)=x$ can be expanded as

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

The integrating factor is $\mathrm{e}^{\int\mathrm{d}x}=\mathrm{e}^x$. Multiplying through by e^x, we get

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^xy\right)=\mathrm{e}^x\frac{\mathrm{d}y}{\mathrm{d}x}+\mathrm{e}^xy=x\mathrm{e}^x.$$

Hence,

$$\mathrm{e}^xy=\int x\mathrm{e}^x\,\mathrm{d}x=(x-1)\,\mathrm{e}^x+C.$$

Thus, the general solution is

$$y=x-1+C\mathrm{e}^{-x}.$$

We have

$$(1+D)^{[2]}(y)=(1+D)^{[1]}\left[(1+D)^{[1]}(y)\right]=x.$$

Using the above solution, we see that

$$(1+D)^{[1]}(y)=x-1+C\mathrm{e}^{-x}.$$

Once again, the integrating factor is e^x. Multiplying it through yields

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^xy\right)=x-1+C\mathrm{e}^{-x}.$$

Thus,

$$\mathrm{e}^x y = \int \left[(x-1)\mathrm{e}^x + C\right]\,\mathrm{d}x = (x-2)\mathrm{e}^x + Cx + D.$$

Hence, the general solution is

$$y = x - 2 + (Cx + D)\,\mathrm{e}^{-x}.$$

Problem 2.

(a) Use the substitution $y = v/t$ to reduce the differential equation

$$2t^2 \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + t\frac{\mathrm{d}y}{\mathrm{d}t} - 3y = 0$$

for $t > 0$ to

$$2t\frac{\mathrm{d}^2 v}{\mathrm{d}t^2} - 3\frac{\mathrm{d}v}{\mathrm{d}t} = 0.$$

(b) Using a suitable substitution, show that the differential equation

$$2t\frac{\mathrm{d}^2 v}{\mathrm{d}t^2} - 3\frac{\mathrm{d}v}{\mathrm{d}t} = 0$$

can be reduced to

$$2t\frac{\mathrm{d}w}{\mathrm{d}t} - 3w = 0.$$

(c) Hence, find a general solution to the differential equation

$$2t^2 \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + t\frac{\mathrm{d}y}{\mathrm{d}t} - 3y = 0.$$

Solution.

Part (a). Note that $y = v/t \implies v = yt$. Differentiating with respect to t,

$$\frac{\mathrm{d}v}{\mathrm{d}t} = t\frac{\mathrm{d}y}{\mathrm{d}t} + y, \quad \frac{\mathrm{d}^2 v}{\mathrm{d}t^2} = t\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + 2\frac{\mathrm{d}y}{\mathrm{d}t}.$$

Since

$$2t^2 \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + t\frac{\mathrm{d}y}{\mathrm{d}t} - 3y = 2t\left(t\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + 2\frac{\mathrm{d}y}{\mathrm{d}t}\right) - 3\left(t\frac{\mathrm{d}y}{\mathrm{d}t} + y\right) = 0,$$

we have

$$2t\frac{\mathrm{d}^2 v}{\mathrm{d}t^2} - 3\frac{\mathrm{d}v}{\mathrm{d}t} = 0$$

as desired.

Part (b). Let $w = \mathrm{d}v/\mathrm{d}t$. Then the DE becomes

$$2t\frac{\mathrm{d}w}{\mathrm{d}t} - 3w = 0.$$

Part (c). Note that the above DE can be written as

$$\frac{1}{w}\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{3}{2t}.$$

Integrating with respect to t,

$$\ln|w| + A = \frac{3}{2}\ln|t| \implies w = Bt^{3/2}.$$

Hence,

$$v = \int w \,\mathrm{d}t = \int Bt^{3/2}\,\mathrm{d}t = Ct^{5/2} + D.$$

Finally,

$$y = \frac{v}{t} = Ct^{3/2} + Dt^{-1}.$$

Tutorial A1.4 Set 2

Problem 1. Reduce the differential equation

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

to a differential equation in x and u, where $u = y/x^2$. Hence, express y in terms of x and an arbitrary constant. Find the equation of the solution curve that passes through the point $(x, y) = (1, \mathrm{e})$.

Solution. Note that $u = y/x^2 \implies y = ux^2$. Differentiating with respect to x,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2ux + x^2\frac{\mathrm{d}u}{\mathrm{d}x}.$$

Substituting this into the given DE, we have

$$\frac{x}{ux^2}\left(2ux + x^2\frac{\mathrm{d}u}{\mathrm{d}x}\right) + \ln\left(\frac{x^2}{ux^2}\right) = 2,$$

which simplifies to

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

Integrating both sides with respect to x yields

$$\ln|\ln u| = \ln|x| + A \implies \ln u = Bx \implies u = \mathrm{e}^{Bx} \implies y = ux^2 = x^2\mathrm{e}^{Bx}.$$

At $(x, y) = (1, \mathrm{e})$, we have $\mathrm{e} = \mathrm{e}^B$, whence $B = 1$. Thus, $y = x^2\mathrm{e}^x$.

* * * * *

Problem 2. $T_n(x)$ is a polynomial of degree n in x defined by

$$T_n(x) = \cos(n \arccos x),$$

so that $T_n(\cos\theta) = \cos(n\theta)$.

(a) Show that

$$T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x.$$

(b) Show that

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

(c) Use the results in (a) and (b) to find $T_4(x)$.

$U_n(x)$ is a polynomial of degree n in x defined by

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}$$

where $x = \cos\theta$.

(d) Show that

$$\frac{\mathrm{d}}{\mathrm{d}x}T_n(x) = nU_{n-1}(x).$$

(e) Show that

$$\frac{\mathrm{d}}{\mathrm{d}x}U_{n-1}(x) = \frac{xU_{n-1}(x) - nT_n(x)}{1 - x^2}.$$

(f) Deduce that $y = T_n(x)$ satisfies the differential equation

$$\left(1-x^2\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - x\frac{\mathrm{d}y}{\mathrm{d}x} + n^2y = 0.$$

Solution.

Part (a). Since

$$T_2(\cos\theta) = \cos 2\theta = 2\cos^2 - 1,$$

it follows that $T_2(x) = 2x^2 - 1$. Similarly, since

$$T_3(\cos\theta) = \cos 3\theta = 4\cos^3 x - 3\cos x,$$

it follows that $T_3(x) = 4x^3 - 3x$.

Part (b). Observe that

$$\begin{gathered} T_{n+1}(\cos\theta) = \cos(n+1)\theta = \cos n\theta\cos\theta - \sin n\theta\sin\theta \\ = 2\cos n\theta\cos\theta - (\cos n\theta\cos\theta + \sin n\theta\sin\theta) = 2\cos n\theta\cos\theta - \cos(n-1)\theta. \end{gathered}$$

Thus, $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$.

Part (c). We have

$$T_4(x) = 2xT_3(x) - T_2(x) = 2x\left(4x^3 - 3x\right) - \left(2x^2 - 1\right) = 8x^4 - 8x^2 + 1.$$

Part (d). Since $x = \cos\theta$, we have $\sin\theta = \sqrt{1-x^2}$. Hence,

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta} = \frac{\sin((n+1)\arccos x)}{\sqrt{1-x^2}}.$$

Thus, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}T_n(x) = \frac{\mathrm{d}}{\mathrm{d}x}\cos(n\arccos x) &= -\sin(n\arccos x)\left(\frac{-n}{\sqrt{1-x^2}}\right) \\ = n\frac{\sin(n\arccos x)}{\sqrt{1-x^2}} &= nU_{n-1}(x). \end{aligned}$$

Part (e). We have

$$\begin{gathered} \frac{\mathrm{d}}{\mathrm{d}x}U_{n-1}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\sin(n\arccos x)}{\sqrt{1-x^2}} \\ = \frac{1}{1-x^2}\left[\sqrt{1-x^2}\left(\frac{-n\cos(n\arccos x)}{\sqrt{1-x^2}}\right) - \sin(n\arccos x)\left(\frac{-x}{\sqrt{1-x^2}}\right)\right] \\ = \frac{1}{1-x^2}\left[-n\cos(n\arccos x) + x\frac{\sin(n\arccos x)}{\sqrt{1-x^2}}\right] = \frac{-nT_n(x) + xU_{n-1}(x)}{1-x^2}. \end{gathered}$$

Part (f). Note that $y = T_n(x)$ implies $y' = nU_{n-1}(x)$. Differentiating once more yields

$$y'' = n\left(\frac{xU_{n-1}(x) - nT_n(x)}{1-x^2}\right) = \frac{nxU_{n-1}(x) - n^2T_n(x)}{1-x^2} = \frac{xy' - n^2y}{1-x^2}.$$

Rearranging, we obtain

$$\left(1-x^2\right)y'' - xy' + n^2y = 0.$$

Analysis 2.1 Limits

Tutorial A2.1

Problem 1. Evaluate the following limits:

(a) $\displaystyle \lim_{n\to\infty} \frac{3+n^2-n^3}{1-3n\sqrt{n}+5n^3}$,

(b) $\displaystyle \lim_{n\to\infty} n^4\left(\sqrt{1+\frac{1}{n^4}}-1\right)$,

(c) $\displaystyle \lim_{n\to\infty} \frac{3^n+(-3)^n}{4^n}$,

(d) $\displaystyle \lim_{n\to\infty} \prod_{k=2}^{n}\left(1-\frac{1}{k^2}\right)$,

(e) $\displaystyle \lim_{n\to\infty} \left(\frac{\sin(n^2+1)}{5}+\frac{\cos(2n)}{4}\right)^n$,

(f) $\displaystyle \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{n^2+3k}$.

* * * * *

Problem 2. For each of the following statements, determine whether it is true or false. If it is true, give a proof. If it is false, provide a counter-example.

(a) If $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ are both divergent, then $\{x_n+y_n\}_{n=1}^{\infty}$ is divergent.

(b) If $\{x_n\}_{n=1}^{\infty}$ is convergent and $\{y_n\}_{n=1}^{\infty}$ is divergent, then $\{x_n+y_n\}_{n=1}^{\infty}$ is divergent.

(c) If $\{x_n\}_{n=1}^{\infty}$ is convergent and $\{y_n\}_{n=1}^{\infty}$ is divergent, then $\{x_ny_n\}_{n=1}^{\infty}$ is divergent.

(d) If $\lim_{n\to\infty}\frac{1}{n}(x_1+x_2+\cdots+x_n)=x$, then $\lim_{n\to\infty}x_n=x$.

Solution.

Part (a). False. Take $x_n=n$ and $y_n=-n$. Clearly, both sequences are divergent (to ∞ and $-\infty$ respectively), but their sum is the zero sequence.

Part (b). True. Seeking a contradiction, suppose $\{x_n\}_{n=1}^{\infty}$ is convergent, $\{y_n\}_{n=1}^{\infty}$ divergent and $\{x_n+y_n\}_{n=1}^{\infty}$ convergent. Then there exists finite constants x and s such that

$$x=\lim_{n\to\infty}x_n \quad \text{and} \quad s=\lim_{n\to\infty}(x_n+y_n).$$

This implies that

$$\lim_{n\to\infty}y_n=\lim_{n\to\infty}(x_n+y_n-x_n)=s-x$$

must also be finite, contradicting the divergence of $\{y_n\}_{n=1}^{\infty}$.

Part (c). False. Take $x_n=0$ (which is clearly convergent) and $y_n=n$ (which diverges to ∞). Then their product $x_ny_n=0$ is clearly convergent.

Part (d). False. Take $x_n=(-1)^n$. Then

$$\left|\lim_{n\to\infty}\frac{x_1+x_2+\cdots+x_n}{n}\right|=\lim_{n\to\infty}\left|\frac{x_1+x_2+\cdots+x_n}{n}\right|\le\lim_{n\to\infty}\frac{1}{n}=0$$

so

$$\lim_{n\to\infty}\frac{x_1+x_2+\cdots+x_n}{n}=0,$$

but

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} (-1)^n$$

does not exist.

* * * * *

Problem 3. Use the result that the sequence $\{e_n\}_{n=1}^{\infty}$ defined by $e_n = (1+1/n)^n$ converges to e as $n \to \infty$ to evaluate the following limits.

(a) $\displaystyle\lim_{n\to\infty}\left(1+\frac{1}{4n+1}\right)^{2n+1}$, (b) $\displaystyle\lim_{n\to\infty}\left(1+\frac{3}{n}\right)^{n}$, (c) $\displaystyle\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{n}$.

Solution.

Part (a). Let

$$L = \lim_{n\to\infty}\left(1+\frac{1}{4n+1}\right)^{2n+1}.$$

Then

$$L^2 = \lim_{n\to\infty}\left(1+\frac{1}{4n+1}\right)^{4n+2} = \lim_{n\to\infty}\left(1+\frac{1}{4n+1}\right)\left(1+\frac{1}{4n+1}\right)^{4n+1} = \mathrm{e},$$

so $L = \sqrt{e}$.

Part (b). Let

$$L = \lim_{n\to\infty}\left(1+\frac{3}{n}\right)^{n} = \lim_{n\to\infty}\left(1+\frac{1}{n/3}\right)^{n}.$$

Then

$$L^{1/3} = \lim_{n\to\infty}\left(1+\frac{1}{n/3}\right)^{n/3} = \mathrm{e},$$

so $L = \mathrm{e}^3$.

Part (c). Let

$$L = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{n}.$$

Then

$$L\mathrm{e} = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{n}\left(1+\frac{1}{n}\right)^{n} = \lim_{n\to\infty}\left(1-\frac{1}{n^2}\right)^{n}.$$

By Bernoulli's inequality, one obtains the bound

$$1-\frac{n}{n^2} \le \left(1-\frac{1}{n^2}\right)^{n} \le 1.$$

Taking limits on both sides, we have

$$1 \le \lim_{n\to\infty}\left(1-\frac{1}{n^2}\right)^{n} \le 1,$$

so by the Squeeze Theorem, $L\mathrm{e} = 1$, which implies $L = 1/\mathrm{e}$.

Problem 4. Evaluate the following limits, without using L'Hôpital's rule.

(a) $\lim_{t\to 6} 8(t-5)(t-7)$,

(b) $\lim_{y\to -3}(5-y)^{4/3}$,

(c) $\lim_{y\to 2} \frac{y+2}{y^2+5y+6}$,

(d) $\lim_{h\to 0} \frac{\sqrt{5h+4}-2}{h}$,

(e) $\lim_{u\to 0} \frac{u^4-1}{u^3-1}$,

(f) $\lim_{x\to 0} \frac{1/(x-1)+1/(x+1)}{x}$,

(g) $\lim_{x\to 4} \frac{4-x}{5-\sqrt{x^2+9}}$.

Solution.

Part (a). Clearly,

$$\lim_{t\to 6} 8(t-5)(t-7) = 8(6-5)(6-7) = -8.$$

Part (b). Clearly,

$$\lim_{y\to -3}(5-y)^{4/3} = (5-(-3))^{4/3} = 16.$$

Part (c). Clearly,

$$\lim_{y\to 2} \frac{y+2}{y^2+5y+6} = \frac{2+2}{2^2+5(2)+6} = \frac{1}{5}.$$

Part (d). Observe that

$$\lim_{h\to 0} \frac{\sqrt{5h+4}-2}{h} = \frac{\mathrm{d}}{\mathrm{d}x}\sqrt{5x+4}\bigg|_{x=0} = \frac{5}{2\sqrt{5x+4}}\bigg|_{x=0} = \frac{5}{4}.$$

Part (e). Clearly,

$$\lim_{u\to 0} \frac{u^4-1}{u^3-1} = \frac{-1}{-1} = 1.$$

Part (f). We have

$$\lim_{x\to 0} \frac{1/(x-1)+1/(x+1)}{x} = \lim_{x\to 0} \frac{2}{(x-1)(x+1)} = \frac{2}{(-1)(1)} = -2.$$

Part (g). Rationalizing, we obtain

$$\lim_{x\to 4} \frac{4-x}{5-\sqrt{x^2+9}} = \lim_{x\to 4} \frac{(4-x)\left(5+\sqrt{x^2+9}\right)}{16-x^2} = \lim_{x\to 4} \frac{5+\sqrt{x^2+9}}{4+x} = \frac{5+\sqrt{4^2+9}}{4+4} = \frac{5}{4}.$$

* * * * *

Problem 5. Suppose that $\lim_{x\to 0} f(x) = 1$ and $\lim_{x\to 0} g(x) = -5$. Find the value of

$$\lim_{x\to 0} \frac{2f(x)-g(x)}{[f(x)+7]^{2/3}}.$$

Solution. Clearly,

$$\lim_{x\to 0} \frac{2f(x)-g(x)}{[f(x)+7]^{2/3}} = \frac{2(1)-(-5)}{(1+7)^{2/3}} = \frac{7}{4}.$$

* * * * *

Problem 6. Suppose that $\lim_{x\to 4} \frac{f(x)-5}{x-2} = 1$. Find the value of $\lim_{x\to 4} f(x)$.

Solution. Trivially,

$$\lim_{x\to 4}\frac{f(x)-5}{x-2}=\frac{\lim_{x\to 4}f(x)-5}{4-2}=1\implies \lim_{x\to 4}f(x)=7.$$

* * * * *

Problem 7. Suppose that $\lim\limits_{x\to 0}\dfrac{f(x)}{x^2}=1$. Find the values of $\lim\limits_{x\to 0}\dfrac{f(x)}{x}$ and $\lim\limits_{x\to 0}f(x)$.

Solution. We have

$$\lim_{x\to 0}\frac{f(x)}{x}=\lim_{x\to 0}x\left[\frac{f(x)}{x^2}\right]=0\cdot 1=0$$

and

$$\lim_{x\to 0}f(x)=\lim_{x\to 0}x^2\left[\frac{f(x)}{x^2}\right]=0^2\cdot 1=0.$$

* * * * *

Problem 8. Find the value of the following one-sided limits.

(a) $\displaystyle\lim_{x\to -2^+}(x+3)\frac{|x+2|}{x+2}$,

(b) $\displaystyle\lim_{x\to 0^+}\frac{1-\cos x}{|\cos x-1|}$,

(c) $\displaystyle\lim_{t\to 4^+}(t-\lfloor t\rfloor)$,

(d) $\displaystyle\lim_{t\to 4^-}(t-\lfloor t\rfloor)$,

(e) $\displaystyle\lim_{h\to 0^+}\frac{\sqrt{6}-\sqrt{5h^2+11h+6}}{h}$,

(f) $\displaystyle\lim_{x\to 0^+}\sqrt{x^3+x^2+x}\sin\frac{\pi}{x}$.

Solution.

Part (a). Observe that

$$\lim_{x\to -2^+}(x+3)\frac{|x+2|}{x+2}=\lim_{x\to -2^+}(x+3)\,\operatorname{sgn}(x+2)=(-2+3)\,\operatorname{sgn}\big(0^+\big)=1.$$

Part (b). Observe that

$$\lim_{x\to 0^+}\frac{1-\cos x}{|\cos x-1|}=\lim_{x\to 0^+}\operatorname{sgn}(1-\cos x)=\operatorname{sgn}\big(0^+\big)=1.$$

Part (c). Clearly,

$$\lim_{t\to 4^+}(t-\lfloor t\rfloor)=4-4=0.$$

Part (d). Clearly,

$$\lim_{t\to 4^-}(t-\lfloor t\rfloor)=4-3=1.$$

Part (e). Observe that

$$\begin{aligned}\lim_{h\to 0^+}\frac{\sqrt{6}-\sqrt{5h^2+11h+6}}{h}&=-\frac{\mathrm{d}}{\mathrm{d}x}\sqrt{5x^2+11x+6}\bigg|_{x=0}\\&=-\frac{10x+11}{2\sqrt{5x^2+11x+6}}\bigg|_{x=0}=-\frac{11}{2\sqrt{6}}.\end{aligned}$$

Part (f). Since $\sin(\pi/x)\in[-1,1]$, it follows that

$$0=-\lim_{x\to 0}\sqrt{x^3+x^2+x}\le\lim_{x\to 0}\sqrt{x^3+x^2+x}\sin\pi t\le\lim_{x\to 0}\sqrt{x^3+x^2+x}=0.$$

Thus, by the Squeeze Theorem, the limit is simply 0.

Problem 9. Find the value of the following limits, without using L'Hôpital's rule.

(a) $\lim\limits_{x\to 0} \dfrac{x \csc 2x}{\cos 5x}$, (b) $\lim\limits_{x\to 0} \dfrac{x + x\cos x}{\sin x \cos x}$,

(c) $\lim\limits_{x\to 0} \dfrac{\sin(1-\cos x)}{1-\cos x}$, (d) $\lim\limits_{x\to 0} \dfrac{x \cot 4x}{\sin^2 x \cot^2 2x}$.

Solution.

Part (a). Observe that

$$\lim_{x\to 0} \frac{x \csc 2x}{\cos 5x} = \lim_{x\to 0} \frac{1}{2} \frac{2x}{\sin 2x} \frac{1}{\cos 5x} = \frac{1}{2}.$$

Part (b). Observe that

$$\lim_{x\to 0} \frac{x + x\cos x}{\sin x \cos x} = \lim_{x\to 0} \frac{x}{\sin x} \frac{1+\cos x}{\cos x} = 2.$$

Part (c). Let $u = 1 - \cos x$. Then

$$\lim_{x\to 0} \frac{\sin(1-\cos x)}{1-\cos x} = \lim_{u\to 0} \frac{\sin u}{u} = 1.$$

Part (d). Observe that

$$\lim_{x\to 0} \frac{x \cot 4x}{\sin^2 x \cot^2 2x} = \lim_{x\to 0} \left(\frac{x}{\sin x}\right)^2 \left(\frac{4x}{\sin 4x}\right) \left(\frac{\sin 2x}{2x}\right)^2 \left(\frac{\cos 4x}{\cos^2 2x}\right) = 1.$$

* * * * *

Problem 10. For what values of a and b is

$$g(x) = \begin{cases} ax + b, & x \le 0, \\ x^2 + 3a - b, & 0 < x \le 2, \\ 3x - 5, & x > 2 \end{cases}$$

continuous at every x?

Solution. Equating the left and right limits at $x = 0$, we get

$$b = \lim_{x\to 0^-} g(x) = \lim_{x\to 0^+} g(x) = 3a - b.$$

Equating the left and right limit at $x = 2$, we get

$$4 + 3a - b = \lim_{x\to 2^-} g(x) = \lim_{x\to 2^+} g(x) = 1.$$

Solving these two linear equations simultaneous, we get $a = b = -3/2$.

* * * * *

Problem 11. For what values of a and b is

$$f(x) = \begin{cases} \frac{(\sin x - a)(\cos x - b)}{e^x - 1}, & x \ne 0, \\ 5, & x = 0 \end{cases}$$

continuous at every x?

Part (a). For the limit to exist and be finite, we require the numerator to be 0 as $x \to 0$. This gives $-a(1-b) = 0$, whence $a = 0$ or $b = 1$.

Case 1. Suppose $a = 0$. Invoking L'Hôpital's rule, we see that

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} \frac{\sin x\,(\cos x - b)}{\mathrm{e}^x - 1} = \lim_{x\to 0} \frac{\cos x\,(\cos x - b) - \sin^2 x}{\mathrm{e}^x} = 1 - b.$$

For $f(x)$ to be continuous, we require $1 - b = 5$. Thus, $a = 0$ and $b = -4$.

Case 2. Suppose $b = 1$. Invoking L'Hôpital's rule, we see that

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} \frac{(\sin x - a)\,(\cos x - 1)}{\mathrm{e}^x - 1} = \lim_{x\to 0} \frac{\cos x\,(\cos x - 1) - \sin x\,(\sin x - a)}{\mathrm{e}^x} = 0.$$

Since $0 \neq 5$, this case yields no solutions.

Hence, the only values of a and b that makes $f(x)$ continuous over $\mathbb{R}$ is $a = 0$ and $b = -4$.

* * * * *

Problem 12. Find the value of the following limits.

(a) $\displaystyle\lim_{x\to 0} \frac{x^2 \sin(1/x)}{\sin x}$, (b) $\displaystyle\lim_{x\to\infty} \frac{2x^3 + \sin\left(x^2\right)}{1 + x^3}$.

Solution.

Part (a). Note that $x\sin(1/x) \in [-x, x]$. Thus, by the Squeeze Theorem,

$$\lim_{x\to 0} x \sin\left(\frac{1}{x}\right) = 0.$$

Hence, the limit in question is simply

$$\lim_{x\to 0} \frac{x^2 \sin(1/x)}{\sin x} = \lim_{x\to 0} \left(\frac{x}{\sin x}\right) \left[x \sin\left(\frac{1}{x}\right)\right] = 0.$$

Part (b). Note that

$$\left|\lim_{x\to\infty} \frac{\sin\left(x^2\right)}{1 + x^3}\right| = \lim_{x\to\infty} \left|\frac{\sin\left(x^2\right)}{1 + x^3}\right| \leq \lim_{x\to\infty} \frac{1}{1 + x^3} = 0.$$

Thus,

$$\lim_{x\to\infty} \frac{2x^3 + \sin\left(x^2\right)}{1 + x^3} = \lim_{x\to\infty} \left(\lim_{x\to\infty} \frac{2x^3}{1 + x^3} + \lim_{x\to\infty} \frac{\sin\left(x^2\right)}{1 + x^3}\right) = 2.$$

* * * * *

Problem 13. Evaluate the limit

$$\lim_{x\to 1} \left(\frac{1}{\ln x} - \frac{x^2}{x - 1}\right).$$

Solution. Let the limit be L. We have

$$L = \lim_{x\to 1} \left(\frac{1}{\ln x} - \frac{x^2}{x - 1}\right) = \lim_{x\to 1} \frac{x - 1 - x^2 \ln x}{(x - 1)\ln x}.$$

Invoking L'Hôpital's rule, we have

$$L = \lim_{x\to 1} \frac{1 - 2x\ln x - x}{(x - 1)/x + \ln x} = \lim_{x\to 1} \frac{1 - 2x\ln x - x}{x - 1 + x\ln x},$$

where we multiplied the denominator by $x = 1$. Invoking L'Hôpital's rule once more, we obtain

$$L = \lim_{x\to 1} \frac{-1-2-2\ln x}{1+\ln x+1} = -\frac{3}{2}.$$

* * * * *

Problem 14. Evaluate the limit

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \sqrt{k}}{\sum_{k=1}^{n} \sqrt{n+k}}.$$

Solution. Observe that

$$\lim_{n\to\infty} n^{-3/2} \sum_{k=1}^{n} \sqrt{k} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{n}\sqrt{\frac{k}{n}} = \int_0^1 \sqrt{x}\,\mathrm{d}x = \left[\frac{2}{3}x^{3/2}\right]_0^1 = \frac{2}{3}.$$

Likewise,

$$\begin{aligned}\lim_{n\to\infty} n^{-3/2} \sum_{k=1}^{n} \sqrt{n+k} &= \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{n}\sqrt{1+\frac{k}{n}} \\ = \int_0^1 \sqrt{1+x}\,\mathrm{d}x &= \left[\frac{2}{3}(1+x)^{3/2}\right]_0^1 = \frac{2}{3}\left(2^{3/2}-1\right).\end{aligned}$$

Thus, the limit in question is simply

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \sqrt{k}}{\sum_{k=1}^{n} \sqrt{n+k}} = \lim_{n\to\infty} \frac{n^{-3/2}\sum_{k=1}^{n} \sqrt{k}}{n^{-3/2}\sum_{k=1}^{n} \sqrt{n+k}} = \frac{2/3}{(2/3)(2^{3/2}-1)} = \frac{1}{2^{3/2}-1}.$$

Analysis 2.2 Sequences and Series

Tutorial A2.2

Problem 1. The geometric progression U has terms $u_1, u_2, u_3, \ldots$, with common ratio r, where $|r| < 1$. It is given that $v_i = u_i^2$ for $i = 1, 2, 3, \ldots$.

(a) Show that

$$\sum_{i=1}^{n} v_i = \frac{u_1}{1+r} \sum_{i=1}^{2n} u_i.$$

It is given further that $w_i = v_i - v_{i+1}$ for $i = 1, 2, 3, \ldots$.

(b) Show that

$$\sum_{i=1}^{n} w_i = u_1 (1-r) \sum_{i=1}^{2n} u_i.$$

Let

$$S_U = \sum_{i=1}^{\infty} u_i, \quad S_V = \sum_{i=1}^{\infty} v_i, \quad S_W = \sum_{i=1}^{\infty} w_i.$$

Show that

(c) $\dfrac{S_U}{S_V} + \dfrac{1}{S_U} = \dfrac{2}{u_1}$,

(d) $S_W = u_1^2$.

Solution.

Part (a). Note that $u_i = r^{i-1} u_1$. Hence,

$$\sum_{i=1}^{2n} u_i = \sum_{i=1}^{2n} r^{i-1} u_1 = u_1 \left(\frac{1 - r^{2n}}{1-r} \right).$$

Meanwhile, we have $v_1 = r^{2i-2} u_1^2$. Thus,

$$\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} r^{2i-2} u_1^2 = u_1^2 \left(\frac{1 - r^{2n}}{1 - r^2} \right) = \frac{u_1}{1+r} \left(u_1 \frac{1 - r^{2n}}{1-r} \right) = \frac{u_1}{1+r} \sum_{i=1}^{2n} u_i.$$

Part (b). We have

$$\sum_{i=1}^{n} w_i = \sum_{i=1} (v_i - v_{i+1}).$$

Quite clearly, this sum telescopes, giving

$$\sum_{i=1}^{n} w_i = v_1 - v_{n+1} = u_1^2 \left(1 - r^{2n} \right) = u_1 (1-r) \left(u_1 \frac{1 - r^{2n}}{1-r} \right) = u_1 (1-r) \sum_{i=1}^{2n} u_i.$$

Part (c). From part (a), we have

$$S_V = \frac{u_1}{1+r} S_U \implies \frac{S_U}{S_V} = \frac{1+r}{u_1}.$$

Note also that

$$S_U = \sum_{i=1}^{\infty} u_i = \sum_{i=1}^{\infty} r^{i-1} u_1 = \frac{u_1}{1-r}.$$

Thus,

$$\frac{S_U}{S_V} + \frac{1}{S_U} = \frac{1+r}{u_1} + \frac{1-r}{u_1} = \frac{2}{u_1}.$$

Part (d). From part (b), we have

$$S_W = u_1 (1-r) S_U = u_1 (1-r) \frac{u_1}{1-r} = u_1^2.$$

* * * * *

Problem 2. For each of the following sequences, determine if it converges or diverges. Take $n = 1, 2, 3, \ldots$.

(a) $u_n = \sin(n\pi/2)$,

(b) $u_n = n^3/3^n$,

(c) $u_n = n^n/n!$,

(d) $u_n = (\sin n)/n$.

Solution.

Part (a). Note that

$$u_n = \begin{cases} 0, & n \equiv 0, 2 \pmod 4, \\ 1, & n \equiv 1 \pmod 4, \\ -1, & n \equiv 3 \pmod 4. \end{cases}$$

Thus, u_n is divergent.

Part (b). Invoking L'Hôpital's rule repeatedly,

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} \frac{n^3}{3^n} = \frac{3}{\ln 3} \lim_{n\to\infty} \frac{n^2}{3^n} = \frac{6}{(\ln 3)^2} \lim_{n\to\infty} \frac{n}{3^n} = \frac{6}{(\ln 3)^3} \lim_{n\to\infty} \frac{1}{3^n} = 0.$$

Part (c). Let the limit be L. Then

$$\ln L = \lim_{n\to\infty} \left(n \ln n - \ln n! \right).$$

By Stirling's approximation, $\ln n! = n \ln n - n + \mathrm{O}(\ln n)$. Thus,

$$\ln L = \lim_{n\to\infty} \left(n - \mathrm{O}(\ln n) \right) = \lim_{n\to\infty} \mathrm{O}(n),$$

which diverges to ∞. Hence, L diverges to ∞.

Part (d). Note that

$$\left| \lim_{n\to\infty} u_n \right| = \lim_{n\to\infty} |u_n| \leq \lim_{n\to\infty} \frac{1}{n} = 0.$$

Thus, the limit converges to 0.

* * * * *

Problem 3. Show that the following series is divergent:

$$5 + \sqrt{5} + \sqrt[3]{5} + \sqrt[4]{5} + \ldots.$$

Solution. We have

$$\sum_{i=1}^{\infty} 5^{1/i} \geq \sum_{i=1}^{\infty} 1^{1/i} = \sum_{i=1}^{\infty} 1,$$

which clearly diverges to ∞.

* * * * *

Problem 4.

(a) Show that

$$\sum_{n=2^{k-1}+1}^{2^k} \frac{1}{n} \geq \frac{1}{2}$$

for every positive integer k.

(b) Show that

$$\sum_{n=1}^{2^k} \frac{1}{n} \geq 1 + k/2$$

for every positive integer k.

(c) Hence, determine if the harmonic series converges or diverges.

Solution.

Part (a). We have

$$\sum_{n=2^{k-1}+1}^{2^k} \frac{1}{n} \geq \sum_{n=2^{k-1}+1}^{2^k} \frac{1}{2^k} = \frac{1}{2^k}\left(2^k - 2^{k-1}\right) = \frac{2^{k-1}}{2^k} = \frac{1}{2}.$$

Part (b). We have

$$\sum_{n=1}^{2^k} \frac{1}{n} = 1 + \sum_{i=1}^{k} \sum_{n=2^{i-1}+1}^{2^i} \frac{1}{n} \geq 1 + \sum_{i=1}^{k} \frac{1}{2} = 1 + \frac{k}{2}.$$

Part (c). Observe that

$$\sum_{n=1}^{\infty} \frac{1}{n} = \lim_{k\to\infty} \sum_{n=1}^{2^k} \frac{1}{n} \geq \lim_{k\to\infty} \left(1 + \frac{k}{2}\right),$$

which clearly diverges to ∞. Hence, the harmonic series diverges.

* * * * *

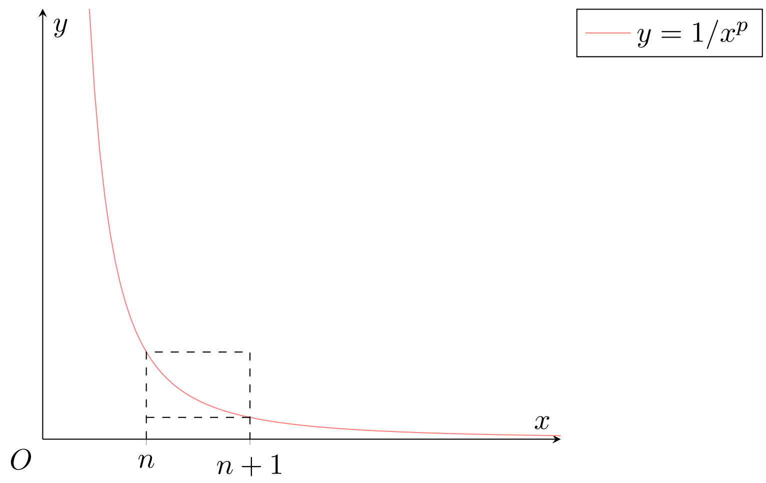
Problem 5. Prove that the following sum of series is less than $3/2$ using the formula for the sum of an infinite geometric series:

$$1 + \frac{1}{3} + \frac{1}{4^2} + \frac{1}{5^3} + \frac{1}{6^4} + \dots.$$

Part (a). Clearly,

$$1 + \frac{1}{3} + \frac{1}{4^2} + \frac{1}{5^3} + \frac{1}{6^4} + \cdots < 1 + \frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \frac{1}{3^4} + \cdots = \frac{1}{1 - 1/3} = \frac{3}{2}.$$

Problem 6. The diagram below shows a sketch of the graph $y = 1/x^p$, where $p > 0$, $p \neq 1$.



By considering the area of appropriate rectangles and the area between the graph and the x-axis, where $n \geq 1$, show that

$$\frac{1}{(n+1)^p} < \frac{1}{1-p}(n+1)^{1-p} - \frac{1}{1-p}n^{1-p} < \frac{1}{n^p}.$$

Deduce that

$$\frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} + \frac{1}{(n+1)^p} < \frac{1}{1-p}\left[(n+1)^{1-p} - 1\right] < 1 + \frac{1}{2^p} + \cdots + \frac{1}{(n-1)^p} + \frac{1}{n^p}.$$

Deduce the convergence of the series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

when

(a) $p = 1/2$, (b) $p = 2$.

Solution. The area of the big rectangle, denoted A_1, is given by

$$A_1 = \text{base} \times \text{height} = \left(\frac{1}{(n+1) - n}\right)\left(\frac{1}{n^p}\right) = \frac{1}{n^p}.$$

Meanwhile, the area under the curve $y = 1/x^p$ from $x = n$ to $n + 1$, denoted A_2, is given by

$$A_2 = \int_n^{n+1} \frac{1}{x^p}\,\mathrm{d}x = \left[\frac{x^{1-p}}{1-p}\right]_n^{n+1} = \frac{1}{1-p}\left[(n+1)^{1-p} - n^{1-p}\right].$$

Lastly, the area of the small rectangle, denoted A_3, is given by

$$A_3 = \text{base} \times \text{height} = \left(\frac{1}{(n+1) - n}\right)\left(\frac{1}{(n+1)^p}\right) = \frac{1}{(n+1)^p}.$$

From the above diagram, it is clear that $A_3 < A_2 < A_1$. Thus,

$$\frac{1}{(n+1)^p} < \frac{1}{1-p}(n+1)^{1-p} - \frac{1}{1-p}n^{1-p} < \frac{1}{n^p}.$$

Summing the above result from $n = 1$ to $n = m$, we get

$$\frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} + \frac{1}{(n+1)^p} < \frac{(n+1)^{1-p} - 1}{1-p} < 1 + \frac{1}{2^p} + \cdots + \frac{1}{(n-1)^p} + \frac{1}{n^p}.$$

Note that the middle term telescopes.

Part (a). When $p = 1/2$, we have

$$\sum_{m=1}^{\infty} \frac{1}{m^p} > \lim_{m\to\infty} \frac{(m+1)^{1/2}-1}{1/2},$$

which diverges to ∞. Thus, the series diverges when $p = 1/2$.

Part (b). When $p = 2$, the sequence $1/x^2$ is strictly positive, so the series is strictly increasing. Since

$$\sum_{m=1}^{\infty} \frac{1}{m^p} = 1 + \lim_{m\to\infty} \left[\frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{m^p} + \frac{1}{(m+1)^p}\right] < 1 + \lim_{m\to\infty} \frac{(m+1)^{-1}-1}{-1} = 2,$$

the series is also bounded above. Thus, the series is convergent.

* * * * *

Problem 7.

(a) Sketch the graph of $y = 1/x$ and hence explain why

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} > \int_1^{n+1} \frac{\mathrm{d}x}{x}.$$

(b) Sketch the graph of $y = \sin x$ and determine the largest constant a such that $ax \le \sin x$ for $0 \le x \le \pi/2$.

(c) Part of a proof of convergence and divergence of series in a textbook is as follows:

Let n be an integer. Then

$$\sum_{i=1}^{n} \sin\frac{1}{i} \ge \frac{2}{\pi} \sum_{i=1}^{n} \frac{1}{i} \ge \frac{2}{\pi} \ln(n+1)\,.$$

$$\sum_{i=1}^{n} \sin^2\frac{1}{i} \le \sum_{i=1}^{n} \frac{1}{i^2} < 1 + \sum_{i=2}^{n} \frac{1}{i(i-1)} < 2.$$

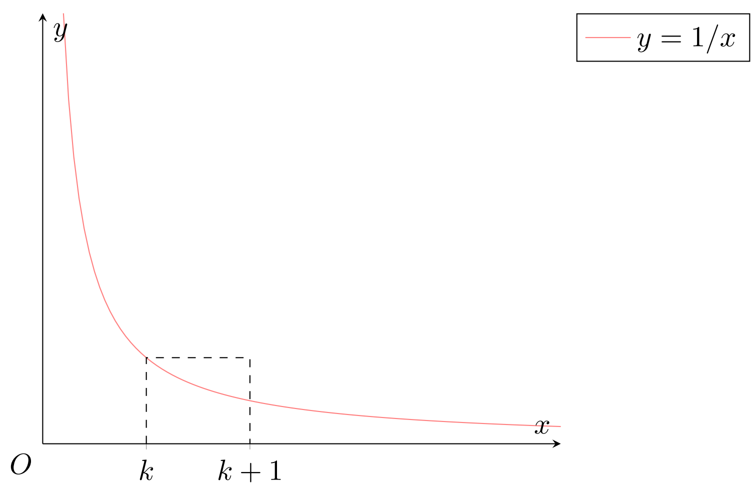
Explain the second line of the proof. Hence, determine for every positive integer k, if the series

$$\sin^k\frac{1}{1} + \sin^k\frac{1}{2} + \sin^k\frac{1}{3} + \ldots$$

is convergent or divergent.

Solution.

Part (a).



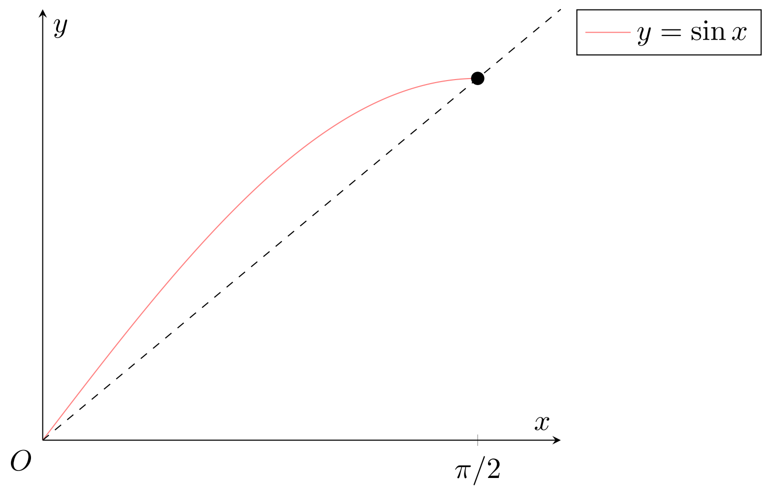
Clearly, the area of the rectangle is greater than the area under the curve $y = 1/x$ over $[k, k+1]$. Hence,

$$\frac{1}{k} > \int_k^{k+1} \frac{1}{x}\,\mathrm{d}x.$$

Summing this result from $k = 1$ to $k = n$, we see that

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} > \int_1^{n+1} \frac{\mathrm{d}x}{x}.$$

Part (b).



From the above figure, the line $y = 2x/\pi$ intersects the maximum point $(\pi/2, 1)$. Hence, the maximum such a is $2/\pi$.

Part (c). From (b), we know that $\sin x \geq 2x/\pi$ for all $x \in [0, \pi/2]$. Since $0 < 1/i < 1 < \pi/2$ for all $i = 1, 2, \ldots, n$, we have

$$\sum_{i=1}^{n} \sin\frac{1}{i} \geq \sum_{i=1}^{n} \frac{2}{\pi}\frac{1}{i} = \frac{2}{\pi}\sum_{i=1}^{n}\frac{1}{i}.$$

From (a), we know that

$$\sum_{i=1}^{n}\frac{1}{i} > \int_1^{n+1}\frac{\mathrm{d}x}{x} = \ln(n+1)\,.$$

Thus,

$$\sum_{i=1}^{n} \sin\frac{1}{i} \geq \frac{2}{\pi}\sum_{i=1}^{n}\frac{1}{i} \geq \frac{2}{\pi}\ln(n+1)\,.$$

When $k = 1$, we have

$$\sum_{i=1}^{\infty}\frac{1}{i} \geq \lim_{n\to\infty}\frac{2}{\pi}\ln(n+1)\,,$$

which diverges. For $k > 2$, by the Cauchy-Schwarz inequality, we have

$$\left(\sum_{i=1}^{n}\sin^2\frac{1}{i}\right)^k \geq \left(\sum_{i=1}^{n}\sin^k\frac{1}{i}\right)^2.$$

Thus,

$$\sum_{i=1}^{n}\sin^k\frac{1}{i} \leq \left(\sum_{i=1}^{n}\sin^2\frac{1}{i}\right)^{k/2} \leq (2)^{k/2}\,.$$

Note also that $\sin^k(1/i) > 0$ for all $i > 1$. Since our series is increasing and bounded above, it is convergent.

* * * * *

Problem 8. A sequence $u_1, u_2, u_3, \ldots$ is given by

$$u_r = \begin{cases} 0, & r = 2, \\ f(r-1) - 2f(r-3), & r \text{ even}, r \neq 2, \\ f(r), & r \text{ odd}. \end{cases}$$

Let

$$S_n = \sum_{r=1}^{n} u_r.$$

(a) Use the method of differences to find S_{2n}.

It is given that $f(r) = \ln((r+1)/r)$.

(b) Use your answer to part (a) to show that

$$S_{2n} = -\ln 2 + 2\ln\left(1 + \frac{1}{2n-1}\right).$$

Hence, state the value of the sum to infinity.

(c) Find the smallest value of n for which S_{2n} is within 10^{-2} of the sum to infinity.

(d) By considering the graph of $y = 1/x$ for $x > 0$, show, with the aid of a sketch, that

$$\frac{1}{2n} < u_{2n-1} < \frac{1}{2n-1}, \quad n \in \mathbb{Z}^+.$$

Solution.

Part (a). We begin by splitting the S_{2n} into odd and even sums:

$$\begin{aligned} S_{2n} = \sum_{m=1}^{n} u_{2m-1} + \sum_{m=2}^{n} u_{2m} &= \sum_{m=1}^{n} f(2m-1) + \sum_{m=2}^{n} \left[f(2m-1) - 2f(2m-3)\right] \\ &= f(1) + 2\sum_{m=2}^{n} \left[f(2m-1) - f(2m-3)\right]. \end{aligned}$$

The resulting sum telescopes, giving

$$S_{2n} = f(1) + 2\left[f(2n-1) - f(1)\right] = 2f(2n-1) - f(1).$$

Part (b). Using (a), we have

$$S_{2n} = 2\ln\left(\frac{2n}{2n-1}\right) - \ln 2 = -\ln 2 + 2\ln\left(1 + \frac{1}{2n-1}\right).$$

As $n \to \infty$, $\ln(1 + 1/(2n-1)) \to \ln 1 = 0$. Thus, the sum to infinity is $-\ln 2$.

Part (c). Consider

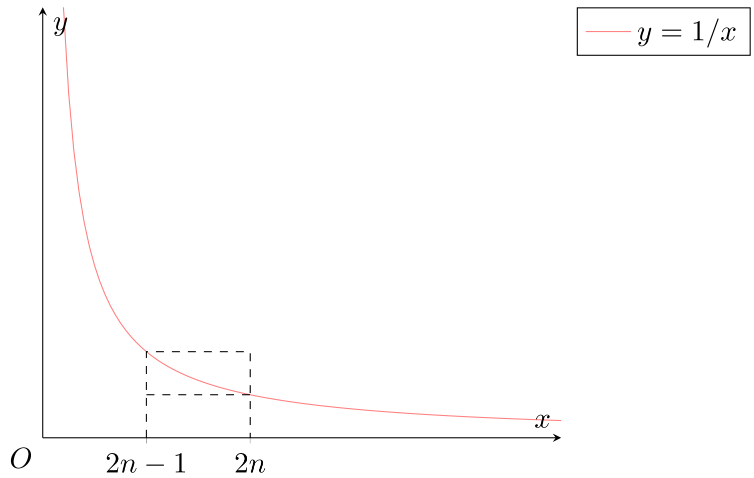
$$|S_{2n} - (-\ln 2)| \leq 10^{-2} \implies 2\ln\left(1 + \frac{1}{2n-1}\right) \leq 10^{-2}.$$

Using G.C., $n \geq 100.25$, so the least n is 101.

Part (d). Observe that

$$u_{2n-1} = f(2n-1) = \ln\left(\frac{2n}{2n-1}\right) = \ln(2n) - \ln(2n-1) = \int_{2n-1}^{2n} \frac{1}{x}\,\mathrm{d}x,$$

which is the area under $y = 1/x$ over $[2n-1, 2n]$. Consider the following diagram:



The area under the curve is larger than the area of the smaller rectangle, but smaller than the area of the larger rectangle. Hence,

$$\frac{1}{2n} < u_{2n-1} < \frac{1}{2n-1}.$$

* * * * *

Problem 9. Use the Binomial Theorem to that for each $n \in \mathbb{N}$,

(a) (i) $\displaystyle\left(1+\frac{1}{n}\right)^n \le \sum_{j=0}^{n} \frac{1}{j!}$,

(ii) $\displaystyle\left(1+\frac{1}{n}\right)^n \ge \sum_{j=0}^{k} \frac{1}{j!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{j-1}{n}\right)$ for any $k < n$.

(b) Deduce that

$$\mathrm{e} = \sum_{j=0}^{\infty} \frac{1}{j!}.$$

(c) Let

$$s_n = \sum_{j=0}^{n} \frac{1}{j!}.$$

Show that for any $m, n \in \mathbb{N}$ such that $m > n$, we have

$$s_m - s_n < \frac{1}{n\,(n!)}.$$

Hence, prove that

$$\mathrm{e} - \sum_{j=0}^{n} \frac{1}{j!} \le \frac{1}{n\,(n!)}.$$

(d) Use (c) to explain why $2 < e < 3$.

(e) Conclude that e is irrational.

Solution.

Part (a).

Part (a)(i). By the Binomial Theorem,

$$\left(1+\frac{1}{n}\right)^n = \sum_{j=0}^{n} \binom{n}{j} \frac{1}{n^j} = \sum_{j=0}^{n} \frac{n!}{(n-j)!j!n^j}.$$

Now observe that

$$\frac{n!}{(n-j)!\,n^j} = \frac{n(n-1)(n-2)\dots(n-j+1)}{n^j} = \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{j-1}{n}\right).$$

A trivial upper bound is

$$\frac{n!}{(n-j)!n^j} \leq 1\cdot 1\dots 1 = 1.$$

Thus,

$$\left(1+\frac{1}{n}\right)^n = \sum_{j=0}^{n} \frac{n!}{(n-j)!j!n^j} \leq \sum_{j=0}^{n} \frac{1}{j!}.$$

Part (a)(ii). Substituting the expression for $\frac{n!}{(n-j)!n^j}$ we found into the sum, we obtain

$$\left(1+\frac{1}{n}\right)^n = \sum_{j=0}^{n} \frac{1}{j!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{j-1}{n}\right).$$

Now observe that all the summands are strictly positive. Thus, the sequence of partial sums is increasing. That is, for all positive integers $k < n$, we have

$$\begin{aligned}\left(1+\frac{1}{n}\right)^n &= \sum_{j=0}^{n} \frac{1}{j!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{j-1}{n}\right)\\ &\geq \sum_{j=0}^{k} \frac{1}{j!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{j-1}{n}\right).\end{aligned}$$

Part (b). From (a)(i) and (ii), we have

$$\sum_{j=0}^{k} \frac{1}{j!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{j-1}{n}\right) \leq \left(1+\frac{1}{n}\right)^n \leq \sum_{j=0}^{n} \frac{1}{j!}.$$

As $k, n \to \infty$, we have

$$\sum_{j=0}^{\infty} \frac{1}{j!} \leq \mathrm{e} \leq \sum_{j=0}^{\infty} \frac{1}{j!}.$$

By the Squeeze Theorem, we have

$$\mathrm{e} = \sum_{j=0}^{\infty} \frac{1}{j!}.$$

Part (c). Observe that

$$\begin{aligned}
s_m - s_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \dots + \frac{1}{m!} \\
&= \frac{1}{(n+1)!}\left[1 + \frac{1}{n+2} + \frac{1}{(n+3)(n+2)} + \dots + \frac{1}{m(m-1)\dots(n+2)}\right] \\
&< \frac{1}{(n+1)!}\left[1 + \frac{1}{(n+1)} + \frac{1}{(n+1)^2} + \dots\right] \\
&= \frac{1}{(n+1)!}\left[\frac{1}{1-1/(n+1)}\right] = \frac{1}{(n+1)!}\left[\frac{n+1}{n}\right] = \frac{1}{n\,(n!)}.
\end{aligned}$$

Note also that $1/j! > 0$ for all natural j, so $s_m > s_n$. This gives the inequality

$$0 < s_m - s_n < \frac{1}{n\,(n!)}.$$

As $m \to \infty$, we have

$$0 < \mathrm{e} - \sum_{j=0}^{n} \frac{1}{j!} < \frac{1}{n\,(n!)}$$

for all $n \in \mathbb{N}$.

Part (d). Taking $n = 1$, the above inequality becomes

$$0 < \mathrm{e} - \left(\frac{1}{0!} + \frac{1}{1!}\right) < \frac{1}{1\,(1!)} \implies 0 < \mathrm{e} - 2 < 1 \implies 2 < \mathrm{e} < 3.$$

Part (e). Seeking a contradiction, suppose e is rational. Write $\mathrm{e} = a/b$, where $a, b \in \mathbb{Z}$ with $b \neq 0$. From part (d), e is not an integer, so $b \neq 1$. Define

$$x = b!\left(\mathrm{e} - \sum_{j=0}^{b} \frac{1}{j!}\right).$$

Note that $x > 0$.

Firstly, observe that we can write

$$x = a\,(b-1)! - \sum_{j=0}^{b} \frac{b!}{j!}.$$

Since $b!/j! \in \mathbb{Z}$ for all $b \geq j$, it follows that $x \in \mathbb{Z}$.

Now, observe that

$$0 < x = b!\left(\mathrm{e} - \sum_{j=0}^{b} \frac{1}{j!}\right) \leq b!\left(\frac{1}{b\,(b!)}\right) = \frac{1}{b} < 1,$$

since $b \neq 1$. This implies that $x \notin \mathbb{Z}$, a contradiction. Thus, e must be irrational.

* * * * *

Problem 10.

(a) Taking x to be positive, expand $\left(2 + x^2\right)^{1/2}$ in a series of decreasing powers of x, as far as the term in x^{-5}, and state the set of values of x for which the expansion is valid.

(b) Given that $f(x) = a + bx + cx^2 - x\left(2 + x^2\right)^{1/2}$, where a, b and c are constants, and that $f(x) \to 0$ as $x \to \infty$, show that $a = 1$, $b = 0$ and $c = 1$.

(c) Replacing a, b and c by the values given in (b), obtain an expression for the area $A(X)$ between the curve $y = f(x)$ and the x-axis for $0 \leq x \leq X$, where $X > 0$.

(d) Determine the limit $A(X)$ as $X \to \infty$.

Solution.

Part (a). We have

$$\begin{aligned}\left(2 + x^2\right)^{1/2} &= x\left(1 + \frac{2}{x}\right)^{1/2} \\ &= x\left[1 + \frac{1}{2}\left(\frac{2}{x^2}\right) + \frac{(1/2)(-1/2)}{2}\left(\frac{2}{x^2}\right)^2 + \frac{(1/2)(-1/2)(-3/2)}{6}\left(\frac{2}{x^2}\right)^3 + \mathrm{O}\left(x^{-8}\right)\right] \\ &= x + x^{-1} - \frac{1}{2}x^{-3} + \frac{1}{2}x^{-5} + \mathrm{O}\left(x^{-7}\right).\end{aligned}$$

The radius of convergence is given by

$$\left|\frac{2}{x^2}\right| = \frac{2}{x^2} < 1 \implies x > \sqrt{2}.$$

Note that we reject $x < -\sqrt{2}$ since $x > 0$.

Part (b). Note that

$$f(x) = a + bx + cx^2 - x\left(2 + x^2\right)^{1/2} = (a - 1) + bx + (c - 1)\,x^2 + \mathrm{O}\left(x^{-2}\right).$$

For $f(x) \to 0$ as $x \to \infty$, we must have $a - 1 = b = c - 1 = 0$, whence $a = 1$, $b = 0$ and $c = 1$.

Part (c). We now have

$$f(x) = 1 + x^2 - x\left(2 + x^2\right)^{1/2}.$$

Note that $f(x)$ is always positive:

$$\begin{aligned}1 + x^2 - x\left(2 + x^2\right)^{1/2} > 0 &\iff \left(1 + x^2\right)^2 > x^2\left(2 + x^2\right) \\ &\iff 1 + 2x^2 + x^4 > 2x^2 + x^4 \\ &\iff 1 > 0.\end{aligned}$$

Thus,

$$\begin{aligned}A(x) &= \int_0^X |f(x)|\;\mathrm{d}x = \int_0^X \left[1 + x^2 - x\left(2 + x^2\right)^{1/2}\right]\mathrm{d}x \\ &= \int_0^X \left(1 + x^2\right)\mathrm{d}x - \frac{1}{2}\int_0^X \sqrt{2 + x^2}\,\mathrm{d}\left(2 + x^2\right) = \left[x + \frac{x^3}{3} - \frac{2}{3}\left(2 + x^2\right)^{3/2}\right]_0^X \\ &= X + \frac{1}{3}X^3 - \frac{1}{3}\left(2 + X^2\right)^{3/2} + \frac{2^{3/2}}{3}.\end{aligned}$$

Part (d). Note that

$$\begin{aligned}\frac{1}{3}\left(2 + X^2\right)^{3/2} &= \frac{1}{3}\left(2 + X^2\right)\left(2 + X^2\right)^{1/2} = \frac{1}{3}\left(2 + X^2\right)\left(X + X^{-1} - \frac{1}{2}X^{-3}\right) \\ &= \frac{1}{3}\left(3X + X^3 + \mathrm{O}\left(X^{-1}\right)\right) = X + \frac{1}{3}X^3 + \mathrm{O}\left(X^{-1}\right).\end{aligned}$$

Thus,

$$A(X) = -\,\mathrm{O}\left(X^{-1}\right) + \frac{2^{3/2}}{3},$$

so

$$\lim_{X\to\infty} A(X) = \frac{2^{3/2}}{3}.$$

* * * * *

Problem 11. Express $\left(x^2+2\right) / \left(x^3+1\right)$ in partial fractions. Hence, or otherwise, show that the coefficient of x^n in the expansion in ascending powers of x of $\left(1-x+x^2\right)^{-1}$ is given by

$$\begin{cases} (-1)^m, & n = 3m, \\ (-1)^m, & n = 3m+1, \\ 0, & n = 3m+2, \end{cases}$$

where m is a non-negative integer.

Solution. Note that

$$\frac{x^2+2}{x^3+1} = \frac{x^2+2}{(x+1)\left(x^2-x+1\right)} = \frac{A}{x+1} + \frac{Bx+C}{x^2-x+1}.$$

By the cover-up rule, we immediately have $A = 1$. Multiplying through by x^3+1, we get

$$x^2+2 = \left(x^2-x+1\right) + (x+1)\left(Bx+C\right).$$

Comparing constant terms, we get $2 = 1 + C \implies C = 1$. Comparing x^2 terms, we have $1 = 1 + B \implies B = 0$. Thus,

$$\frac{x^2+2}{x^3+1} = \frac{1}{x+1} + \frac{1}{x^2-x+1}.$$

Now observe that

$$\frac{1}{x+1} = \sum_{n=0}^{\infty}(-x)^n = \sum_{n=0}^{\infty}(-1)^n x^n$$

and

$$\begin{aligned} \frac{x^2+2}{x^3+1} &= \left(x^2+2\right)\sum_{m=0}^{\infty}\left(-x^3\right)^m = \left(x^2+2\right)\sum_{m=0}^{\infty}(-1)^m x^{3m} \\ &= \sum_{m=0}^{\infty}(-1)^m x^{3m+2} + \sum_{m=0}^{\infty}2(-1)^m x^{3m}. \end{aligned}$$

Hence,

$$\frac{1}{x^2-x+1} = \frac{x^2+2}{x^3+1} - \frac{1}{x+1} = \underbrace{\sum_{m=0}^{\infty}(-1)^m x^{3m+2}}_{S_1} + \underbrace{\sum_{m=0}^{\infty}2(-1)^m x^{3m}}_{S_2} - \underbrace{\sum_{n=0}^{\infty}(-1)^n x^n}_{S_3}.$$

Case 1. Suppose $n = 3m$ for some $m \in \mathbb{N}_0$. S_1 does not contribute anything, while S_2 contributes a coefficient of $2(-1)^m$ and S_3 contributes a coefficient of $-(-1)^n$. Thus, the coefficient of x^n is

$$2(-1)^m - (-1)^n = 2(-1)^m - (-1)^{3m} = 2(-1)^m - (-1)^m = (-1)^m.$$

Case 2. Suppose $n = 3m + 1$ for some $m \in \mathbb{N}_0$. S_1 and S_2 do not contribute anything, while S_3 contributes a coefficient of $-(-1)^n$. Thus, the coefficient of x^n is

$$-(-1)^n = -(-1)^{3m+1} = (-1)^{3m+2} = (-1)^m.$$

Case 3. Suppose $n = 3m+2$ for some $M \in \mathbb{N}_0$. S_1 contributes a coefficient of $(-1)^m$. S_2 does not contribute anything. S_3 contributes a coefficient of $-(-1)^n$. Thus, the coefficient of x^n is

$$(-1)^m - (-1)^n = (-1)^m - (-1)^{3m+2} = (-1)^m - (-1)^m = 0.$$

* * * * *

Problem 12. The expansion of $E = (1 + x)^n/(1 - x)$ in ascending powers of x, for small values of x and for any real value of n, is denoted by

$$1 + p_1(n)x + p_2(n)x^2 + \cdots + p_r(n)x^r + \ldots.$$

Show that $p_r(n)$ is a polynomial in n of degree r, given by

$$p_r(n) = 1 + n + \frac{n(n-1)}{2} + \frac{n(n-1)(n-2)}{3!} + \cdots + \frac{n(n-1)(n-2)\ldots(n-r+1)}{r!}.$$

By putting $n = -1$ in E and its expansion, and by using the factor theorem, or otherwise, show that

(a) when r is odd, $(n + 1)$ is a factor of the polynomial $p_r(n)$.

(b) when r is even, $(n + 1)$ is a factor of the polynomial $p_r(n) - 1$.

(c) Deduce that if $f(n)$ and $g(n)$ are any polynomials in n such that

$$F(n) = [f(n) + g(n)]\, p_r(n) - g(n),$$

then $F(-1) = -g(-1)$ if r is odd and $F(-1) = f(-1)$ if r is even.

(d) Prove that, if N is a positive integer, $p_r(N) = 2^N$ for all $r \geq N$.

Solution. Observe that

$$E = \frac{(1+x)^n}{1-x} = \left(\sum_{i=0}^{n}\binom{n}{i}x^i\right)\left(\sum_{j=0}^{\infty}x^j\right) = \sum_{r=0}^{\infty}\left(\sum_{i+j=r}\binom{n}{i}\right)x^r.$$

Thus,

$$\begin{aligned} p_r(n) &= \sum_{i+j=r}\binom{n}{i} = \sum_{i=0}^{r}\binom{n}{i} = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{r} \\ &= 1 + n + \frac{n(n-1)}{2} + \frac{n(n-1)(n-2)}{3!} + \cdots + \frac{n(n-1)(n-2)\ldots(n-r+1)}{r!}. \end{aligned}$$

Observe that

$$p_r(-1) = 1 - 1 + \frac{(-1)(-2)}{2!} + \cdots + \frac{(-1)(-2)\ldots(-r)}{r!} = 1 - 1 + 1 + 1 + \cdots \pm 1.$$

Since $p_r(n)$ has degree r, there are $r + 1$ terms above.

Part (a). When k is odd, there are an even number of terms, so $p_r(-1) = 0$. By the Factor Theorem, it follows that $x + 1$ is a factor of $p_r(n)$.

Part (b). When k is even, there are an odd number of terms, so $p_r(-1) = 1$, whence $p_r(-1) - 1 = 0$. By the Factor Theorem, it follows that $x + 1$ is a factor of $p_r(n) - 1$.

Part (c). If r is odd, we have

$$F(-1) = [f(-1) + g(-1)]\underbrace{p_r(-1)}_{0} - g(-1) = -g(-1).$$

If r is even,

$$F(-1) = [f(-1) + g(-1)]\underbrace{p_r(-1)}_{1} - g(-1) = f(-1) + g(-1) - g(-1) = f(-1).$$

Part (d). For all $r \geq N$,

$$p_r(N) = \sum_{i=0}^{r} \binom{N}{i} = \sum_{i=0}^{N} \binom{N}{i} + \sum_{i=N+1}^{r} \binom{N}{i} = 2^N + 0 = 2^N.$$

Analysis 3 Inequalities

Tutorial A3

Problem 1. If a, b, c are sides of a triangle, show that

$$\frac{a}{b+c-a}+\frac{b}{a+c-b}+\frac{c}{a+b-c}\geq 3.$$

Proof. Let

$$x=\frac{a+b-c}{2},\quad y=\frac{b+c-a}{2},\quad z=\frac{c+a-b}{2}.$$

By the triangle inequality, we have $x,y,z>0$. Hence, by AM-GM,

$$2\sqrt{xy}\leq x+y,\quad 2\sqrt{yz}\leq y+z,\quad 2\sqrt{zx}\leq z+x,$$

so

$$8xyz\leq(x+y)(y+z)(z+x).$$

Replacing x,y,z with their corresponding definitions in a,b,c, we get

$$\frac{(x+y)(y+z)(z+x)}{8xyz}=\frac{abc}{(b+c-a)(a+c-b)(a+b-c)}\geq 1.$$

Thus, by AM-GM,

$$\frac{a}{b+c-a}+\frac{b}{a+c-b}+\frac{c}{a+b-c}\geq\frac{3abc}{(b+c-a)(a+c-b)(a+b-c)}\geq 3$$

as desired. □

* * * * *

Problem 2.

(a) For some positive integer n, let $x_1\leq x_2\leq\cdots\leq x_n$ and $y_1\leq y_2\leq\cdots\leq y_n$ be real numbers. By considering the sum of all n^2 terms of the form $(x_i-x_j)(y_i-y_j)$, prove that

$$\sum_{i=1}^{n}x_iy_i\geq\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right).$$

(b) Let a triangle have angles A, B, C and let the lengths of the opposite sides be a, b, c. By applying the result of (a), prove that $aA+bB+cC\geq\frac{1}{3}\pi(a+b+c)$.

(c) Let a, b, c be three positive numbers such that $a^2+b^2+c^2=1$. By applying the result of part (a) with

$$\{x_i\}=\left\{\frac{a+b}{c},\frac{c+a}{b},\frac{b+c}{a}\right\},$$

find the maximum possible value of

$$\frac{(a+b)\left(a^2+b^2\right)}{c}+\frac{(c+a)\left(c^2+a^2\right)}{b}+\frac{(b+c)\left(b^2+c^2\right)}{a}.$$

Solution.

Part (a). Because $x_1 \le x_2 \le \cdots \le x_n$ and $y_1 \le y_2 \le \cdots \le y_n$, we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_j)(y_i - y_j) \ge 0.$$

Meanwhile, we can manipulate the sum as

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_j)(y_i - y_j) &= \sum_{i=1}^{n}\sum_{j=1}^{n}(x_iy_i + x_jy_j - x_iy_j - x_jy_i)\\ &= 2\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_i - 2\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_j\\ &= 2n\sum_{i=1}^{n}x_iy_i - 2\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{j=1}^{n}y_j\right).\end{aligned}$$

Thus,

$$2n\sum_{i=1}^{n}x_iy_i - 2\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{j=1}^{n}y_j\right) \ge 0,$$

and our inequality follows immediately.

Part (b). Without loss of generality, suppose $a \le b \le c$. Consider the sine rule:

$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}.$$

We now show that $A \le B \le C$.

Case 1. If $C \ge \pi/2$, then $A, B < \pi/2$. Since $\sin\theta$ is increasing on $[0, \pi/2]$, it is obvious by the sine rule that we must have $A \le B$, so $A \le B < \pi/2 \le C$.

Case 2. If $C < \pi/2$, then $A, B, C \le \pi/2$. Since $\sin\theta$ is increasing on $[0, \pi/2]$, it is obvious by the sine rule that $A \le B \le C$.

We thus have $a \le b \le c$ and $A \le B \le C$. Applying (a), we have

$$aA + bB + cC \ge \frac{1}{3}(a+b+c)(A+B+C) = \frac{\pi}{3}(a+b+c),$$

since the sum of angles in a triangle is $A + B + C = \pi$.

Part (c). Without loss of generality, assume $a \le b \le c$. Then

$$\frac{a+b}{c} \le \frac{c+a}{b} \le \frac{b+c}{a}$$

and

$$a^2 + b^2 \le c^2 + a^2 \le b^2 + c^2.$$

Applying (a), we have

$$\begin{aligned}&\frac{(a+b)\left(a^2+b^2\right)}{c} + \frac{(c+a)\left(c^2+a^2\right)}{b} + \frac{(b+c)\left(b^2+c^2\right)}{a}\\ &= \frac{1}{3}bp\frac{a+b}{c} + \frac{c+a}{b} + \frac{b+c}{a}\left[\left(a^2+b^2\right) + \left(c^2+a^2\right) + \left(b^2+c^2\right)\right]\\ &= \frac{2}{3}\left(\frac{a+b}{c} + \frac{c+a}{b} + \frac{b+c}{a}\right)\left(a^2+b^2+c^2\right)\\ &= \frac{2}{3}\left(\frac{a+b}{c} + \frac{c+a}{b} + \frac{b+c}{a}\right).\end{aligned}$$

By AM-GM,

$$\begin{aligned}
\frac{a+b}{c}+\frac{c+a}{b}+\frac{b+c}{a} &\geq 3\sqrt[3]{\frac{(a+b)(c+a)(b+c)}{abc}} \\
&\geq 3\sqrt[3]{\frac{\left(2\sqrt{ab}\right)(2\sqrt{ca})\left(2\sqrt{bc}\right)}{abc}} \\
&= 3\cdot 2 = 6.
\end{aligned}$$

Thus,

$$\frac{(a+b)\left(a^2+b^2\right)}{c}+\frac{(c+a)\left(c^2+a^2\right)}{b}+\frac{(b+c)\left(b^2+c^2\right)}{a} \geq \frac{2}{3}\cdot 6 = 4.$$

The minimum value of 4 is attained when $a=b=c=\sqrt[3]{1/3}$.

* * * * *

Problem 3. By considering $a_i = \sqrt[n]{x_i}$ for $i=1,\ldots,n$, show that for all positive real numbers $x_1, x_2, \ldots, x_n$ such that $x_1 x_2 \ldots x_n = 1$, the following inequality holds:

$$\frac{1}{n-1+x_1}+\frac{1}{n-1+x_2}+\cdots+\frac{1}{n-1+x_n} \leq 1.$$

Proof. Observe that $a_1 \ldots a_n = \sqrt[n]{x_1 \ldots x_n} = 1$, so

$$x_i = \frac{a_i^n}{a_1 \ldots a_n} = \frac{a_i^{n-1}}{a_1 \ldots a_{i-1} a_{i+1} \ldots a_n}.$$

By AM-GM, we have

$$\begin{aligned}
a_1 \ldots a_{i-1} a_{i+1} \ldots a_n &= \sqrt[n-1]{a_1^{n-1} \ldots a_{i-1}^{n-1} a_{i+1}^{n-1} \ldots a_n^{n-1}} \\
&\leq \frac{a_1^{n-1}+\cdots+a_{i-1}^{n-1}+a_{i+1}^{n-1}+\cdots+a_n^{n-1}}{n-1}.
\end{aligned}$$

Thus,

$$x_i \geq \frac{(n-1)\, a_i^{n-1}}{a_1^{n-1}+\cdots+a_{i-1}^{n-1}+a_{i+1}^{n-1}+\cdots+a_n^{n-1}}.$$

It follows that

$$\begin{aligned}
n-1+x_i &\geq n-1+\frac{(n-1)\, a_i^{n-1}}{a_1^{n-1}+\cdots+a_{i-1}^{n-1}+a_{i+1}^{n-1}+\cdots+a_n^{n-1}} \\
&= \frac{(n-1)\left(a_1^{n-1}+\cdots+a_n^{n-1}\right)}{a_1^{n-1}+\cdots+a_{i-1}^{n-1}+a_{i+1}^{n-1}+\cdots+a_n^{n-1}}.
\end{aligned}$$

Reciprocating, we see that

$$\frac{1}{n-1+x_i} \leq \frac{a_1^{n-1}+\cdots+a_{i-1}^{n-1}+a_{i+1}^{n-1}+\cdots+a_n^{n-1}}{(n-1)\left(a_1^{n-1}+\cdots+a_n^{n-1}\right)} = \frac{\sum_{k=1}^n a_k - a_i}{(n-1)\sum_{k=1}^n a_k}.$$

Summing over $i=1,\ldots,n$, we finally get

$$\frac{1}{n-1+x_1}+\cdots+\frac{1}{n-1+x_n} \leq \sum_{i=1}^n \frac{\sum_{k=1}^n a_k - a_i}{(n-1)\sum_{k=1}^n a_k} = \frac{(n-1)\sum_{k=1}^n a_k}{(n-1)\sum_{k=1}^n a_k} = 1.$$

□

Problem 4.

(a) For all positive real numbers x, y, z, prove that

$$\left(\frac{x}{y}\right)^2+\left(\frac{y}{z}\right)^2+\left(\frac{z}{x}\right)^2 \geq \frac{x}{z}+\frac{y}{x}+\frac{z}{y} \geq 3.$$

(b) (i) Let $\mathbf{a}=(a_1,\, a_2,\, a_3)^\mathsf{T}$ and $\mathbf{b}=(b_1,\, b_2,\, b_3)^\mathsf{T}$ be two non-zero vectors. By considering the scalar product of $\mathbf{a}$ and $\mathbf{b}$, or otherwise, prove that

$$\left(\sum_{i=1}^{3} a_i b_i\right)^2 \leq \left(\sum_{i=1}^{3} a_i^2\right)\left(\sum_{i=1}^{3} b_i^2\right)$$

and state the necessary condition for equality to hold.

(ii) Hence, for all positive real numbers x, y, z, prove that

$$x+y+z \leq 2\left(\frac{x^2}{y+z}+\frac{y^2}{z+x}+\frac{z^2}{x+y}\right).$$

Solution.

Part (a). We begin with the first inequality. Clearly,

$$a^2+b^2 \geq 2ab, \quad b^2+c^2 \geq 2bc, \quad c^2+a^2 \geq 2ca.$$

Adding these three inequalities and dividing by two yields

$$a^2+b^2+c^2 \geq ab+bc+ca.$$

Now, let

$$a=\frac{x}{y}, \quad b=\frac{y}{z}, \quad c=\frac{z}{x}.$$

Then

$$\left(\frac{x}{y}\right)^2+\left(\frac{y}{z}\right)^2+\left(\frac{z}{x}\right)^2 \geq \frac{x}{z}+\frac{y}{x}+\frac{z}{y}$$

as desired.

We now prove the second inequality. By AM-GM,

$$\frac{x}{z}+\frac{y}{x}+\frac{z}{y} \geq 3\sqrt[3]{\frac{xyz}{zxy}}=3.$$

Putting both inequalities together, we have

$$\left(\frac{x}{y}\right)^2+\left(\frac{y}{z}\right)^2+\left(\frac{z}{x}\right)^2 \geq \frac{x}{z}+\frac{y}{x}+\frac{z}{y} \geq 3.$$

Part (b).

Part (b)(i). Observe that

$$(\mathbf{a}\cdot\mathbf{b})^2=(a_1b_1+a_2b_2+a_3b_3)^2=\left(\sum_{i=1}^{3} a_i b_i\right)^2.$$

We also have

$$(\mathbf{a}\cdot\mathbf{b})^2=\left(|\mathbf{a}|\,|\mathbf{b}|\cos\theta\right)^2,$$

where θ is the angle between the two vectors. Since $\cos^2 t \in [0,1]$, it follows that

$$\left(\sum_{i=1}^{3} a_i b_i\right)^2 = (\mathbf{a}\cdot\mathbf{b})^2 \le |\mathbf{a}|^2\,|\mathbf{b}|^2 = \left(a_1^2+a_2^2+a_3^2\right)\left(b_1^2+b_2^2+b_3^2\right) = \left(\sum_{i=1}^{3} a_i^2\right)\left(\sum_{i=1}^{3} b_i^2\right).$$

Equality holds when $\cos\theta = \pm 1$, i.e. when $\mathbf{a}$ is parallel to $\mathbf{b}$. Equivalently, $a_i = kb_i$ for all $i = 1, 2, 3$, where k is a constant.

Part (b)(ii). From part (b)(i), we have

$$[(y+z)+(z+x)+(x+y)]\left(\frac{x^2}{y+z}+\frac{y^2}{z+x}+\frac{z^2}{x+y}\right) \ge (x+y+z)^2.$$

The first term on the LHS is simply $2(x+y+z)$, so dividing both sides by $x+y+z$ yields the desired inequality:

$$2\left(\frac{x^2}{y+z}+\frac{y^2}{z+x}+\frac{z^2}{x+y}\right) \ge x+y+z.$$

* * * * *

Problem 5. Prove by induction, the AM-GM inequality for general n.

Proof. Let $P(n)$ be the statement

$$\frac{1}{n}\sum_{i=1}^{n} x_i \ge \prod_{i=1}^{n} x_i^{1/n}$$

for all $x_1, \ldots, x_n > 0$.

Note that $P(1)$ is trivially true:

$$\frac{1}{1}\sum i = 1^1 x_i = x_1 = \prod_{i=1}^{1} x_i^{1/1}.$$

$P(2)$ can also be easily proven:

$$(x_1-x_2)^2 \ge 0 \implies x_1^2+x_2^2-2x_1x_2 \ge 0 \implies \frac{x_1^2+2x_1x_2+x_2^2}{4} \ge x_1x_2$$
$$\implies \left(\frac{x_1+x_2}{2}\right)^2 \ge x_1x_2 \implies \frac{x_1+x_2}{2} \ge \sqrt{x_1x_2}.$$

Suppose $P(k)$ is true for some $k \in \mathbb{N}$. We now show that $P(k) \implies P(2k)$. By our inductive hypothesis,

$$\frac{1}{2n}\sum_{i=1}^{2n} x_i = \frac{1}{2}\left[\frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum_{i=n+1}^{2n} x_i\right] \ge \frac{1}{2}\left(\prod_{i=1}^{n} x_i^{1/n} + \prod_{i=n+1}^{2n} x_i^{1/n}\right).$$

By $P(2)$, we have

$$\frac{1}{2n}\sum_{i=1}^{2n} x_i \ge \sqrt{\prod_{i=1}^{n} x_i^{1/n}\prod_{i=n+1}^{2n} x_i^{1/n}} = \prod_{i=1}^{2n} x_i^{1/2n}.$$

Thus, $P(k) \implies P(2k)$.

Suppose now that $P(k+1)$ is true for some $k+1 \in \mathbb{N}$. We now show that $P(k+1) \implies P(k)$. Taking

$$x_{k+1} = \frac{x_1 + \cdots + x_k}{k},$$

by $P(k+1)$, we have

$$\frac{x_1 + \cdots + x_k + \frac{x_1+\cdots+x_k}{k}}{k+1} \geq \prod_{i=1}^{k} x_i^{1/(k+1)} \left(\frac{x_1 + \cdots + x_k}{k}\right)^{1/(k+1)}.$$

The LHS simplifies to

$$\frac{x_1 + \cdots + x_k}{k} \geq \prod_{i=1}^{k} x_i^{1/(k+1)} \left(\frac{x_1 + \cdots + x_k}{k}\right)^{1/(k+1)},$$

so

$$\left(\frac{x_1 + \cdots + x_k}{k}\right)^{k/k+1} \geq \prod_{i=1}^{k} x_i^{1/(k+1)},$$

Raising both sides to the $(k+1)/k$th power, we finally have

$$\frac{x_1 + \cdots + x_k}{k} \geq \prod_{i=1}^{k} x_i^{k},$$

so $P(k+1) \implies P(k)$.

From the base cases $P(1)$ and $P(2)$, along with the results $P(k) \implies P(2k)$ and $P(k+1) \implies P(k)$, it stands to reason that $P(k)$ is true for all $k \in \mathbb{N}$. □

* * * * *

Problem 6 (Nesbitt's Inequality). For positive real numbers a, b, c, prove that

$$\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} \geq \frac{3}{2}$$

using

(a) AM-GM inequality,

(b) Cauchy-Schwarz inequality.

Proof of (a). Let

$$x = a+b, \quad y = b+c, \quad z = c+a.$$

By AM-GM,

$$\frac{z+x}{y} + \frac{y+z}{x} + \frac{x+y}{z} \geq 6\sqrt[6]{\frac{zxyzxy}{yyxxzz}} = 6.$$

Substituting a, b, c back in, we see that

$$\frac{2a+b+c}{b+c} + \frac{a+2b+c}{c+a} + \frac{a+b+2c}{a+b} = 3 + \frac{2a}{b+c} + \frac{2b}{c+a} + \frac{2c}{a+b} \geq 6.$$

Our desired inequality follows immediately. □

Proof of (b). By Cauchy-Schwarz,

$$[(b+c)+(c+a)+(a+b)]\left(\frac{1}{b+c}+\frac{1}{c+a}+\frac{1}{a+b}\right) \geq (1+1+1)^2 = 9,$$

so

$$(a+b+c)\left(\frac{1}{b+c}+\frac{1}{c+a}+\frac{1}{a+b}\right) \geq \frac{9}{2},$$

from which it immediately follows that

$$\frac{a}{b+c}+\frac{b}{c+a}+\frac{c}{a+b} \geq \frac{9}{2}-3 = \frac{3}{2}.$$

□

* * * * *

Problem 7 (Carlson's Inequality). Consider n arbitrary real numbers $x_1, x_2, \ldots, x_n$. Show that

$$(x_1+x_2+x_3+\cdots+x_n)^2 \leq \frac{\pi^2}{6}\left(x_1^2+4x_2^2+9x_3^3+\cdots+n^2x_n^2\right).$$

You may use the well-known result $\sum_{r=1}^{\infty} 1/r^2 = \pi^2/6$.

Proof. By Cauchy-Schwarz,

$$\left(\sum_{i=1}^{n} x_i\right)^2 \leq \left(\sum_{i=1}^{n} \frac{1}{i^2}\right)\left(\sum_{i=1}^{n} i^2x_i^2\right) \leq \left(\sum_{i=1}^{\infty} \frac{1}{i^2}\right)\left(\sum_{i=1}^{n} i^2x_i^2\right),$$

so

$$(x_1+x_2+x_3+\cdots+x_n)^2 \leq \frac{\pi^2}{6}\left(x_1^2+4x_2^2+9x_3^3+\cdots+n^2x_n^2\right)$$

as desired. □

* * * * *

Problem 8. Suppose $x, y, z > 0$ and $x+y+z=1$. Show that

$$\frac{1}{x}+\frac{4}{y}+\frac{9}{z} \geq 36.$$

Proof. By Cauchy-Schwarz,

$$(x+y+z)\left(\frac{1}{x}+\frac{4}{y}+\frac{9}{z}\right) \geq (1+2+3)^2.$$

Since $x+y+z=1$, we have our desired result. □

* * * * *

Problem 9 (🌶️). The Archimedean property of the real numbers states that for any $x \in \mathbb{R}$, there exists $n \in \mathbb{N}$ such that $x < n$.

The well-ordering principle states that any non-empty subset of positive integers $S \subseteq \mathbb{Z}^+$ has a smallest element.

(a) Let $x, y \in \mathbb{R}$ be such that $1 < x < y$. Use the Archimedean property to show that there exists $n \in \mathbb{N}$ such that $ny > 1 + nx$.

(b) Let $S = \{m \in \mathbb{N} : m > nx\}$. Use the well-ordering principle to show that there exists $m_0 \in S$ such that $m_0 - 1 \notin S$, $m_0 - 1 \in \mathbb{N}$ and $m_0 > nx$.

(c) Deduce that there exists a rational number $r \in \mathbb{Q}$ such that $x < r < y$.

(d) Hence, prove that for any two real numbers $x, y \in \mathbb{R}$ satisfying $x < y$, there exists a rational number $t \in \mathbb{Q}$ such that $x < t < y$.

(e) Deduce that for any two real numbers $x, y \in \mathbb{R}$ satisfying $x < y$, there exists an irrational number $u \notin \mathbb{Q}$ such that $x < u < y$.

Solution.

Part (a). Since $x < y$, we have $y - x \neq 0$, so $1/(y - x)$ is real. By the Archimedean property, there exists some $n \in \mathbb{N}$ such that

$$n > \frac{1}{y-x} \implies n(y-x) > 1 \implies ny > 1 + nx,$$

which is what we wanted.

Part (b). Since $nx > 0$, we clearly have $S \subseteq \mathbb{Z}^+$. By the well-ordering principle, there exists a smallest element of S. Let m_0 be this smallest element. Clearly, $m_0 \in S$ so $m_0 > nx$. Further, $m_0 - 1 \notin S$; if it were, then $m_0 - 1$ would be the smallest element, contradicting the minimality of m_0.

Part (c). From (b), we have $m_0 - 1 \leq nx$. Adding 1 on both sides and invoking (a),

$$m_0 \leq nx + 1 < ny \implies \frac{m_0}{n} < y.$$

Meanwhile, from (b), we also have

$$nx < m_0 \implies x < \frac{m_0}{n}.$$

Putting the two inequalities together, we see that

$$x < \frac{m_0}{n} < y.$$

Taking $r = m_0/n$, which is rational ($m_0, n \in \mathbb{N}$), we are done.

Part (d). *Case 1*. If $1 < x < y$, we are done by (c).

Case 2. Suppose $x < y < -1$. Then $1 < -y < -x$. Applying (c), there exists a rational number $-r$ such that $-y < -r < -x$, so $x < r < y$ and we are done.

Case 3. If $x < 1 < y$, we simply take $r = 1$.

Part (e). *Case 1*. Suppose x or y is irrational. Without loss of generality, we take x to be irrational. Taking

$$u = \frac{x + m_0/n}{2},$$

which is clearly irrational, we observe that

$$x < \frac{x + m_0/n}{2} < \frac{m_0}{n} < y,$$

and we are done.

Case 2. Suppose both x and y are rational. Then we take

$$u = x + (y - x)\frac{\sqrt{2}}{2},$$

which is clearly irrational and strictly between x and y since $\left|\sqrt{2}/2\right| < 1$.

Problem 10 (🌶).

(a) Let $f : [0, \infty) \to \mathbb{R}$ be a strictly increasing differentiable function such that $f(0) = 0$. With the aid of a diagram, prove Young's inequality for increasing functions:

$$ab \leq \int_0^a f(x)\,\mathrm{d}x + \int_0^b f^{-1}(x)\,\mathrm{d}x$$

for all $a, b > 0$.

(b) Prove Young's inequality for products:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

where $1/p + 1/q = 1$. Deduce that

$$\frac{|ab|}{cd} \leq \frac{|a|^p}{c^p p} + \frac{|b|^q}{d^q q}$$

for any $a, b \in \mathbb{R}$ and $c, d > 0$.

Hölder's inequality states that if $p, q > 1$ is such that $1/p + 1/q = 1$, then

$$\sum_{i=1}^{n} |a_i b_i| \leq \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p} \left(\sum_{i=1}^{n} |b_i|^q \right)^{1/q}$$

for any $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \in \mathbb{R}$ and any $n \in \mathbb{N}$.

(c) Let $r \geq 1$, $n \in \mathbb{N}$ and let $c_1, c_2, \ldots, c_m, d_1, d_2, \ldots, d_m \in \mathbb{R}$.

(i) Explain why

$$\sum_{i=1}^{m} |c_i + d_i|^p \leq \sum_{i=1}^{m} \Big[(|c_i| + |d_i|)\, |c_i + d_i|^{p-1} \Big].$$

(ii) Show that $q(p-1) = p$.

(iii) Using Hölder's inequality, prove the Minkowski inequality:

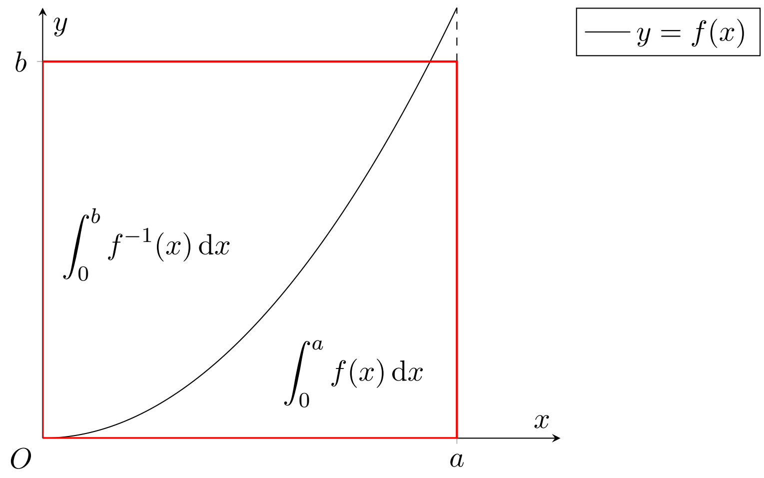
$$\sum_{i=1}^{m} |c_i + d_i|^p \leq \left[\left(\sum_{i=1}^{m} |c_i|^p \right)^{1/p} + \left(\sum_{i=1}^{m} |d_i|^p \right)^{1/p} \right]^p.$$

(iv) Hence, show that the following weighted inequality holds for all $w_1, \ldots, w_n > 0$:

$$\left(\sum_{i=1}^{m} |c_i + d_i|^p \, w_i \right)^{1/p} \leq \left(\sum_{i=1}^{m} |c_i|^p \, w_i \right)^{1/p} + \left(\sum_{i=1}^{m} |d_i|^p \, w_i \right)^{1/p}.$$

Solution.

Part (a).



The sum of the areas represented by the two integrals is bigger than the area of the rectangle, ab.

Part (b). Take $f(x) = x^{p-1}$, which is strictly increasing with $f(0) = 0$. Note that $f^{-1}(x) = x^{1/(p-1)}$. By (a), for all $a, b > 0$,

$$ab \leq \int_0^a x^{p-1}\,\mathrm{d}x + \int_0^b x^{\frac{1}{p-1}}\,\mathrm{d}x = \frac{a^p}{p} + \frac{p-1}{p} b^{\frac{p}{p-1}} = \frac{a^p}{p} + \frac{a^q}{q},$$

where $q = p/(p-1)$. Note that we indeed have $1/p + 1/q = 1$:

$$q = \frac{p}{p-1} \implies \frac{1}{q} + \frac{p-1}{p} = 1 - \frac{1}{p} \implies \frac{1}{p} + \frac{1}{q} = 1.$$

Under the transformations $a \mapsto |a|/c$, $b \mapsto |b|/d$, we see that

$$\frac{|ab|}{cd} \leq \frac{|a|^p}{c^p p} + \frac{|b|^q}{d^q q}$$

for all $a, b \in \mathbb{R}$ and $c, d > 0$.

Part (c).

Part (c)(i). The triangle inequality states that $|c_i| + |d_i| \geq |c_i + d_i|$. So

$$\sum_{i=1}^m \left[(|c_i| + |d_i|)\,|c_i + d_i|^{p-1}\right] \geq= \sum_{i=1}^m |c_i + d_i|\,|c_i + d_i|^{p-1} = \sum_{i=1}^m |c_i + d_i|^p\,.$$

Part (c)(ii). Since $q = p/(p-1)$, we trivially have $q(p-1) = p$.

Part (c)(iii). From (a),

$$\begin{aligned}\sum_{i=1}^m |c_i + d_i|^p &\leq \sum_{i=1}^m \left[(|c_i| + |d_i|)\,|c_i + d_i|^{p-1}\right] \\ &= \sum_{i=1}^m |c_i|\,|c_i + d_i|^{p-1} + \sum_{i=1}^m |d_i|\,|c_i + d_i|^{p-1}\,.\end{aligned}$$

Applying Hölder's inequality on both sums, we see that

$$
\begin{aligned}
&\sum_{i=1}^{m} |c_i + d_i|^p \\
&\quad \le \left(\sum_{i=1}^{m} |c_i|^p\right)^{1/p} \left(\sum_{i=1}^{m} |c_i + d_i|^{(p-1)q}\right)^{1/q} + \left(\sum_{i=1}^{m} |d_i|^p\right)^{1/p} \left(\sum_{i=1}^{m} |c_i + d_i|^{(p-1)q}\right)^{1/q} \\
&\quad = \left[\left(\sum_{i=1}^{m} |c_i|^p\right)^{1/p} + \left(\sum_{i=1}^{m} |d_i|^p\right)^{1/p}\right] \left(\sum_{i=1}^{m} |c_i + d_i|^p\right)^{1/q}
\end{aligned}
$$

Dividing both sides by $(\sum_{i=1}^{m} |c_i + d_i|^p)^{1/q}$, we have our desired inequality:

$$
\left(\sum_{i=1}^{m} |c_i + d_i|^p\right)^{1/p} = \left(\sum_{i=1}^{m} |c_i + d_i|^p\right)^{1-1/q} \le \left(\sum_{i=1}^{m} |c_i|^p\right)^{1/p} + \left(\sum_{i=1}^{m} |d_i|^p\right)^{1/p}.
$$

Part (c)(iv). Take $c \mapsto cw^{1/p}$, $d \mapsto dw^{1/p}$ and invoke (c)(ii).

Analysis 4 Complex Numbers

Tutorial A4

Problem 1. Show that if $n, m \in \mathbb{Z}^+$, then the equations $z^n = 1 + \mathrm{i}$ and $z^m = 2 - \mathrm{i}$ have no common solutions for $z \in \mathbb{C}$.

Solution. Seeking a contradiction, suppose there exists some $z \in \mathbb{C}$ that satisfies the given equations. Then we have

$$|z|^n = |1 + \mathrm{i}| = \sqrt{2}$$

and

$$|z|^m = |2 - \mathrm{i}| = \sqrt{5},$$

so

$$|z|^{2mn} = 2^m = 5^n,$$

but this is a clear contradiction of the Fundamental Theorem of Algebra, so such a z cannot exist, i.e. the two equations have no common solutions.

* * * * *

Problem 2. Let z and w be complex numbers such that $|z| = |w| = r > 0$. Show that

$$\mathrm{Re}\left(\frac{z+w}{z-w}\right) = 0.$$

Solution. Let $z = r\mathrm{e}^{\mathrm{i}\theta}$ and $w = r\mathrm{e}^{\mathrm{i}\varphi}$. Then

$$\frac{z+w}{z-w} = \frac{\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{\mathrm{i}\varphi}}{\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{\mathrm{i}\varphi}} = \frac{\mathrm{e}^{\mathrm{i}(\theta-\varphi)/2} + \mathrm{e}^{-\mathrm{i}(\theta-\varphi)/2}}{\mathrm{e}^{\mathrm{i}(\theta-\varphi)/2} - \mathrm{e}^{-\mathrm{i}(\theta-\varphi)/2}} = \frac{2\cos\left(\frac{\theta-\varphi}{2}\right)}{2\mathrm{i}\sin\left(\frac{\theta-\varphi}{2}\right)} = -\mathrm{i}\cot\left(\frac{\theta-\varphi}{2}\right),$$

which is purely imaginary, so the claim holds.

* * * * *

Problem 3. Suppose that the complex number z satisfies the equation $5\,(z+\mathrm{i})^n = (4+3\mathrm{i})\,(1+\mathrm{i}z)^n$. Show that z is purely real.

Solution. Taking the modulus on both sides, we see that

$$5\,|z+\mathrm{i}|^n = 5\,|1+\mathrm{i}z|^n \implies |z+\mathrm{i}| = |1+\mathrm{i}z| = |z-\mathrm{i}|\,.$$

The locus of z is the perpendicular bisector of the points representing i and $-$i, i.e. the real axis. Hence z is purely real.

* * * * *

Problem 4. Given that $w + z = w\,|z|$, where $w, z \in \mathbb{C}$ with $|w| > 1$, show that

$$\left(|z| + \frac{|w|}{1-|w|}\right)\left(|z| - \frac{|w|}{1+|w|}\right) = 0.$$

Solution. Solving for w, we get

$$w = \frac{z}{|z|-1} \implies |w| = \frac{|z|}{||z|-1|}.$$

Note that $|z| \neq 1$, for we would get $w + z = w$, whence $z = 0$, a contradiction. We thus have two cases to examine:

Case 1: $|z| - 1 > 0$. Then

$$|w| = \frac{|z|}{|z|-1} \implies |z| = \frac{|w|}{|w|-1} \implies |z| - \frac{|w|}{|w|-1} = 0.$$

Case 2: $|z| - 1 < 0$. Then

$$|w| = \frac{|z|}{1-|z|} \implies |z| = \frac{|w|}{1+|w|} \implies |z| - \frac{|w|}{1+|w|} = 0.$$

Putting both cases together, we see that

$$\left(|z| + \frac{|w|}{1-|w|}\right)\left(|z| - \frac{|w|}{1+|w|}\right) = 0$$

as desired.

* * * * *

Problem 5. Let z be a complex number such that $|z| = 1$. Find the minimum and maximum value of

$$\left|z^5 + (z^*)^3 + 6z\right| - 2\left|z^4 + 1\right|.$$

Solution. Let $z = \mathrm{e}^{\mathrm{i}\theta}$, where $\theta \in [0, 2\pi)$. Then

$$\begin{aligned}\left|z^5 + (z^*)^3 + 6z\right| &= \left|\mathrm{e}^{5\mathrm{i}\theta} + \mathrm{e}^{-3\mathrm{i}\theta} + 6\mathrm{e}^{\mathrm{i}\theta}\right| = \left|\mathrm{e}^{4\mathrm{i}\theta} + \mathrm{e}^{-4\mathrm{i}\theta} + 6\right| \\ &= |2\cos(4\theta) + 6| = \left|4\cos^2(2\theta) + 4\right| = 4\cos^2(2\theta) + 4\end{aligned}$$

and

$$\left|z^4 + 1\right| = \left|\mathrm{e}^{4\mathrm{i}\theta} + 1\right| = \left|\mathrm{e}^{2\mathrm{i}\theta} + \mathrm{e}^{-2\mathrm{i}\theta}\right| = |2\cos(2\theta)|.$$

Thus, the given expression is equal to

$$\left|z^5 + (z^*)^3 + 6z\right| - 2\left|z^4 + 1\right| = 4\cos^2(2\theta) - 4|\cos(2\theta)| + 4.$$

Case 1: $\cos(2\theta) \geq 0$. Then our expression becomes

$$4\cos^2(2\theta) - 4\cos(2\theta) + 4 = 4\left(\cos(2\theta) - \frac{1}{2}\right)^2 + 3,$$

which attains a maximum of 4 (when $\cos(2\theta) = 1$) and a minimum of 3 (when $\cos(2\theta) = 1/2$).

Case 2: $\cos(2\theta) < 0$. Then our expression becomes

$$4\cos^2(2\theta) + 4\cos(2\theta) + 4 = 4\left(\cos(2\theta) + \frac{1}{2}\right)^2 + 3,$$

which once again attains a maximum of 4 (when $\cos(2\theta) = -1$) and a minimum of 3 (when $\cos(2\theta) = -1/2$).

Thus, the global maximum is 4, while the global minimum is 3.

Problem 6.

(a) If z_1^*, z_2^* denote the conjugates of the complex numbers z_1, z_2, show that $(z_1 + z_2)^* = z_1^* + z_2^*$ and $(z_1 z_2)^* = z_1^* z_2^*$.

(b) Given the equation $az + bz^* = c$, where $a, b, c \in \mathbb{C}$, deduce the equation $a^* z^* + b^* z = c^*$. If $|a| \neq |b|$, show that
$$z = \frac{a^* c - bc^*}{|a|^2 - |b|^2}.$$

(c) If $|a| = |b|$, show that no solution for z exists unless $c = \lambda\,(a + b)$ or $c = b\mu\mathrm{i}$ for some $\lambda, \mu \in \mathbb{R}$.

(d) Find the solution for z of the equation
$$(7 + \mathrm{i})\,z + 5\,(1 - \mathrm{i})\,z^* = 3 - \mathrm{i}$$
in the form $z = p + tq$, where p and q are fixed complex numbers and t is a real parameter.

Solution.

Part (a). Let $z_1 = a + b\mathrm{i}$ and $z_2 = c + d\mathrm{i}$, where $a, b, c, d \in \mathbb{R}$. Then
$$(z_1 + z_2)^* = [(a + c) + (b + d)\,\mathrm{i}]^* = (a + c) - (b + d)\,\mathrm{i} = (a - b\mathrm{i}) + (c - d\mathrm{i}) = z_1^* + z_2^*,$$
and
$$(z_1 z_2)^* = [(ac - bd) + (ad + bc)\,\mathrm{i}]^* = (ac - bd) - (ad + bc)\,\mathrm{i} = (a - b\mathrm{i})\,(c - d\mathrm{i}) = z_1^* z_2^*.$$

Part (b). Taking conjugates, we get
$$a^* z^* + b^* z = c^*.$$
We thus get the system of equations
$$\begin{cases} az + bz^* = c, \\ a^* z^* + b^* z = c^* \end{cases}.$$
This can be represented with the following matrix equation:
$$\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} \begin{pmatrix} z \\ z^* \end{pmatrix} = \begin{pmatrix} c \\ c^* \end{pmatrix}.$$
Assuming that $|a| \neq |b|$, the determinant of the matrix is
$$\det \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} = aa^* - bb^* = |a|^2 - |b|^2 \neq 0,$$
so we can invert it to get
$$\begin{pmatrix} z \\ z^* \end{pmatrix} = \frac{1}{|a|^2 - |b|^2} \begin{pmatrix} a^* & -b \\ -b^* & a \end{pmatrix} \begin{pmatrix} c \\ c^* \end{pmatrix} \implies z = \frac{a^* c - bc^*}{|a|^2 - |b|^2}.$$

Part (c). If $|a| = |b|$, then the determinant of the above matrix is zero, so it is not invertible. We hence turn to Gaussian elimination to solve the matrix equation. Let $a = r\mathrm{e}^{\mathrm{i}\theta}$ and $b = r\mathrm{e}^{\mathrm{i}\varphi}$. Then
$$\left(\begin{array}{cc|c} a & b & c \\ b^* & a^* & c^* \end{array}\right) = \left(\begin{array}{cc|c} r\mathrm{e}^{\mathrm{i}\theta} & r\mathrm{e}^{\mathrm{i}\varphi} & c \\ r\mathrm{e}^{-\mathrm{i}\varphi} & r\mathrm{e}^{-\mathrm{i}\theta} & c^* \end{array}\right) \to_{\mathrm{e}^{\mathrm{i}(\theta+\varphi)} R_2} \left(\begin{array}{cc|c} r\mathrm{e}^{\mathrm{i}\theta} & r\mathrm{e}^{\mathrm{i}\varphi} & c \\ r\mathrm{e}^{\mathrm{i}\theta} & r\mathrm{e}^{\mathrm{i}\varphi} & c^* \mathrm{e}^{\mathrm{i}(\theta+\varphi)} \end{array}\right).$$

For there to be solutions, we require

$$c = c^* e^{i(\theta+\varphi)} \implies c^2 = |c|^2 e^{i(\theta+\varphi)} = \frac{|c|^2}{r^2} ab \implies c = \lambda\sqrt{ab}$$

for some real λ. All that remains is to show that $\sqrt{ab}$ can always be expressed as a real multiple of either $a+b$ or bi.

Let $\alpha = \arg(a)$ and $\beta = \arg(b)$, with $|a| = |b| = r$.

Case 1. Suppose $\alpha \neq \beta \pmod{\pi}$. Then

$$a+b = r\left(e^{i\alpha} + e^{i\beta}\right) = re^{i(\alpha+\beta)/2}\left[e^{i(\alpha-\beta)/2} + e^{-i(\alpha-\beta)/2}\right] = 2r\cos\left(\frac{\alpha-\beta}{2}\right)e^{i(\alpha+\beta)/2},$$

so

$$\arg(a+b) = \frac{\alpha+\beta}{2} = \arg\left(\sqrt{ab}\right),$$

so $\sqrt{ab}$ is a real multiple of $a+b$ in this case.

Case 2. Suppose $\alpha = \beta \pmod{\pi}$. Since $|a| = |b|$, we are left with two possibilities: either $a = b$ or $a = -b$. If $a = b$, then we very clearly have

$$\frac{\sqrt{ab}}{a+b} = \frac{a}{2a} = \frac{1}{2} \in \mathbb{R},$$

so $\sqrt{ab}$ is a real multiple of $a+b$. Likewise, if $a = -b$, then

$$\frac{\sqrt{ab}}{bi} = \frac{\sqrt{-b^2}}{bi} = \frac{bi}{bi} = 1 \in \mathbb{R},$$

so $\sqrt{ab}$ is a real multiple of bi.

Part (d). By inspection,

$$3 - i = \frac{1}{4}\left[(7+i) + (5-5i)\right],$$

so $z = 1/4$ is a particular solution to the above equation. Now consider the associated homogeneous equation:

$$(7+i)\,z + (5-5i)\,z^* = 0.$$

Multiplying by z yields

$$z^2\,(7+i) + |z|^2\,(5-5i) = 0 \implies z^2 = \frac{|z^2|}{5}(-3+4i) = \frac{|z|^2}{5}(1+2i)^2.$$

Thus, the complementary solution is $z = t\,(1+2i)$, where t is a real parameter (namely, $t = |z|/\sqrt{5}$). The general solution is thus

$$z = \frac{1}{4} + t\,(1+2i).$$

* * * * *

Problem 7. The point P represents the complex number p in an Argand diagram, and the points Q and R represent the roots of the equation $(z+p)^2 + n^2p^2 = 0$, where n is a real positive real constant. Show that

(a) $PQ = PR = (4+n^2)^{1/2}\,|p|$,

(b) the angle between the lines PQ and PR is $2\arctan(n/2)$.

The point S is such that $PQSR$ is a rhombus.

(c) Find, in terms of p, the complex number represented by S.

(d) Find the value of n for which $PQSR$ is a square.

Solution.

Part (a). Manipulating the given equation, we see that

$$(z+p)^2 + n^2p^2 = 0 \implies z = -p \pm \mathrm{i}np.$$

Without loss of gerenality, take $q = -p + \mathrm{i}np$ and $r = -p - \mathrm{i}np$. Then

$$PQ = |q-p| = |-2p + \mathrm{i}np| = |p|\,|-2+\mathrm{i}n|\,.$$

and

$$PR = |r-p| = |-2p - \mathrm{i}np| = |p|\,|2+\mathrm{i}n|\,.$$

Since $-2+\mathrm{i}n$ and $2+\mathrm{i}n$ are reflections of each other in the imaginary axis, their magnitudes are clearly equal, so $PQ = PR$. But

$$PR = |p|\,|2+\mathrm{i}n| = |p|\,\sqrt{2^2+n^2} = |p|\,\sqrt{4+n^2},$$

so

$$PQ = PR = |p|\left(4+n^2\right)^{1/2}.$$

Part (b). Observe that

$$\arg(q-p) = \arg(-2p+\mathrm{i}np) = \arctan\left(\frac{np}{-2}\right) = -\arctan\left(\frac{n}{2}\right),$$

and

$$\arg(r-p) = \arg(-2p-\mathrm{i}np) = \arctan\left(\frac{-np}{-2p}\right) = \arctan\left(\frac{n}{2}\right).$$

Thus, the difference in the angle between the lines PQ and PR is

$$|\arg(q-p) - \arg(r-p)| = \left|-\arctan\left(\frac{n}{2}\right) - \arctan\left(\frac{n}{2}\right)\right| = 2\arctan\left(\frac{n}{2}\right).$$

Part (c). Since $PQSR$ is a rhombus,

$$\overrightarrow{PS} = \overrightarrow{PR} + \overrightarrow{RS} = \overrightarrow{PR} + \overrightarrow{PQ}.$$

Rewriting in complex number form, we have

$$s - p = (r-p) + (q-p) \implies s = r + q - p = (-p+\mathrm{i}np) + (-p-\mathrm{i}np) - p = 3p.$$

Part (d). For $PQSR$ to be a square, the angle between PQ and PR must be $\pi/2$. Thus,

$$2\arctan\left(\frac{n}{2}\right) = \frac{\pi}{2} \implies n = 2\tan\left(\frac{\pi}{4}\right) = 2.$$

* * * * *

Problem 8. The complex number z has real part x and imaginary part y. Show that

(a) $\tan(\arg(z-1)) = \frac{y}{x-1}$,

(b) $\tan(\arg(1-4/z)) = \frac{4y}{x^2+y^2-4x}$.

The points P and Q in an Argand diagram represent z and $4/z$ respectively.

(c) Given that the line PQ passes through the point representing $1+0\mathrm{i}$, show that

$$y = 0 \qquad \text{or} \qquad (x-4)^2 + y^2 = 12.$$

Sketch the locus of P.

(d) Show that, for any position of P on this locus for which $y \neq 0$,

$$\left|1 - \frac{1}{z}\right| = \frac{\sqrt{3}}{2}.$$

Solution.

Part (a). We have

$$\tan(\arg(z-1)) = \tan(\arg((x-1)+\mathrm{i}y)) = \tan\arctan\left(\frac{y}{x-1}\right) = \frac{y}{x-1}.$$

Part (b). We have

$$\frac{1}{z} = \frac{z^*}{|z|^2} = \frac{x-\mathrm{i}y}{x^2+y^2},$$

so

$$\tan\left(\arg\left(1-\frac{4}{z}\right)\right) = \tan\left(\arg\left(\frac{x^2+y^2-4x+4\mathrm{i}y}{x^2+y^2}\right)\right) = \tan\arctan\left(\frac{4y}{x^2+y^2-4x}\right)$$
$$= \frac{4y}{x^2+y^2-4x}.$$

Part (c). Note that P and Q are on opposite sides of the real axis (since $\arg(z) = -\arg(4/z)$). Thus, we have

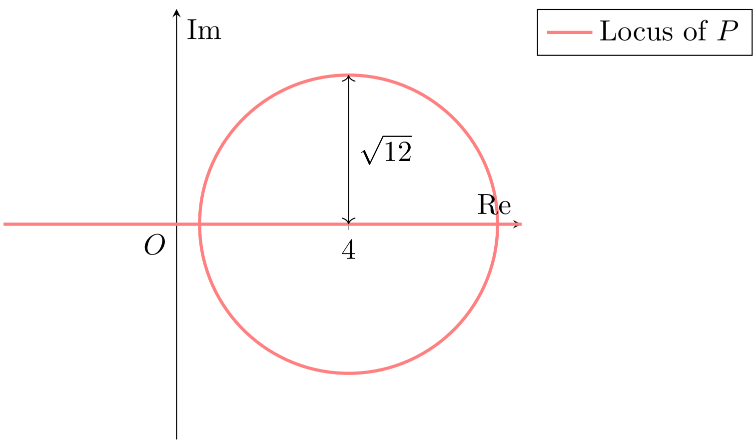
$$\arg\left(1-\frac{4}{z}\right) = \arg(z-1)\,.$$

From the previous parts, we obtain

$$\frac{y}{x-1} = \frac{4y}{x^2+y^2-4x}.$$

Thus, $y=0$ is clearly a solution. If $y \neq 0$, then further maniupulation yileds

$$x^2+y^2-4x = 4x-4 \implies x^2-8x+16+y^2 = 12 \implies (x-4)^2+y^2 = 12.$$



Part (d). Let $R(z)$ be the statement "the points representing z, $4/z$ and 1 are collinear". Note that this is equivalent to the statement "the points representing $4/z$, $4/(4/z)$ and 1 are collinear". Thus,

$$R(z) \iff R\left(\frac{4}{z}\right).$$

From the previous part, if $z \in \mathbb{C} \setminus \mathbb{R}$, then

$$R(z) \implies |z-4| = \sqrt{12} = 2\sqrt{3}.$$

As such, if $R(z)$ is true (i.e. z lies on the locus of P), then

$$R(z) \implies R\left(\frac{4}{z}\right) \implies \left|\frac{4}{z} - 4\right| = 2\sqrt{3} \implies \left|1 - \frac{1}{z}\right| = \frac{1}{2}\sqrt{3}$$

as desired.

* * * * *

Problem 9. The complex numbers a, b, c are represented in an Argand diagram by the points A, B and C respectively. The triangle ABC is equilateral and is labelled anticlockwise. Show that

(a) $(c-a)\mathrm{e}^{\mathrm{i}\pi/3} = c - b$,

(b) $(c-a)\mathrm{e}^{-\mathrm{i}\pi/3} = b - a$,

(c) $a^2 + b^2 + c^2 = bc + ca + ab$.

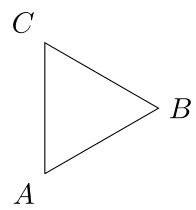
Given that $a = -4 + 4\mathrm{i}$ and $b = 4 - 2\mathrm{i}$, use the result in (c) to show that

$$(c - \mathrm{i})^2 = k\,(3 + 4\mathrm{i})^2$$

where k is real. Hence, find c in exact form.

Solution.

Part (a). Consider the figure below.



Clearly, rotating $\overrightarrow{AC}$ 60° clockwise results in $\overrightarrow{BC}$, so

$$(c-a)\,\mathrm{e}^{\mathrm{i}\pi/3} = c - b.$$

Part (b). Referring to the above figure, we also see that rotating $\overrightarrow{AC}$ 60° counterclockwise results in $\overrightarrow{AB}$, so

$$(c-a)\,\mathrm{e}^{-\mathrm{i}\pi/3} = b - a.$$

Part (c). Multiplying the above two results, we get

$$c^2 - 2ac + a^2 = (c-a)^2 = (c-b)\,(b-a) = cb - ac - b^2 + ab \implies a^2 + b^2 + c^2 = ab + bc + ac.$$

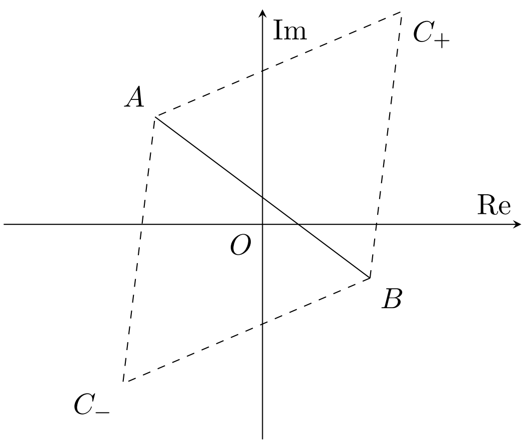
Using (c), we have

$$(-4+4\mathrm{i})^2 + (4-2\mathrm{i})^2 + c^2 = (4-2\mathrm{i})\,c + (-4+4\mathrm{i})\,c + (-4+4\mathrm{i})\,(4-2\mathrm{i})\,,$$

which simplifies to

$$(c-\mathrm{i})^2 = c^2 - 2\mathrm{i}c - 1 = -21 + 72\mathrm{i} = 3\,(-7+24\mathrm{i}) = 3\,(3+4\mathrm{i})\,.$$

Thus, $k = 3$. Taking square roots, we see that $c = \mathrm{i} \pm \sqrt{3}\,(3+4\mathrm{i})$. Drawing both possibilities on an Argand diagram, along with a and b, we reject the negative branch, so $c = \mathrm{i} + \sqrt{3}\,(3+4\mathrm{i})$.



* * * * *

Problem 10. The point Z in an Argand diagram represents the variable complex number z, and a is a fixed non-zero complex number.

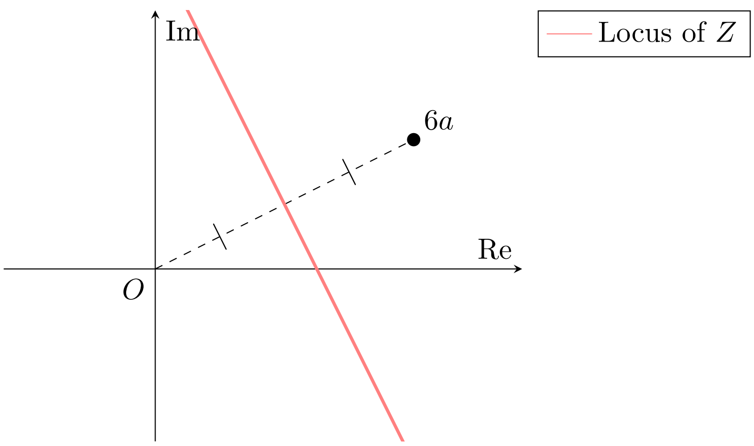
(a) Given that $|z| = |z - 6a|$, sketch the locus of Z.

(b) Given that $|z| = 2\,|z - 3a|$, show that $|z - 4a| = 2\,|a|$, and hence sketch the locus of Z.

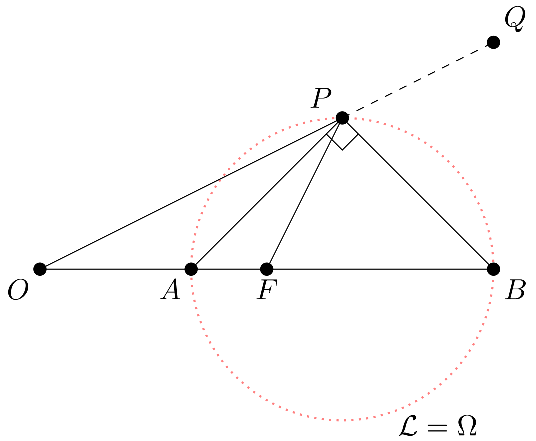
P and Q, represent the complex numbers p and q, are the common points of the loci in (a) and (b). Find $\arg(p/q)$, given that it is positive.

Solution.

Part (a).



Part (b). Let $\mathcal{L}$ be the locus of Z and P an arbitrary point on $\mathcal{L}$. Define $F(3a)$, $A = \mathcal{L} \cap OF$, $B = \mathcal{L} \cap (OF \text{ extended})$ and Q on OP extended.



Since $P, A \in \mathcal{L}$, we have

$$\frac{OP}{PF} = \frac{OA}{AF} = \frac{|z|}{|z-3a|} = 2.$$

Thus, by the angle bisector theorem, it follows that $\angle APF$ bisects $\angle OPF$.

Similarly, since $P, B \in \mathcal{L}$, we have

$$\frac{OP}{PF} = \frac{OB}{BF} = 2.$$

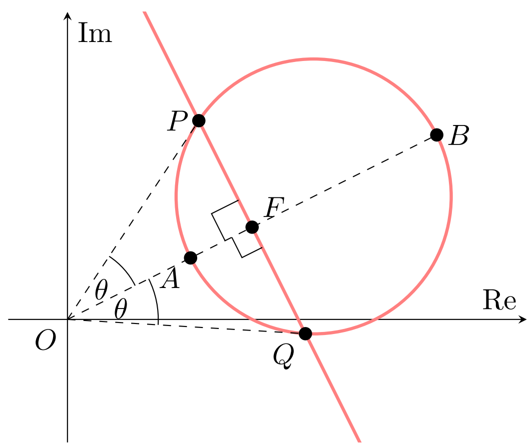
Thus, by the angle bisector theorem, it follows that $\angle FPB$ bisects $\angle FPQ$.

Hence,

$$\angle APB = \angle APF + \angle FPB = \frac{\angle OPF + \angle FPQ}{2} = 90°.$$

Since P is arbitrary, by the converse of Thale's theorem, it follows that $\mathcal{L}$ lies on a circle. By symmetry, OB contains the diameter of the circle. Elementary calculations show that $A(2a)$ and $B(6a)$, so the center of the circle is $(2a + 6a)/2 = 4a$ and the radius is $|4a - 2a| = |2a|$. Call this circle Ω, so $\mathcal{L} \subseteq \Omega$.

By inverting the above argument and invoking the converse of the angle bisector theorem (and Thale's theorem), we see that $\Omega \subseteq \mathcal{L}$, so we must have $\mathcal{L} = \Omega$. The complex equation for Ω (and thus $\mathcal{L}$) is given by $|z - 4a| = 2\,|a|$ as desired.



Let $\theta = \angle POF = \angle QOF$. Then

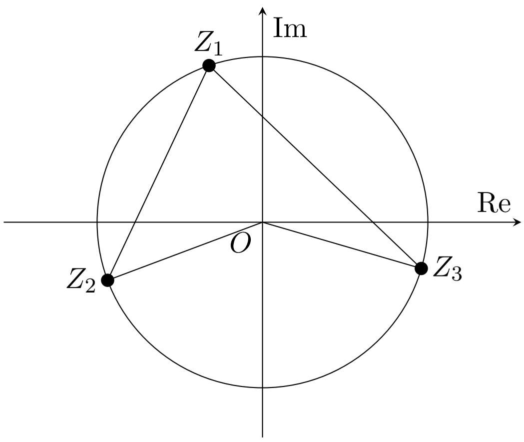
$$\sin\theta = \frac{PF}{OF} = \frac{|z-3a|}{|z|} = \frac{1}{2} \implies \theta = \frac{\pi}{3}.$$

Thus,

$$\arg\left(\frac{p}{q}\right) = \arg(p) - \arg(q) = 2\theta = \frac{\pi}{3}.$$

* * * * *

Problem 11.



The complex numbers $z_1 = x_1 + \mathrm{i}y_1$, $z_2 = x_2 + \mathrm{i}y_2$, $z_3 = x_3 + \mathrm{i}y_3$, where $x_k, y_k \in \mathbb{R}$ and $|z_k| = 1$ for $k = 1, 2, 3$ are represented by the points Z_1, Z_2 and Z_3 on the Argand diagram above.

(a) Express the complex numbers z_1, z_2, z_3 in exponential form, where $\alpha = \arg(z_1)$, $\beta = \arg(z_2)$, $\gamma = \arg(z_3)$.

(b) Using part (a), show that $\arg(z_1 z_2) = \alpha + \beta$.

(c) Without using the property that $\angle Z_2 O Z_3 = 2\angle Z_2 Z_1 Z_3$, find the value of $\angle Z_2 Z_1 Z_3$ in terms of β and γ.

(d) Hence, or otherwise, prove that $\angle Z_2 O Z_3 = 2\angle Z_2 Z_1 Z_3$.

Solution.

Part (a). We have

$$z_1 = \mathrm{e}^{\mathrm{i}\alpha}, \quad z_2 = \mathrm{e}^{\mathrm{i}\beta}, \quad z_3 = \mathrm{e}^{\mathrm{i}\gamma}.$$

Part (b). We have

$$\arg(z_1 z_2) = \arg\left(\mathrm{e}^{\mathrm{i}\alpha}\mathrm{e}^{\mathrm{i}\beta}\right) = \arg\left(\mathrm{e}^{\mathrm{i}(\alpha+\beta)}\right) = \alpha + \beta.$$

Part (c). Observe that

$$\arg(z_2 - z_1) = \arg\left(\mathrm{e}^{\mathrm{i}\beta} - \mathrm{e}^{\mathrm{i}\alpha}\right) = \arg\left(\mathrm{e}^{\mathrm{i}\frac{\alpha+\beta}{2}} \underbrace{\left(\mathrm{e}^{\mathrm{i}\frac{\beta-\alpha}{2}} - \mathrm{e}^{-\mathrm{i}\frac{\beta-\alpha}{2}}\right)}_{2\mathrm{i}\,\mathrm{Im}\,\mathrm{e}^{\mathrm{i}\frac{\beta-\alpha}{2}}}\right) = \frac{\alpha+\beta+\pi}{2}.$$

Similarly,

$$\arg(z_3 - z_1) = \arg\left(\mathrm{e}^{\mathrm{i}\gamma} - \mathrm{e}^{\mathrm{i}\alpha}\right) = \frac{\alpha+\gamma+\pi}{2}.$$

Thus,

$$\angle Z_2 Z_1 Z_3 = |\arg(z_2 - z_1) - \arg(z_3 - z_1)| = \left|\frac{\alpha+\beta+\pi}{2} - \frac{\alpha+\gamma+\pi}{2}\right| = \left|\frac{\beta-\gamma}{2}\right|.$$

Part (d). It is clear that

$$\angle Z_2 O Z_3 = |\arg z_2 - \arg z_3| = |\beta - \gamma| = 2\left|\frac{\beta - \gamma}{2}\right| = 2\angle Z_2 Z_1 Z_3.$$

* * * * *

Problem 12. The complex numbers $z_1, z_2, \ldots, z_6$ are represented by six distinct points $P_1, P_2, \ldots, P_6$ in the Argand diagram. Express the following statements in terms of complex numbers:

(a) $\overrightarrow{P_1P_2} = \overrightarrow{P_5P_4}$ and $\overrightarrow{P_2P_3} = \overrightarrow{P_6P_5}$;

(b) $\overrightarrow{P_2P_4}$ is perpendicular to $\overrightarrow{P_3P_6}$.

(c) If (a) holds, show that $\overrightarrow{P_3P_4} = \overrightarrow{P_1P_6}$.

Suppose that (a) and (b) both hold, and that $z_1 = 0$, $z_2 = 1$, $z_3 = z$, $z_5 = \mathrm{i}$ and $z_6 = w$.

(d) Show that if $P_1P_2P_3P_4P_5P_6$ forms a convex hexagon, then $\mathrm{Re}(z) + \mathrm{Re}(w) = 1$ with $\mathrm{Re}(z) > 1$ and $\mathrm{Re}(w) < 0$.

(e) Find the distance between P_3 and P_6 when $\tan \angle P_3P_2P_6 = -2/3$.

Solution.

Part (a). We have

$$z_2 - z_1 = z_4 - z_5 \quad \text{and} \quad z_3 - z_2 = z_5 - z_6.$$

Part (b). We have that $z_4 - z_2$ is perpendicular to $z_6 - z_3$. Hence,

$$\begin{pmatrix}\mathrm{Re}(z_4 - z_2)\\ \mathrm{Re}(z_4 - z_2)\end{pmatrix} \cdot \begin{pmatrix}\mathrm{Re}(z_6 - z_3)\\ \mathrm{Im}(z_6 - z_3)\end{pmatrix} = \mathrm{Re}(z_4 - z_2)\,\mathrm{Re}(z_6 - z_3) + \mathrm{Im}(z_4 - z_2)\,\mathrm{Im}(z_6 - z_3) = 0.$$

Part (c). Adding both equations in (a), we have $z_3 - z_1 = z_4 - z_6$, so $z_4 - z_3 = z_6 - z_1$, implying $\overrightarrow{P_3P_4} = \overrightarrow{P_1P_6}$.

Part (d).

$$w = z_6 = z_5 + z_2 - z_3 = \mathrm{i} + 1 - z \implies w + z = 1 + \mathrm{i}.$$

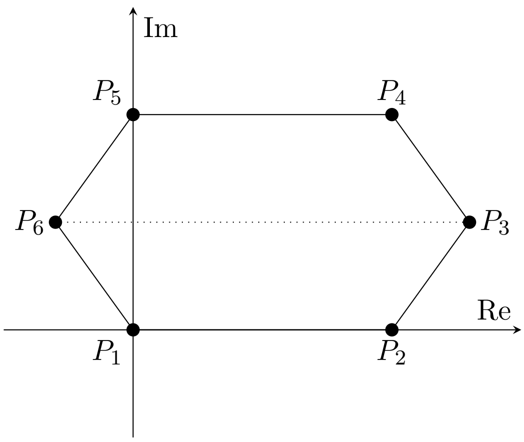
Comparing real parts, we have

$$\mathrm{Re}(w) + \mathrm{Re}(w) = 1.$$

Now observe that

$$z_4 = z_2 - z_1 + z_5 = 1 - 0 + \mathrm{i} = 1 + \mathrm{i}.$$

Further, because $\overrightarrow{P_2P_4}$ is perpendicular to $\overrightarrow{P_3P_6}$, we must have $\mathrm{Im}(z) = \mathrm{Im}(w) = 1/2$. We can represent this on the following Argand diagram:



Clearly, for $P_1P_2P_3P_4P_5P_6$ to be a convex hexagon, we must have $\mathrm{Re}(w) < \mathrm{Re}(z_1) = 0$ and $\mathrm{Re}(z) > \mathrm{Re}(z_2) = 1$.

Part (e). Since $z_4 - z_3 = z_6 - z_1$, we have $\mathrm{Re}(z_4 - z_3) = \mathrm{Re}(z_6 - z_1)$. Let this modulus of this common value be a. Then $z_6 = -a + \mathrm{i}/2$ and $z_3 = 1 + a + \mathrm{i}/2$. It is hence clear that

$$\tan\arg z_6 = \frac{1/2}{-a} \quad \text{and} \quad \tan\arg z_3 = \frac{1/2}{a+1}.$$

Thus,

$$-\frac{2}{3} = \tan\angle P_3P_2P_6 = \tan(\arg z_6 - \arg z_3) = \frac{\tan\arg z_6 - \tan\arg z_3}{1 + \tan\arg z_6 \tan\arg z_3} = \frac{\frac{1/2}{-a} - \frac{1/2}{a+1}}{1 + \frac{1/2}{-a}\frac{1/2}{a+1}}.$$

Solving, we get

$$2a^2 - a - 2 = 0 \implies a = \frac{1+\sqrt{17}}{4},$$

where we reject the negative branch since $a \geq 0$. Thus,

$$P_3P_6 = 2a + 1 = \frac{3+\sqrt{17}}{2}.$$

* * * * *

Problem 13. Let z, w be complex numbers such that $w = (z-1)/(z+1)$.

(a) Prove that w lies within the unit circle in the Argand diagram if and only if $\mathrm{Re}\, z \geq 0$.

(b) (i) Suppose that w lies within the unit circle in the Argand diagram. Show that $|1-w|^2 \,\mathrm{Re}\, z = 1 - |w|^2$.

(ii) Hence, prove that if $\mathrm{Re}\, z \neq -1$, then w satisfies the equation

$$\left| w - \frac{\mathrm{Re}\, z}{\mathrm{Re}\, z + 1} \right| = \left| \frac{1}{\mathrm{Re}\, z + 1} \right|.$$

(iii) Show instead that if $\mathrm{Re}\, z = -1$, then w satisfies the equation $w + w^* - 2 = 0$. Hence, explain why the locus representing w is a vertical line passing through $x = 1$.

Solution.

Part (a). For w to lie within the unit circle, we require $|w| \leq 1$, but

$$|w| = \frac{|z-1|}{|z+1|} \leq 1 \iff |z-1| \leq |z+1|.$$

But this is a standard locus, corresponding precisely to the region of the Argand diagram where the real part is non-negative, so $|w| \leq 1$ if and only if $\mathrm{Re}(z) \geq 0$.

Part (b).

Part (b)(i). Note that

$$|z+1|^2 = (z+1)(z^*+1) = |z|^2 + z + z^* + 1,$$

and

$$|z-1|^2 = (z-1)(z^*-1) = |z|^2 - z - z^* + 1.$$

Hence,

$$\frac{1-|w|^2}{|1-w|^2} = \frac{1-\left|\frac{z-1}{z+1}\right|^2}{\left|1-\frac{z-1}{z+1}\right|^2} = \frac{|z+1|^2-|z-1|^2}{|(z+1)-(z-1)|^2}$$
$$= \frac{\left(|z|^2+z+z^*+1\right)-\left(|z|^2-z-z^*+1\right)}{4} = \frac{z+z^*}{2} = \operatorname{Re} z,$$

and we immediately get our desired identity upon clearing denominators.

Part (b)(ii). From the previous part, we have $\operatorname{Re}(z)\,|1-w|^2 = 1-|w|^2$, so

$$\operatorname{Re}(z)\left(1-w-w^*+|w|^2\right) = 1-|w|^2.$$

Expanding and simplifying, we have

$$(1+\operatorname{Re}(z))\,|w|^2 - \operatorname{Re}(z)\,w - \operatorname{Re}(z)\,w^* = 1-\operatorname{Re}(z).$$

Dividing throughout by $1+\operatorname{Re}(z)$ yields

$$|w|^2 - \frac{\operatorname{Re}(z)}{\operatorname{Re}(z)+1}w - \frac{\operatorname{Re}(z)}{\operatorname{Re}(z)+1}w^* = \frac{1-\operatorname{Re}(z)}{\operatorname{Re}(z)+1}.$$

Hence,

$$\left|w-\frac{\operatorname{Re}(z)}{\operatorname{Re}(z)+1}\right|^2 = \frac{1-\operatorname{Re}(z)}{\operatorname{Re}(z)+1} + \left(\frac{\operatorname{Re}(z)}{\operatorname{Re}(z)+1}\right)^2 = \frac{1}{(\operatorname{Re}(z)+1)^2}.$$

The desired result follows immediately.

Part (b)(iii). If $\operatorname{Re}(z) = -1$, we have

$$|w|^2 - 1 = |1-w|^2 = 1-w-w^*+|w| \implies w+w^*-2 = 0.$$

Rewriting, we have

$$2\operatorname{Re}(w) = 2 \implies \operatorname{Re}(w) = 1.$$

Hence, the locus of w is the vertical line passing through $x = 1$.

* * * * *

Problem 14. Let z_1, z_2 be complex numbers such that $\arg(z_1), \arg(z_2) \in (0, \pi/4)$.

(a) Show that the function $A(x) = \arctan x$ is an increasing function on $\mathbb{R}$.

(b) Suppose that a, b, c, d are positive real numbers such that $a/b < c/d$. Prove that

$$\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}.$$

(c) Prove that

$$\min\{\arg(z_1), \arg(z_2)\} < \arg(z_1+z_2) < \max\{\arg(z_1), \arg(z_2)\}.$$

Solution.

Part (a). We have

$$\frac{\mathrm{d}}{\mathrm{d}x}A(x) = \frac{1}{1+x^2} > 0$$

for all $x \in \mathbb{R}$. Hence, $A(x)$ is increasing on $\mathbb{R}$.

Part (b). Since $a/b < c/d$, we have

$$a < \frac{cb}{d} \implies a + c < \frac{cb + cd}{d} \implies \frac{a+c}{b+d} < \frac{c}{d}.$$

Similarly, we have

$$\frac{ad}{b} < c \implies \frac{ad + ab}{b} < a + c \implies \frac{a}{b} < \frac{a+c}{b+d}.$$

Chaining both inequalities yields

$$\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}.$$

Part (c). Let $z_1 = b + a\mathrm{i}$, $z_2 = d + c\mathrm{i}$. If $a/b = c/d$, then z_1 and z_2 are scalar multiples of each other, so

$$\arg(z_1 + z_2) = \arg(z_1) \in \left(0, \frac{\pi}{4}\right).$$

Else, without loss of generality, take $a/b < c/d$. From (b), we know that

$$\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}.$$

Further, from (a), we can apply arctan on all sides of the inequality to get

$$\arctan \frac{a}{b} < \arctan \frac{a+c}{b+d} < \arctan \frac{c}{d}.$$

But

$$\arg(z_1) = \arctan \frac{a}{b}, \quad \arg(z_2) = \arctan \frac{c}{d} \quad , \arg(z_1 + z_2) = \arctan \frac{a+c}{b+d},$$

so

$$\min\{\arg(z_1), \arg(z_2)\} = \arg(z_1) < \arg(z_1 + z_2) < \arg(z_2) = \max\{\arg(z_1), \arg(z_2)\}.$$

* * * * *

Problem 15. The complex number $z \neq \mathrm{i}$ satisfies the equation $|z - \mathrm{i}/2| = 1/2$, and $w = \mathrm{i}z/(\mathrm{i} - z)$.

(a) Show that w is real.

(b) Show that the points representing w, z and i on the Argand diagram are collinear.

Illustrate (a) and (b) on an Argand diagram.

Solution.

Part (a). Let $z = x + \mathrm{i}y$ where $x, y \in \mathbb{R}$. The locus of z is a circle centered at $(0, 1/2)$ with radius $1/2$, so

$$x^2 + \left(y - \frac{1}{2}\right)^2 = \frac{1}{2^2} \implies |z|^2 = x^2 + y^2 = y.$$

Hence,

$$w = \frac{\mathrm{i}z}{\mathrm{i} - z} = \frac{\mathrm{i}z\,(-\mathrm{i} - z^*)}{|\mathrm{i} - z|^2} = \frac{z - |z|^2\,\mathrm{i}}{|\mathrm{i} - z|^2} = \frac{(x + \mathrm{i}y) - y\mathrm{i}}{|\mathrm{i} - z|^2} = \frac{x}{|\mathrm{i} - z|^2} \in \mathbb{R}.$$

Part (b). Observe that

$$|\mathrm{i}-z|^2 = (\mathrm{i}-z)(-\mathrm{i}-z^*) = 1+|z|^2+\mathrm{i}z-\mathrm{i}z^* = 1+y+\mathrm{i}(x+\mathrm{i}y)-\mathrm{i}(x-\mathrm{i}y) = 1-y.$$

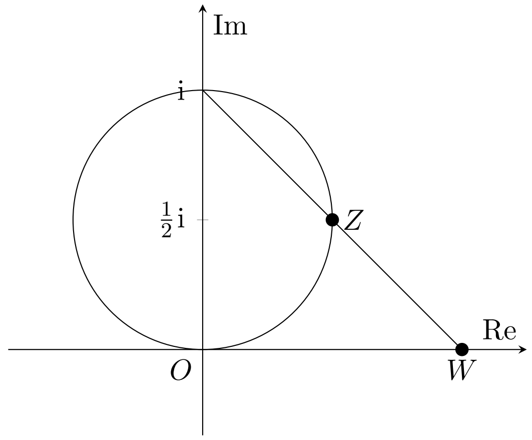
Hence,

$$\arg(i-w) = -\frac{x}{1-y}+\mathrm{i} = \arctan\frac{y-1}{x}.$$

But

$$\arg(\mathrm{i}-z) = \arg(-x+\mathrm{i}(1-y)) = \arctan\frac{y-1}{x},$$

so $\arg(\mathrm{i}-w) = \arg(\mathrm{i}-z)$, from which it follows that the points representing w, z and i on the Argand diagram are collinear.



* * * * *

Problem 16. Let

$$z_k = \cos\left(\frac{2k\pi}{n}\right)+\mathrm{i}\sin\left(\frac{2k\pi}{n}\right),$$

where $k \neq 0$. Show that the set $\{(z_k)^t : t = 1, 2, \ldots, n\}$ has exactly n elements if and only if k and n do not have any common prime factors.

Solution. Let

$$\mathcal{Z} = \{(z_k)^t : t = 1, 2, \ldots, n\} = \left\{\exp\left(2\pi\mathrm{i}\frac{kt}{n}\right) : t = 1, 2, \ldots, n\right\}.$$

Note that the $n = 1$ case is trivial, so we take $n > 1$. It suffices to show that $|\mathcal{Z}| < n \iff \gcd(n, k) > 1$.

We begin with the forwards direction. Suppose $|\mathcal{Z}| < n$. Then there exist distinct $t_1, t_2 \in \{1, \ldots, n\}$ such that $z_k^{t_1} = z_k^{t_2}$. Without loss of generality, suppose $t_1 > t_2$. Then

$$\exp\left(2\pi\mathrm{i}\frac{kt_1}{n}\right) = \exp\left(2\pi\mathrm{i}\frac{kt_2}{n}+2\pi m\mathrm{i}\right)$$

for some integer m. Taking logarithms, we have

$$2\pi\mathrm{i}\frac{kt_1}{n} = 2\pi\mathrm{i}\frac{kt_2}{n}+2\pi m\mathrm{i} \implies k(t_1-t_2) = mn.$$

Seeking a contradiction, suppose $\gcd(k, n) = 1$. Then we necessarily have

$$n \mid t_1 - t_2.$$

But this is impossible, since $1 \leq t_1 - t_2 \leq n-1$. Thus, we conclude that $\gcd(k,n) > 1$ as desired.

Now, suppose that $\gcd(k,n) > 1$. Define $d = \gcd(n,k)$. Then $n = dn'$ and $k = dk'$ for some integers n' and k'. Note that

$$d > 1 \implies n' = \frac{n}{d} < n \implies n' + 1 \leq n.$$

Additionally, $n' + 1 > 1$. Hence, we can select $t = n' + 1$, whence we get

$$\begin{aligned} z_k^{n'+1} &= \exp\left(2\pi \mathrm{i} \frac{k(n'+1)}{n}\right) = \exp\left(2\pi \mathrm{i} \frac{k'(n'+1)}{n'}\right) = \exp\left(2\pi \mathrm{i} k' + 2\pi \mathrm{i} \frac{k'}{n'}\right) \\ &= \exp\left(2\pi \mathrm{i} \frac{k'}{n'}\right) = \exp\left(2\pi \mathrm{i} \frac{k}{n}\right) = z_k^1. \end{aligned}$$

But $n' + 1 \neq 1$, so there is a duplicate element in $\mathcal{Z}$ and we immediately have $|\mathcal{Z}| < n$.

* * * * *

Problem 17.

(a) Show that for any $\theta \in \mathbb{R}$, we have

$$\sin\theta = \frac{2\tan\frac{\theta}{2}}{1+\tan^2\frac{\theta}{2}} \quad \text{and} \quad \cos\theta = \frac{1-\tan^2\frac{\theta}{2}}{1+\tan^2\frac{\theta}{2}}.$$

(b) Let $z \in \mathbb{C}$ be a complex number such that $|z| = 1$ and $z \neq -1$. By writing z in trigonometric form, prove that $z = (1+\mathrm{i}t)/(1-\mathrm{i}t)$ for some $t \in \mathbb{R}$.

Solution.

Part (a). We have

$$\frac{2\tan\frac{\theta}{2}}{1+\tan^2\frac{\theta}{2}} = \frac{2\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{\cos^2\frac{\theta}{2}+\sin^2\frac{\theta}{2}} = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2} = \sin\theta$$

and

$$\frac{1-\tan^2\frac{\theta}{2}}{1+\tan^2\frac{\theta}{2}} = \frac{\cos^2\frac{\theta}{2}-\sin^2\frac{\theta}{2}}{\cos^2\frac{\theta}{2}+\sin^2\frac{\theta}{2}} = \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} = \cos\theta.$$

Part (b). We have $z = \cos\theta + \mathrm{i}\sin\theta$ for some $\theta \in [0, 2\pi) \setminus \{\pi\}$. Let $t = \tan^2(\theta/2)$. Then

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

* * * * *

Problem 18. Let z be a complex number satisfying the equation $z^n = z + z^*$ for some positive integer n. By expressing z^n in polar form, show that there are exactly n solutions to the equation if and only if $n \not\equiv 1 \pmod 4$.

Solution. Let $z = r\mathrm{e}^{\mathrm{i}\theta}$, where $r \geq 0$ and $\theta \in [0, 2\pi)$. Then the given equation can be rewritten as

$$r^n \mathrm{e}^{\mathrm{i}n\theta} = 2r\cos\theta.$$

If $r = 0$, we obtain the trivial solution $z = 0$. For the rest of the proof, we simply take $z \neq 0$, i.e. $r > 0$. Then

$$r^{n-1}\mathrm{e}^{\mathrm{i}n\theta} = 2\cos\theta,$$

which immediately implies $e^{in\theta} \in \mathbb{R}$, so $n\theta = k\pi$ for some integer k, whence

$$r^{n-1} = 2\,(-1)^k \cos\frac{k\pi}{n}.$$

For k to yield a solution, we require the RHS to be positive, so k must satisfy

$$(-1)^k \cos\frac{k\pi}{n} > 0. \tag{$*$}$$

If k fulfils this condition, then the solution it generates is given by

$$z_k = \sqrt[n-1]{2\,(-1)^k \cos\frac{k\pi}{n}}\, e^{ik\pi/n}.$$

Observe that if z satisfies the given equation, then so must z^*. Hence, the number of solutions given by $\theta \in (0,\pi)$ is equal to the number of solutions given by $\theta \in (\pi, 2\pi)$. Note also that $\theta = \pi/2$ corresponds to the trivial solution $z = 0$. Thus, for the remainder of this proof, we consider only the case where $\theta = 0, \pi$ (i.e. $z \in \mathbb{R}$), $\theta \in (0, \pi/2)$ and $\theta \in (\pi/2, \pi)$.

Case 1: $n \equiv 0 \pmod 4$. Let $n = 4m$. $\theta = 0$ gives the real solution $z = \sqrt[n-1]{2}$, while $\theta = \pi$ does not yield any solution (it does not satisfy $(*)$). Consider now $\theta \in (0, \pi/2)$, i.e. $k = 1, 2, \ldots, 2m-1$. For $(*)$ to hold, we must have k even, so we have $m-1$ valid values of k:

$$k = 2, 4, \ldots, 2m-2.$$

Consider now $\theta \in (\pi/2, \pi)$, i.e. $k = 2m+1, 2m+2, \ldots, 4m-1$. For $(*)$ to hold, we must have k odd, so we have another m valid values of k:

$$k = 2m+1, 2m+3, \ldots, 4m-1.$$

Now observe that because the first set of k's is even and the second set of k's is odd, the symmetry around $\theta = \pi/2$ ($k = 2m$) is broken, so these two sets of k's must yield distinct z_k. Altogether, we have a total of $4m = n$ distinct solutions:

θ	Unique solutions in interval	Remarks
0	1	$z = \sqrt[n-1]{2}$
$(0, \pi/2)$	$m-1$	
$\pi/2$	1	$z = 0$
$(\pi/2, \pi)$	m	
π	0	
$(\pi, 3\pi/2)$	m	Conjugate of $(\pi/2, \pi)$
$3\pi/2$	0	Counted already ($z = 0$)
$(3\pi/2, 2\pi)$	$m-1$	Conjugate of $(0, \pi/2)$

Case 1: $n \equiv 2 \pmod 4$. Let $n = 4m+2$. $\theta = 0$ gives the real solution $z = \sqrt[n-1]{2}$, while $\theta = \pi$ does not yield any solution. Consider now $\theta \in (0, \pi/2)$, i.e. $k = 1, 2, \ldots, 2m$. For $(*)$ to hold, we must have k even, so we have m valid values of k:

$$k = 2, 4, \ldots, 2m.$$

Consider now $\theta \in (\pi/2, \pi)$, i.e. $k = 2m+2, 2m+3, \ldots, 4m+1$. For $(*)$ to hold, we must have k odd, so we have another m valid values of k:

$$k = 2m+3, 2m+5, \ldots, 4m+1.$$

Once again, the difference in parity between the two sets of k's ensures that they correspond to different z_k's. Altogether, we have a total of $4m+2 = n$ distinct solutions:

θ	Unique solutions in interval	Remarks
0	1	$z = \sqrt[n-1]{2}$
$(0, \pi/2)$	m	
$\pi/2$	1	$z = 0$
$(\pi/2, \pi)$	m	
π	0	
$(\pi, 3\pi/2)$	m	Conjugate of $(\pi/2, \pi)$
$3\pi/2$	0	Counted already $(z = 0)$
$(3\pi/2, 2\pi)$	m	Conjugate of $(0, \pi/2)$

Case 1: $n \equiv 3 \pmod 4$. Let $n = 4m + 3$. $\theta = 0, \pi$ give the real solutions $z = \pm \sqrt[n-1]{2}$. Consider now $\theta \in (0, \pi/2)$, i.e. $k = 1, 2, \ldots, 2m+1$. For $(*)$ to hold, we must have k even, so we have m valid values of k:

$$k = 2, 4, \ldots, 2m.$$

Consider now $\theta \in (\pi/2, \pi)$, i.e. $k = 2m + 2, 2m + 3, \ldots, 4m + 2$. For $(*)$ to hold, we must have k odd, so we have another m valid values of k:

$$k = 2m + 3, 2m + 5, \ldots, 4m + 1.$$

Again, the difference in parity between the two sets of k's ensures that they correspond to different z_k's. Altogether, we have a total of $4m + 3 = n$ distinct solutions:

θ	Unique solutions in interval	Remarks
0	1	$z = \sqrt[n-1]{2}$
$(0, \pi/2)$	m	
$\pi/2$	1	$z = 0$
$(\pi/2, \pi)$	m	
π	1	$z = -\sqrt[n-1]{2}$
$(\pi, 3\pi/2)$	m	Conjugate of $(\pi/2, \pi)$
$3\pi/2$	0	Counted already $(z = 0)$
$(3\pi/2, 2\pi)$	m	Conjugate of $(0, \pi/2)$

Case 1: $n \equiv 1 \pmod 4$. Let $n = 4m + 1$. $\theta = 0, \pi$ give the real solutions $z = \pm \sqrt[n-1]{2}$. Consider now $\theta \in (0, \pi/2)$, i.e. $k = 1, 2, \ldots, 2m$. For $(*)$ to hold, we must have k even, so we have m valid values of k:

$$k = 2, 4, \ldots, 2m.$$

Consider now $\theta \in (\pi/2, \pi)$, i.e. $k = 2m+1, 2m+3, \ldots, 4m$. For $(*)$ to hold, we must have k odd, so we have another m valid values of k:

$$k = 2m + 1, 2m + 5, \ldots, 4m - 1.$$

Again, the difference in parity between the two sets of k's ensures that they correspond to different z_k's. Altogether, we have a total of $4m + 3 = n + 2$ distinct solutions:

θ	Unique solutions in interval	Remarks
0	1	$z = \sqrt[n-1]{2}$
$(0, \pi/2)$	m	
$\pi/2$	1	$z = 0$
$(\pi/2, \pi)$	m	
π	1	$z = -\sqrt[n-1]{2}$
$(\pi, 3\pi/2)$	m	Conjugate of $(\pi/2, \pi)$
$3\pi/2$	0	Counted already ($z = 0$)
$(3\pi/2, 2\pi)$	m	Conjugate of $(0, \pi/2)$

* * * * *

Problem 19. Let

$$w = \frac{z - s}{1 - s^* z},$$

where $z = \mathrm{e}^{2\pi \mathrm{i}/3}$ and s is a complex number such that $\pi/2 < \arg s < 2\pi/3$ and $|s| = |z - s|$. Show that $|w| = 1$ and $2\pi/3 < \arg(w) < \pi$.

Solution. We note that $z^* = 1/z$, so

$$w = \frac{z - s}{1 - s^* z} = \frac{z - s}{z\,(1/z - s^*)} = \frac{z - s}{z\,(z^* - s^*)} = \frac{z - s}{z\,(z - s)^*}.$$

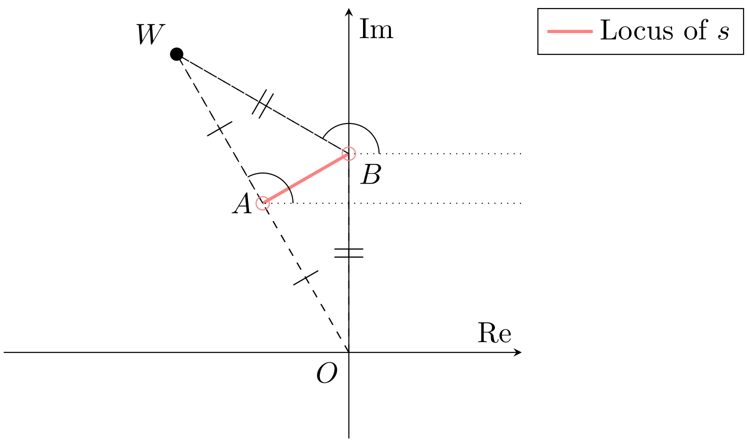
From here, we easily conclude that

$$|w| = \frac{|z - s|}{|z|\,|(z - s)^*|} = \frac{|z - s|}{|(z - s)^*|} = 1.$$

Next, we observe that

$$\arg w = \arg(z - s) - \arg(z) - \arg((z - s)^*) = 2\arg(z - s) - \frac{2\pi}{3}.$$

Consider now the locus of s. Let $W(w)$, $A(w/2)$ and $B(\mathrm{i}/\sqrt{3})$. One can verify that $w/2$ and $\mathrm{i}/\sqrt{3}$ are the solutions to $\arg(s) = \arg(z - s)$ when $\arg s = 2\pi/3, \pi/2$ respectively.



From the above diagram, it is clear that $\inf \arg(z - s)$ is attained at A, so

$$\inf \arg(z - s) = \frac{2\pi}{3}.$$

Meanwhile, $\sup \arg(z - s)$ is attained at B. Using the property that $|s| = |z - s|$, we see that $\triangle OBW$ is isosceles, with

$$\angle BOW = \angle BWO = \arg w - \frac{\pi}{2} = \frac{\pi}{6}.$$

Hence, $\angle WBO = \pi - 2(\pi/6) = 2\pi/3$, whence

$$\sup \arg(z - s) = 2\pi - \frac{2\pi}{3} - \frac{\pi}{2} = \frac{5\pi}{6}.$$

Thus,

$$\arg(z - s) \in \left(\frac{2\pi}{3}, \frac{5\pi}{6}\right) \implies \arg w = 2\arg(z - s) - \frac{2\pi}{3} \in \left(\frac{2\pi}{3}, \pi\right).$$

Part XII

Examinations

9758 H2 Mathematics

9758 JC1 Weighted Assessment 1

Problem 1. The curve C has parametric equations

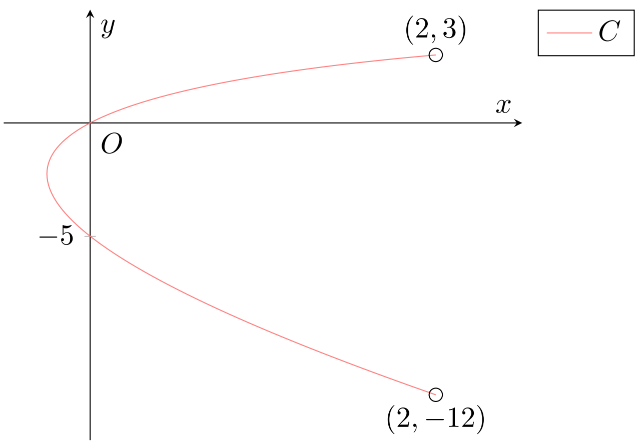
$$x = t^2 + t,\ y = 4t - t^2,\ -2 < t < 1.$$

(a) Sketch C, indicating the coordinates of the end-points and the axial intercepts (if any) of this curve.

(b) Find the coordinates of the point(s) of intersection between C and the line $8y - 12x = 5$.

Solution.

Part (a).



Part (b).

$$28y - 12x = 8(4t - t^2) - 12(t^2 + t) = 5 \implies -20t^2 + 20t - 5 = 0$$
$$\implies t^2 - t + \frac{1}{4} = 0 \implies \left(t - \frac{1}{2}\right)^2 = 0 \implies t = \frac{1}{2}.$$

When $t = 1/2$, we have that $x = 3/4$ and $y = 7/4$. Thus, C and the line $8y - 12x = 5$ intersect at $(3/4, 7/4)$.

* * * * *

Problem 2.

(a) Without using a calculator, solve $\frac{4}{3+2x-x^2} \leq 1$.

(b) Hence, solve $\frac{4}{3+2|x|-x^2} \leq 1$.

Solution.

Part (a).

$$\frac{4}{3+2x-x^2} \leq 1 \implies \frac{4}{x^2-2x-3} \geq -1$$
$$\implies \frac{4}{(x-3)(x+1)} + 1 = \frac{4+(x-3)(x+1)}{(x-3)(x+1)} = \frac{(x-1)^2}{(x-3)(x+1)} \geq 0$$

We thus have that $x = 1$ is a solution. In the case when $(x-1)^2 > 0$,

$$\frac{1}{(x-3)(x+1)} \geq 0 \implies (x-3)(x+1) \geq 0$$

whence $x < 1$ or $x > 3$. Putting everything together, we have

$$x < -1 \text{ or } x = 1 \text{ or } x > 3.$$

Part (b).

$$\frac{4}{3+2|x|-x^2} \leq 1 \implies \frac{4}{3+2|x|-|x|^2} \leq 1.$$

From part (a), we have that $|x| < -1$, $|x| = 1$ or $|x| > 3$.

Case 1: $|x| < -1$. Since $|x| \geq 0$ this case yields no solutions.

Case 2: $|x| = 1$. We have $x = 1$ or $x = -1$.

Case 3: $|x| > 3$. We have $x > 3$ or $x < -3$.

Thus,

$$x < -3 \text{ or } x = -1 \text{ or } x = 1 \text{ or } x > 3.$$

* * * * *

Problem 3. The curve C_1 has equation

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

(a) Sketch the graph of C_1, stating the equations of any asymptotes and the coordinates of any axial intercepts and/or turning points.

The curve C_2 has equation

$$\frac{(x-a)^2}{1^2} + \frac{(y-6)^2}{b^2} = 1$$

where $b > 0$. It is given that C_1 and C_2 have no points in common for all $a \in \mathbb{R}$.

(b) By adding an appropriate curve in part (a), state the range of values of b, explaining your answer.

(c) The function f is defined by

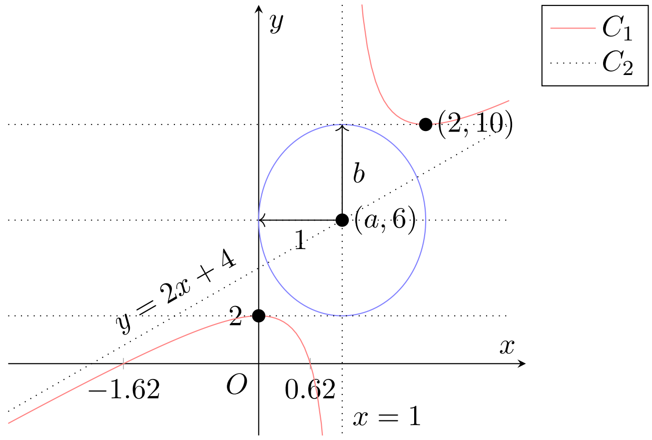
$$f(x) = \frac{2x^2+2x-2}{x-1}, \; x < 1.$$

(i) By using the graph in part (a) or otherwise, explain why the inverse function f^{-1} does not exist.

(ii) The domain of f is restricted to $[c, 1)$ such that c is the least value for which the inverse function f^{-1} exists. State the value of c and define f^{-1} clearly.

Solution.

Part (a).



Part (b). Observe that C_2 describes an ellipse with vertical radius b and horizontal radius 1. Furthermore, the ellipse is centred at $(a, 6)$. Since C_1 and C_2 have no points in common for all $a \in \mathbb{R}$, the maximum y-value of the ellipse corresponds to the y-value of the minimum point $(2, 10)$ of C_1. Similarly, the minimum y-value of the ellipse corresponds to the y-value of the maximum point $(0, 2)$ of C_1. Thus, $2 < y < 10$, whence $b < \min\{|6-2|, |6-10|\} = 4$. Thus, $0 < b < 4$.

Part (c).

Part (c)(i). Observe that $f(-1.62) = f(0.618) = 0$. Hence, there exist two different values of x in D_f that have the same image under f. Thus, f is not one-one. Hence, f^{-1} does not exist.

Part (c)(ii). Clearly, $c = 0$. We now find f^{-1}.

$$f(x) = \frac{2x^2 + 2x - 2}{x - 1} \implies xf(x) - f(x) = 2x^2 + 2x - 2$$

$$\implies 2x^2 + \left[2 - f(x)\right]x + \left[f(x) - 2\right] = 0 \implies x = \frac{f(x) - 2 \pm \sqrt{f(x)^2 - 12f(x) + 20}}{4}.$$

Replacing $x \mapsto f^{-1}(x)$, we get

$$f^{-1}(x) = \frac{x - 2 \pm \sqrt{x^2 - 12x + 20}}{4}.$$

Note that $D_f = R_{f^{-1}} = [0, 1)$. We thus take the positive root. Also note that $R_f = D_{f^{-1}} = (-\infty, 2]$. Hence,

$$f^{-1} : x \mapsto \frac{x - 2 + \sqrt{x^2 - 12x + 20}}{4}, \; x \in \mathbb{R},\; x \leq 2.$$

9758 JC1 Weighted Assessment 2

Problem 1. Differentiate $\arccos\left(\sqrt{1-4x}\right)$ with respect to x, simplifying your answer.

Solution.

$$\frac{\mathrm{d}}{\mathrm{d}x}\arccos\left(\sqrt{1-4x}\right) = -\frac{1}{\sqrt{1-(1-4x)}}\left(\frac{-4}{2\sqrt{1-4x}}\right) = \frac{2}{\sqrt{4x}\sqrt{1-4x}} = \frac{1}{\sqrt{x-4x^2}}.$$

* * * * *

Problem 2. It is given that x and y satisfy the equation $xy^2 = \ln\left(x^2 \mathrm{e}^y\right) - \frac{2\mathrm{e}}{x}$.

(a) Verify that $(\mathrm{e}, 0)$ satisfies the equation.

(b) Hence, show that at $y = 0$, $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{k}{\mathrm{e}}$, where k is a constant to be determined.

Solution.

Part (a). Substituting $x = \mathrm{e}$ and $y = 0$ into the given equation,

$$\text{LHS} = \mathrm{e}\cdot 0^2 = 0, \quad \text{RHS} = \ln\left(\mathrm{e}^2 \cdot e^0\right) - \frac{2\mathrm{e}}{\mathrm{e}} = 2 - 2 = 0.$$

Since the LHS is equal to the RHS, $(\mathrm{e}, 0)$ satisfies the equation.

Part (b). From the given equation, we have

$$xy^2 = 2\ln x + y - \frac{2\mathrm{e}}{x}.$$

Implicitly differentiating yields

$$x\left(2y\,\frac{\mathrm{d}y}{\mathrm{d}x}\right) + y^2 = \frac{2}{x} + \frac{\mathrm{d}y}{\mathrm{d}x} + \frac{2\mathrm{e}}{x^2}.$$

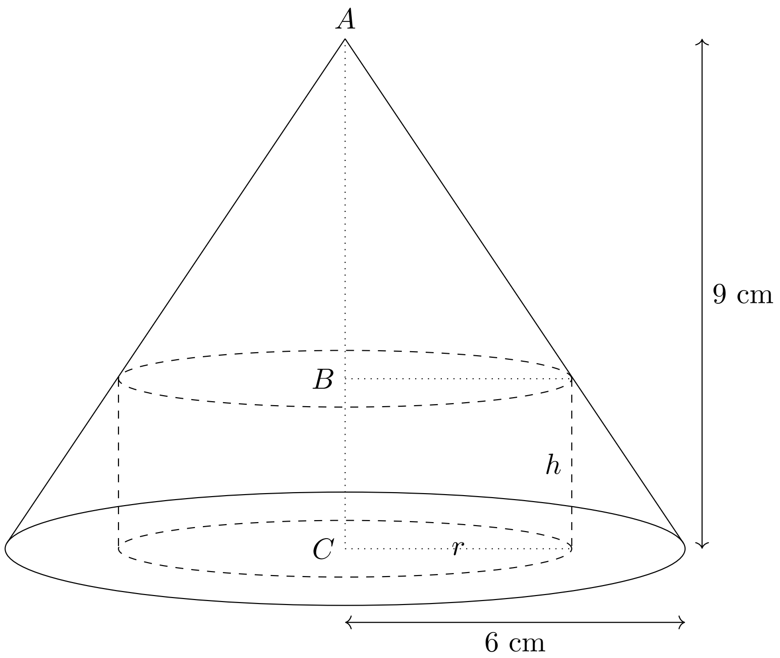
Substituting $x = e$ and $y = 0$ into the above equation gives

$$0 = \frac{2}{\mathrm{e}} + \frac{\mathrm{d}y}{\mathrm{d}x} + \frac{2\mathrm{e}}{\mathrm{e}^2} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{-4}{\mathrm{e}}.$$

Thus, $k = -4$.

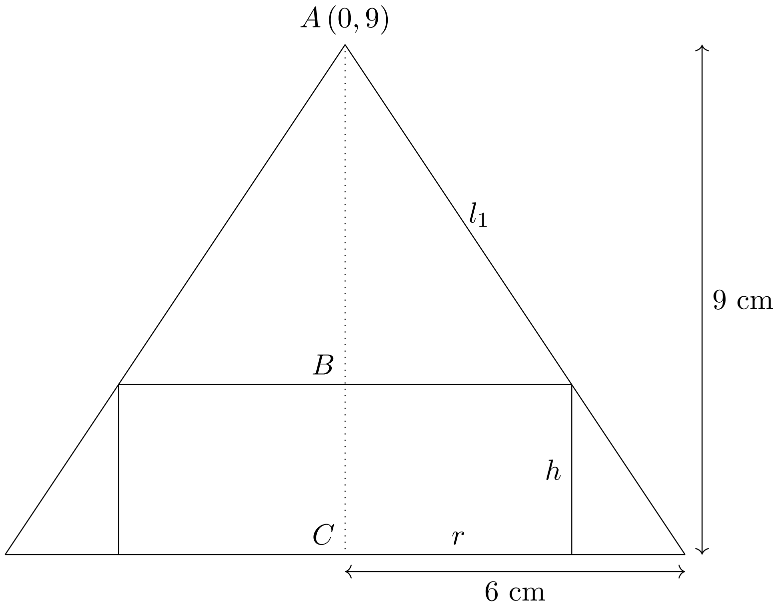
* * * * *

Problem 3. A toy manufacturer wants to make a toy in the shape of a right circular cone with a cylinder drilled out, as shown in the diagram below. The cylinder is inscribed in the cone. The circumference of the top of the cylinder is in contact with the inner surface of the cone and the base of the cylinder is level with the base of the cone. The base radius of the cylinder is r cm and the base radius of the cone is 6 cm. The height of the cylinder, BC, is h cm and the height of the cone, AC is 9 cm.



Using differentiation, find the minimum volume of the toy, V cm^3, in terms of π.

Solution.



Consider the diagram above. Let C be the origin. Note that l_1 has gradient $-\frac{9}{6} = -\frac{3}{2}$. Hence, l_1 has equation

$$l_1 : y = 9 - \frac{3}{2}x.$$

When $x = r$, we have $y = 9 - \frac{3}{2}r$. Thus, the height of the cylinder is $\left(9 - \frac{3}{2}r\right)$ cm. Let the volume of the cylinder be V_1 cm^3.

$$V_1 = \pi r^2 h = \pi r^2 \left(9 - \frac{3}{2}r\right) = 9\pi r^2 - \frac{3}{2}\pi r^3.$$

For stationary points, $\frac{\mathrm{d}V_1}{\mathrm{d}r} = 0$.

$$\frac{\mathrm{d}V_1}{\mathrm{d}r} = 0 \implies 18\pi r - \frac{9}{2}\pi r^2 = 0 \implies \frac{9}{2}\pi r(4 - r) = 0.$$

Hence, V_1 has a stationary point when $r = 4$. Note that we reject $r = 0$ since $r > 0$.

r	4^-	4	4^+
$\frac{\mathrm{d}V_1}{\mathrm{d}r}$	+ve	0	−ve

By the first derivative test, V_1 attains a maximum when $r = 4$. Hence,

$$\min V = \text{Volume of cone} - \max V_1 = \left[\frac{1}{3}\pi\left(6^2\right)(9)\right] - \left[9\pi\left(4^2\right) - \frac{3}{2}\pi\left(4^3\right)\right] = 60\pi.$$

Thus, the minimum volume of the toy is 60π cm^3.

* * * * *

Problem 4. A curve C has parametric equations

$$x = 2\theta + \sin 2\theta,\ y = \cos 2\theta,\ 0 \leq \theta \leq \pi.$$

(a) Find $\frac{\mathrm{d}y}{\mathrm{d}x}$, expressing your answer in terms of only a single trigonometric function.

(b) Hence, find the coordinates of point Q, on C, whose tangent is parallel to the y-axis.

Solution.

Part (a). Note that $\frac{\mathrm{d}x}{\mathrm{d}\theta} = 2 + 2\cos 2\theta$ while $\frac{\mathrm{d}y}{\mathrm{d}\theta} = -2\sin 2\theta$. Hence,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = \frac{-2\sin 2\theta}{2 + 2\cos 2\theta} = -\frac{\sin 2\theta}{1 + \cos 2\theta}$$
$$= -\frac{2\sin\theta\cos\theta}{1 + (2\cos^2\theta - 1)} = -\frac{2\sin\theta\cos\theta}{2\cos^2\theta} = -\frac{\sin\theta}{\cos\theta} = -\tan\theta.$$

Part (b). Since the tangent at Q is parallel to the y-axis, the derivative $\mathrm{d}y/\mathrm{d}x = -\tan\theta$ must be undefined there. Hence, $\cos\theta = 0 \implies \theta = \pi/2$. Substituting $\theta = \pi/2$ into the given parametric equations, we obtain $x = \pi$ and $y = -1$, whence $Q(\pi, -1)$.

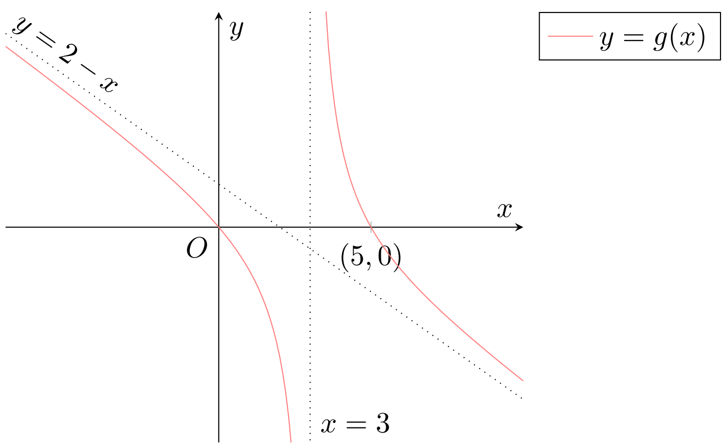
* * * * *

Problem 5.

(a) A function is defined as $f(x) = a(2 - x)^2 - b$, where a and b are positive constants such that $a < 1$ and $b > 4$.

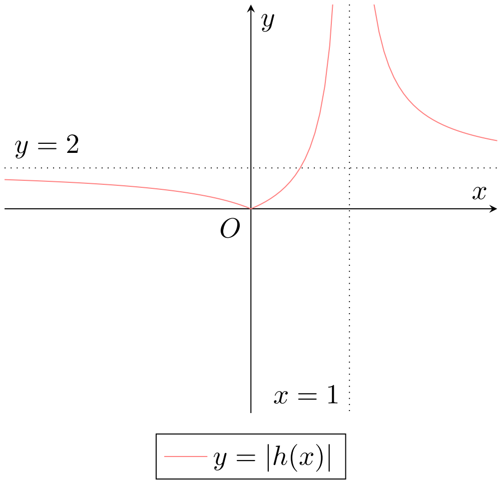
State a sequence of transformations that will transform the curve with equation $y = x^2$ on to the curve with equation $y = f(x)$.

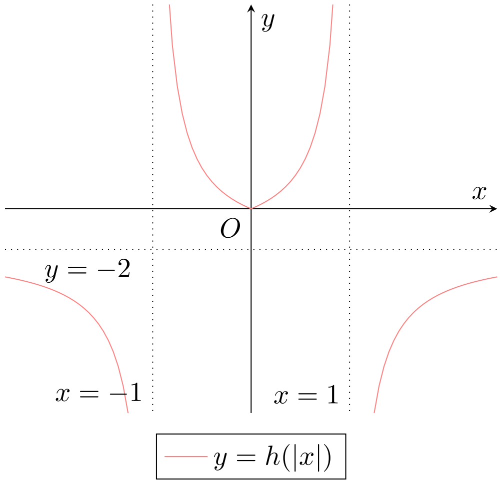
(b) The diagram shows the graph of $y = g(x)$. The lines $x = 3$ and $y = 2 - x$ are asymptotes to the curve, and the graph passes through the points $(0, 0)$ and $(5, 0)$.



Sketch the graph of $y = \frac{1}{g(x)}$, indicating clearly the coordinates of any axial intercepts (where applicable) and the equations of any asymptotes.

(c) Given the graphs of $y = |h(x)|$ and $y = h(|x|)$ below, sketch the two possible graphs of $y = h(x)$.



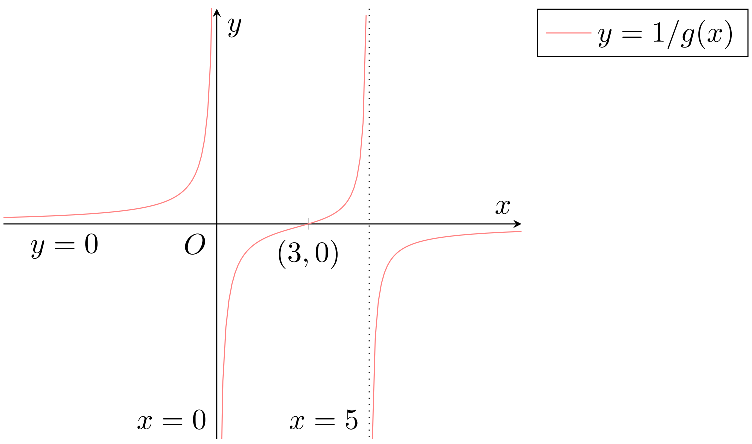


Solution.

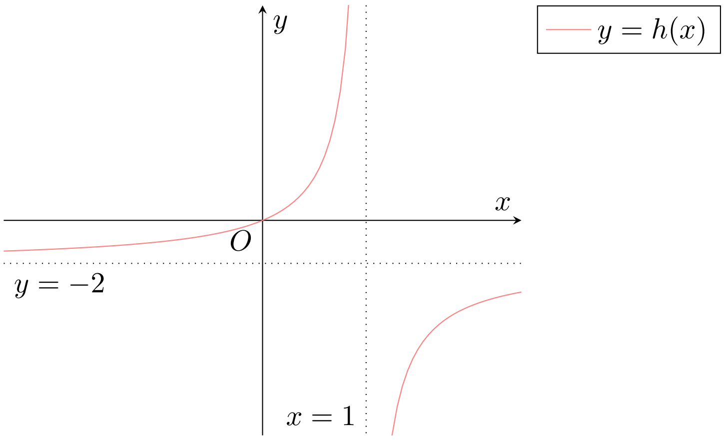
Part (a).

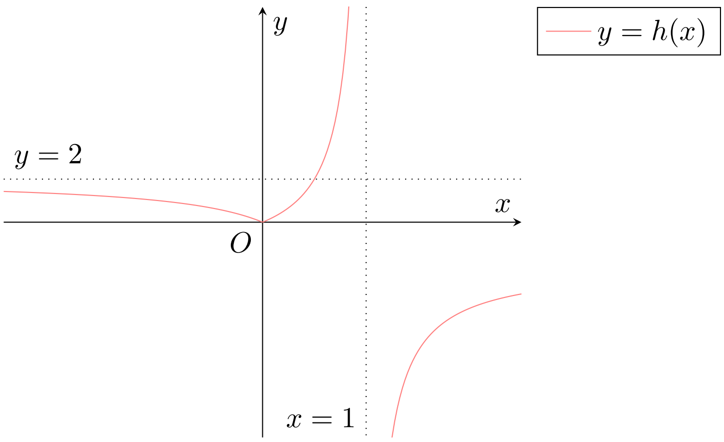
1. Translate the graph 2 units in the positive x-direction.
2. Scale the graph by a factor of a parallel to the y-axis.
3. Translate the graph b units in the negative y-direction.

Part (b).



Part (c).





9758 JC1 Promotional Examination

Problem 1. A dietitian in a hospital is to arrange a special diet meal composed of Food A, Food B and Food C. The diet is to include exactly 7800 units of calcium, 80 units of iron and 7.5 units of vitamin A. The number of units of calcium, iron and vitamin A for each gram of the foods is indicated in the table.

	Units Per Gram		
	Food A	Food B	Food C
Calcium	15	20	24
Iron	0.2	0.15	0.28
Vitamin A	0.015	0.02	0.02

Find the total weight of the foods, in grams, of this special diet.

Solution. Let a, b, and c represent the weight of Food A, B, and C respectively, in grams. We have the following system of equations:

$$\begin{cases} 15a + 20b + 24c = 7800 \\ 0.2a + 0.15b + 0.28c = 80 \\ 0.015a + 0.02b + 0.02c = 7.5 \end{cases}$$

Using G.C., $a = 160$, $b = 180$ and $c = 75$. Hence, the total weight of the foods is $160 + 180 + 75 = 415$ grams.

Problem 2. By expressing $\frac{3x^2+2x-12}{x-1}-(x+2)$ as a single simplified fraction, solve the inequality

$$\frac{3x^2+2x-12}{x-1} \geq x+2,$$

without using a calculator.

Solution.

$$\frac{3x^2+2x-12}{x-1}-(x+2)=\frac{3x^2+2x-12-(x+2)(x-1)}{x-1}$$
$$=\frac{3x^2+2x-12-(x^2+x-2)}{x-1}=\frac{2x^2+x-10}{x-1}=\frac{(x-2)(2x+5)}{x-1}.$$

Consider the inequality.

$$\frac{3x^2+2x-12}{x-1} \geq x+2 \implies \frac{3x^2+2x-12}{x-1}-(x+2) \geq 0$$
$$\implies \frac{(x-2)(2x+5)}{x-1} \geq 0 \implies (x-2)(2x+5)(x-1) \geq 0$$

Hence, $-5/2 \leq x < 1$ or $x \geq 2$.

* * * * *

Problem 3.

(a) Given that $\sum_{r=1}^{n} r^2 = \frac{n}{6}(n+1)(2n+1)$, evaluate $\sum_{r=-n}^{n}(r+1)(r+3)$ in terms of n.

(b) Using standard series from the List of Formulae (MF27), find the range of values of x for which the series $\sum_{r=1}^{\infty} \frac{(-1)^{r+1}x^r}{r2^r}$ converges. State the sum to infinity in terms of x.

Solution.

Part (a).

$$\sum_{r=-n}^{n}(r+1)(r+3)=\sum_{r=0}^{2n}(r-n+1)(r-n+3)=\sum_{r=0}^{2n}\left[r^2+r(4-2n)+(n^2-4n+3)\right]$$
$$=\frac{(2n)(2n+1)(4n+1)}{6}+(4-2n)\frac{(2n)(2n+1)}{2}+(n^2-4n+3)(2n+1)$$
$$=(2n+1)\left[\frac{n(4n+1)}{3}+(4-2n)(n)+(n^2-4n+3)\right]=\frac{2n+1}{3}\left(n^2+n+9\right).$$

Part (b).

$$\sum_{r=1}^{\infty}\frac{(-1)^{r+1}x^r}{r2^r}=\sum_{r=1}^{\infty}\frac{(-1)^{r+1}(x/2)^r}{r}=\ln\left(1+\frac{x}{2}\right).$$

Range of convergence: $-1 < x/2 \leq 1 \implies -2 < x \leq 2$.

* * * * *

Problem 4. The curve C has parametric equations

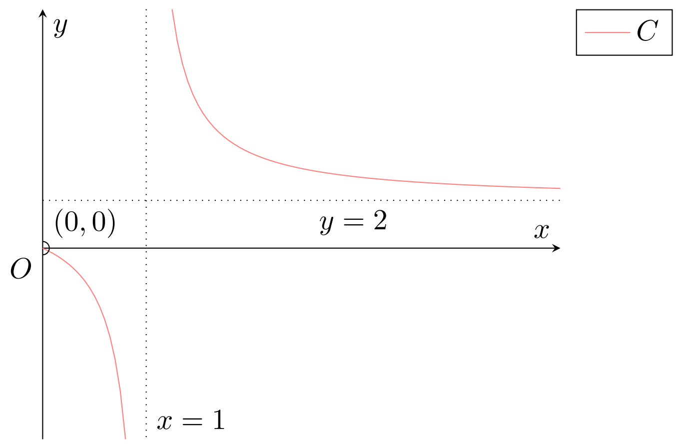
$$x=-\frac{2}{t-1}, \quad y=\frac{4}{t+1}, \quad t<1, \quad t \neq \pm 1.$$

(a) Sketch a clearly labelled diagram of C, indicating any axial intercepts and asymptotes (if any) of this curve.

(b) Find also its Cartesian equation, stating any restrictions where applicable.

Solution.

Part (a).



Part (b). Note that $x = -\frac{2}{t-1} \implies t = -\frac{2}{x} + 1$. Hence,

$$y = \frac{4}{(-2/x+1)+1} = \frac{4x}{-2+2x} = \frac{2x}{x-1}.$$

Thus,

$$y = \frac{2x}{x-1}, \quad x \neq 1, x > 0.$$

* * * * *

Problem 5.

(a) Find, using an algebraic method, the exact roots of the equation $\left|3x^2+5x-8\right| = 4-x$.

(b) On the same axes, sketch the curves with equations $y = \left|3x^2+5x-8\right|$ and $y = 4-x$. Hence, solve exactly the inequality $\left|3x^2+5x-8\right| < 4-x$.

Solution.

Part (a). *Case 1*: $3x^2+5x-8 = 4-x$.

$$3x^2+5x-8 = 4-x \implies 3x^2+6x-12 = 0 \implies x^2+2x-4 = 0.$$

By the quadratic formula, we get

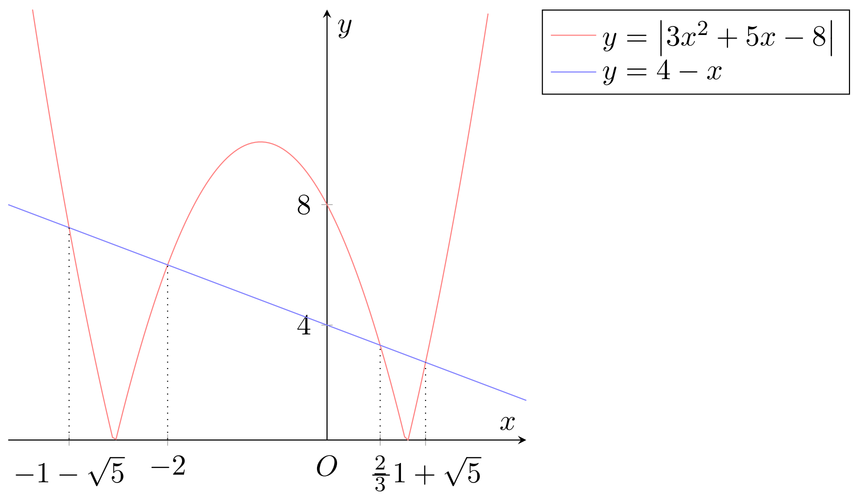
$$x = -1 \pm \sqrt{5}.$$

Case 2: $3x^2+5x-8 = -(4-x)$.

$$3x^2+5x-8 = -4+x \implies 3x^2+4x-4 = (3x-2)(x+2) = 0 \implies x = \frac{2}{3} \text{ or } -2.$$

Hence, the roots are $x = -1 \pm \sqrt{5}$, $2/3$ and -2.

Part (b).



From the graph, $-1-\sqrt{5} < x < -2$ or $\frac{2}{3} < x < -1+\sqrt{5}$.

* * * * *

Problem 6.

(a) The transformations A, B and C are given as follows:

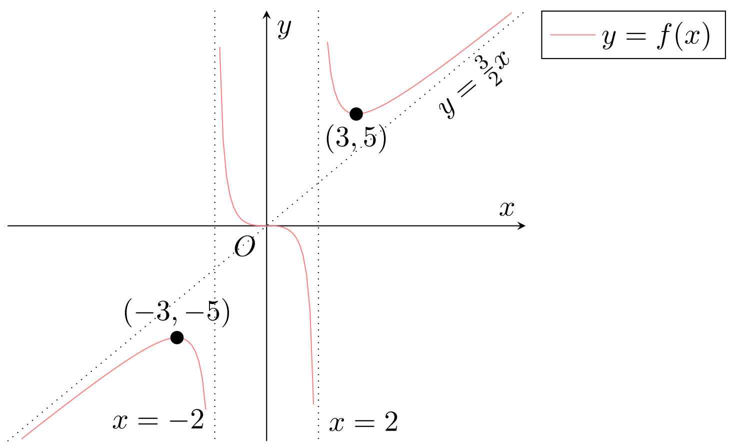
A: A reflection in the x-axis.

B: A translation of 1 unit in the positive y-direction.

C: A translation of 2 units in the negative x-direction.

A curve undergoes in succession, the transformations A, B and C, and the equation of the resulting curve is $y = \frac{2x+1}{2x+2}$. Determine the equation of the curve before the transformations, expressing your answer as a single fraction.

(b) The diagram shows the curve $y = f(x)$. The lines $x = -2$, $x = 2$ and $y = \frac{3}{2}x$ are asymptotes to the curve. The curve has turning points at $(-3, -5)$ and $(3, 5)$. It also has a stationary point of inflexion at the origin O.



(i) State the range of values of x for which the graph is concave downwards.

(ii) Sketch the graph of $y = \frac{1}{f(x)}$.

(iii) Sketch the graph of $y = f'(x)$.

Solution.

Part (a). Observe that

$$\begin{aligned} A : y \mapsto -y &\implies A^{-1} : y \mapsto -y \\ B : y \mapsto y-1 &\implies B^{-1} : y \mapsto y+1 \\ C : x \mapsto x+2 &\implies C^{-1} : x \mapsto x-2 \end{aligned}$$

Hence,

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

Thus,

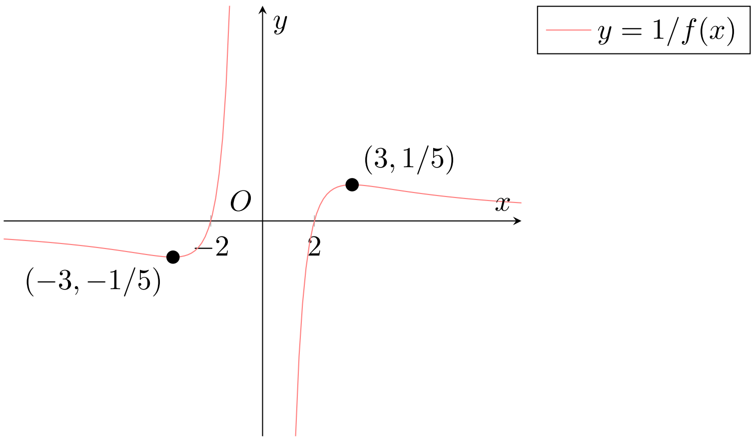
$$y = 1 - \frac{2x-3}{2x-2} = \frac{1}{2x-2}.$$

Part (b).

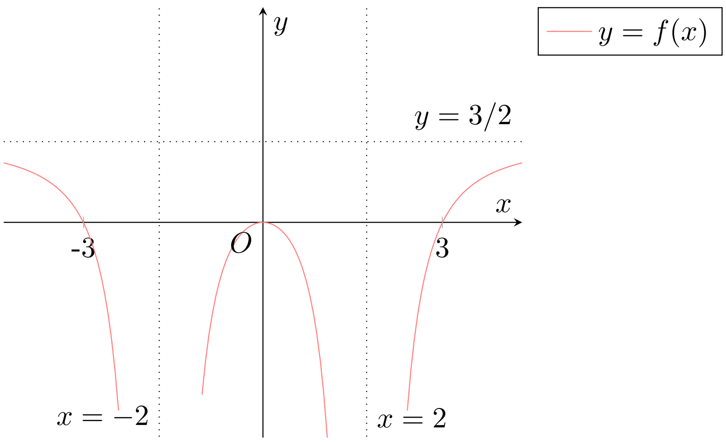
Part (b)(i). From the graph, we clearly have

$$x < -2 \text{ or } 0 < x < 2.$$

Part (b)(ii).



Part (b)(iii).



* * * * *

Problem 7. A curve has parametric equations

$$x = 3u^2, \quad y = 6u.$$

(a) Find the equations of the normal to the curve at the point $P\left(3p^2, 6p\right)$, where p is a non-zero constant.

(b) The normal meets the x-axis at Q and the y-axis at R. Find the coordinates of Q and of R.

(c) Find two possible expressions for the area bounded by the normal and the axes in terms of p, stating the range of values of p in each case.

(d) Given that p is positive and increasing at a rate of 2 units/s, find the rate of change of the area of the triangle in terms of p.

Solution.

Part (a). Note that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}u}{\mathrm{d}x/\mathrm{d}u} = \frac{6}{6u} = \frac{1}{u}.$$

At $u = p$, the gradient of the normal is hence $-\frac{1}{1/p} = -p$. Thus, the equation of the normal at P is

$$y - 6p = -p\left(x - 3p^2\right).$$

Part (b). At Q, $y = 0$. Hence, $x = 6 + 3p^2$, whence $Q(6 + 3p^2, 0)$. At R, $x = 0$. Hence, $y = 6p + 3p^2$, whence $R(0, 6p + 3p^2)$.

Part (c). When $p > 0$, the area of the triangle is $\frac{1}{2}(6p+3p^3)(6+3p^2)$ units^2. When $p < 0$, the area of the triangle is $-\frac{1}{2}(6p + 3p^3)(6 + 3p^2)$ units^2.

Part (d). Let the area of the triangle be A unit^2. Since $p > 0$, we have

$$A = \frac{1}{2}(6p + 3p^3)(6 + 3p^2).$$

Thus,

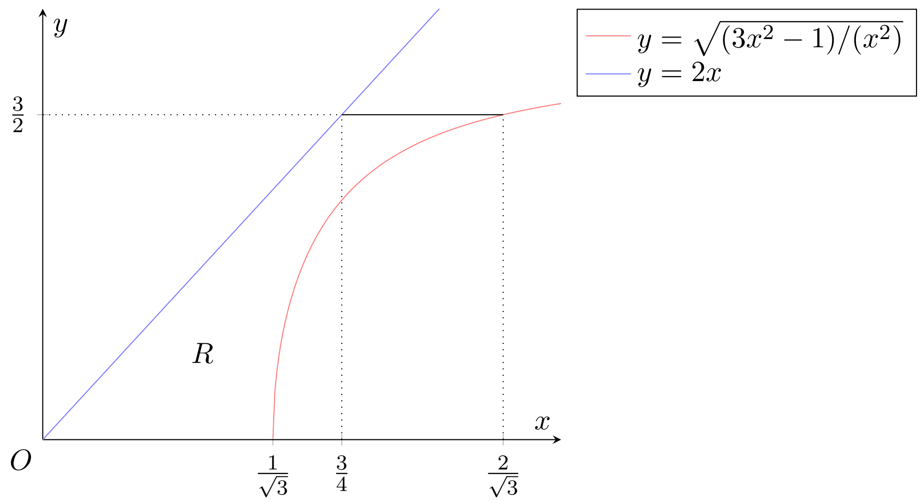
$$\frac{\mathrm{d}A}{\mathrm{d}p} = \frac{1}{2}\left[(6p + 3p^3)(6p) + (6 + 3p^2)(6 + 9p^2)\right] = \frac{9}{2}\left(5p^4 + 12p^2 + 4\right).$$

Hence, the rate of change of area of the triangle is

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\mathrm{d}A}{\mathrm{d}p} \cdot \frac{\mathrm{d}p}{\mathrm{d}t} = 9\left(5p^4 + 12p^2 + 4\right) \text{ units}^2/\text{s}.$$

* * * * *

Problem 8. The shaded region R is bounded by the lines $y = 2x$, $y = \frac{3}{2}$, the x-axis and the curve $y = \sqrt{\frac{3x^2-1}{x^2}}$.



(a) By using the substitution $x = \frac{1}{\sqrt{3}}\sec\theta$, find the exact value of $\int_{1/\sqrt{3}}^{2/\sqrt{3}} \sqrt{\frac{3x^2-1}{x^2}}\,\mathrm{d}x$.

(b) Hence, find the exact area of the shaded region R.

(c) Find the volume of the solid generated when R is rotated through 2π radians about the x-axis, giving your answer in 3 decimal places.

Solution.

Part (a). Note that

$$x = \frac{1}{\sqrt{3}}\sec\theta \implies \mathrm{d}x = \frac{1}{\sqrt{3}}\sec\theta\tan\theta\,\mathrm{d}\theta,$$

with the bounds of integration going from $\theta = 0$ to $\pi/3$.

$$\int_{1/\sqrt{3}}^{2/\sqrt{3}} \sqrt{\frac{3x^2-1}{x^2}}\,\mathrm{d}x = \int_0^{\pi/3} \sqrt{\frac{\tan^2\theta}{\frac{1}{3}\sec^2\theta}}\frac{1}{\sqrt{3}}\sec\theta\tan\theta\,\mathrm{d}\theta = \int_0^{\pi/3} \tan^2\theta\,\mathrm{d}\theta$$

$$= \int_0^{\pi/3}\left(\sec^2\theta - 1\right)\,\mathrm{d}\theta = [\tan\theta - t]_0^{\pi/3} = \sqrt{3} - \frac{\pi}{3}.$$

Part (b).

$$\text{Area}\, R = \underbrace{\frac{1}{2}\left[\frac{2}{\sqrt{3}} + \left(\frac{2}{\sqrt{3}} - \frac{3}{4}\right)\right]\frac{3}{2}}_{\text{area of trapezium}} - \int_{1/\sqrt{3}}^{2/\sqrt{3}} \sqrt{\frac{3x^2-1}{x^2}}\,\mathrm{d}x = \frac{\pi}{3} - \frac{9}{16}\text{ units}^2.$$

Part (c).

$$\begin{aligned}\text{Volume} &= \underbrace{\frac{1}{3}\pi\left(\frac{3}{2}\right)^2\left(\frac{3}{4}\right)}_{\text{volume of cone}} + \underbrace{\pi\left(\frac{2}{\sqrt{3}} - \frac{3}{4}\right)\left(\frac{3}{2}\right)^2}_{\text{volume of cylinder}} - \pi\int_{1/\sqrt{3}}^{2/\sqrt{3}}\left(\sqrt{\frac{3x^2-1}{x^2}}\right)^2\mathrm{d}x \\ &= 1.907\text{ units}^3.\end{aligned}$$

* * * * *

Problem 9. It is given that $y = \arccos(2x)\arcsin(2x)$, where $-0.5 \leq x \leq 0.5$, and $\arccos(2x)$ and $\arcsin(2x)$ denote their principal values.

(a) Show that $\left(1-4x^2\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 4x\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right) + 8 = 0$. Hence, find the MacLaurin series for y up to and including the term in x^3, giving the coefficients in exact form.

(b) Use your expansion from part (a) and integration to find an approximate value for $\int_{0.1}^{0.2}\left(\frac{2}{x}\right)^3\arccos(2x)\arcsin(2x)\,\mathrm{d}x$, correct to 4 decimal places.

(c) A student, Adam, claims that the approximation in part (b) is accurate. Without performing any further calculations, justify whether Adam's claim is valid.

(d) Suggest one way to improve the accuracy of the approximated value obtained in part (b).

Solution.

Part (a). Differentiating y with respect to x,

$$y' = 2\left(\frac{\arccos(2x) - \arcsin(2x)}{\sqrt{1-4x^2}}\right).$$

Differentiating once more,

$$\begin{aligned}y'' &= \frac{2}{\sqrt{1-4x^2}}\left(-\frac{4}{\sqrt{1-4x^2}}\right) + \frac{2\left[\arccos(2x) - \arcsin(2x)\right](-8x)}{-2\left(1-4x^2\right)^{3/2}} \\ &= \frac{1}{1-4x^2}\left[-8 + 4x\cdot\frac{2\left[\arccos(2x)-\arcsin(2x)\right]}{\sqrt{1-4x^2}}\right] = \frac{1}{1-4x^2}\left(-8+4xy'\right).\end{aligned}$$

Hence,

$$\left(1-4x^2\right)y'' - 4xy' + 8 = 0.$$

Differentiating with respect to x, we get

$$\left(1-4x^2\right)y''' - 12xy'' - 4y' = 0.$$

Evaluating y, y', y'' and y''' at $x=0$, we get

$$y(0) = 0, \quad y'(0) = \pi, \quad y''(0) = -8, \quad y'''(0) = 4\pi.$$

Hence,

$$y = \pi x - 4x^2 + \frac{2}{3}\pi x^3 + \cdots.$$

.

Part (b).

$$\int_{0.1}^{0.2}\left(\frac{2}{x}\right)^3 \arccos(2x)\arcsin(2x)\ \mathrm{d}x = 8\int_{0.1}^{0.2} x^{-3}\arccos(2x)\arcsin(2x)\ \mathrm{d}x$$

$$\approx 8\int_{0.1}^{0.2} x^{-3}\left[\pi x - 4x^2 + \frac{2}{3}\pi x^3\right]\ \mathrm{d}x = 8\int_{0.1}^{0.2}\left(\pi x^{-2} - 4x^{-1} + \frac{2}{3}\pi x\right)\ \mathrm{d}x$$

$$= 8\left[-\frac{\pi}{x} - 4\ln|x| + \frac{2}{3}\pi x\right]_{0.1}^{0.2} = 105.1585 \text{ (4 d.p.)}.$$

Part (c). Adam's claim is valid. Since the approximation for $\arccos(2x)\arcsin(2x)$ is accurate for x near 0, and we are integrating over $(0.1, 0.2)$ (which is close to 0), the integral approximation should also be accurate.

Part (d). Consider more terms in the MacLaurin series of $y = \arccos(2x)\arcsin(2x)$ and use the improved series in the approximation for the integral.

* * * * *

Problem 10. The function f is defined by

$$f : x \mapsto 3\sin\left(2x - \frac{1}{6}\pi\right), \quad 0 \leq x \leq k.$$

(a) Show that the largest exact value of k such that f^{-1} exists is $\frac{1}{3}\pi$. Find $f^{-1}(x)$.

(b) It is given that $k = \frac{1}{3}\pi$. In a single diagram, sketch the graphs of $y = f(x)$ and $y = f^{-1}(x)$, labelling your graphs clearly.

The function h is defined by $h : x \mapsto 3\sin\left(2x - \frac{1}{6}\pi\right)$, $0 \leq x \leq \frac{1}{3}\pi$, $x = \frac{1}{12}\pi$.
Another function g is defined by $g : x \mapsto \left|3 - \frac{1}{x}\right|$, $-3 \leq x \leq 3$, $x \neq 0$.

(c) Explain clearly why gh exists. Find $gh(x)$ and its range.

(d) Supposing $(gh)^{-1}$ exists for some restriction, find the exact value of x for which $(gh)^{-1}(x) = 0$. Show your working clearly.

Solution.

Part (a). For f^{-1} to exist, f must be one-one. Since $\frac{1}{3}\pi$ is the first maximum point of f, it is the largest value of k.

Let $y = f(x)$.

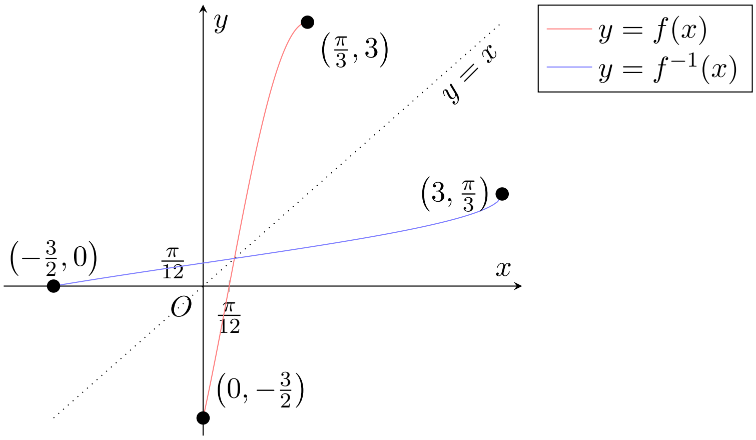
$$3\sin\left(2x - \frac{1}{6}\pi\right) = y \implies \sin\left(2x - \frac{1}{6}\pi\right) = \frac{y}{3}$$

$$\implies 2x - \frac{1}{6}\pi = \arcsin\left(\frac{y}{3}\right) \implies x = \frac{1}{2}\arcsin\left(\frac{y}{3}\right) + \frac{1}{12}\pi.$$

Hence,

$$f^{-1}(x) = \frac{1}{2}\arcsin\left(\frac{x}{3}\right) + \frac{1}{12}\pi.$$

Part (b).



Part (c). Since $R_h = [-\frac{3}{2}, 0) \cup (0, 3]$ and $D_g = [-3, 0) \cup (0, 3]$, we have $R_h \subseteq D_g$, whence gh exists.

$$gh(x) = g\left[3\sin\left(2x - \frac{1}{6}\pi\right)\right] = \left|3 - \frac{1}{3\sin\left(2x - \frac{1}{6}\pi\right)}\right|.$$

When $h(x) = \frac{1}{3}$, $gh(x) = 0$. When $x \to \frac{1}{12}\pi$, $gh(x) \to \infty$. Hence, $R_{gh} = [0, \infty)$.

Part (d).

$$gh^{-1}(x) = 0 \implies x = gh(0) = \left|3 - \frac{1}{3\sin\left(-\frac{1}{6}\pi\right)}\right| = \frac{11}{3}.$$

* * * * *

Problem 11.

(a) (i) Express $\frac{2x}{(x+1)(x^2+1)} = \frac{A}{x+1} + \frac{Bx+c}{x^2+1}$, where A, B and C are constants to be found.

(ii) Evaluate $\int_0^1 \frac{\ln\left(1+x^2\right)}{(x+1)^2} \, dx$, giving your answer in the form $a\pi - \ln b$, where a and b are positive constants to be found.

(b) Find $\int \sin^3(kx) \, dx$, where k is a constant.

Solution.

Part (a).

Part (a)(i). Clearing denominators, we have

$$2x = A(x^2 + 1) + (Bx + C)(x + 1) = (A + B)x^2 + (B + C)x + (A + C).$$

Comparing coefficients of x^2, x and constant terms, we have

$$\begin{cases} A + B = 0 \\ B + C = 2 \\ A + C = 0 \end{cases}$$

Hence, $A = -1$, $B = 1$ and $C = 1$, giving

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

Part (a)(ii). Note that
$$\frac{x+1}{x^2+1} = \frac{1}{2}\left(\frac{2x}{x^2+1} + \frac{2}{x^2+1}\right).$$
Hence,
$$\begin{aligned}\int_0^1 \frac{2x}{(x+1)(x^2+1)}\,\mathrm{d}x &= \int_0^1 \left[-\frac{1}{x+1} + \frac{1}{2}\left(\frac{2x}{x^2+1} + \frac{2}{x^2+1}\right)\right]\mathrm{d}x \\ &= \left[-\ln|x+1| + \frac{1}{2}\left(\ln\left(x^2+1\right) + 2\arctan x\right)\right]_0^1 = -\frac{\ln 2}{2} + \frac{\pi}{4}.\end{aligned}$$
Thus,
$$\begin{aligned}\int_0^1 \frac{\ln\left(1+x^2\right)}{(x+1)^2}\,\mathrm{d}x &= \left[-\frac{\ln\left(1+x^2\right)}{x+1}\right]_0^1 + \int_0^1 \frac{2x}{(x+1)(x^2+1)}\,\mathrm{d}x \\ &= -\frac{\ln 2}{2} + \left(-\frac{\ln 2}{2} + \frac{\pi}{2}.\right) = -\ln 2 + \frac{\pi}{4}.\end{aligned}$$
Hence, $a = \frac{1}{4}$ and $b = 2$.

Part (b). Note that $\sin 3u = 3\sin u - 4\sin^3 u$, whence $\sin^3 u = \frac{3\sin u - \sin 3u}{4}$.
$$\int \sin^3(kx)\,\mathrm{d}x = \int \frac{3\sin(kx) - \sin(3kx)}{4}\,\mathrm{d}x = \frac{-3\cos(kx) + \frac{1}{3}\cos(3kx)}{4k} + C.$$

* * * * *

Problem 12. Alan wants to sign up for a triathlon competition which requires him to swim for 1.5 km, cycle for 30 km and run for 10 km. He plans a training programme as follows: In the first week, Alan is to swim 400 m, cycle 1 km and run 400 m. Each subsequent week, he will increase his swimming distance by 50 m, his cycling distance by 15% and his running distance by r%.

(a) Given that Alan will run 829.44 m in Week 5, show that $r = 20$. Hence, determine the distance that Alan will run in Week 20, giving your answer to the nearest km.

(b) Determine the week when Alan first achieves the distances required for all three categories of the competition.

During a particular running practice, Alan plans to run q metres in the first minute. The distance he covers per minute will increase by 80 m for the next four minutes. Subsequently, he will cover 6% less distance in a minute than that in the previous minute.

(c) Show that the distance, in metres, Alan will cover in the sixth minute is $0.94q + 300.8$, and hence find the minimum value of q, to the nearest metre, such that he can eventually complete 10 km.

While training, Alan suffers from inflammation and needs medication. The concentration of the medicine in the bloodstream after administration can be modelled by the recurrence relation
$$C_{n+1} = C_n \mathrm{e}^{-\left(p + \frac{1}{n+100}\right)},$$
where n represents the number of complete hours from which the medicine is first taken and p is the decay constant.

(d) The dosage of the medicine prescribed for Alan is 200 mg and the concentration of the medicine dropped to approximately 168 mg one hour later. It is given that his pain will be significant once the concentration falls below 60 mg. Determine after which complete hour he would feel significant pain and should take his medicine again.

Solution.

Part (a). Let R_n m be the distance ran in the nth week. We have $R_1 = 400$ and $R_{n+1} = \left(1 + \frac{r}{100}\right) R_n$, whence

$$R_n = 400 \left(1 + \frac{r}{100}\right)^{n-1}.$$

Since $R_5 = 829.44$, we have

$$400 \left(1 + \frac{r}{100}\right)^4 = 829.44 \implies \left(1 + \frac{r}{100}\right)^4 = 2.0736 \implies 1 + \frac{r}{100} = 1.2 \implies r = 20.$$

Hence,

$$R_20 = 400 \cdot 1.2^19 = 12779.2 = 13000,$$

rounded to the nearest thousand. Hence, Alan will run 13 km in week 20.

Part (b). Let S_n, C_n be the distance swam and cycled in week n, respectively, in metres.

We have $S_1 = 400$ and $S_{n+1} = S_n + 50$, whence

$$S_n = 400 + (n-1)50.$$

Consider $S_n \geq 1500$. Then $n \geq 23$.

We have $C_1 = 1000$ and $C_{n+1} = 1.15C_n$, whence

$$C_n = 1000 \left(1.15^{n-1}\right).$$

Consider $C_n \geq 30000$. Then $n \geq \log_{1.15} 30 = 25.3$.

Consider $R_n \geq 10000$. Then $n \geq 1 + \log_{1.2} 25 = 18.7$.

Hence, the minimum n required is 26. Thus, in week 26, Alan will achieve all distances required.

Part (c). The distance Alan will run in the 6th minute is $(q+4\cdot 80)(1-0.06) = 0.94q+300.8$.

Let d_n be the distance travelled in the nth minute, where $n \geq 6$. We have $d_6 = 0.94q + 300.8$ and $d_{n+1} = (1 - 0.06)d_n$, whence

$$d_n = 0.94^{n-6} d_6.$$

The total distance Alan will eventually travel is thus given by

$$5q + 80(4+3+2+1) + \sum_{n=6}^{\infty} (0.94)^{n-6} d_6 = 5q + 800 + \frac{0.94q + 300.8}{1 - 0.94}.$$

Let the above expression be greater than 10000. Then $q \geq 202.581$. Hence, $\min q = 203$.

Part (d). We have $C_0 = 200$ and $C_1 = 168$. We hence have

$$168 = 200\mathrm{e}^{-(p+\frac{1}{100})},$$

whence $p = 0.16435$. Using G.C., the first time $C_n \leq 60$ occurs when $n = 7$. Thus, after 7 complete hours, Alan will feel significant pain and should take his medicine again.

9758 JC2 Weighted Assessment 1

Problem 1. A study of the ant population living in the forest is being conducted. It is suggested that the population of ants, x thousand at time t years, can be modelled by the differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = kx(4-x),$$

where k is a constant. When $t = 0$, then population size is 1000. It is also known that $\mathrm{d}x/\mathrm{d}t = 0.75$ when the population size is 1500.

(a) Show that $k = 0.2$.

(b) Find an expression for x in terms of t.

(c) Explain what happens to the ant population in the long run.

Solution.

Part (a). We have

$$\left.\frac{\mathrm{d}x}{\mathrm{d}t}\right|_{x=1.5} = k(1.5)(4-1.5) = 0.75 \implies k = \frac{0.75}{1.5(4-1.5)} = 0.2.$$

Part (b). Rearranging the given DE, we have

$$\frac{x}{4-x}\frac{\mathrm{d}x}{\mathrm{d}t} = 0.2.$$

Integrating both sides with respect to t,

$$\int \frac{1}{x(4-x)}\,\mathrm{d}x = \int 0.2\,\mathrm{d}t = 0.2t + C_1.$$

Splitting the LHS using partial fractions yields

$$\int \frac{1}{x(4-x)}\,\mathrm{d}x = \frac{1}{4}\int\left(\frac{1}{x}+\frac{1}{4-x}\right)\mathrm{d}x = \frac{1}{4}\left(\ln x - \ln(4-x)\right) + C_2 = \frac{1}{4}\ln\frac{x}{4-x} + C_2.$$

Note that x has an unstable equilibrium at $x = 0$ and a stable equilibrium at $x = 4$. Since $x(0) = 1 \in (0,4)$, it follows that $x(t) \in (0,4)$ for all $t \geq 0$. Thus, $x > 0$ and $4 - x > 0$. Hence,

$$\frac{1}{4}\ln\frac{x}{4-x} + C_2 = 0.2t + C_1 \implies \ln\frac{x}{4-x} = \frac{4}{5}t + C_3 \implies \frac{x}{4-x} = C_4\mathrm{e}^{4t/5}.$$

Since $x(0) = 1$, we have $C_4 = 1/3$. Simple algebraic manipulation yields the final expression

$$x = \frac{4\mathrm{e}^{4t/5}}{3+\mathrm{e}^{4t/5}}.$$

Part (c). Observe that

$$\lim_{t\to\infty} x = \lim_{t\to\infty}\frac{4\mathrm{e}^{4t/5}}{3+\mathrm{e}^{4t/5}} = \lim_{t\to\infty}\frac{4}{3\mathrm{e}^{-4t/5}+1} = \frac{4}{1} = 4.$$

Thus, in the long run, the ant population will approach 4 thousand.

* * * * *

Problem 2. The equations of the lines l_1 and l_2 are

$$\mathbf{r} = \begin{pmatrix}5\\1\\3\end{pmatrix} + s\begin{pmatrix}-2\\1\\-4\end{pmatrix},\ s \in \mathbb{R} \quad \text{and} \quad \mathbf{r} = \begin{pmatrix}-3\\0\\6\end{pmatrix} + t\begin{pmatrix}c\\d\\2\end{pmatrix},\ t \in \mathbb{R}$$

respectively, where c and d are real constants.

(a) Find c and d if the lines l_1 and l_2 are parallel.

(b) For $c = 0$ and $d = 2$,

(i) determine if lines l_1 and l_2 intersect. Find the acute angle between them.

(ii) find the shortest distance of the point $(-3, 0, 6)$ from l_1.

Solution.

Part (a). If l_1 and l_2 are parallel, so are their direction vectors. Thus,

$$\begin{pmatrix} c \\ d \\ 2 \end{pmatrix} = k \begin{pmatrix} -2 \\ 1 \\ -4 \end{pmatrix}$$

for some $k \in \mathbb{R}$. Comparing the z-coordinates, we see that $2 = -4k$, whence $k = -1/2$. Thus, $c = -(1/2)(-2) = 1$ and $d = -(1/2)(1) = -1/2$.

Part (b).

Part (b)(i). Equating the two, we have

$$\begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix} + s \begin{pmatrix} -2 \\ 1 \\ -4 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 6 \end{pmatrix} + t \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} \implies \begin{pmatrix} -2 \\ 1 \\ -4 \end{pmatrix} + t \begin{pmatrix} 0 \\ -2 \\ -2 \end{pmatrix} = \begin{pmatrix} -8 \\ -1 \\ 3 \end{pmatrix}.$$

This gives the system

$$\begin{cases} -2s \phantom{{}-2t} = -8 \\ s - 2t = -1 \\ -4s - 2t = 3 \end{cases}.$$

Using G.C., we see that this system has no solutions. Thus, l_1 and l_2 do not have a common point, whence they do not intersect.

Let θ be the acute angle between l_1 and l_2. Then

$$\cos\theta = \frac{\left|(-2,\ 1,\ -4)^{\mathsf{T}} \cdot (0,\ 2,\ 2)^{\mathsf{T}}\right|}{\left|(-2,\ 1,\ -4)^{\mathsf{T}}\right| \left|(0,\ 2,\ 2)^{\mathsf{T}}\right|} = \frac{6}{\sqrt{21}\sqrt{8}}.$$

Thus, $\theta = 62.4^\circ$ (1 d.p.).

Part (b)(ii). Let $\overrightarrow{OA} = (5,\ 1,\ 3)^{\mathsf{T}}$ and $\overrightarrow{OB} = (-3,\ 0,\ 6)^{\mathsf{T}}$. Then

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \begin{pmatrix} -3 \\ 0 \\ 6 \end{pmatrix} - \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -8 \\ -1 \\ 3 \end{pmatrix}.$$

Let N be the foot of perpendicular from B to l_1. The shortest distance is thus given by

$$BN = \left| \overrightarrow{AB} \times \frac{(-2,\ 1,\ -4)^{\mathsf{T}}}{\left|(-2,\ 1,\ -4)^{\mathsf{T}}\right|} \right| = \frac{1}{\sqrt{21}} \left| \begin{pmatrix} 1 \\ -38 \\ -10 \end{pmatrix} \right| = \frac{\sqrt{1545}}{\sqrt{21}} = 8.58 \text{ units (3 s.f.)}.$$

* * * * *

Problem 3. $OABC$ is a rhombus, where $\overrightarrow{OA} = \mathbf{a}$, $\overrightarrow{OC} = \mathbf{c}$ and O is the origin. The point M lies on AB, between A and B, such that $\overrightarrow{AM} = k\overrightarrow{AB}$, where k is a positive constant. The point N lies on BC, between B and C, such that $\overrightarrow{NC} = s\overrightarrow{BC}$, where s is a positive constant. The area of triangle OAM is λ times the area of triangle OMN.

(a) Show that OB is perpendicular to AC.

(b) By finding the area of triangle OAM and OMN in terms of $\mathbf{a}$ and $\mathbf{c}$, find λ in terms of k and s.

(c) The point P divides MN in the ratio $3:2$. It is given that $\overrightarrow{OP} = 3\overrightarrow{PB}$. Find the values of k and s.

Solution.

Part (a). Note that

$$\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB} = \overrightarrow{OA} + \overrightarrow{OC} = \mathbf{a} + \mathbf{c}.$$

Further,

$$\overrightarrow{AC} = \overrightarrow{OC} - \overrightarrow{OA} = \mathbf{c} - \mathbf{a}.$$

Thus,

$$\overrightarrow{OB} \cdot \overrightarrow{AC} = (\mathbf{a}+\mathbf{c}) \cdot (\mathbf{c}-\mathbf{a}) = \mathbf{a}\cdot\mathbf{a} - \mathbf{a}\cdot\mathbf{a} + \mathbf{c}\cdot\mathbf{c} - \mathbf{c}\cdot\mathbf{a} = \mathbf{c}\cdot\mathbf{c} - \mathbf{a}\cdot\mathbf{a} = |\mathbf{c}|^2 - |\mathbf{a}|^2 = 0,$$

where in the last step we used the fact that $|\mathbf{a}| = |\mathbf{c}|$ since $OABC$ is a rhombus.

Part (b). Note that

$$\overrightarrow{OM} = \overrightarrow{OA} + \overrightarrow{AM} = \overrightarrow{OA} + k\overrightarrow{AB} = \mathbf{a} + k\mathbf{c}.$$

Also,

$$\overrightarrow{ON} = \overrightarrow{OC} - \overrightarrow{NC} = \overrightarrow{OC} - s\overrightarrow{BC} = \overrightarrow{OC} - s\overrightarrow{AO} = \mathbf{c} + s\mathbf{a}.$$

Thus, we have

$$[\triangle OAM] = \frac{1}{2}\left|\overrightarrow{OA} \times \overrightarrow{OM}\right| = \frac{1}{2}\left|\mathbf{a} \times (\mathbf{a}+k\mathbf{c})\right| = \frac{1}{2}\left|\mathbf{a}\times\mathbf{a} + k\mathbf{a}\times\mathbf{c}\right| = \frac{k}{2}\left|\mathbf{a}\times\mathbf{c}\right|.$$

Similarly,

$$[\triangle OMN] = \frac{1}{2}\left|\overrightarrow{OM} \times \overrightarrow{ON}\right| = \frac{1}{2}\left|(\mathbf{a}+k\mathbf{c}) \times (\mathbf{c}+s\mathbf{a})\right|$$

$$= \frac{1}{2}\left|\mathbf{a}\times\mathbf{c} + s\mathbf{a}\times\mathbf{a} + k\mathbf{c}\times\mathbf{c} + ks\mathbf{c}\times\mathbf{a}\right| = \frac{1}{2}\left|\mathbf{a}\times\mathbf{c} - ks\mathbf{a}\times\mathbf{c}\right| = \left|\frac{1-ks}{2}\right|\left|\mathbf{a}\times\mathbf{c}\right|.$$

Because $k, s \in (0,1)$, it follows that $1-ks \in (0,1)$ too. We can hence remove the modulus sign, whence we obtain

$$\frac{k}{2}\left|\mathbf{a}\times\mathbf{c}\right| = \lambda\left(\frac{1-ks}{2}\left|\mathbf{a}\times\mathbf{c}\right|\right) \implies \lambda = \frac{k}{1-ks}.$$

9758 JC2 Weighted Assessment 2

Problem 1. Futsal is a type of indoor football game. A futsal club has 12 players of which 8 are local players and 4 are foreign players. Each of the 12 players is either a goalkeeper, defender, winger or striker.

Of the 8 local players, 1 is a goalkeeper, 2 are defenders, 2 are wingers and 3 are strikers. Of the 4 foreign players, 1 is a goalkeeper, 1 is a defender and 2 are wingers.

(a) All the players of the club stand in a line for photo taking. Find the number of different possible arrangements if no two foreign players are next to each other.

(b) All the players of the club sit in a circle for a pre-game talk. Find the number of different possible arrangements if the players of the same position are seated together, and a goalkeeper must not sit next to a winger.

(c) A team comprises 1 goalkeeper, 1 defender, 2 wingers are 1 striker. Find the number of ways the futsal club can form a team if it includes exactly 3 foreign players.

Solution.

Part (a). First, arrange the local players in a line. There are $8!$ ways to do so. Next, slot in the foreign players into the 9 available slots. There are ${}^9P_4$ ways to do so. Thus, the required answer is $8! \cdot {}^9P_4 = 121927680$.

Part (b). First, group players by their position. Next, we arrange the non-goalkeepers in a circle. There are $(3-1)!$ ways to do so. Next, we slot in the goalkeeper group. Note that there is only 1 way to do so. Now, note that there are 2 goalkeepers, 3 defenders, 4 wingers and 3 strikers. There are hence $2!\,3!\,4!\,3!$ ways to arrange the players within their groups. Altogether, the required answer is $(3-1)! \cdot 1 \cdot 2!\,3!\,4!\,3! = 3436$.

Part (c). We split our analysis into three cases, depending on which foreign players are selected.

Case 1: *1 goalkeeper, 1 defender, 1 winger.* There are ${}^1C_1 \cdot {}^1C_1 \cdot {}^2C_1$ ways to select the foreign players. There are a further ${}^2C_1 \cdot {}^3C_1$ ways to select the remaining winger and striker from the local players. In this case, we have a total of

$$ {}^1C_1 \cdot {}^1C_1 \cdot {}^2C_1 \cdot {}^2C_1 \cdot {}^3C_1 = 12 $$

ways to choose the players.

Case 2: *1 goalkeeper, 2 wingers.* There are ${}^1C_1 \cdot {}^2C_2$ ways to select the foreign players. There are a further ${}^2C_1 \cdot {}^3C_1$ ways to select the remaining defender and striker from the local players. In this case, we have a total of

$$ {}^1C_1 \cdot {}^2C_2 \cdot {}^2C_1 \cdot {}^3C_1 = 6 $$

ways to choose the players.

Case 3: *1 defender, 2 wingers.* There are ${}^1C_1 \cdot {}^2C_2$ ways to select the foreign players. There are a further ${}^1C_1 \cdot {}^3C_1$ ways to select the remaining goalkeeper and striker from the local players. In this case, we have a total of

$$ {}^1C_1 \cdot {}^2C_2 \cdot {}^1C_1 \cdot {}^3C_1 = 3 $$

ways to choose the players.

Altogether, there are $12 + 6 + 3 = 21$ ways to choose the players.

Problem 2. Do not use a calculator in answering this question.

(a) Find the complex numbers w and z which satisfy the following simultaneous equations.

$$2\mathrm{i}z - w = -7 \quad \text{and} \quad (2-\mathrm{i})\,z - 3\mathrm{i}w^* = 10\mathrm{i}.$$

Give your answers in the form $a + b\mathrm{i}$, where a and b are real numbers.

(b) One of the roots of the equation $3x^3 - 28x^2 + 58x + c = 0$, where c is real, is $5 + \mathrm{i}$.

(i) Find the other roots of the equation and the value of c.

(ii) Deduce the roots of the equation $-3\mathrm{i}w^3 + 28w^2 + 58\mathrm{i}w + c = 0$, where c takes the value obtained in (i).

(iii) Suppose the points P and Q represent the complex numbers $5 + \mathrm{i}$ and $-5 - \mathrm{i}$ respectively. State the geometrical relationship between the points P and Q.

Solution.

Part (a). Let $z = x + \mathrm{i}y$, where $x, y \in \mathbb{R}$. From the first equation, we get

$$w = 2\mathrm{i}z + 7 = 2\mathrm{i}\,(x + \mathrm{i}y) + 7 = (7 - 2y) + \mathrm{i}\,(2x)\,.$$

Substituting this into the second equation, we obtain

$$(2-\mathrm{i})\,(x+\mathrm{i}y) - 3\mathrm{i}\,(7 - 2y - 2\mathrm{i}x) = 10\mathrm{i} \implies (-4x + y) + \mathrm{i}\,(-x + 8y) = 31\mathrm{i}.$$

Comparing real and imaginary parts, we have

$$\begin{cases} -4x + y = 0 \\ {-x} + 8y = 31 \end{cases},$$

whence $x = 1$ and $y = 4$. Thus, $z = 1 + 4\mathrm{i}$ and $w = -1 + 2\mathrm{i}$.

Part (b).

Part (b)(i). Since the coefficients of the polynomial are real, by the conjugate root theorem, $(5 + \mathrm{i})^* = 5 - \mathrm{i}$ is also a root. Let α be the other root. By Vieta's formula,

$$(5+\mathrm{i}) + (5-\mathrm{i}) + \alpha = \frac{28}{3} \implies \alpha = -\frac{2}{3}.$$

Invoking Vieta's formula once more,

$$(5+\mathrm{i})\,(5-i)\left(-\frac{2}{3}\right) = -\frac{c}{3} \implies c = 52.$$

Part (b)(ii). Observe that the given polynomial is equivalent to

$$-3\mathrm{i}w^3 + 28w^2 + 58\mathrm{i}w + c = 3\,(\mathrm{i}w)^3 - 28\,(\mathrm{i}w)^2 + 58\,(\mathrm{i}w) + 52.$$

Thus,

$$\mathrm{i}w = 5+\mathrm{i},\ 5-\mathrm{i},\ -\frac{2}{3} \implies w = 1 - 5\mathrm{i},\ -1-5\mathrm{i},\quad \frac{2}{3}\mathrm{i}.$$

Part (b)(iii). The points are reflections of each other in the origin.

* * * * *

Problem 3. The line l_1 has equation

$$-\frac{x+1}{2} = 8 - y = \frac{3-z}{5}$$

and the line l_2 passes through the points $A(-1,-2,3)$ and $B(-9,4,1)$. The plane p_1 has equation $ax + by = 17$, where a and b are real constants. It is given that l_1 lies on p_1.

(a) Show that $a = -1$ and $b = 2$.

(b) The line l_2 meets the xz-plane at point C. Find the coordinates of C.

(c) Another plane p_2 is parallel to p_1 and contains C. Find, in scalar product form, an equation of the third plane p_3 which is a reflection of p_1 in p_2.

Solution.

Part (a). The Cartesian equation of l_1 can be rewritten as

$$\frac{x+1}{-2} = \frac{y-8}{-1} = \frac{z-3}{-5},$$

so l_1 has vector equation

$$l_1 : \mathbf{r} = \begin{pmatrix} -1 \\ 8 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Further, p_1 has equation

$$p_1 : \mathbf{r} \cdot \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} = 17.$$

Since l_1 lies on p_1, we have

$$\left[\begin{pmatrix} -1 \\ 8 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix} \right] \cdot \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} = 17 \implies (-a + 8b) + \lambda (2a + b) = 17.$$

For this to hold true for all $\lambda \in \mathbb{R}$, we require

$$\begin{cases} -a + 8b = 17 \\ 2a + b = 0 \end{cases},$$

which gives $a = -1$ and $b = 2$.

Part (b). Note that

$$\overrightarrow{AB} = \begin{pmatrix} -9 \\ 4 \\ 1 \end{pmatrix} - \begin{pmatrix} -1 \\ -2 \\ 3 \end{pmatrix} = -2 \begin{pmatrix} 4 \\ -3 \\ 1 \end{pmatrix}.$$

Thus, l_2 has vector equation

$$l_2 : \mathbf{r} = \begin{pmatrix} -1 \\ -2 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ -3 \\ 1 \end{pmatrix}, \quad \mu \in \mathbb{R}.$$

Let $C(x, 0, z)$. Since C lies on the xz-plane, its y-component is 0, so C corresponds to $\mu = -2/3$, which gives $C(-11/3, 0, 7/3)$.

Part (c). Let F and G be a point on p_1 and p_3 respectively such that

$$\overrightarrow{OC} = \frac{\overrightarrow{OF} + \overrightarrow{OG}}{2}.$$

Then

$$\overrightarrow{OG} \cdot \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \left(2\overrightarrow{OC} - \overrightarrow{OF} \right) \cdot \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -11/3 \\ 0 \\ 7/3 \end{pmatrix} - 17 = -\frac{29}{3}.$$

Thus, the equation of p_3 is

$$p_3 : \mathbf{r} \cdot \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = -\frac{29}{3}.$$

9649 H2 Further Mathematics

9649 JC1 Weighted Assessment 1

Problem 1. Show that

$$\sum_{r=1}^{(n-1)^2+3} \left(3^{2r} - n + 1\right) = \frac{a}{b}\left(729 \cdot 9^{(n-1)^2} - 1\right) - c(n-1)^3 - d(n-1)$$

where a, b, c and d are constants to be determined.

Solution.

$$\begin{aligned}\sum_{r=1}^{(n-1)^2+3} \left(3^{2r} - n + 1\right) &= \sum_{r=1}^{(n-1)^2+3} 9^r - \sum_{r=1}^{(n-1)^2+3} (n-1) \\ &= \frac{9\left(9^{(n-1)^2+3} - 1\right)}{9-1} - (n-1)\left[(n-1)^2 + 3\right] \\ &= \frac{9}{8}\left(729 \cdot 9^{(n-1)^2} - 1\right) - (n-1)^3 - 3(n-1).\end{aligned}$$

* * * * *

Problem 2. **Do not use a calculator in answering this question.**

The sequence of positive numbers, u_n, satisfies the recurrence relation:

$$u_{n+1} = \sqrt{2u_n + 3}, \qquad n = 1, 2, 3, \ldots.$$

(a) If the sequence converges to m, find the value of m.

(b) By using a graphical approach, explain why $m < u_{n_1} < u_n$ when $u_n > u_m$. Hence, determine the behaviour of the sequence when $u_1 > m$.

Solution.

Part (a). Observe that

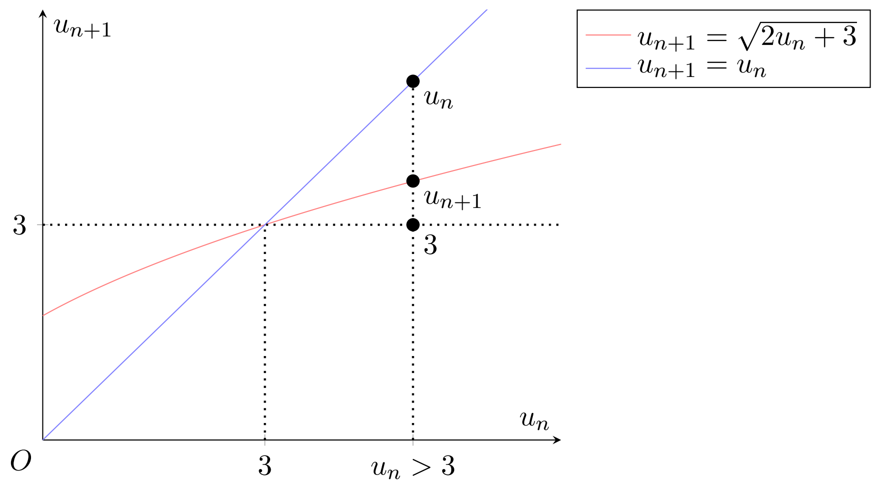
$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} \sqrt{2u_{n-1} + 3} = \sqrt{2 \lim_{n\to\infty} u_{n-1} + 3} = \sqrt{2 \lim_{n\to\infty} u_n + 3}.$$

Since the sequence converges to m, we have $\lim_{n\to\infty} u_n = m$. Thus,

$$m = \sqrt{2m+3} \implies m^2 = 2m + 3 \implies m^2 - 2m - 3 = (m-3)(m+1) = 0.$$

Thus, $m = 3$ or $m = -1$. Since u_n is always positive, we take $m = 3$.

Part (b).



From the graph, if $u_n > 3$, then $3 < u_{n+1} < u_n$. Hence, the sequence decreases and converges to 3.

* * * * *

Problem 3. Two expedition teams are to climb a vertical distance of 8100 m from the foot to the peak of a mountain. Team A plans to cover a vertical distance of 400 m on the first day. On each subsequent day, the vertical distance covered is 5 m less than the vertical distance covered in the previous day. Team B plans to cover a vertical distance of 800 m on the first day. On each subsequent day, the vertical distance covered is 90% of the vertical distance covered in the previous day.

(a) Find the number of days required for Team A to reach the peak.

(b) Explain why Team B will never be able to reach the peak.

(c) At the end of the 15th day, Team B decided to modify their plan, such that on each subsequent day, the vertical distance covered is 95% of the vertical distance covered in the previous day. Which team will be the first to reach the peak of the mountain? Justify your answer.

Solution.

Part (a). The vertical distance Team A plans to cover in a day can be described as a sequence in arithmetic progression with first term 400 and common difference -5. In order to reach the peak, the total vertical distance covered by Team A has to be greater than 8100 m. Hence,

$$\frac{n}{2}\left(2(400) + (n-1)(-5)\right) \geq 8100.$$

Using G.C., $n \geq 24$. Hence, Team A requires 24 days to reach the peak.

Part (b). The vertical distance Team B plans to cover in the nth day can be described by the sequence U_n in geometric progression with first term 800 and common ratio $r = 0.9$. Let S_n^U be the nth partial sum of U_n. Since $|r| < 1$, the sum to infinity of exists and is equal to

$$S_\infty^U = \frac{800}{1 - 0.9} = 8000.$$

Hence, Team B will never be able to surpass 8 km in height. Thus, they will not reach the peak no matter how long they take.

Part (c). The new vertical distance covered by Team B after Day 15 can be described by the sequence V_n in geometric progression with first term U_{15} and common ratio $r = 0.95$. Let S_n^V be the nth partial sum of V_n. Then,

$$S_n^V = \frac{U_{15} \cdot 0.95 \left[1 - (0.95)^n\right]}{1 - 0.95}.$$

Note that

$$S_n^U = \frac{800 \left[1 - (0.9)^n\right]}{1 - 0.9}.$$

Hence, after Day 15, Team B has to climb another $8000 - S_{15}^U = 1747.13$ metres. Since $U_{15} = 183.01$, we have the inequality

$$\frac{183.01 \cdot 0.95 \left[1 - (0.95)^n\right]}{1 - 0.95} \geq 1747.13.$$

Using G.C., $n \geq 14$. Hence, Team B will need at least $15 + 14 = 29$ days to reach the peak. Thus, team A will reach the peak first.

* * * * *

Problem 4. The function f is given by $f(x) = x^2 - 3x + 2 - \mathrm{e}^{-x}$. It is known from graphical work that this equation has 2 roots $x = \alpha$ and $x = \beta$, where $\alpha < \beta$.

(a) Show that $f(x) = 0$ has at least one root in the interval $[0, 1]$.

It is known that there is exactly one root in $[0, 1]$ where $x = \alpha$.

(b) Starting with $x_0 = 0.5$, use an iterative method based on the form

$$x_{n+1} = p\left(x_n^2 + q - \mathrm{e}^{-x_n}\right)$$

where p and q are real numbers to be determined, to find the value of α correct to 3 decimal places. You should demonstrate that your answer has the required accuracy.

It is known that the other root $x = \beta$ lies in the interval $[2, 3]$.

(c) With the aid of a clearly labelled diagram, explain why the method in (b) will fail to obtain any reasonable approximation to β, where x_0 is chosen such that $x_0 \in [2, 3]$, $x_0 \neq \beta$.

To obtain an approximation to β, another approach is used.

(d) Use linear interpolation once in the interval $[2, 3]$ to find a first approximation to β, giving your answer to 2 decimal places. Explain whether this approximate is an overestimate or underestimate.

(e) With your answer in (d) as the initial approximate, use the Newton-Raphson method to obtain β correct to 3 decimal places. Your process should terminate when you have two successive iterates that are equal when rounded to 3 decimal places.

Solution.

Part (a). Observe that $f(0) = 1 > 0$ and $f(1) = -e^{-1} < 0$. Since f is continuous and $f(0)f(1) < 0$, there must be at least one root to $f(x) = 0$ in the interval $[0, 1]$.

Part (b). Let $f(x) = 0$. Then,

$$x^2 - 3x + 2 - \mathrm{e}^{-x} = 0 \implies x^2 + 2 - \mathrm{e}^{-x} = 3x \implies x = \frac{1}{3}\left(x^2 + 2 - \mathrm{e}^{-x}\right).$$

Hence, we should use an iterative method based on the form

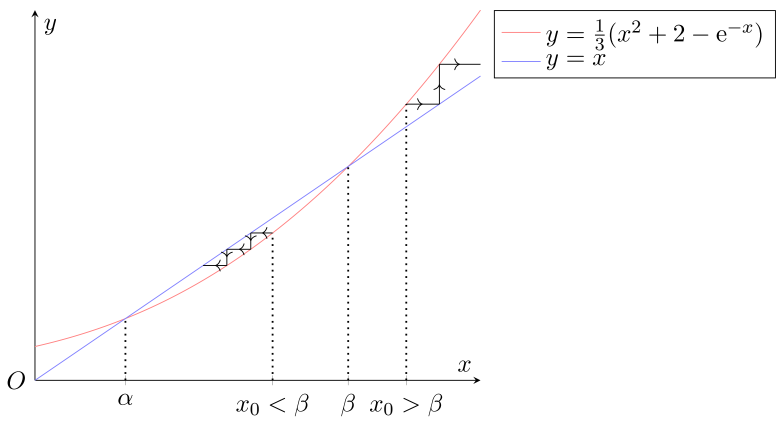
$$x_{n+1} = \frac{1}{3}\left(x_n^2 + 2 - \mathrm{e}^{-x_n}\right).$$

Starting with $x_0 = 0.5$,

n	x_n	n	x_n
0	0.5	6	0.60494
1	0.54782	7	0.60662
2	0.57396	8	0.60759
3	0.58871	9	0.60817
4	0.59718	10	0.60851
5	0.60208	11	0.60870

Since $f(0.6085) = 0.000606 > 0$ and $f(0.6095) = -0.000632 < 0$, we have that $\alpha \in (0.6085, 0.6095)$. Thus, $\alpha = 0.609$ (3 d.p.).

Part (c).



From the diagram, we see that whether we chose $x_0 < \beta$ or $x_0 > \beta$, the approximates move away from the root β. In fact, if we choose $x_0 < \beta$, the approximates converge to the root α instead.

Part (d). Using linear interpolation on the interval $[2, 3]$,

$$x_1 = \frac{2f(3) - 3f(2)}{f(3) - f(2)} = 2.06 \text{ (2 d.p.)}.$$

Thus, $\beta = 2.06$ (2 d.p.).

Observe that $f(2.06) = -0.039 < 0$ and $f(3) = 1.950 > 0$. Hence, $\beta \in (2.06, 3)$. Thus, $\beta = 2.06$ is an underestimate.

Part (e). Observe that $f'(x) = 2xx - 3 + \mathrm{e}^{-x}$. Using the Newton-Raphson method with $x_1 = 2.06$,

n	x_n
1	2.06
2	2.11118
3	2.10935
4	2.10935

Hence, $\beta = 2.109$ (3 d.p.).

9649 JC1 Weighted Assessment 2

Problem 1. Referred to the origin O, points A and B have position vectors $\mathbf{a}$ and $\mathbf{b}$ respectively where $\mathbf{a}$ and $\mathbf{b}$ are non-zero and non-parallel vectors. The point C is such that $\overrightarrow{OC} = m\overrightarrow{OA}$ where m is a constant. The point D lies on AB produced such that B divides AD in the ratio $1:2$.

(a) Express the area of triangle ADC in the form $k\left|\mathbf{a}\times\mathbf{b}\right|$, where k is an expression in terms of m. Show your working clearly.

(b) If $\overrightarrow{AC}$ is a unit vector, give a geometrical interpretation of the value of $\left|\mathbf{b}\times\overrightarrow{AC}\right|$ and find the possible values of m in terms of $|\mathbf{a}|$.

Solution.

Part (a).

$$\overrightarrow{OC} = m\overrightarrow{OA} = m\mathbf{a} \implies \overrightarrow{AC} = \overrightarrow{OC} - \overrightarrow{OA} = m\mathbf{a} - \mathbf{a} = (m-1)\mathbf{a}.$$

By the Ratio Theorem,

$$\overrightarrow{OB} = \frac{\overrightarrow{OD} + 2\cdot\overrightarrow{OA}}{1+2}.$$

Hence,

$$\overrightarrow{OD} = 3\overrightarrow{OB} - 2\overrightarrow{OA} = 3\mathbf{b} - 2\mathbf{a} \implies \overrightarrow{AD} = \overrightarrow{OD} - \overrightarrow{OA} = 3\mathbf{b} - 3\mathbf{a}.$$

Thus,

$$\begin{aligned}\text{Area}\,\triangle ADC &= \frac{1}{2}\left|\overrightarrow{AC}\times\overrightarrow{AD}\right| = \frac{1}{2}\left|(m-1)\mathbf{a}\times(3\mathbf{b}-3\mathbf{a})\right| = \frac{3}{2}\left|m-1\right|\left|\mathbf{a}\times(\mathbf{b}-\mathbf{a})\right| \\ &= \frac{3}{2}\left|m-1\right|\left|\mathbf{a}\times\mathbf{b} - \mathbf{a}\times\mathbf{a}\right| = \frac{3}{2}\left|m-1\right|\left|\mathbf{a}\times\mathbf{b}\right|\end{aligned}$$

whence $k = \frac{3}{2}\left|m-1\right|$.

Part (b). Since $\overrightarrow{AC}$ is parallel to $\mathbf{a}$, if $\overrightarrow{AC}$ is a unit vector, then $\overrightarrow{AC} = \hat{\mathbf{a}}$. Hence, $\left|\mathbf{b}\times\overrightarrow{AC}\right| = \left|\mathbf{b}\times\hat{\mathbf{a}}\right|$ is the shortest distance from B to the line OA.

Since $\overrightarrow{AC}$ is a unit vector, we have

$$\left|\overrightarrow{AC}\right| = \left|(m-1)\mathbf{a}\right| = 1 \implies \left|m-1\right| = \frac{1}{|\mathbf{a}|} \implies m = 1 \pm \frac{1}{|\mathbf{a}|}.$$

* * * * *

Problem 2. Marine biologist experts calculated that when the concentration of chemical X in a sea inlet reaches 6 milligrams per litre (mg/l), the level of pollution endangers marine life. A factory wishes to release waste containing chemical X into the inlet. It claimed that the discharge will not endanger the marine life, and they provided the local authority with the following information:

- There is no presence of chemical X in the sea inlet at present.

- The plain is to discharge chemical X on a weekly basis into the sea inlet. The discharge, which will be done at the beginning of each week, will result in an increase in concentration of 2.3 mg/l of chemical X in the inlet.
- The tidal streams will remove 7% of chemical X from the inlet at the end of every day.

(a) Form a recurrence relation for the concentration level of chemical X, u_n, at the beginning of week n. Hence, find the concentration at the beginning of week n.

(b) Should the local authority allow the factory to go ahead with the discharge if they are concerned with the marine life at the sea inlet? Justify your answer.

Solution.

Part (a). We have

$$u_n = 0.93^7 u_{n-1} + 2.3, \quad u_0 = 0.$$

Let k be the constant such that $u_n + k = 0.93^7(u_{n-1} + k)$. Then $k = \frac{2.3}{0.93^7 - 1}$. Hence,

$$u_n - \frac{2.3}{1-0.93^7} = 0.93^7\left(u_{n-1} - \frac{2.3}{1-0.93^7}\right) = 0.93^{7n}\left(u_0 - \frac{2.3}{1-0.93^7}\right) = -\frac{2.3 \cdot 0.93^{7n}}{1-0.93^7}$$

$$\implies u_n = \frac{2.3}{1-0.93^7} - \frac{2.3 \cdot 0.93^{7n}}{1-0.93^7}.$$

Part (b).

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty}\left(\frac{2.3}{1-0.93^7} - \frac{2.3 \cdot 0.93^{7n}}{1-0.93^7}\right) = \frac{2.3}{1-0.93^7} = 5.77 \text{ (3 s.f.)}.$$

Since $5.77 < 6$, if the local authority's only concern is marine life, they should allow the factory to go ahead with the discharge.

* * * * *

Problem 3. Referred to the origin O, the position vector of the point A is $3\mathbf{i} - 2\mathbf{j} - 6\mathbf{k}$ and the Cartesian equation of the line l_1 is $x - 1 = 2 - y = 2z + 6$.

(a) Find the position vector of the foot of perpendicular from A to l_1.

Line l_2 has the vector equation $\mathbf{r} = (-1,\, 6,\, -1)^\mathsf{T} + \mu\,(-6,\, 6,\, -3)^\mathsf{T}$, where $\mu \in \mathbb{R}$.

(b) Find the shortest distance between l_1 and l_2.

(c) Given that l_2 is the reflection of l_1 about the line l_3, find the vector equation of the line l_3.

Solution.

Part (a). Note that l_1 has vector equation

$$l_1 : \mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix},\ \lambda \in \mathbb{R}$$

Let F be the foot of perpendicular from A to l_1. Since F is on l_1, $\overrightarrow{OF} = (1,\, 2,\, -3)^\mathsf{T} + \lambda\,(2,\, -2,\, 1)^\mathsf{T}$ for some $\lambda \in \mathbb{R}$. Thus,

$$\overrightarrow{AF} = \overrightarrow{OF} - \overrightarrow{OA} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} - \begin{pmatrix} 3 \\ -2 \\ -6 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}.$$

Note also that $\overrightarrow{AF}$ is perpendicular to l_1. Hence,

$$\left[\begin{pmatrix}-2\\4\\3\end{pmatrix}+\lambda\begin{pmatrix}2\\-2\\1\end{pmatrix}\right]\cdot\begin{pmatrix}2\\-2\\1\end{pmatrix}=0 \implies -9+9\lambda=0 \implies \lambda=1.$$

Thus,

$$\overrightarrow{OF}=\begin{pmatrix}1\\2\\-3\end{pmatrix}+\begin{pmatrix}2\\-2\\1\end{pmatrix}=\begin{pmatrix}3\\0\\-2\end{pmatrix}.$$

Part (b). Note that $(-6, 6, -3)^\mathsf{T} \parallel (2, -2, 1)^\mathsf{T}$. Hence, l_2 is parallel to l_1. Hence, the shortest distance between l_1 and l_2 is the perpendicular distance from a point on l_1 to l_2, which is

$$\frac{\left|\left[(1, 2, -3)^\mathsf{T}-(-1, 6, -1)^\mathsf{T}\right]\times(2, -2, 1)^\mathsf{T}\right|}{\left|(2, -2, 1)^\mathsf{T}\right|}=\frac{1}{\sqrt{9}}\left|\begin{pmatrix}2\\-4\\-2\end{pmatrix}\times\begin{pmatrix}2\\-2\\1\end{pmatrix}\right|$$
$$=\frac{2}{3}\left|\begin{pmatrix}-1\\2\\1\end{pmatrix}\times\begin{pmatrix}2\\-2\\1\end{pmatrix}\right|=\frac{2}{3}\left|\begin{pmatrix}4\\3\\-2\end{pmatrix}\right|=\frac{2}{3}\sqrt{29}\text{ units.}$$

Part (c). Observe that l_3 passes through the midpoint of $(1, 2, -3)^\mathsf{T}$ and $(-1, 6, -1)^\mathsf{T}$, which evaluates to

$$\frac{1}{2}\left[\begin{pmatrix}1\\2\\-3\end{pmatrix}+\begin{pmatrix}-1\\6\\-1\end{pmatrix}\right]=\begin{pmatrix}0\\4\\-2\end{pmatrix}.$$

l_3 is also parallel to both l_1 and l_2. Hence,

$$l_3:\mathbf{r}=\begin{pmatrix}0\\4\\-2\end{pmatrix}+\nu\begin{pmatrix}2\\-2\\1\end{pmatrix},\ \nu\in\mathbb{R}.$$

* * * * *

Problem 4. A first order recurrence relation is given as

$$u_{n+1}\left[u_n+\left(\frac{1}{2}\right)^n\right]+u_n\left[\left(\frac{1}{2}\right)^{n+1}-10\right]=10\left(\frac{1}{2}\right)^n-\left(\frac{1}{2}\right)^{2n+1}-16$$

where $u_1 = 1$.

(a) Using the substitution $u_n = \frac{v_n}{v_{n-1}} - \left(\frac{1}{2}\right)^n$ where $v_{n-1} \neq 0$, show that the recurrence relation can be expressed as a second order recurrence relation of the form $v_{n+1} + av_n + 16v_{n-1} = 0$, where a is a constant to be found.

(b) By first solving the second order recurrence relation in (a), find an expression for u_n in terms of n.

(c) Describe what happens to the value of u_n for large values of n.

Solution.

Part (a). Substituting in v_n for u_n into the LHS of the recurrence relation, we get

$$\begin{aligned} & u_{n+1}\left[u_n+\left(\frac{1}{2}\right)^n\right]+u_n\left[\left(\frac{1}{2}\right)^{n+1}-10\right] \\ &= \left[\frac{v_{n+1}}{v_n}-\left(\frac{1}{2}\right)^{n+1}\right]\left[\frac{v_n}{v_{n-1}}\right]+\left[\frac{v_n}{v_{n-1}}-\left(\frac{1}{2}\right)^n\right]\left[\left(\frac{1}{2}\right)^{n+1}-10\right] \\ &= \frac{v_{n+1}}{v_{n-1}}-10\left(\frac{v_n}{v_{n-1}}\right)-\left(\frac{1}{2}\right)^{2n+1}+10\left(\frac{1}{2}\right)^n. \end{aligned}$$

Cancelling terms from the RHS, we get

$$\frac{v_{n+1}}{v_{n-1}}-10\left(\frac{v_n}{v_{n-1}}\right)=-16 \implies v_{n+1}-10v_n+16v_{n-1}=0.$$

Hence, $a=-10$.

Part (b). Consider the characteristic equation of v_n.

$$x^2-10x+16=(x-2)(x-8)=0.$$

Hence, 2 and 8 are the roots of the characteristic equation. Thus,

$$v_n=A\cdot 2^n+B\cdot 8^n.$$

Consider u_1.

$$u_1=\frac{v_1}{v_0}-\frac{1}{2}=1 \implies \frac{2A+8B}{A+B}=\frac{3}{2} \implies \frac{4A+16B}{A+B}=3 \implies A=-13B.$$

Now observe that

$$\begin{aligned} \frac{v_n}{v_{n-1}} &= \frac{A\cdot 2^n+B\cdot 8^n}{A\cdot 2^{n-1}+B\cdot 8^{n-1}}=\frac{-13\cdot 2^n+8^n}{-13\cdot 2^{n-1}+8^{n-1}}=8\left(\frac{-13\cdot 2^n+8^n}{-52\cdot 2^n+8^n}\right) \\ &= 8\left(1+\frac{39\cdot 2^n}{-52\cdot 2^n+8^n}\right)=8-\frac{312\cdot 2^n}{52\cdot 2^n-8^n}. \end{aligned}$$

Thus,

$$u_n=\frac{v_n}{v_{n-1}}-\left(\frac{1}{2}\right)^n=8-\frac{312\cdot 2^n}{52\cdot 2^n-8^n}-\left(\frac{1}{2}\right)^n.$$

Part (c).

$$\lim_{n\to\infty} u_n=\lim_{n\to\infty}\left[8-\frac{312\cdot 2^n}{52\cdot 2^n-8^n}-\left(\frac{1}{2}\right)^n\right]=8.$$

Thus, u_n converges to 8 for large values of n.

9649 JC1 Promotional Examination

Problem 1. Given that $z=f(x,y)$ is a differentiable function and $f(x,y)=k$ is the level curve of f that passes through the point P, show that the gradient vector ∇f is perpendicular to the tangent of the level curve at P.

Solution. Let x and y be parametrized t. Implicitly differentiating $f(x, y) = k$ with respect to t, we obtain

$$\frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = 0 \implies \begin{pmatrix}\partial f/\partial x\\ \partial f/\partial y\end{pmatrix} \cdot \begin{pmatrix}\mathrm{d}x/\mathrm{d}t\\ \mathrm{d}y/\mathrm{d}t\end{pmatrix} = 0.$$

Observe that $(\partial f/\partial x,\ \partial f/\partial y)^{\mathsf{T}}$ is exactly ∇f. Taking $t = x$, we also have $(\mathrm{d}x/\mathrm{d}t,\ \mathrm{d}y/\mathrm{d}t)^{\mathsf{T}} = (1,\ \mathrm{d}y/\mathrm{d}x)^{\mathsf{T}}$, which is the tangent vector of the level curve at P. Since the dot product of the two vectors is 0, they must be perpendicular.

* * * * *

Problem 2. A harvesting model is given by $\frac{\mathrm{d}P}{\mathrm{d}t} = (20 - P)\left(P^2 - 30P + h\right)$ where P is the population of the resource at time t and h is the constant harvest rate. It is given that at $t = 0$, $P = 50$. Find the range of values of h such that $P = 20$ in the long run.

Solution. Let $f(P) = (20 - P)\left(P^2 - 30P + h\right)$. For $P = 20$ in the long run, we need $f(P) < 0$ for all $P \in (20, 50]$. Observe that $20 - P < 0$ for all $P \in (20, 50]$. We hence need $P^2 - 30P + h > 0$ for all $P \in (20, 50]$. However, observe that $P^2 - 30P + h$ is strictly increasing after $P > 15$. Thus, we only require the quadratic to be non-negative at $P = 20$, whence $20^2 - 30(20) + h \geq 0$, giving $h \geq 200$.

* * * * *

Problem 3.

(a) Describe the locus given by $|2\mathrm{i} + 1 + \mathrm{i}z| = |4\mathrm{i} - 5 - z|$.

S is the set of complex numbers z for which

$$|2\mathrm{i} + 1 + \mathrm{i}z| \geq |4\mathrm{i} - 5 - z| \text{ and } |z + 2 - 3\mathrm{i}| \leq \sqrt{13}.$$

(b) On a single Argand diagram, shade the region corresponding to S.

(c) Find the set of values of $\arg(z - 8\mathrm{i})$.

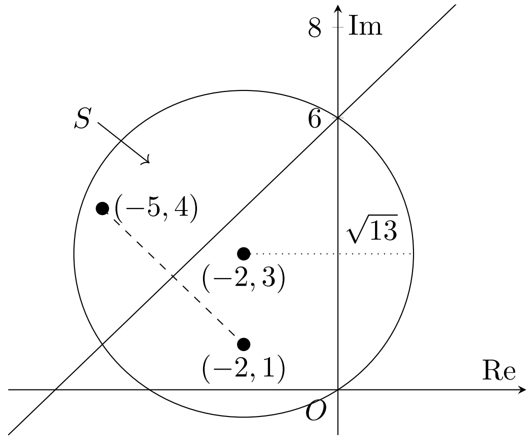
Solution.

Part (a). Note that $|2\mathrm{i} + 1 + \mathrm{i}z| = |2 - \mathrm{i} + z| = |z - (-2 + \mathrm{i})|$. We hence have

$$|z - (-2 + \mathrm{i})| = |z - (-5 + 4\mathrm{i})|,$$

which describes the perpendicular bisector of the line segment joining $(-2, 1)$ and $(-5, 4)$.

Part (b).



Part (c). Clearly, $\min \arg(z - 8\mathrm{i}) = -\frac{\pi}{2}$ (where $z = 6\mathrm{i}$). Let $A(-2, 3)$ and $B(0, 8)$. Let C be the point on the circle such that BC is tangent to the circle. We have $\angle ABO = \arctan \frac{2}{5}$. Since $AB = \sqrt{5^2 + 2^2} = \sqrt{29}$, we also have $\angle CBA = \arcsin \frac{\sqrt{13}}{\sqrt{29}}$. Thus,

$$\max \arg(z - 8\mathrm{i}) = -\frac{\pi}{2} - \arctan \frac{2}{5} - \arcsin \frac{\sqrt{13}}{\sqrt{29}} = -2.68,$$

whence

$$-2.68 \leq \arg(z - 8\mathrm{i}) \leq -\frac{\pi}{2}.$$

* * * * *

Problem 4. A tuition agency is designing a revision programme to help students to prepare for their A-level examinations. There are three types of revisions that the agency can run each day: Basic, Intermediate and Challenging. Due to resource constraints, the Challenging revision cannot be run consecutively. The revision programme consists of n days. Let a_n be the number of possible programmes for the duration.

(a) Explain why $a_n = 2(a_{n-1} + a_{n-2})$.

(b) Solve the recurrence relation and find a_n in terms of n.

(c) For a 10 days revision programme, given that both the 1st and 10th days consist of the Challenging revision, find the number of possible programmes.

Solution.

Part (a). Suppose the nth day ran Basic or Intermediate. The agency was hence free to run any programme on the $n-1$th day, thus contributing $2 \cdot a_{n-1}$ total programmes towards a_n.

Now suppose the nth day ran Challenging. The agency could have only ran Basic or Intermediate on the $n-1$th day. This hence contributes $1 \cdot 2 \cdot a_{n-2}$ total programmes towards a_n.

Hence, $a_n = 2a_{n-1} + 2a_{n-2} = 2(a_{n-1} + a_{n-2})$.

Part (b). Consider the characteristic polynomial of the second-order recurrence relation:

$$x^2 - 2x - 2 = 0 \implies x = 1 \pm \sqrt{3}.$$

Hence,

$$a_n = A\left(1+\sqrt{3}\right)^n + B\left(1-\sqrt{3}\right)^n$$

for some constants A and B.

Observe that $a_0 = 1$ (since there is only one way to do nothing). This gives $A+B=1$.
Also observe that $a_1 = 3$. Hence, $(A+B)+\sqrt{3}(A-B)=3$.
Solving, we get $A = \frac{1}{2}+\frac{1}{\sqrt{3}}$ and $B = \frac{1}{2}-\frac{1}{\sqrt{3}}$. Thus,

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

Part (c). Since the first day ran Challenging, there are only two options for the second day (Basic and Intermediate). Likewise, since the tenth day ran Challenging, there are only two options for the ninth day. The remaining six days are free. This gives a total of $2 \cdot 2 \cdot a_6 = 1792$ possibilities.

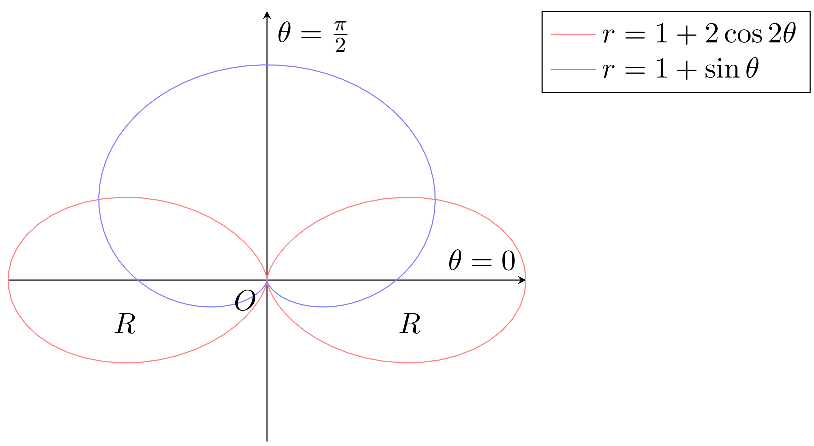
* * * * *

Problem 5. The curve C_1 has equation

$$\left(x^2+y^2\right)^{3/2} = 2x^2.$$

(a) Show that the polar equation of C_1 is $r = 1 + 2\cos 2\theta$.

The curve C_2 has polar equation $r = 1 + \sin\theta$.
The diagram below shows the region R enclosed by C_1 and C_2.



(b) Find the exact area of R.

(c) Find the perimeter of R.

Solution.

Part (a).

$$\left(x^2+y^2\right)^{3/2} = 2x^2 \implies \left(r^2\right)^{3/2} = 2(r\cos\theta)^2 \implies r = 2\cos^2\theta = 1 + 2\cos 2\theta.$$

Part (b). Observe that $t = -\frac{\pi}{2}$ is tangent to the pole to both C_1 and C_2. Now consider the intersections of C_1 and C_2.

$$2\cos^2\theta = 2 - 2\sin^2\theta = 1 + \sin\theta \implies 2\sin^2\theta + \sin\theta - 1 = (2\sin\theta - 1)(\sin\theta + 1) = 0.$$

We hence have $\sin\theta = \frac{1}{2}$ or $\sin\theta = -1$, whence $\theta = \frac{1}{6}\pi, \frac{5}{6}\pi, -\frac{1}{2}\pi$.

$$\begin{aligned}
\text{Area } R &= 2\left(\frac{1}{2}\int_{-\pi/2}^{\pi/6}\left[(1+\cos 2\theta)^2-(1+\sin\theta)^2\right]\,\mathrm{d}\theta\right)\\
&= \int_{-\pi/2}^{\pi/6}\left(1+2\cos 2\theta+\cos^2 2\theta-1-2\sin\theta-\sin^2\theta\right)\,\mathrm{d}\theta\\
&= \int_{-\pi/2}^{\pi/6}\left(2\cos 2\theta+\frac{1+\cos 4\theta}{2}-2\sin\theta-\frac{1-\cos 2\theta}{2}\right)\,\mathrm{d}\theta\\
&= \int_{-\pi/2}^{\pi/6}\left(\frac{5}{2}\cos 2\theta+\frac{1}{2}\cos 4\theta-2\sin\theta\right)\,\mathrm{d}\theta\\
&= \left[\frac{5}{4}\sin 2\theta+\frac{1}{8}\sin 4\theta+2\cos\theta\right]_{-\pi/2}^{\pi/6} = \frac{27}{16}\sqrt{3}\text{ units}^2.
\end{aligned}$$

Part (c). Observe that for C_1, $\frac{\mathrm{d}r}{\mathrm{d}\theta}=-2\sin 2\theta$, while for C_2, $\frac{\mathrm{d}r}{\mathrm{d}\theta}=\cos\theta$. Hence, the perimeter of R is

$$2\int_{-\pi/2}^{\pi/6}\left(\sqrt{(1+\cos 2\theta)^2+(-2\sin 2\theta)^2}+\sqrt{(1+\sin\theta)^2+\cos^2\theta}\right)\,\mathrm{d}\theta = 11.7\text{ units.}$$

* * * * *

Problem 6. A civil engineer is designing a decorative water feature for a garden. The profile of the water feature is modelled by the curve $y=\sin\left(x^2\right)$, for $-\sqrt{\pi}\le x\le\sqrt{\pi}$. The region R is bounded by this curve and the x-axis.

(a) To create the actual water feature, the region R is rotated by π radians about the y-axis forming a symmetrical, bowl-shaped structure. Find the exact volume of the water feature.

(b) Water is poured into the bowl of the water feature at a rate of 2 units^3 per second. Given that the bowl is initially empty, find the rate of change of the depth of the water when the depth is at $\frac{\sqrt{3}}{2}$ units.

(c) The engineer also needs to know the total length of the curved surface of the profile of the water feature. Estimate, to 4 decimal places, the arc length of the profile curve from $x=0$ to $x=1.25$ using Simpson's rule with 4 strips.

Solution.

Part (a).

$$\text{Volume} = 2\pi\int_0^{\sqrt{\pi}} x\sin\left(x^2\right)\,\mathrm{d}x = \pi\left[-\cos\left(x^2\right)\right]_0^{\sqrt{\pi}} = 2\pi\text{ units}^3.$$

Part (b). Let the depth of the water be h. We have $V=\pi\int_0^h \arcsin(y)\,\mathrm{d}y$, whence $\frac{\mathrm{d}V}{\mathrm{d}h}=\pi\arcsin(h)$. Hence,

$$\frac{\mathrm{d}h}{\mathrm{d}t}=\frac{\mathrm{d}V/\mathrm{d}t}{\mathrm{d}V/\mathrm{d}h}=\frac{2}{\pi\arcsin(h)}.$$

When $h=\frac{\sqrt{3}}{2}$, we have $\frac{\mathrm{d}h}{\mathrm{d}t}=\frac{2}{\pi\cdot\pi/3}$. Thus, the rate of change of the depth of water is $\frac{6}{\pi^2}$ units/s.

Part (c). Note that $\frac{\mathrm{d}y}{\mathrm{d}x} = 2x\cos\left(x^2\right)$. Hence, the arc length is given by

$$\int_0^{1.25} \sqrt{1 + (2x\cos(x^2))^2}\,\mathrm{d}x.$$

Let $f(x) = \sqrt{1 + (2x\cos(x^2))^2}$. By Simpson's rule, the arc length is approximately

$$\frac{1.25 - 0}{3 \cdot 4}\left[f(0) + f\left(\frac{5}{16}\right) + 2f\left(\frac{10}{16}\right) + 4f\left(\frac{15}{16}\right) + f(1.25)\right] = 1.6671 \text{ units (4 d.p.)}.$$

* * * * *

Problem 7. It is given that $y = f(x)$ satisfies the following differential equation:

$$\left(x^3 + 1\right) y \frac{\mathrm{d}y}{\mathrm{d}x} + 3x^2y^2 = 2, \quad \text{where } y = 2 \text{ when } x = 0.$$

(a) By using the substitution $z = y^2$, solve the differential equation and find the value of y when $x = 1$.

(b) A point on the graph, initially at $x = 1$, varies such that x is increasing at a rate of $\mathrm{e}^{\sqrt{t}}$ units/s, where t represents time in seconds. Show that $\frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{19}{12}\mathrm{e}^{\sqrt{t}}$ at that instance and use the Euler's method with step length 0.2 to find an approximation of the value of y when $t = 1$.

(c) Explain whether the approximation in part (b) is an underestimation or an overestimation of the true value.

Solution.

Part (a). Note that $z' = 2y \cdot y'$. Hence,

$$\left(x^3 + 1\right) y \frac{\mathrm{d}y}{\mathrm{d}x} + 3x^2y^2 = \left(x^3 + 1\right) \frac{1}{2} z' + 3x^2 z = 2 \implies z' + \frac{6x^2}{x^3 + 1} z = \frac{4}{x^3 + 1}.$$

The integrating factor is hence

$$\text{I. F.} = \exp \int \frac{6x^2}{x^3 + 1}\,\mathrm{d}x = \exp\left(2\ln\left(x^3 + 1\right)\right) = \left(x^3 + 1\right)^2.$$

Multiplying through, we get

$$\begin{aligned}
&\left(x^3 + 1\right)^2 z' + 6x^2\left(x^3 + 1\right) z = \frac{\mathrm{d}}{\mathrm{d}x}\left[\left(x^3 + 1\right)^2 z\right] = 4\left(x^3 + 1\right) \\
\implies &\left(x^3 + 1\right)^2 z = \int 4\left(x^3 + 1\right)\,\mathrm{d}x = x^4 + 4x + C \implies y^2 = \frac{x^4 + 4x + C}{\left(x^3 + 1\right)^2}.
\end{aligned}$$

When $x = 0$, $y = 2$, giving $C = 4$. Hence,

$$y = \sqrt{\frac{x^4 + 4x + 4}{\left(x^3 + 1\right)^2}}.$$

When $x = 1$, $y = \frac{3}{2}$.

Part (b).

$$\left(x^3+1\right)y\frac{\mathrm{d}y}{\mathrm{d}x}+3x^2y^2=\left(1^3+1\right)\left(\frac{3}{2}\right)\frac{\mathrm{d}y}{\mathrm{d}x}+3\left(1^2\right)\left(\frac{3}{2}\right)^2=2$$
$$\implies \frac{\mathrm{d}y}{\mathrm{d}x}=-\frac{19}{12} \implies \frac{\mathrm{d}y}{\mathrm{d}t}=-\frac{19}{12}\cdot\frac{\mathrm{d}x}{\mathrm{d}t}=-\frac{19}{12}\mathrm{e}^{\sqrt{t}}.$$

Let $f(y,t)=-\frac{19}{12}\mathrm{e}^{\sqrt{t}}$, $t_0=0$, $y_0=1.5$. Using Euler's method with $h=0.2$,

$$y_{n+1}=y_n+0.2f(t_n,y_n)$$

Using G.C.,

$$y_1=1.1833,\quad y_2=0.68808,\quad y_3=0.09204,\quad y_4=-0.59503,\quad y_5=-1.36958.$$

Hence, when $t=1$, $y\approx-1.37$.

Part (c). Observe that $\sqrt{t}>0$ and $\mathrm{e}^{\sqrt{t}}>0$ for all $t\in\mathbb{R}$. Thus,

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2}=-\frac{19}{12}\frac{1}{2\sqrt{t}}\mathrm{e}^{\sqrt{t}}<0,$$

whence y is concave downwards, making the approximation an overestimate.

* * * * *

Problem 8.

(a) Use De Moivre's theorem to find a polynomial expression for $\cos 5\theta$ in terms of u, where $u=\cos\theta$.

(b) Write down the five values of θ, $0\le\theta\le\pi$, for which $\cos 5\theta=0$.

(c) Find in trigonometric form, the roots of the equation $16z^4-20z^2+5=0$.

(d) Express the roots found in part (c) in exact surd form. Hence, find the value of

$$\sin^2\frac{\pi}{10}+\sin^2\frac{3\pi}{10}+\sin^2\frac{7\pi}{10}+\sin^2\frac{9\pi}{10}.$$

Solution.

Part (a).

$$\begin{aligned}\cos 5\theta &= \operatorname{Re}\left(\cos\theta+\mathrm{i}\sin\theta\right)^5\\ &= \cos^5\theta-10\cos^3\theta\sin^2\theta+5\cos\theta\sin^4\theta\\ &= \cos^5\theta-10\cos^3\theta\left(1-\cos^2\theta\right)+5\cos\theta\left(1-\cos^2\theta\right)^2\\ &= \cos^5\theta-10\cos^3\theta\left(1-\cos^2\theta\right)+5\cos\theta\left(1-2\cos^2\theta+\cos^4\theta\right)\\ &= 16\cos^5\theta-20\cos^3\theta+5\cos\theta=16u^5-20u^3+5u.\end{aligned}$$

Part (b).

$$\theta=\frac{\pi}{10},\frac{3\pi}{10},\frac{5\pi}{10},\frac{7\pi}{10},\frac{9\pi}{10}.$$

Part (c).

$$16z^4-20z^2+5=0 \implies 16z^5-20z^3+5z=0,\quad z\neq 0.$$

Let $z=\cos\theta$. Then $\cos 5\theta=0$, whence $\theta=\frac{\pi}{10},\frac{3\pi}{10},\frac{5\pi}{10},\frac{7\pi}{10},\frac{9\pi}{10}$. Hence,

$$z=\cos\frac{\pi}{10},\cos\frac{3\pi}{10},\cos\frac{7\pi}{10},\cos\frac{9\pi}{10}.$$

Note that we reject $z=\cos\frac{5\pi}{10}$ since $z\neq 0$.

Part (d). By the quadratic formula,

$$z^2 = \frac{5 \pm \sqrt{5}}{8} \implies z = \pm\sqrt{\frac{5 \pm \sqrt{5}}{8}}.$$

The corresponding trigonometric forms are

$$\cos\frac{\pi}{10} = \sqrt{\frac{5+\sqrt{5}}{8}}, \quad \cos\frac{3\pi}{10} = \sqrt{\frac{5-\sqrt{5}}{8}}$$
$$\cos\frac{7\pi}{10} = -\sqrt{\frac{5-\sqrt{5}}{8}}, \quad \cos\frac{9\pi}{10} = -\sqrt{\frac{5+\sqrt{5}}{8}}.$$

Hence,

$$\begin{aligned}\sin^2\frac{\pi}{10} + \sin^2\frac{3\pi}{10} + \sin^2\frac{7\pi}{10} + \sin^2\frac{9\pi}{10} &= 4 - \left(\cos^2\frac{\pi}{10} + \cos^2\frac{3\pi}{10} + \cos^2\frac{7\pi}{10} + \cos^2\frac{9\pi}{10}\right) \\ &= 4 - \left(\frac{5+\sqrt{5}}{8} + \frac{5-\sqrt{5}}{8} + \frac{5-\sqrt{5}}{8} + \frac{5+\sqrt{5}}{8}\right) = 4 - \frac{20}{8} = \frac{3}{2}.\end{aligned}$$

* * * * *

Problem 9. Suppose the complex number w is a root of the equation $z^9 - 1 = 0$.

(a) (i) Express all the roots of this equation in the form w^n, $n \in \mathbb{Z}$, $0 \leq n \leq 8$, where w is a complex number to be determined.

(ii) Show that $\sum_{r=0}^{8} w^r = 0$.

(iii) Show that $w^2 + w^7 = 2\cos\frac{4\pi}{9}$.

(iv) Using the results above, deduce that $16\cos\frac{2\pi}{9}\cos\frac{4\pi}{9}\cos\frac{6\pi}{9}\cos\frac{8\pi}{9} = 1$.

Let point O be the origin, and points A, B and C represent the complex numbers w^2, $2\mathrm{i}w^2$ and $\frac{w}{u^*}$ respectively, where $u = \frac{1}{3}\left(\cos\frac{5\pi}{18} - \mathrm{i}\sin(5\pi)\,18\right)$.

(b) (i) Find the modulus and arguments of the complex numbers $\frac{w}{u^*}$ and $2iw^2$, and illustrate the points A, B and C on a clearly labelled Argand diagram.

(ii) Find the area of triangle ABC.

Solution.

Part (a).

Part (a)(i).

$$z^9 - 1 = 0 \implies z^9 = \mathrm{e}^{2\pi n\mathrm{i}} \implies z = \mathrm{e}^{2\pi n\mathrm{i}/9}.$$

Hence, the roots are w^n, where $w = \mathrm{e}^{2\pi\mathrm{i}/9}$.

Part (a)(ii).

$$\sum_{r=0}^{8} w^r = \frac{w^9 - 1}{w - 1} = \frac{1-1}{w-1} = 0$$

Part (a)(iii).

$$w^2 + w^7 = w^2 + w^{-2} = 2\cos\left(2 \cdot \frac{2\pi}{9}\right) = 2\cos\frac{4\pi}{9}$$

Part (a)(iv).

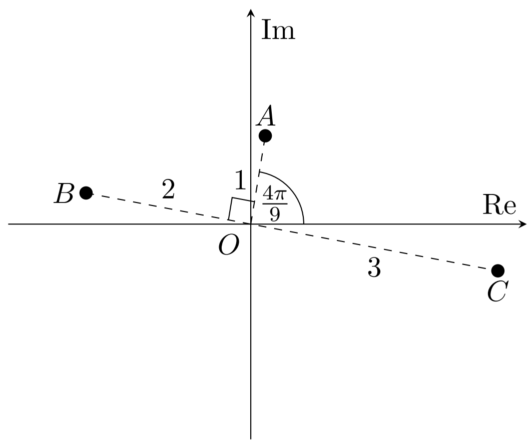
$$\begin{aligned}
&16\cos\frac{2\pi}{9}\cos\frac{4\pi}{9}\cos\frac{6\pi}{9}\cos\frac{8\pi}{9}\\
&=\left(w+w^8\right)\left(w^2+w^7\right)\left(w^3+w^6\right)\left(w^4+w^5\right)\\
&=\left(w^5+w^6+w^{12}+w^{13}\right)\left(w^5+w^8+w^{10}+w^{13}\right)\\
&=\left(w^{10}+w^{13}+w^{15}+w^{18}\right)+\left(w^{11}+w^{14}+w^{16}+w^{19}\right)\\
&\quad+\left(w^{17}+w^{20}+w^{22}+w^{25}\right)+\left(w^{18}+w^{21}+w^{23}+w^{26}\right)\\
&=\left(w+w^4+w^6+1\right)+\left(w^2+w^5+w^7+w\right)\\
&\quad+\left(w^8+w^2+w^4+w^7\right)+\left(1+w^3+w^5+w^8\right)\\
&=2\left(1+w+w^2+w^3+w^4+w^5+w^6+w^7+w^8\right)-\left(w^3+w^6\right)\\
&=-2\cos\frac{6\pi}{9}=1.
\end{aligned}$$

Part (b).

Part (b)(i). Note that $w^2=e^{4\pi i/9}$. Hence, $\left|w^2\right|=1$ and $\arg w^2=\frac{4\pi}{9}$. Likewise, $\left|2iw^2\right|=2$ and $\arg\left(2iw^2\right)=\frac{4\pi}{9}+\frac{\pi}{2}=\frac{17}{18}\pi$. Lastly, note that

$$\frac{w}{u^*}=\frac{uw}{uu^*}=\frac{1/3\cdot e^{-5\pi i/18}\cdot e^{2\pi i/9}}{1/9}=3e^{-i\pi/18},$$

whence $\left|\frac{w}{u^*}\right|=3$ and $\arg\left(\frac{w}{u^*}\right)=-\frac{\pi}{18}$.

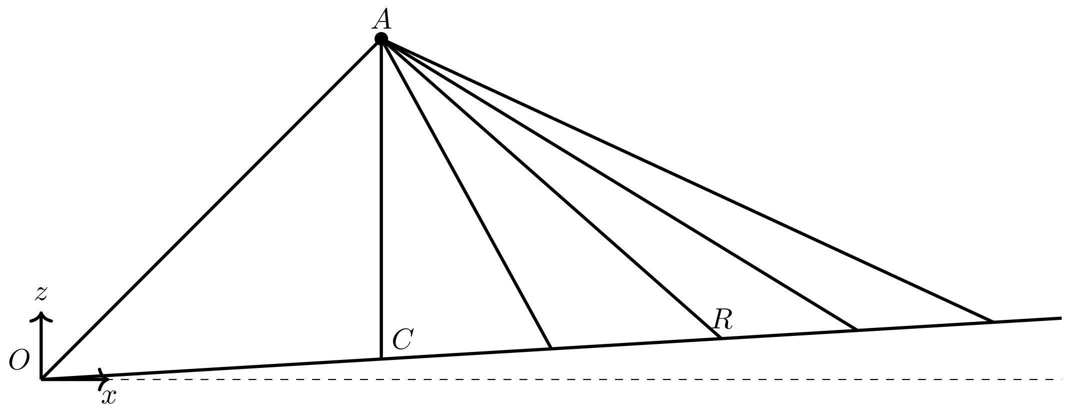


Part (b)(ii). Observe that B, O and C are collinear. Hence,

$$[\triangle ABC]=\frac{1}{2}(1)(2+3)=\frac{5}{2}\text{ units}^2.$$

* * * * *

Problem 10. The diagram below shows the elevation view of a single vertical tower cable-stayed-inclined bridge which stretches across a river. The bridge deck is supported by the tower, a main cable, and four smaller cables. Points are defined relative to an origin O, the point of intersection between the main cable and the deck. The x-, y- and z-axes are in the directions east, north and vertically upwards respectively, with units in metres. The deck of the bridge is modelled as a plane. Points P and Q are on this plane and have coordinates $(20,0,1)$ and $(40,4,2)$ respectively.



(a) Find the Cartesian equation of the plane.

Point A is at the top of the vertical tower and has coordinates $(20, 1, 20)$. Point C is the intersection of the tower and the deck. The tower and the five cables are attached on the deck along the line passing through Points O and C.

(b) The bridge is considered stable if the distance between C and the foot of perpendicular from A to the deck does not exceed 1 m. Comment whether the bridge is stable. Show your working clearly.

(c) One of the cables, which is installed at a point R, has the same length as the main cable. Find the coordinates of R.

(d) Find the acute angle that the deck makes with the horizontal plane.

Solution.

Part (a). Observe that $(20,\ 0,\ 1)^\mathsf{T}$ and $(40,\ 4,\ 2)^\mathsf{T}$ are parallel to the plane. Note that

$$\begin{pmatrix}20\\0\\1\end{pmatrix} \times \begin{pmatrix}40\\4\\2\end{pmatrix} = \begin{pmatrix}-4\\0\\80\end{pmatrix} \parallel \begin{pmatrix}-1\\0\\20\end{pmatrix}.$$

Hence, the vector equation of the plane is

$$\mathbf{r} \cdot \begin{pmatrix}-1\\0\\20\end{pmatrix} = 0,$$

whence the Cartesian equation is $-x + 20z = 0, y \in \mathbb{R}$.

Part (b). Observe that $\overrightarrow{AC} = k\,(0,\ 0,\ 1)^\mathsf{T}$ for some $k \in \mathbb{R}$. Hence, $\overrightarrow{OC} = (20,\ 1,\ 20-k)^\mathsf{T}$. Since C lies on the deck, we have

$$\begin{pmatrix}20\\1\\20-k\end{pmatrix} \cdot \begin{pmatrix}-1\\0\\20\end{pmatrix} = 0,$$

whence $k = 19$ and $\overrightarrow{OC} = (20,\ 1,\ 1)^\mathsf{T}$.

Let F be the foot of perpendicular of A to the deck. Note that $\overrightarrow{OF} \cdot (-1,\ 0,\ 20)^\mathsf{T} = 0$ and $\overrightarrow{AF} = \lambda\,(-1,\ 0,\ 20)^\mathsf{T}$ for some $\lambda \in \mathbb{R}$. Thus,

$$\left[\begin{pmatrix}20\\1\\20\end{pmatrix} + \lambda\begin{pmatrix}-1\\0\\20\end{pmatrix}\right] \cdot \begin{pmatrix}-1\\0\\20\end{pmatrix} = 0, \implies \lambda = -\frac{380}{401} \implies \overrightarrow{OF} = \frac{1}{401}\begin{pmatrix}8400\\401\\420\end{pmatrix}.$$

Hence,

$$\overrightarrow{FC} = \begin{pmatrix} 20 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{401}\begin{pmatrix} 8400 \\ 401 \\ 420 \end{pmatrix} \implies \left|\overrightarrow{FC}\right| = \sqrt{0.948^2 + 0^2 + (-0.0474)^2} = 0.949 < 1.$$

The bridge is thus stable.

Part (c). We have $\left|\overrightarrow{AR}\right| = \left|\overrightarrow{OA}\right| = \sqrt{20^2 + 1^2 + 20^1} = \sqrt{801}$. Since O, C and R are collinear, we also have $\overrightarrow{OR} = \mu\,(20,\ 1,\ 1)^{\mathsf{T}}$ for some $\mu \in \mathbb{R}$. Thus,

$$\left|\overrightarrow{AR}\right| = \left|\mu\begin{pmatrix} 20 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 20 \\ 1 \\ 20 \end{pmatrix}\right| = \sqrt{(20\mu - 20)^2 + (\mu - 1)^2 + (\mu - 20)^2} = \sqrt{801}.$$

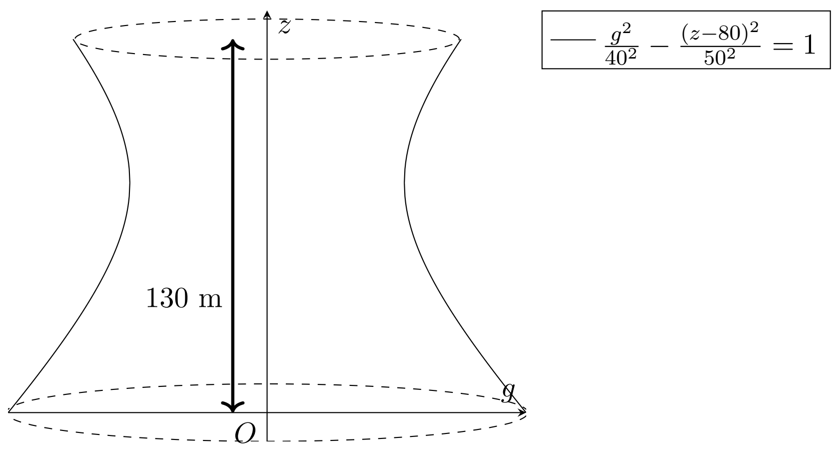
Using G.C., $\mu = 2.09$, whence $R(41.9, 2.09, 2.09)$.

Part (d). Let θ be the acute angle between the deck and the horizontal. Note that the horizontal plane has normal vector $(0,\ 0,\ 1)^{\mathsf{T}}$. Thus,

$$\cos\theta = \frac{\left|(0,\ 0,\ 1)^{\mathsf{T}} - (-1,\ 0,\ 20)^{\mathsf{T}}\right|}{\left|(0,\ 0,\ 1)^{\mathsf{T}}\right|\left|(-1,\ 0,\ 20)^{\mathsf{T}}\right|} = \frac{20}{\sqrt{401}} \implies \theta = 2.9^\circ \text{ (2 d.p.)}.$$

* * * * *

Problem 11. A nuclear reactor plant is built to meet the increased demand for electricity due to a particular country's economic developments. The cooling tower of the nuclear reactor is as shown in the figure below. The curved surface area of the cooling tower is modelled by rotating the region enclosed by a part of a hyperbolic curve about an axis. The height of the tower is 130 m.



The equation of the hyperbolic curve is given as $\frac{g^2}{40^2} - \frac{(z-80)^2}{50^2} = 1$ where g is the axis that represents the ground and z is the axis that represents the height of the reactor. The curve surface area of the tower is formed by rotating the region bounded by the hyperbolic curve, the line $z = 130$ and the g axis about the z-axis by π radians. The external curved surface area of the tower is to be painted with weather resistant paint.

(a) Find the external curved surface area of the tower. Leave your answer to the nearest m^2.

The ground is now represented by the x-y plane.

(b) Find the Cartesian equation that models the surface of the tower in terms of z, y and z.

Before the paint can be applied, a robot is programmed to go around the tower to clean and polish its surface. Assuming that the robot is negligible compared to the tower, it can be viewed as a point on the curved surface of the tower.

(c) Given that the robot is at $(40, 40, 30)$ and is about the move in the direction of $(3,\ -4)^{\mathsf{T}}$ parallel to the x-y plane, determine whether the robot will be ascending or descending in height.

The robot is now at $(40, 40, 130)$ on the surface of the tower. A signal needs to be transmitted from the ground to the robot such that the signal travels in a straight line and its direction must be normal to the surface of the tower where the robot is at.

(d) Find the coordinates on the ground where the signal can be transmitted to the robot.

Solution.

Part (a). Note that $g = \sqrt{40^2\left[\frac{(z-80)^2}{50^2}+1\right]}$. Using G.C.,

$$\text{Area} = 2\pi\int_0^{130} g\sqrt{1+\left(\frac{\mathrm{d}g}{\mathrm{d}z}\right)^2}\,\mathrm{d}z \approx 45552 \text{ units}^2.$$

Part (b). For every constant value of z, we will have the value of g such that $x^2+y^2=g^2$. Hence,

$$\frac{x^2+y^2}{40^2} - \frac{(z-80)^2}{50^2} = 1.$$

Part (c). Implicitly differentiating the above expressing with respect to x and y, we have

$$\frac{\partial z}{\partial x} = \left(\frac{5}{4}\right)^2 \frac{x}{z-80}, \quad \frac{\partial z}{\partial y} = \left(\frac{5}{4}\right)^2 \frac{y}{z-80}.$$

Evaluating at $(40, 40, 30)$, we have that $\nabla z = -\frac{5}{4}\,(1,\ 1)^{\mathsf{T}}$. Hence,

$$\nabla z \cdot \frac{1}{\sqrt{3^2+5^2}}\begin{pmatrix}3\\-4\end{pmatrix} = \frac{1}{4} > 0.$$

Thus, the robot is ascending.

Part (d). At $(40, 40, 130)$, we have $\nabla z = (5/4,\ 5/4)^{\mathsf{T}}$. The equation of the tangent plane at that point is hence

$$z = 130 + \frac{5}{4}(x-40) + \frac{5}{4}(y-40),$$

which has vector equation

$$\mathbf{r}\cdot\begin{pmatrix}5\\5\\-4\end{pmatrix} = -120.$$

The line of the signal is hence given by

$$\mathbf{r} = \begin{pmatrix}40\\40\\130\end{pmatrix} + \lambda\begin{pmatrix}5\\5\\-4\end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Setting $z = 0$, we have $\lambda = 130/4$, whence $x = y = 405/2$. The required coordinates are thus $\left(\frac{405}{2}, \frac{405}{2}, 0\right)$.

9649 JC2 Weighted Assessment 1

Problem 1. A popular food chain Dunman Fried Chicken (DFC) is giving out 9 types of LaTuTu dolls as a promotional item. Each customer will randomly get any one of the 9 types of dolls when they purchase a meal. Suppose a customer has already collected r different types of dolls where $r = 1, 2, 3, \ldots, 8$. Let X_r be the random variable denoting the additional number of dolls that needs to be collected by the customer until he gets a different type of doll from his current collection of r types of dollts. It can be assumed that the food chain has many dolls for each type.

(a) Find $\mathbb{P}[X_r \leq n]$ in terms of n.

(b) Given that it took 6 attempts for the customer to collect the $(r+1)$th doll, find the probability that the customer takes either less than 9 or more than 12 attempts in total to collect the $(r+1)$th and $(r+2)$th dolls.

(c) If the customer gets a type of doll that he already possessed, he will sell the doll at a loss of \$2, after taking into account the cost of purchasing a meal for DFC. There is no loss if the customer gets a different type of doll when he purchases a meal. Find the expected amount of money that he will lose for him to collect the entire collection of dolls.

Solution.

Part (a). The probability of getting a different doll is $(9-r)/9 = 1 - r/9$. Thus, $X_r \sim \text{Geo}(1 - r/9)$. Hence,

$$\mathbb{P}[X_r > n] = \mathbb{P}[\text{first } n \text{ trials all failures}] = \left(\frac{r}{9}\right)^n.$$

Thus,

$$\mathbb{P}[X_r \leq n] = 1 - \mathbb{P}[X_r > n] = 1 - \left(\frac{r}{9}\right)^n.$$

Part (b). The desired probability is

$$\begin{aligned}\mathbb{P}[X_r + X_{r+1} < 9 \text{ or } X_r + X_{r+1} > 12 \mid X_r = 6] &= \mathbb{P}[X_{r+1} < 3 \text{ or } X_{r+1} > 6] \\ = \mathbb{P}[X_{r+1} < 3] + \mathbb{P}[X_{r+1} > 6] &= 1 - \left(\frac{r+1}{9}\right)^2 + \left(\frac{r+1}{9}\right)^6.\end{aligned}$$

Part (c). Note that $\mathbb{E}[X_r] = 1/(1 - r/9)$. Hence, the expected number of attempts is given by

$$1 + \sum_{r=1}^{8} \mathbb{E}[X_r] = 1 + \sum_{r=1}^{8} \frac{1}{1 - r/9} = \frac{7129}{280}.$$

Hence, the expected amount of money that he will lose is

$$2\left(\frac{7129}{280} - 8\right) = \$32.92 \text{ (2 d.p.)}.$$

* * * * *

Problem 2. A car salesman receives \$1000 commission for each new car that he sells and \$600 for each used car that he sells. The weekly number of new cars that he sells has a Poisson distribution with mean 3 and, independently, the number of used cars that he sells has a Poisson distribution with mean 2.

(a) Find the probability that his commission in a week is exactly \$3000.

(b) Calculate the mean and variance of the salesman's weekly commission and determine whether the commission has a Poisson distribution.

(c) The salesman sold a total of 16 cars in 4 weeks. Find the probability that he sold less than half of it in the first week.

Solution. Let the number of new and used cards sold in a week be N and U respectively. Then $N \sim \text{Po}(3)$ and $U \sim \text{Po}(5)$.

Part (a). For his weekly commission to be exactly \$3000, he must either sell only 3 new cars, or only 5 used cards. The required probability is thus

$$\mathbb{P}[N = 3 \mid U = 0] + \mathbb{P}[N = 0]\,\mathbb{P}[U = 5] = 0.0321 \text{ (3 s.f.)}.$$

Part (b). Let his weekly commission be C. Then $C = 1000N + 600U$. The mean is

$$\mathbb{E}[C] = 1000\,\mathbb{E}[N] + 600\,\mathbb{E}[U] = 1000(3) + 600(2) = 4200.$$

However, the variance is

$$\text{Var}[C] = 1000^2\,\text{Var}[N] + 600^2\,\text{Var}[U] = 1000^2(3) + 600^2(2) = 3720000.$$

Since $\mathbb{E}[C] \neq \text{Var}[C]$, it follows that C does not follow a Poisson distribution.

Part (c). Let X be the number of cars sold in the first week. Then $X \sim \text{B}(16, 1/4)$. The required probability is thus

$$\mathbb{P}[X < 8] = 0.973 \text{ (3 s.f.)}.$$

* * * * *

Problem 3. The variables x, y and z are related by the two differential equations

$$\frac{\mathrm{d}y}{\mathrm{d}x} - 2y + z = 4\sin(2x)\,, \tag{1}$$

$$\frac{\mathrm{d}z}{\mathrm{d}x} - 8y + 2z = 2\cos(2x)\,. \tag{2}$$

When $x = 0$, $y = z = 0$.

(a) Show that the system of differential equations can be reduced to the second-order differential equation

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 4y = 6\cos(2x) + 8\sin(2x)\,.$$

(b) Hence, solve the differential equation in (a) to find y in terms of x. Hence, find z in terms of x.

Solution.

Part (a). Differentiating (1) with respect to x,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 2\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\mathrm{d}z}{\mathrm{d}x} = 8\cos(2x)\,.$$

Taking (2) - 2(1) and rearranging, we have

$$\frac{\mathrm{d}z}{\mathrm{d}x} = 4y + 2\frac{\mathrm{d}y}{\mathrm{d}x} + 2\cos(2x) - 8\sin(2x)\,.$$

Substituting this into (3) yields

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 2\frac{\mathrm{d}y}{\mathrm{d}x} + \left(4y + 2\frac{\mathrm{d}y}{\mathrm{d}x} + 2\cos(2x) - 8\sin(2x)\right) = 8\cos(2x)\,.$$

Rearranging, we get

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 4y = 6\cos(2x) + 8\sin(2x)\,.$$

Part (b). The characteristic equation of the DE is $m^2+4=0$, whence its roots are $m\pm 2\mathrm{i}$. The complementary solution is thus

$$y_c = A\sin(2x) + B\cos(2x)\,.$$

For the particular solution, we try $y_p = x\,[C\sin(2x) + D\cos(2x)]$. Differentiating, we get

$$\frac{\mathrm{d}y_p}{\mathrm{d}x} = 2x\,[C\cos(2x) - D\sin(2x)] + [C\sin(2x) + D\cos(2x)]\,.$$

Differentiating once more, we get

$$\frac{\mathrm{d}^2y_p}{\mathrm{d}x^2} = -4x\,[C\sin(2x) + D\cos(2x)] + 4\,[C\cos(2x) - D\sin(2x)]\,.$$

Substituting this into the DE, we obtain

$$\begin{aligned} &-4x\,[C\sin(2x) + D\cos(2x)] + 4\,[C\cos(2x) - D\sin(2x)] \\ &\qquad + 4x\,[C\sin(2x) + D\cos(2x)] = 6\cos(2x) + 8\sin(2x)\,. \end{aligned}$$

Simplifying,

$$C\cos(2x) - D\sin(2x) = \frac{3}{2}\cos(2x) + 2\sin(2x)\,.$$

Comparing coefficients, we see that $C = 3/2$ and $D = -2$. Thus, the general solution for y is

$$y = y_c + y_p = A\sin(2x) + B\cos(2x) + \frac{3}{2}x\sin(2x) - 2x\cos(2x)\,.$$

When $x = 0$, $y = 0$. Thus, $B = 0$.

Differentiating y, we get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2A\cos(2x) + 3x\cos(2x) + \frac{3}{2}\sin(2x) + 4x\sin(2x) - 2\cos(2x)\,.$$

Substituting the initial conditions into (1), we see that $\mathrm{d}y/\mathrm{d}x = 0$ when $x = 0$. Thus, $2A - 2 = 0$, whence $A = 1$. Thus,

$$y = \sin(2x) + \frac{3}{2}x\sin(2x) - 2x\cos(2x)\,.$$

Substituting our expressions for y and $\mathrm{d}y/\mathrm{d}x$ into (1) and simplifying, we see that

$$z = \frac{9}{2}\sin(2x) - 7x\cos(2x) - x\sin(2x)\,.$$

9649 JC2 Weighted Assessment 1

Problem 1. The matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{pmatrix} 8 & 3 & -12 \\ -5 & -2 & 8 \\ 10 & 4 & -15 \end{pmatrix}.$$

(a) By performing elementary row operations on the matrix $\left(\mathbf{A} \mid \mathbf{I}\right)$, showing all necessary working, find $\mathbf{A}^{-1}$.

(b) Solve the equation

$$\begin{pmatrix} x & y & z & w \end{pmatrix} \begin{pmatrix} 8 & 3 & -12 \\ -1 & 2 & k \\ -5 & -2 & 8 \\ 10 & 4 & -15 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 \end{pmatrix},$$

where k is a real constant, leaving your answers in terms of k where appropriate.

Solution.

Part (a). We have

$$\left(\begin{array}{ccc|ccc} 8 & 3 & -12 & 1 & 0 & 0 \\ -5 & -2 & 8 & 0 & 1 & 0 \\ 10 & 4 & -15 & 0 & 0 & 1 \end{array}\right) \to \begin{array}{r} {\scriptstyle 2R_1+3R_2} \\ {\scriptstyle -5R_1+4R_3} \\ {\scriptstyle 2R_2+R_3} \end{array} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 3 & 0 \\ 0 & 1 & 0 & -5 & 0 & 4 \\ 0 & 0 & 1 & 0 & 2 & 1 \end{array}\right),$$

so

$$\mathbf{A}^{-1} = \begin{pmatrix} 2 & 3 & 0 \\ -5 & 0 & 4 \\ 0 & 2 & 1 \end{pmatrix}.$$

Part (b). Rearranging rows, we see that

$$\begin{pmatrix} x & z & w & y \end{pmatrix} \begin{pmatrix} & \mathbf{A} & \\ -1 & 2 & k \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 \end{pmatrix}.$$

Post-multiplying by $\mathbf{A}^{-1}$, we get

$$\begin{pmatrix} x & z & w & y \end{pmatrix} \begin{pmatrix} & \mathbf{I} & \\ -12 & 2k-3 & k+8 \end{pmatrix} = \begin{pmatrix} 14 & 8 & -7 \end{pmatrix}.$$

We thus get the system of equations

$$\begin{cases} x - 12y = 14 \\ (2k-3)\,y + z = 8 \\ (k+8)\,y + w = -7 \end{cases}.$$

Let $y = t \in \mathbb{R}$. The solution to the equation is thus

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 14 + 12t \\ t \\ 8 - (2k-3)t \\ -7 - (k+8)t \end{pmatrix}.$$

Problem 2.

(a) A square matrix $\mathbf{A}$ of order n is said to be skew symmetric if $\mathbf{A}^{\mathsf{T}} = -\mathbf{A}$. Prove that a skew symmetric matrix is not invertible if n is odd.

(b) Let T be the linear transformation such that

$$T : \mathbb{R}^3 \to \mathbb{R}^3 \quad \text{and} \quad T(\mathbf{x}) = \mathbf{A}\mathbf{x},$$

where

$$\mathbf{A} = \begin{pmatrix} 0 & -1 & -5 \\ 1 & 0 & a \\ 5 & b & 0 \end{pmatrix},$$

where a and b are real numbers. It is given that the nullity of T is 1.

(i) Show that $\mathbf{A}$ must be a skew symmetric matrix.

(ii) Find the kernel of T in terms of a.

(iii) State a basis, in terms of a where appropriate, for the range space of T and give a geometrical interpretation of your answer of the range space of T.

Solution.

Part (a). Note that

$$\det \mathbf{A} = \det \mathbf{A}^{\mathsf{T}} = \det(-A) = (-1)^n \det \mathbf{A}.$$

For odd n, $\det \mathbf{A} = -\det \mathbf{A}$ so $\det \mathbf{A} = 0$, whence $\mathbf{A}$ is not invertible.

Part (b).

Part (b)(i). Since the nullity of T, is 1, the rows of $\mathbf{A}$ must be linearly dependent. That is, there exist $\lambda, \mu \in \mathbb{R}$ such that

$$\lambda \begin{pmatrix} 0 \\ -1 \\ -5 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ a \end{pmatrix} = \begin{pmatrix} 5 \\ b \\ 0 \end{pmatrix}.$$

From the first and second rows, we immediately have $\lambda = -b$ and $\mu = 5$. Substituting this into the third row, we get $a = -b$. Thus,

$$\mathbf{A} = \begin{pmatrix} 0 & -1 & -5 \\ 1 & 0 & a \\ 5 & -a & 0 \end{pmatrix}$$

and is skew symmetric.

Part (b)(ii). Consider $\mathbf{A}\mathbf{v} = \mathbf{0}$, where $\mathbf{v} = (x, y, z)^{\mathsf{T}} \in \mathbb{R}^3$. Then

$$\begin{pmatrix} 0 & -1 & -5 \\ 1 & 0 & a \\ 5 & -a & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{cases} y + 5z = 0 \\ x + az = 0 \\ 5x - ay = 0 \end{cases}.$$

Let $z = \lambda \in \mathbb{R}$. From the first two equations, we see that $y = -5\lambda$ and $x = -a\lambda$. Thus, $\mathbf{v}$ is of the form $\lambda(-a, -5, 1)^{\mathsf{T}}$, whence the kernel of T is given by

$$\ker T = \left\{ \mathbf{v} \in \mathbb{R}^3 : \mathbf{v} = \lambda \begin{pmatrix} -a \\ -5 \\ 1 \end{pmatrix}, \lambda \in \mathbb{R} \right\}.$$

Part (b)(iii). A basis is

$$\left\{ \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} -5 \\ a \\ 0 \end{pmatrix} \right\}.$$

Note that the range and null spaces of T are orthogonal. Thus, the range space of T is the plane normal to $(-a, -5, 1)^\mathsf{T}$ that passes through the origin.

* * * * *

Problem 3. A company invests its funds in 3 interdependent sectors of Manufacturing, Research and Development, and Services. While these sectors had typically generated good profits of 200 million each monthly, the company is concerned that recent and rapid changes in the economic conditions in the country will very quickly negatively affect the returns from these sectors.

To model the possible profits from the sectors given the current conditions, the company's analyst suggested the following systems of equations:

$$\begin{aligned} M_{n+1} &= \frac{2}{3}M_n - \frac{1}{3}R_n + \frac{1}{3}S_n, \\ R_{n+1} &= \frac{1}{2}M_n - \frac{1}{6}R_n - \frac{1}{2}S_n, \\ S_{n+1} &= \frac{5}{6}M_n - \frac{5}{6}R_n + \frac{1}{6}S_n, \end{aligned}$$

where M_n, R_n and S_n are the profits (in millions) earned from the Manufacturing, Research and Development, and Services respectively after n months.

The system of equations may be written in the form $\mathbf{P}_{n+1} = \mathbf{A}\mathbf{P}_n$, where $P_n = (M_n, R_n, S_n)^\mathsf{T}$ and $\mathbf{A}$ is an appropriate matrix.

(a) Evaluate $\mathbf{P}_1$ and $\mathbf{P}_2$.

(b) Determine the eigenvalues and the corresponding eigenvectors of $\mathbf{A}$.

(c) Use your answers above to explain why the company's concern is valid.

The company decide to re-strategise its position in the 3 sectors, following which the analyst revises the model to the following:

$$\begin{aligned} M_{n+1} &= 2M_n - R_n + S_n, \\ R_{n+1} &= \frac{3}{2}M_n - \frac{1}{2}R_n - \frac{3}{2}S_n, \\ S_{n+1} &= \frac{5}{2}M_n - \frac{5}{2}R_n + \frac{1}{2}S_n, \end{aligned}$$

which may be expressed in the form $\mathbf{P}_{n+1} = \mathbf{B}\mathbf{P}_n$, where $\mathbf{B}$ is an appropriate matrix.

(d) Write down the eigenvalues of $\mathbf{B}$.

(e) Explain, with appropriate working, the profit trend the company will see arising from the 3 sectors after re-strategising.

Solution.

Part (a). Note that

$$\mathbf{A} = \frac{1}{6}\begin{pmatrix} 4 & -2 & 2 \\ 3 & -1 & -3 \\ 5 & -5 & 1 \end{pmatrix}.$$

The first few values of $\mathbf{P}_n$ are

$$\mathbf{P}_0 = \begin{pmatrix} 200 \\ 200 \\ 200 \end{pmatrix}, \quad \mathbf{P}_1 = \begin{pmatrix} 133.33 \\ -33.333 \\ 33.333 \end{pmatrix}, \quad \mathbf{P}_2 = \begin{pmatrix} 111.11 \\ 55.556 \\ 144.44 \end{pmatrix}.$$

Part (b). Let the characteristic polynomial of $\mathbf{A}$ be $\chi(\lambda) = \lambda^3 - a_2\lambda^2 + a_1\lambda - a_0$. Then

$$\begin{aligned} a_0 &= |\mathbf{A}| = -\frac{2}{9}, \\ a_1 &= \frac{1}{6^2}\left(\begin{vmatrix} 4 & -2 \\ 3 & -1 \end{vmatrix} + \begin{vmatrix} -1 & -3 \\ -5 & 1 \end{vmatrix} + \begin{vmatrix} 4 & 2 \\ 5 & 1 \end{vmatrix}\right) = \frac{-5}{9}, \\ a_2 &= \frac{1}{6}\left(|4| + |-1| + |1|\right) = \frac{2}{3}. \end{aligned}$$

Thus, $\chi(\lambda) = \lambda^3 - \frac{2}{3}\lambda^2 - \frac{5}{9}\lambda + \frac{2}{9}$, which has roots $\lambda = 1$, $\lambda = -2/3$ and $\lambda = 1/3$.

Let $\mathbf{x}$ be a non-zero eigenvector of $\mathbf{A}$.

Case 1: $\lambda = 1$. Consider

$$\left[\frac{1}{6}\begin{pmatrix} 4 & -2 & 2 \\ 3 & -1 & -3 \\ 5 & -5 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]\mathbf{x} = \mathbf{0}.$$

Using G.C., $\mathbf{x} = (1, 0, 1)^\mathsf{T}$.

Case 2: $\lambda = -2/3$. Consider

$$\left[\frac{1}{6}\begin{pmatrix} 4 & -2 & 2 \\ 3 & -1 & -3 \\ 5 & -5 & 1 \end{pmatrix} + \frac{2}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]\mathbf{x} = \mathbf{0}.$$

Using G.C., $\mathbf{x} = (0, 1, 1)^\mathsf{T}$.

Case 3: $\lambda = 1/3$. Consider

$$\left[\frac{1}{6}\begin{pmatrix} 4 & -2 & 2 \\ 3 & -1 & -3 \\ 5 & -5 & 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]\mathbf{x} = \mathbf{0}.$$

Using G.C., $\mathbf{x} = (1, 1, 0)^\mathsf{T}$.

Thus, $\mathbf{A} = \mathbf{QDQ}^{-1}$, where

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}.$$

Part (c). Note that

$$\lim_{n\to\infty} \mathbf{P}_n = \lim_{n\to\infty} A^n\mathbf{P}_0 = \lim_{n\to\infty} \mathbf{QD}^n\mathbf{Q}^{-1}\mathbf{P}_0 = \mathbf{Q}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\mathbf{Q}^{-1}\mathbf{P}_0 = \begin{pmatrix} 100 \\ 0 \\ 100 \end{pmatrix}.$$

Thus, in the long run, profits from Manufacturing and Services will fall from \$200 million to \$100 million, while profits from Research and Development will completely vanish. Hence, the companies concern is valid.

Part (d). Clearly, $\mathbf{B} = 3\mathbf{A}$, so its eigenvalues are 3, -2 and 1.

Part (e). Note that $\mathbf{B} = \mathbf{Q}(3\mathbf{D})\mathbf{Q}^{-1}$. Thus,

$$\begin{aligned}
\mathbf{P}_n &= A^n\mathbf{P}_0 = \mathbf{Q}(3\mathbf{D})^n\mathbf{Q}^{-1}\mathbf{P}_0 \\
&= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3^n & 0 & 0 \\ 0 & (-2)^n & 0 \\ 0 & 0 & 1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix} 200 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
&= 200 \begin{pmatrix} 3^n + 1 \\ (-2)^n + 1 \\ 3^n + (-2)^n \end{pmatrix}.
\end{aligned}$$

Profits from Manufacturing and Services will increase greatly over time. However, Research and Development will fluctuate between earning large profits and incurring large losses.

9820 H3 Mathematics

9820 Timed Practice 1

Problem 1. At a school charity bazaar, students sell chicken nuggets in boxes of 5 or 11. Customers can buy any combination of boxes of 5 or 11 chicken nuggets.

(a) Show that it is possible to buy exactly k chicken nuggets, for all integer values of k between 40 and 44 inclusive.

(b) Use mathematical induction to show that it is possible to buy exactly k chicken nuggets for all integer values of k greater than or equal to 40.

(c) Determine if it is possible to buy exactly 39 chicken nuggets.

(d) In the general case where they sell chicken nuggets in boxes of p_1 or p_2, where p_1 and p_2 are distinct primes, show that it is not possible to buy exactly $p_1p_2 - p_1 - p_2$ chicken nuggets.

Solution.

Part (a).

k	Boxes of 5	Boxes of 11
40	8	0
41	6	1
42	4	2
43	2	3
44	0	4

Part (b). Let $P(5n + r)$ be the statement "it is possible to buy exactly $5n + r$ chicken nuggets", where $n \in \mathbb{Z}$ and $r \in \{0, 1, 2, 3, 4\}$. From part (a), the statement holds for $n = 8$ and $r \in \{0, 1, 2, 3, 4\}$. We now induct on n. Suppose $P(5k + r)$ is true for some $k \in \mathbb{Z}$. Then there exist positive integers x and y such that

$$5x + 11y = 5k + r.$$

It follows that

$$5\,(k + 1) + r = 5\,(x + 1) + 11y.$$

Hence, taking $(x+1)$ boxes of 5 chicken nuggets and y boxes of chicken nuggets, we obtain exactly $5(k + 1) + r$ chicken nuggets. Hence, $P(5k + r) \implies P(5(k + 1) + r)$. This closes the induction. We thus conclude that we can get exactly m chicken nuggets for all $m \geq 8 \cdot 5 = 40$.

Part (c). By way of contradiction, suppose we can buy exactly 39 chicken nuggets. Then there exist positive integers x and y such that

$$5x + 11y = 39.$$

Reducing modulo 5, we see that

$$y \equiv 4 \pmod{5},$$

so $\min y = 4$. Thus,

$$\min (5x + 11y) \geq 11 \min y = 44 > 39.$$

Thus, such a y cannot exist. Therefore, we cannot buy exactly 39 nuggets.

Part (d). By way of contradiction, suppose we can buy exactly $p_1 p_2 - p_1 - p_2$ nuggets. Then there exist positive integers x and y such that

$$p_1 x + p_2 y = p_1 p_2 - p_1 - p_2.$$

Reducing modulo p_1, we see that

$$p_2 y \equiv -p_2 \pmod{p_1} \implies p_2 (y + 1) \equiv 0 \pmod{p_1}.$$

Because p_1 and p_2 are distinct, they must be coprime, so

$$y + 1 \equiv 0 \pmod{p_1}.$$

It follows that $\min y = p_1 - 1$. Thus,

$$\min (p_1 x + p_2 y) \geq p_2 \min y = p_2 (p_1 - 1) = p_1 p_2 - p_2 > p_1 p_2 - p_1 - p_2.$$

Thus, such a y cannot exist. Therefore, we cannot buy exactly $p_1 p_2 - p_1 - p_2$ nuggets.

* * * * *

Problem 2. Let $g(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ be a polynomial with integer coefficients, that is, $a_0, a_1, \ldots, a_n \in \mathbb{Z}$, with $a_0, a_n \neq 0$.

(a) (i) By writing

$$g(x) = x^n \left(\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \frac{a_2}{x^{n-1}} + \cdots + \frac{a_{n-1}}{x} + a_n \right),$$

explain why as $x \to \infty$, either $g(x) \to \infty$ or $g(x) \to -\infty$.

(ii) Let

$$P = \{ p : p \text{ is a prime and } \exists\, k \in \mathbb{Z} \text{ such that } p \mid g(k) \}.$$

Show that P contains infinitely many elements.

(b) Show that there exists an integer $m > 2025$ such that $|g(m)|$ is not prime.

Solution.

Part (a).

Part (a)(i). Note that for all $k \in \mathbb{R}$ and $r \in \mathbb{N}$, we have

$$\lim_{x \to \infty} \frac{k}{x^r} = 0.$$

Thus,

$$\lim_{x \to \infty} g(x) = \lim_{x \to \infty} x^n \left(\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \frac{a_2}{x^{n-1}} + \cdots + \frac{a_{n-1}}{x} + a_n \right) = \lim_{x \to \infty} a_n x^n,$$

which approaches ∞ if $a_n > 0$, or $-\infty$ if $a_n < 0$.

Part (a)(ii). By way of contradiction, suppose P is finite, say $P = \{p_1, \ldots, p_k\}$. Then for all $m \in \mathbb{Z}$, either $g(m) = 0$ or, by the Fundamental Theorem of Arithmetic,

$$g(m) = \pm p_1^{a_1} p_2^{a_2} \ldots p_k^{a_k},$$

where $a_i \in \mathbb{Z}_0^+$. By part (a)(i), there exists some $M \in \mathbb{N}$ such that $g(m) \geq 0$ for all $m \geq M$. Define

$$S = \sum_{m=M}^{\infty} \frac{1}{g(m)^{1/n}}.$$

Since g is a polynomial of degree n, it can take on a given value at most n times. Thus, for a given set of integers $a_1, a_2, \ldots, a_k$,

$$g(m)^{1/n} = p_1^{a_1/n} p_2^{a_2/n} \ldots p_k^{a_k/n}$$

for at most n values of m. We thus obtain the following estimate:

$$S \leq \sum_{a_1, \ldots, a_k = 0}^{\infty} \frac{n}{p_1^{a_1/n} \ldots p_k^{a_k/n}} = n \prod_{r=1}^{k} \sum_{a_r=0}^{\infty} p_r^{-a_r/n} = n \prod_{r=1}^{k} \frac{1}{1 - p_r^{-1/n}},$$

which is finite. But this is a clear contradiction, since

$$\frac{1}{g(m)^{1/n}} \sim \frac{1}{a_n^{1/n} x},$$

so S diverges like the harmonic series. Thus, P must be infinite.

Part (b). By way of contradiction, suppose $|g(m)|$ for all integers $m > 2025$. Let $p = |g(m)|$, which is prime, and $t \in \mathbb{N}$. Observe that

$$g(m + tp) = \sum_{i=0}^{n} a_i \left(m + tp\right)^i \equiv \sum_{i=0}^{n} a_i m^i = g(m) \equiv 0 \pmod{p}.$$

Since $m + tp > 2025$, by our assumption, $g(m + tp)$ must also be prime. Thus,

$$\pm p = g(m) = g(m + p) = g(m + 2p) = g(m + 3p) = \ldots,$$

so there are infinitely many solutions to the equation $g(x) = \pm p$. However, because g is a polynomial of degree n, there are at most $2n$ solutions to $g(x) = \pm p$, a contradiction. Thus, there must exist some $m > 2025$ such that $|g(m)|$ is not prime.

9820 Timed Practice 1

Problem 1.

(a) Given that $x \geq 1$, $y \geq 1$, $x \geq 1$, and $x^{-1} + y^{-1} + z^{-1} = 2$, by using the Cauchy-Schwarz inequality, prove that

$$\sqrt{x+y+z} \geq \sqrt{x-1} + \sqrt{y-1} + \sqrt{z-1}.$$

(b) If w, x, y, z are positive integers, using the AM-GM inequality, find the maximum possible value of

$$\frac{wxyz}{(w+x+y)(x+y+z)(y+z+w)(z+w+x)}.$$

Solution.

Part (a). By the Cauchy-Schwarz inequality,

$$\left(\sqrt{x-1} + \sqrt{y-1} + \sqrt{z-1}\right)^2 \leq \left(\frac{x-1}{x} + \frac{y-1}{y} + \frac{z-1}{z}\right)(x+y+z).$$

From the given condition, we see that

$$\frac{x-1}{x} + \frac{y-1}{y} + \frac{z-1}{z} = \left(1 - \frac{1}{x}\right) + \left(1 - \frac{1}{y}\right) + \left(1 - \frac{1}{z}\right) = 3 - 2 = 1.$$

Thus,

$$\sqrt{x-1} + \sqrt{y-1} + \sqrt{z-1} \leq \sqrt{1}\sqrt{x+y+z} = \sqrt{x+y+z}$$

as desired.

Part (b). By the AM-GM inequality, $w + x + y \geq 3\sqrt[3]{wxy}$. Thus,

$$\frac{wxyz}{\prod_{\text{cyc}}(w+x+y)} \leq \frac{wxyz}{\prod_{\text{cyc}} 3\sqrt[3]{wxy}} = \frac{wxyz}{3^4\sqrt[3]{w^3x^3y^3z^3}} = \frac{1}{3^4}.$$

Thus, the maximum is 3^{-4}, which occurs when $w = x = y = z$.

* * * * *

Problem 2. Prove that, for any positive integers n and r,

$$\frac{1}{{}^{n+r}C_{r+1}} = \frac{r+1}{r}\left(\frac{1}{{}^{n+r-1}C_r} - \frac{1}{{}^{n+r}C_r}\right).$$

Hence, determine

$$\sum_{n=1}^{\infty} \frac{1}{{}^{n+r}C_{r+1}}$$

in terms of r, and use the result obtained to deduce that

$$\sum_{n=2}^{\infty} \frac{1}{{}^{n+2}C_3} = \frac{1}{2}.$$

Solution. Observe that

$$\begin{aligned}
\frac{r+1}{r}\left(\frac{{}^{n+r}C_{r+1}}{{}^{n+r-1}C_r} - \frac{{}^{n+r}C_{r+1}}{{}^{n+r}C_r}\right) &= \frac{r+1}{r}\left(\frac{\frac{(n+1)!}{(r+1)!(n-1)!}}{\frac{(n+r-1)!}{r!(n-1)!}} - \frac{\frac{(n+r)!}{(r+1)!(n-1)!}}{\frac{(n+r)!}{r!n!}}\right) \\
&= \frac{r+1}{r}\left(\frac{n+r}{r+1} - \frac{n}{r+1}\right) \\
&= \frac{n+r}{r} - \frac{n}{r} \\
&= 1.
\end{aligned}$$

Dividing through by ${}^{n+r}C_{r+1}$, we get

$$\frac{1}{{}^{n+r}C_{r+1}} = \frac{r+1}{r}\left(\frac{1}{{}^{n+r-1}C_r} - \frac{1}{{}^{n+r}C_r}\right).$$

Hence, the required sum telescopes to

$$\sum_{n=1}^{\infty} \frac{1}{{}^{n+r}C_{r+1}} = \frac{r+1}{r}\sum_{n=1}^{\infty}\left(\frac{1}{{}^{n+r-1}C_r} - \frac{1}{{}^{n+r}C_r}\right)$$

$$= \frac{r+1}{r}\left(\frac{1}{{}^{r}C_r} - \frac{1}{{}^{r+1}C_r} + \frac{1}{{}^{r+1}C_r} - \frac{1}{{}^{r+2}C_r} + \dots\right) = \frac{r+1}{r}\frac{1}{{}^{r}C_r} = \frac{r+1}{r}.$$

Thus, we have

$$\sum_{n=2}^{\infty} \frac{1}{{}^{n+2}C_3} = \sum_{n=1}^{\infty} \frac{1}{{}^{n+2}C_3} - \frac{1}{{}^{3}C_3} = \frac{2+1}{2} - 1 = \frac{1}{2}.$$

* * * * *

Problem 3.

(a) By using a suitable substitution or otherwise, find

$$\int \frac{1}{\mathrm{e}^x + 1}\,\mathrm{d}x,$$

expressing your answer as a single logarithm.

(b) By considering the geometric series $1 - t + t^2 - t^3 + \dots$, show that for $|t| < 1$,

$$\ln(1+t) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}t^k}{k}.$$

(c) Explain why for all $n \in \mathbb{Z}^+$,

$$\lim_{x\to\infty} x\mathrm{e}^{nx} = 0.$$

(d) Prove that

$$\int_0^{\infty} \frac{x}{\mathrm{e}^x + 1}\,\mathrm{d}x = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2},$$

showing clearly at which step of your working you have used part (c). You may interchange the summation and integral without justification.

Solution.

Part (a). Dividing the integrand by e^x, we see that

$$\int \frac{1}{\mathrm{e}^x+1}\,\mathrm{d}x = \int \frac{\mathrm{e}^{-x}}{1+\mathrm{e}^{-x}}\,\mathrm{d}x = -\ln\bigl(1+\mathrm{e}^{-x}\bigr)+C = \ln\left(\frac{1}{1+\mathrm{e}^{-x}}\right)+C.$$

Part (b). Note that

$$1-t+t^2-t^3+\cdots = \sum_{k=0}^{\infty}(-t)^k = \frac{1}{1+t}.$$

Integrating with respect to t,

$$\ln(1+t) = \int \sum_{k=0}^{\infty}(-t)^k\,\mathrm{d}t = \sum_{k=0}^{\infty}\int(-t)^k\,\mathrm{d}t = \sum_{k=0}^{\infty}\frac{-(-t)^{k+1}}{k+1}+C = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}t^k}{k}+C.$$

At $t=0$, we see that $C=0$, so

$$\ln(1+t) = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}t^k}{k}.$$

Part (c). By L'Hôpital's rule,

$$\lim_{x\to\infty} x\mathrm{e}^{-nx} = \lim_{x\to\infty}\frac{x}{\mathrm{e}^{nx}} = \lim_{x\to\infty}\frac{1}{n\mathrm{e}^{nx}} = 0.$$

Part (d). Integrating by parts, we see that

$$\int_0^\infty \frac{x}{\mathrm{e}^x+1}\,\mathrm{d}x = \int_0^\infty \frac{x\mathrm{e}^{-x}}{1+\mathrm{e}^{-x}}\,\mathrm{d}x = \bigl[-x\ln\bigl(1+\mathrm{e}^{-x}\bigr)\bigr]_0^\infty + \int_0^\infty \ln\bigl(1+\mathrm{e}^{-x}\bigr)\,\mathrm{d}x.$$

Notice that

$$\lim_{x\to\infty} -x\ln\bigl(1+\mathrm{e}^{-x}\bigr) = \lim_{x\to\infty} -x\sum_{k=1}^{\infty}\frac{(-1)^{k+1}\mathrm{e}^{-kx}}{k} = \sum_{k=1}^{\infty}\frac{(-1)^k}{k}\lim_{x\to\infty} x\mathrm{e}^{-kx} = 0,$$

where we used part (b) in the first equality, and part (c) in the last equality.

Our target integral is hence just

$$\begin{aligned}\int_0^\infty \frac{x}{\mathrm{e}^x+1}\,\mathrm{d}x &= \int_0^\infty \ln\bigl(1+\mathrm{e}^{-x}\bigr)\,\mathrm{d}x = \int_0^\infty \sum_{k=1}^{\infty}\frac{(-1)^{k+1}\mathrm{e}^{-kx}}{k}\,\mathrm{d}x \\ &= \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\int_0^\infty \mathrm{e}^{-kx}\,\mathrm{d}x = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\left[-\frac{1}{k}\mathrm{e}^{-kx}\right]_0^\infty = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k^2}.\end{aligned}$$