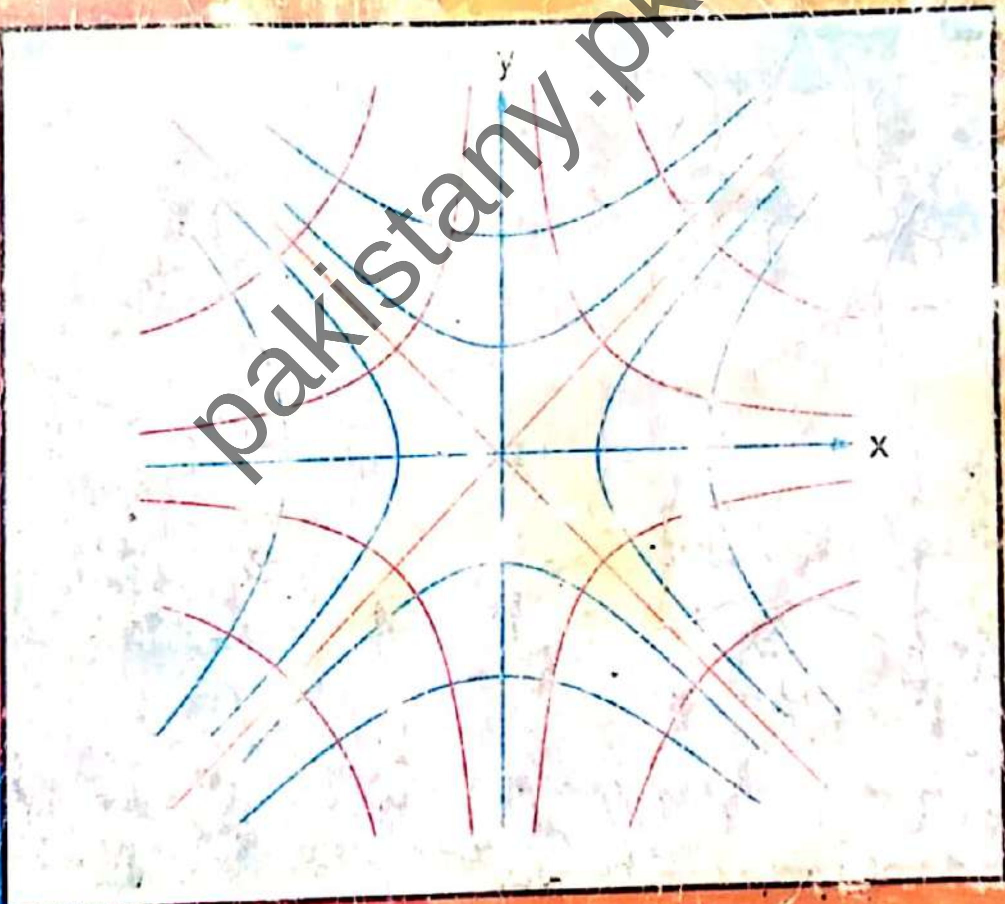
MATHEMATICAL METHODS

S.M. Yusuf
Abdul Majeed
Muhammad Amin



ILMI KITAB KHANA
KABIR STREET, URDU BAZAR, LAHORE - 54000

CONTENTS

Section I
chp# 1, ch# 8
& Vector

CHAPTER 8 INFINITE SERIES 325

CHAPTER 9 FIRST-ORDER DIFFERENTIAL EQUATIONS 377

Chapter 1

TRIGONOMETRY

COMPLEX NUMBERS

The reader is already familiar with complex numbers and a few of their properties. We now define a complex number in a formal manner.

(1.1) Definition. A complex number is an element (x, y) of the set

$$R^2 = \{(x, y) : x, y \in R\}$$

Complex no.s
The no. in the form of
a+ib, a,b∈R

obeying the following rules of addition and multiplication.

For $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2) \in R^2$, we put

A1: $z_2 + z_2 = (x_1 + x_2, y_1 + y_2)$ **(Addition)**

M1: $z_1 z_2 = (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2)$ **(Multiplication)**

Equality of Complex Numbers. Two complex numbers $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$ are equal if and only if $x_1 = x_2$, $y_1 = y_2$.

We then write $z_1 = z_2$.

(1.2) Properties of Complex Numbers. Some important consequences of the definitions of addition (A1) and multiplication (M1) of complex numbers are as follows:

Properties of Addition Defined by (AI):

A2: For all $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$, $z_3 = (x_3, y_3) \in R^2$

$(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$ **(Associative Law of Addition)**

Here, using **(A1)**, we have

$$\begin{aligned}(z_1 + z_2) + z_3 &= (x_1 + x_2, y_1 + y_2) + (x_3, y_3)\\ &= ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3)\\ &= (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3)). \quad \textbf{(Associative Law in } \boldsymbol{R}\textbf{)}\\ &= (x_1, y_1) + (x_2 + x_3, y_2 + y_3)\end{aligned}$$

A3: There is a complex number $\mathit{0} = (0, 0) \in R^2$, called the **additive identity** such that for all $z = (x, y) \in R^2$,

$$\mathit{0} + z = z + \mathit{0} = z \qquad \textbf{(Additive Identity)}$$

A4: For each $z = (x, y) \in R^2$, there is a $-z = (-x, -y) \in R^2$, called the **additive inverse of z**, such that

$$z + (-z) = (0, 0) = \mathit{0} \qquad \textbf{(Additive Inverse)}$$

A5: For all $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in R^2$

$$z_1 + z_2 = z_2 + z_1 \qquad \textbf{(Commutative Law of Addition)}$$

Here

$$\begin{aligned}z_1 + z_2 &= (x_1 + x_2, y_1 + y_2)\\ &= (x_2 + x_1, y_2 + y_1) \qquad \textbf{(Commutative Law in } \boldsymbol{R}\textbf{)}\\ &= z_2 + z_1\end{aligned}$$

The properties **(A1 – A5)** in R^2 show that R^2, under addition, is an abelian group.

Properties of Multiplication Defined by (M1):

M2: For all $z_1 = (x_1, y_1), z_2 = (x_2, y_2), z_3 = (x_3, y_3) \in R^2$

$$(z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3) \qquad \textbf{(Associative Law of Multiplication)}$$

Here

$$\begin{aligned}(z_1 \cdot z_2) \cdot z_3 &= (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2)(x_3, y_3)\\ &= ((x_1x_2 - y_1y_2)x_3 - (x_1y_2 + y_1x_2)y_3,\\ &\qquad (x_1x_2 - y_1y_2)y_3 + (x_1y_2 + y_1x_2)x_3)\\ &= (x_1(x_2x_3 - y_2y_3) - y_1(x_2y_3 + y_2y_3),\\ &\qquad x_1(x_2y_3 + y_2y_3) + y_1(x_2y_3 - y_2x_3))\\ &= (x_1, y_1)(x_2x_3 - y_2y_3, x_2y_3 + y_2x_3)\\ &= z_1 \cdot (z_2 \cdot z_3)\end{aligned}$$

M3: The complex number $\underline{1} = (1, 0) \in R^2$ and it satisfies the condition

$$\underline{1} \cdot z = z \cdot \underline{1} = z \qquad \textbf{(Multiplicative Identity)}$$

for all $z = (x, y) \in R^2$

Here

$$\begin{aligned} \underline{1} \cdot z &= (1, 0)(x, y) \\ &= (1 \cdot x - 0 \cdot y, 1 \cdot y + 0 \cdot x) \\ &= (x, y) \\ &= z \end{aligned}$$

M4: For each nonzero $z = (x, y) \in R^2$, $x \neq 0$ or $y \neq 0$, the complex number

$$z^* = \left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right)$$

is such that

$$\begin{aligned} zz^* &= (x, y)\left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right) = \left(\frac{x^2+y^2}{x^2+y^2}, \frac{-xy}{x^2+y^2} + \frac{xy}{x^2+y^2}\right) \\ &= (1, 0) = \underline{1} \end{aligned}$$

The complex number z^* is called the **multiplicative inverse** of z. It is usually denoted by z^{-1}. Thus

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

has the property that

$$z \cdot z^{-1} = \underline{1}. \qquad \textbf{(Multiplicative Inverse)}$$

for all $z = (x, y) \in R^2$.

M5: For any $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2) \in R^2$

$$\begin{aligned} z_1 \cdot z_2 &= (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2) \\ &= (x_2x_1 - y_2y_1, x_2y_1 + y_2x_1) \\ &= (x_2, y_2) \cdot (x_1, y_1) \\ &= z_2 \cdot z_1 \qquad \textbf{(Commutative Law of Multiplication)} \end{aligned}$$

The conditions (**M1 – M5**) show that the nonzero elements of R^2 form an abelian group under multiplication.

There are two more properties satisfied by the complex numbers. These are called the **distributive laws** and are defined as follows.

For all $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$, $z_3 = (x_3, y_3) \in R^2$.

D1: $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$ **(Left Distributive Law)**

D2: $(z_1 + z_2) \cdot z_3 = z_1 \cdot z_3 + z_2 \cdot z_3$ **(Right Distributive Law)**

Verification of these laws is left as an exercise.

(1.3) Remarks

1. Usually we write $z_1 + (-z_2)$ as $z_1 - z_2$ and $z_1 z_2^{-1}$ as $\frac{z_1}{z_2}$.

2. The complex number $(0, 1) \in R^2$. Also $R^2 : R \to R$

$$\begin{aligned}(0, 1) . (0, 1) &= (0.0 - 1.1,\ 0.1 + 1.0)\\ &= (-1, 0)\\ &= -(1, 0)\\ &= -\underline{1}\end{aligned}$$

Hence

$$\sqrt{-\underline{1}} = (0, 1).$$

The complex number $(0, 1)$ or $\sqrt{-\underline{1}}$ is called **purely imaginary number.** We denote the complex number $(0, 1) = \sqrt{-\underline{1}}$ by i.

Henceforth we shall not distinguish between the complex number $\underline{1}$ and the real number 1, the complex number $\boldsymbol{0}$ and the real number 0. So we write 1 for $(1, 0)$ and 0 for $(0, 0)$.

3. Let $z = (x, y) \in R^2$ and $\alpha \in R$. Then we define αz, called the scalar multiple of z by α, as:

S: $\alpha z = (\alpha x, \alpha y)$ **(Scalar Multiplication)**

It is easy to verify the following properties of scalar multiplication.

For all $z, z_1, z_2 \in R^2$, $\alpha, \beta \in R$.

S1: $\alpha(z_1 + z_2) = \alpha z_1 + \alpha z_2$

S2: $(\alpha + \beta) z = \alpha z + \beta z$

S3: $(\alpha\beta) z = \alpha(\beta z)$

S4: $1 \cdot z = z$, $1 \in R$ is the multiplicative identity.

4. **Algebraic form of a Complex Number:** Algebraic

Let $z = (x, y) \in R^2$. Then

$$\begin{aligned} z &= (x, y)\\ &= (x, 0) + (0, y) && \text{by (A1)}\\ &= x(1, 0) + y(0, 1) && \text{by (S)}\\ &= x \cdot 1 + iy, \quad 1 = (1, 0),\ i = (0, 1) = \sqrt{-1}\\ &= x + iy.\end{aligned}$$

So every complex number $z = (x, y)$ can be written as $x + iy$ and conversely.

The expression $z = x + iy, x, y \in R$ is called the **algebraic (or Cartesian) form** of z.

In $z = x + iy$, x is called the **real part** and y the **imaginary part** of z and we write $\text{Re}\, z = x$, $\text{Im}\, z = y$.

The equations (A1), (M1) and (S) then assume the following forms:

For $z_1 = (x_1, y_1) = x_1 + iy_1$, $z_2 = (x_2, y_2) = x_2 + iy_2$, we have:

$$\begin{aligned}
\textbf{A'1:}\quad z_1 + z_2 &= x_1 + iy_1 + x_2 + iy_2 \\
&= x_1 + x_2 + i(y_1 + y_2), \\
\textbf{M'1:}\quad z_1 \cdot z_2 &= (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2) \\
&= (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2) \\
\textbf{S'1:}\quad \alpha z &= (\alpha x, \alpha y) \\
&= \alpha x + i\alpha y
\end{aligned}$$

The equations (A'1), (M'1) and (S') also define addition, multiplication and scalar multiplication of complex numbers respectively.

Let us denote the set

$$\{x + iy : x, y \in R\}$$

by C. Then the correspondence

$$(x, y) \longleftrightarrow x + iy$$

between the set R^2 and C is obviously one to one correspondence. **Henceforth we use the concept of a complex number interchangeably as the ordered pair (x, y) or $x + iy$ together with the corresponding operations.**

We call C the **complex plane.**

(1.4) Definition. **(Conjugate of a Complex Number)**

Let $z = x + iy$ be a complex number. The complex number $\bar{z} = x - iy = x + i(-y)$ is called the **conjugate of** z.

One can see that the conjugate of a complex number is also a complex number. Some properties of conjugate complex numbers are stated in the following theorem.

(1.5) Theorem. For any complex numbers z, z_1, z_2 where $z = x + iy$, $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, the following hold:

(i) $\bar{\bar{z}} = z$.

(ii) $\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2$

(iii) $\overline{z_1 z_2} = \bar{z}_1 . \bar{z}_2$

(iv) $\overline{\left[\dfrac{z_1}{z_2}\right]} = \dfrac{\bar{z}_1}{\bar{z}_2}$

Proof. Let $z = x + iy$. Then

(i) $\quad \bar{z} = x - iy, \ \bar{\bar{z}} = x - i(-y) = x + iy = z.$

(ii) $$\overline{z_1 + z_2} = \overline{x_1 + x_2 + i(y_1 + y_2)} = x_1 + x_2 - i(y_1 + y_2) = (x_1 - iy_1) + (x_1 - y_2)$$
$$= \bar{z}_1 + \bar{z}_2$$

Similarly for $\overline{z_1 - z_2}$.

(iii) $$\overline{z_1 z_2} = \overline{x_1x_2 - y_1y_2 + i(x_1y_2 + y_1x_2)}$$
$$= x_1x_2 - y_1y_2 - i(x_1y_2 + y_1x_2)$$
$$= (x_1 - iy_1)(x_2 - iy_2) = \bar{z}_1 \cdot \bar{z}_2$$

(iv) $$\overline{\left[\frac{z_1}{z_2}\right]} = \overline{\left[\frac{x_1 + iy_1}{x_2 + iy_2}\right]}$$
$$= \overline{\left[\frac{x_1 + iy_1}{x_2 + iy_2} \cdot \frac{x_2 - iy_2}{x_2 - iy_2}\right]}$$
$$= \overline{\left[\frac{x_1x_2 + y_1y_2 - i(x_1y_2 - y_1x_2)}{x_2^2 + y_2^2}\right]}$$
$$= \frac{x_1x_2 + y_1y_2 + i(x_1y_2 - y_1x_2)}{x_2^2 + y_2^2} \qquad (1)$$

Also $$\frac{\bar{z}_1}{\bar{z}_2} = \frac{x_1 - iy_1}{x_2 - iy_2}$$
$$= \frac{(x_1 - iy_1)(x_2 + iy_2)}{(x_2 - iy_2)(x_2 + iy_2)}$$
$$= \frac{x_1x_2 + y_1y_2 - i(x_1y_2 - y_1x_2)}{x_2^2 + y_2^2} \qquad (2)$$

From (1) and (2), we have

$$\overline{\left[\frac{z_1}{z_2}\right]} = \frac{\bar{z}_1}{\bar{z}_2}$$

THE ARGAND'S DIAGRAM

(1.6) Representation of a Complex Number in Plane: Because of the one-one correspondence between R^2 and the complex plane C, and the corresponding binary operations (A1) and (M1) in R^2 and (A1′), (M1′) in C, there is a functional and useful way to describe a complex number in the complex plane.

For this, the complex number $z = x + iy = (x, y)$ is represented in the complex plane C or R^2 as follows:

The point (x, y) is represented in the Cartesian plane R^2 having coordinates x along the x-axis (called the **real axis**) and y along the y-axis (called the **imaginary axis**). This is shown in Figure 1.1.

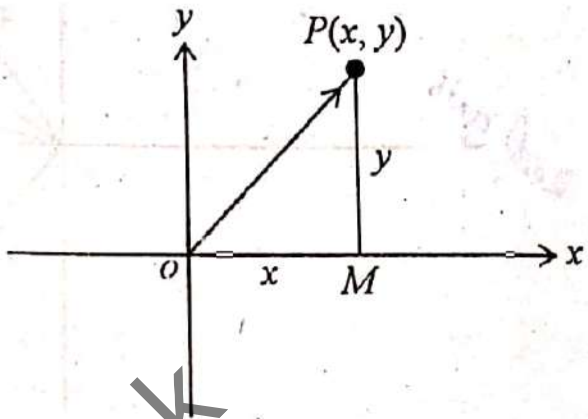


Figure 1.1

Join the origin o to the point P. Then, from the right angled triangle oPM.

$$|\overline{oP}|^2 = |\overline{oM}|^2 + |\overline{MP}|^2 = x^2 + y^2$$

For $z = x + iy$, we write

$$|z| = \sqrt{x^2 + y^2}$$

and call $|z|$ the **modulus** of z.

This idea of representing a complex number by a point in R^2 is due to a French mathematician J.A. Argand (1769 – 1822).

The diagram representing a complex number in this manner is called the **Argand's diagram**. The Argand's diagram helps us understand the relationship between the points z_1, z_2 and $z_1 \pm z_2$. This is explained by the following construction.

(1.7) Representation of $z_1 + z_2 = z$ by Argand's Diagram. Let $z_1 = x_1 + iy_1 = (x_1, y_1)$, $z_2 = x_2 + iy_2 = (x_2, y_2)$ be two complex numbers. Let the points $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ represent these complex numbers by the Argand's diagram as given in Figure 1.2. Draw lines P_2R and P_1R parallel to oP_1 and oP_2 respectively meeting at the point $R(x, y)$. Then the point R represents the complex number $z_1 + z_2$. Here, in Figure 1.2,

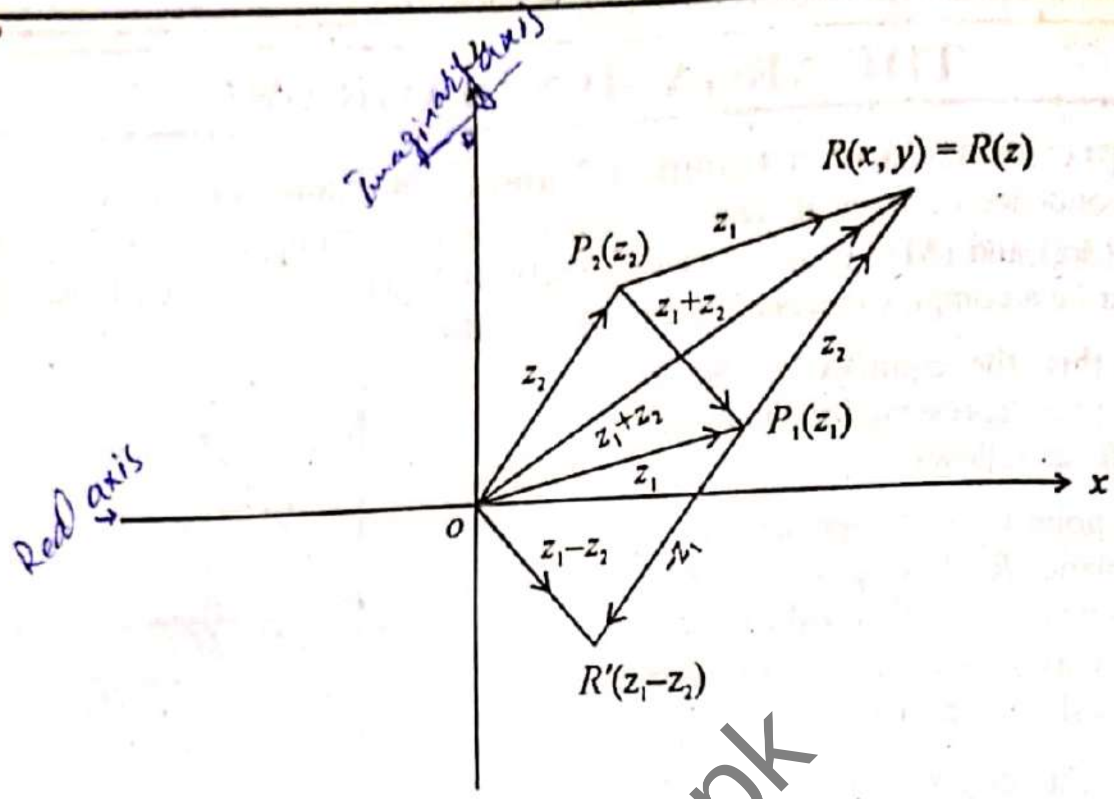


Figure 1.2

the vectors $\overrightarrow{oP_1} = (x_1, y_1)$, $\overrightarrow{oP_2} = (x_2, y_2)$ correspond to z_1 and z_2 respectively whil $\overrightarrow{oR} = (x, y)$ represents the diagonal $\overrightarrow{oR}$ of the parallelogram $oP_1 RP_2$. Hence the vector $\overrightarrow{oR}$ is $(x_1 + x_2, y_1 + y_2)$. So

$$z_1 + z_2 = (x_1 + x_2) + i\,(y_1 + y_2)$$

From the Figure 1.2, we also see that $z_1 - z_2$ is represented by the diagonal P_2P_1 o the point $R'(z_1 - z_2)$. So

$$z_1 - z_2 = (x_1 - x_2) + i\,(y_1 - y_2)$$

(1.8) Trigonometric or Polar Form of a Complex Number.

Let $z = x + iy$ be a complex number. Suppose that the point $P(z)$ represents z in the complex plane by the Argand's diagram.

Let the diagonal oP of the right-angled triangle oPM make an angle θ with the positive direction of the x-axis (real axis). Then

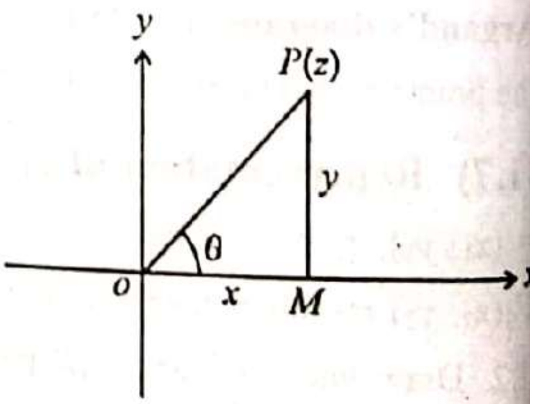


Figure 1.3

$$|\overline{oM}| = x, \quad |\overline{PM}| = y \quad \text{and}$$

$$|\overline{oP}| = \sqrt{x^2 + y^2} = r$$

Also

$$\tan\theta = \frac{y}{x} \quad \text{or} \quad \tan^{-1}(y/x) = \theta.$$

The measure of θ is called the **argument** (or **amplitude**) of z and is denoted by arg z. Then, as stated earlier, r is the modulus of z and is denoted by $|z|$. Note that the argument of the complex number $\boldsymbol{0} = 0 + i0$ is not defined. However $|\boldsymbol{0}| = 0$.

Also, from the triangle oPM,

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

so that

$$x = r\cos\theta, \quad y = r\sin\theta, \quad \theta = \tan^{-1}(y/x) \qquad (1)$$

Hence

$$\boxed{z = x + iy}$$ Complex

$$= r\cos\theta + i\,r\sin\theta$$

$$= r(\cos\theta + i\sin\theta) \qquad (2)$$ ✓(2) Polar form

Thus the point $P(x, y)$ can also be considered as the point $P(r\cos\theta, r\sin\theta)$ or simply $P(r, \theta)$. The pair (r, θ) where r, θ are related to x and y by (1) are called the polar coordinates of z.

We call (2) the **trigonometric or polar** form of the complex number $z = x + iy$.

We also write

$$z = r(\cos\theta + i\sin\theta)$$

as $\quad z = r\,\text{cis}\,\theta$

In the equation

$$\theta = \tan^{-1}\frac{y}{x}, \qquad (3)$$ 3

there are an infinite number of values of θ for which (3) is satisfied. For example if $\theta' = \theta + 2n\pi$, $n \in \boldsymbol{Z}$, then

$$\tan\theta' = \tan(\theta + 2n\pi)$$

$$= \tan\theta = \frac{y}{x}$$

One of the values of θ is such that $-\pi < \theta \leq \pi$. This value of θ is uniquely determined because of the one-one correspondence between the line segment OP and the polar coordinates (r, θ).

This value of θ is called **Principal Argument** of z and is written as Arg z in contrast to the general value arg z.

The angle arg z is usually taken in one of the half intervals

$$(2k-1)\,\pi < \arg z \le (2k+1)\,\pi,\ k = 0, \pm 1, \pm 2, \ldots$$

Thus, for $k = 0$ in the above inequalities,

$$\tan(\text{Arg } z) = \frac{y}{x},\ -\pi < \text{Arg } z \le \pi.$$

Example 1. The complex numbers

(i) $3+4i$, (ii) $-3+4i$, (iii) $-3-4i$ and (iv) $3-4i$

are represented by the Argand's diagrams as follows:

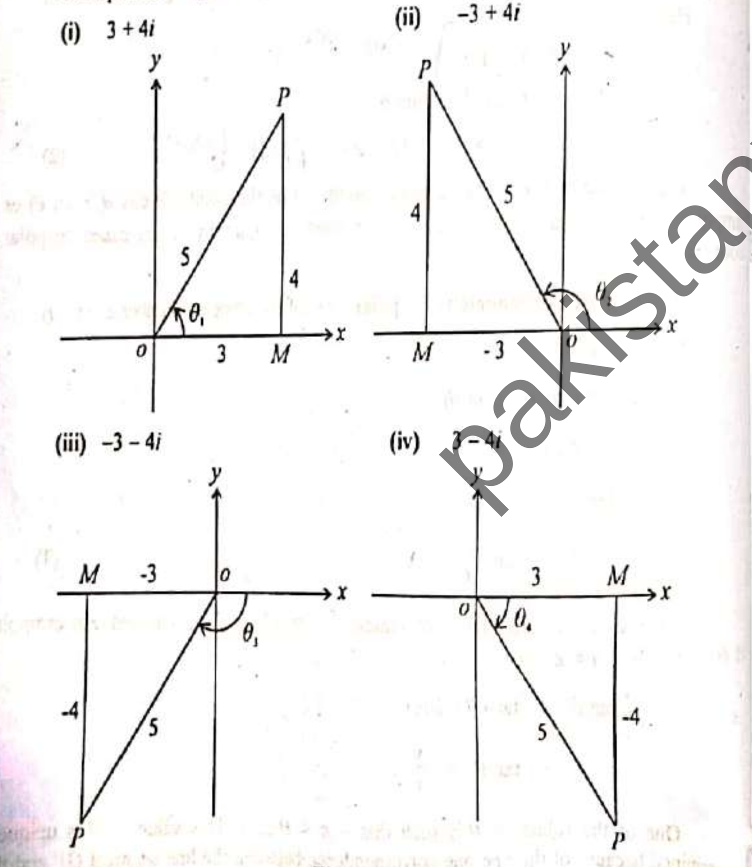


Figure 1.4

Here the modulus of each number is 5.

The polar forms of z_1, z_2, z_3 and z_4 are as follows:

$$z_1 = 5(\cos\theta_1 + i\sin\theta_1)\ ;\ r = 5,\ \text{Arg } z_1 = \theta_1 = 53°\,8'$$
$$z_2 = 5(\cos\theta_2 + i\sin\theta_2)\ ;\ r = 5,\ \text{Arg } z_2 = \theta_2 = 126°\,52'$$
$$z_3 = 5(\cos\theta_3 + i\sin\theta_3)\ ;\ r = 5,\ \text{Arg } z_3 = \theta_3 = -126°\,52'$$
$$z_4 = 5(\cos\theta_4 + i\sin\theta_4)\ ;\ r = 5,\ \text{Arg } z_4 = \theta_4 = -53°\,8'$$

Example 2. Express each of the following complex numbers in the polar form

(i) $1 - i\sqrt{3}$ (ii) 2 (iii) $-5i$ (iv) $-2+2i$

Solution.

(i) For $z_1 = 1 - i\sqrt{3}$,

$$r = \sqrt{1 + \left(-\sqrt{3}\right)^2} = \sqrt{1+3} = 2$$

$$\tan\theta = -\sqrt{3}$$

So $\text{Arg } z_1 = \theta = \tan^{-1}\left(-\sqrt{3}\right) = -\frac{\pi}{3}$

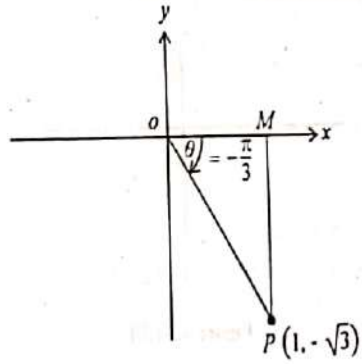


Figure 1.5 (i)

Hence $z_1 = 1 - i\sqrt{3} = 2\left(\cos\left(-\frac{\pi}{3}\right) + i\sin\left(-\frac{\pi}{3}\right)\right)$

Using the properties of trigonometric functions, we have

$$z_1 = 2\left(\cos\frac{\pi}{3} - i\sin\frac{\pi}{3}\right)$$

(ii) For $z_2 = 2$, $r = 2$ and $\theta = \text{Arg } z_2 = \tan^{-1}(0) = 0.$

$$z_2 = 2(\cos 0 + i\sin 0)$$

Its location is shown by Figure 1.5 (ii).

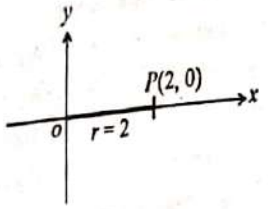


Figure 1.5 (ii)

(iii) For $z_3 = -5i$

$$r = \sqrt{0+5^2} = 5,\ \theta = \text{Arg } z_3 = \tan^{-1}\left(-\frac{5}{0}\right) = \tan^{-1}(-\infty) = -\frac{\pi}{2}$$

Hence $$z_3 = 5\left(\cos\left(-\frac{\pi}{2}\right) + i\sin\left(-\frac{\pi}{2}\right)\right) = 5\left(\cos\frac{\pi}{2} - i\sin\frac{\pi}{2}\right)$$

The point z_3 is represented by Figure 1.5 (iii).

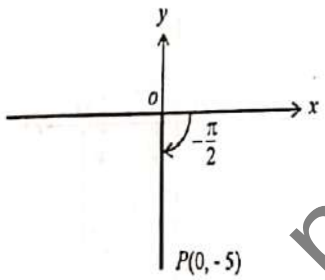


Figure 1.5 (iii)

(iv) For $z_4 = -2 + 2i$

$$|z_4| = \sqrt{4+4} = 2\sqrt{2}$$

$$\theta = \text{Arg } z_4 = \tan^{-1}\left(\frac{y}{x}\right) = \tan^{-1}(-1) = \frac{3\pi}{4}$$

So

$$z_4 = (2\sqrt{2})\left(\cos\frac{3\pi}{4} + i\sin\frac{3\pi}{4}\right)$$

The location of z_4 is shown in Figure 1.5 (iv)

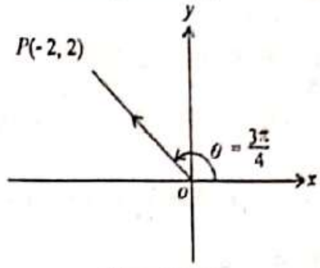


Figure 1.5 (iv)

Example 3. Find z such that $|z| = 2$ and Arg $z = \frac{\pi}{4}$

Solution. Let $z = x + iy$. Then

$$r = |z| = 2,\ \theta = \text{Arg } z = \frac{\pi}{4}$$

$$x = r\cos\theta = 2\cos\frac{\pi}{4} = 2.\frac{1}{\sqrt{2}} = \sqrt{2}$$

$$y = r\sin\theta = 2\sin\frac{\pi}{4} = 2.\frac{1}{\sqrt{2}} = \sqrt{2}$$

Therefore,

$$z = x + yi = \sqrt{2} + \sqrt{2}\,i = \sqrt{2}(1+i)$$

is the required complex number.

(1.9) Product and Quotient of Complex Numbers in Polar Form.

In the previous classes we have already learnt about the sine and cosine formulas for the sum and difference of angles. In this section we use these formulas to find the polar form of the product and quotient of two complex numbers.

Let $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$ be two complex numbers. Then their product is

$$z_1z_2 = r_1r_2(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2)$$

$$= r_1r_2\left[\cos\theta_1\cos\theta_2 + i\sin\theta_1\cos\theta_2 + i\sin\theta_2\cos\theta_1 + i^2\sin\theta_1\sin\theta_2\right]$$

$$= r_1r_2\left[\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2)\right]$$

$$= r_1r_2\left[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\right]$$

Thus $|z_1z_2| = r_1r_2 = |z_1||z_2|$

and $\arg(z_1z_2) = \theta_1 + \theta_2 = \arg z_1 + \arg z_2$

Again, $\dfrac{z_1}{z_2} = \dfrac{r_1(\cos\theta_1 + i\sin\theta_1)}{r_2(\cos\theta_2 + i\sin\theta_2)}$

$$= \frac{r_1}{r_2}\left[\frac{(\cos\theta_1 + i\sin\theta_1)}{(\cos\theta_2 + i\sin\theta_2)} \cdot \frac{(\cos\theta_2 - i\sin\theta_2)}{(\cos\theta_2 - i\sin\theta_2)}\right]$$

$$= \frac{r_1}{r_2}\left[\frac{\cos\theta_1\cos\theta_2 - i\cos\theta_1\sin\theta_2 + i\sin\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2}{\cos^2\theta_2 + \sin^2\theta_2}\right]$$

$$= \frac{r_1}{r_2}\left[\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 + i(\sin\theta_1\cos\theta_2 - \cos\theta_1\sin\theta_2)\right]$$

$$= \frac{r_1}{r_2}\left[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)\right]$$

Hence $\left|\dfrac{z_1}{z_2}\right| = \dfrac{|z_1|}{|z_2|}$,

and $\arg\left(\dfrac{z_1}{z_2}\right) = \theta_1 - \theta_2 = \text{Arg } z_1 - \text{Arg } z_2$.

Example 4. Let z_1, z_2 be two complex numbers. Determine the greatest and least values of $|z_1 + z_2|$.

Solution. We know that, for any two complex numbers z_1 and z_2, represented by $\overrightarrow{oA}$ and $\overrightarrow{oB}$, their sum $z_1 + z_2$ is represented by the diagonal $\overrightarrow{oC}$, of the parallelogram $oACB$.

Now

$$|z_1| = |\overline{oA}|, \quad |z_2| = |\overline{AC}|$$

and

$$|z_1 + z_2| = |\overline{oC}|$$

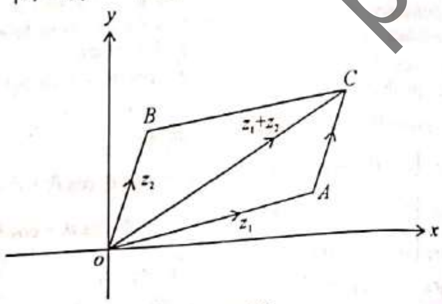


Figure 1.6

In any triangle the sum of the lengths of any two sides is greater than the length of the third side. Thus in triangle oAC

$$|\overline{oA}| + |\overline{AC}| > |\overline{oC}|$$

i.e., $$|z_1| + |z_2| > |z_1 + z_2| \qquad (1)$$

In the special case when $\overrightarrow{oB}$ and $\overrightarrow{oA}$ are themselves parallel (*i.e.*, Arg z_2 = Arg z_1) the parallelogram $oACB$ becomes a straight line with $\overrightarrow{oB}$ coinciding with $\overrightarrow{oA}$. In this case

$$|\overline{oA}| + |\overline{AC}| = |\overline{oC}|$$

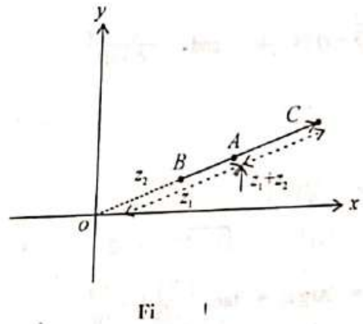
becomes $|\overline{oA}| + |\overline{oB}| = |\overline{oC}|$.

Here $\overrightarrow{oB}$ represents z_2, $\overrightarrow{oA}$ represents z_1, $\overrightarrow{AC}$ also represents z_2. Also $|\overline{oA}| + |\overline{AC}|$ and $|\overline{oC}|$ represent $|z_1 + z_2|$. Thus

$$|\overline{oA}| + |\overline{AC}| = |\overline{oC}|$$

That is

$$|z_1| + |z_2| = |z_1 + z_2| \qquad (2)$$



Fi

Thus, from (1) and (2), we have:

$$|z_1 + z_2| \le |z_1| + |z_2| \qquad (3)$$

i.e., the greatest possible value of $|z_1 + z_2|$ is $|z_1| + |z_2|$.

Returning to the triangle oAC with $\overrightarrow{oB} = \overrightarrow{AC}$ two more inequalities are admissible.

$|\overline{oC}| + |\overline{oA}| > |\overline{AC}| = |\overline{oB}|$

i.e., $|\overline{oC}| + |\overline{oB}| > |\overline{oA}|$ and $|z_1 + z_2| + |z_1| > |z_2|$

or $|z_1 + z_2| + |z_2| > |z_1|$ and $|z_1| - |z_2| > -|z_1 + z_2|$

or $|z_1 + z_2| > |z_1| - |z_2|$

i.e., $-|z_1 + z_2| < |z_1| - |z_2| < |z_1 + z_2|$

These, together with the extreme case when o, A, B and C are collinear, give the result

$$\big|\,|z_1| - |z_2|\,\big| \leq |z_1 + z_2| \tag{4}$$

Thus the least possible value of $|z_1 + z_2|$ is $\big|\,|z_1| - |z_2|\,\big|$.

Combining (3) and (4), we have

$$\big|\,|z_1| - |z_2|\,\big| \leq |z_1 + z_2| \leq |z_1| + |z_2|$$

Example 5. Write down the modulus and argument of

(i) $-\sqrt{3} + i$ (ii) $4 + 4i$.

Hence express the complex numbers

$$(-\sqrt{3} + i)(4 + 4i) \quad \text{and} \quad \frac{-\sqrt{3} + i}{4 + 4i}$$

in polar forms.

Solution.

(i) Here, for $z_1 = -\sqrt{3} + i$

$$r_1 = |z_1| = \sqrt{3 + 1} = 2$$

and $\theta_1 = \text{Arg } z_1 = \tan^{-1}\left(\dfrac{-1}{\sqrt{3}}\right) = \dfrac{5\pi}{6}$

So, polar form of z_1 is

$$z_1 = 2\left(\cos\frac{5\pi}{6} + i\sin\frac{5\pi}{6}\right)$$

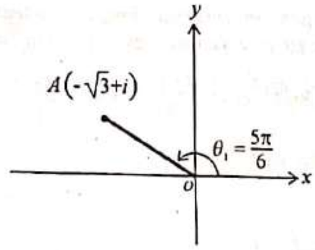


Figure 1.8 (i)

(ii) For $z_2 = 4 + 4i$

$$|z_2| = \sqrt{16 + 16} = 4\sqrt{2}$$

and $\theta_2 = \text{Arg } z_2 = \tan^{-1}\left(\dfrac{4}{4}\right) = \tan^{-1}(1) = \dfrac{\pi}{4}$

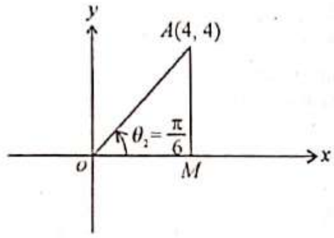


Figure 1.8 (ii)

Hence polar form of z_2 is:

$$z_2 = (4\sqrt{2})\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$$

Now

$$z_1 z_2 = r_1 r_2 (\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2))$$

$$= (8\sqrt{2})\left(\cos\left(\frac{5\pi}{6} + \frac{\pi}{4}\right) + i\sin\left(\frac{5\pi}{6} + \frac{\pi}{4}\right)\right)$$

$$= (8\sqrt{2})\left(\cos\left(\frac{13\pi}{12}\right) + i\sin\left(\frac{13\pi}{12}\right)\right)$$

$$= (8\sqrt{2})\left(\cos\left(\frac{13\pi}{12} - 2\pi\right) + i\sin\left(\frac{13\pi}{12} - 2\pi\right)\right)$$

The value of the expression remains unchanged with the addition or subtraction of 2π. (sine and cosine are periodic functions with period 2π)

$$= (8\sqrt{2})\left(\cos\left(-\frac{11\pi}{12}\right) + i\sin\left(-\frac{11\pi}{12}\right)\right)$$

having $\frac{-11\pi}{12}$ as the Arg $(z_1 z_2)$

Similarly,

$$\frac{z_1}{z_2} = \frac{-\sqrt{3}+i}{4+4i}$$

$$= \frac{r_1}{r_2}[\cos(\theta_1-\theta_2) + i\sin(\theta_1-\theta_2)]$$

$$= \left(\frac{2}{4\sqrt{2}}\right)\left[\cos\left(\frac{5\pi}{6}-\frac{\pi}{4}\right) + i\sin\left(\frac{5\pi}{6}-\frac{\pi}{4}\right)\right]$$

$$= \frac{1}{2\sqrt{2}}\left[\cos\frac{7\pi}{12} + i\sin\frac{7\pi}{12}\right]$$

Note. In case it is difficult to know the exact value of arguments of complex number without using the tables, one just writes

$$\theta_1 = \tan^{-1}\left(\frac{y_1}{x_1}\right), \quad \theta_2 = \tan^{-1}\left(\frac{y_2}{x_2}\right)$$

and $$\theta' = \theta_1 - \theta_2 = \tan^{-1}\left(\frac{y_1}{x_1}\right) - \tan^{-1}\left(\frac{y_2}{x_2}\right)$$

(1.10) Locus of a Complex Number.

Let $P(z)$ be a property satisfied by a complex number $z = x + iy$.

For example a complex number may satisfy the condition $|z| = 2$.

The set

$$S = \{z : |z| = 2\}$$

$$= \left\{(x, y) : |z| = \sqrt{x^2+y^2} = 2\right\}$$

$$= \{(x, y) : x^2 + y^2 = 4\}$$

is called the locus of the complex number z satisfying $|z| = 2$.

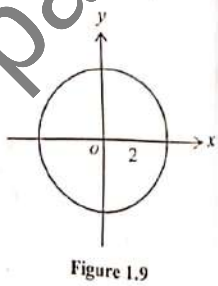


Figure 1.9

Here the locus of z represents a circle with centre at (0, 0) and radius 2.

Example 6. Find the locus of the complex number $z = x + iy$, such that for a fixed point $z_1 = x_1 + iy_1$, $|z - z_1| = a$

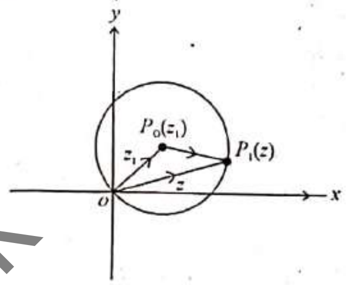


Figure 1.10

Here, from Figure 1.10, we have $\overrightarrow{P_0P_1} = z - z_1$ so that $|\overrightarrow{P_0P_1}| = |z - z_1| = a$. Hence the locus of z satisfying $|z - z_1| = a$ is the set

$$S = \{z : |z - z_1| = a\}$$

$$= \{(x, y) : |(x - x_1) + i(y - y_1)| = a\}$$

$$= \{(x, y) : (x - x_1)^2 + (y - y_1)^2 = a^2\}$$

which is a circle with centre at $z_1 = (x_1, y_1)$ and radius a.

Example 7. Let $z_1 = x_1 + iy_1$ be a fixed complex number. Find the locus of all $z = x + iy$ such that $\text{Arg}(z - z_1) = \frac{\pi}{4}$

Solution. Here $z - z_1 = (x - x_1) + i(y - y_1)$

So $$\theta = \text{Arg}(z - z_1) = \tan^{-1}\left(\frac{y-y_1}{x-x_1}\right) = \frac{\pi}{4}$$

But $$1 = \tan\frac{\pi}{4} = \frac{y-y_1}{x-x_1}$$

So $$x - x_1 = y - y_1 \quad \text{or} \quad x - y = x_1 - y_1$$

Thus the locus of z is the set of points on the half line $\overrightarrow{AP}$ which makes an angle of $\frac{\pi}{4}$ with the line $\overrightarrow{oA}$.

The locus is shown in Figure 1.11.

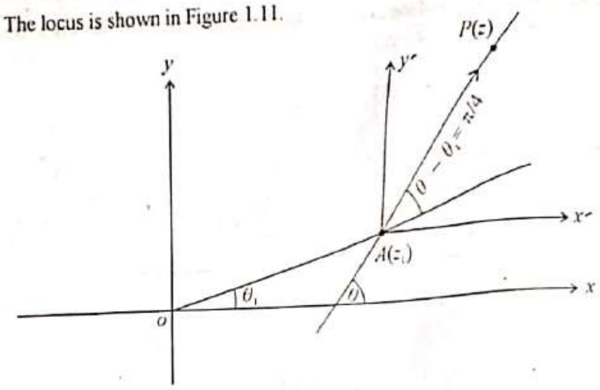


Figure 1.11

EXERCISE 1.1

Express each of the following complex numbers in the polar form (**Problems 1 – 6**):

1. $-\sqrt{3}+i$
2. $-i$
3. $-1-\sqrt{3}\,i$
4. $-1+i$
5. $(-2+2i)(1-i)$
6. $\dfrac{-34\,i}{5-3i}$

Express the given complex number in Cartesian form and plot on an Argand diagram (**Problems 7 – 10**):

7. $2\,\text{cis}\left(\dfrac{\pi}{6}\right)$
8. $5\,\text{cis}\left(\dfrac{3\pi}{4}\right)$
9. $\sqrt{3}\,\text{cis}\left(\dfrac{7\pi}{6}\right)$
10. $\dfrac{5\,\text{cis}\,(\pi/3)}{5\,\text{cis}\,(\pi/2)}$

11. Find $|z|$, where

 (i) $z=-2i\,(1+i)(2+4i)(3+i)$ (ii) $z=\dfrac{(3+4i)(-1+2i)}{(-1-i)(3-i)}$

12. Show that $z=a+ib$ is

 (i) real if and only if $z=\bar{z}$

 (ii) pure imaginary if and only if $z=-\bar{z}$.

13. Prove analytically that for complex numbers z_1, z_2

$$\big|\,|z_1|-|z_2|\,\big| \le |z_1+z_2| \le |z_1|+|z_2|$$

14. Let $z_1 = 24+7i$ and $|z_2| = 6$. Find the greatest and least values of $|z_1+z_2|$.

15. If z_1, z_2 are complex numbers, show that

$$|z_1+z_2|^2+|z_1-z_2|^2 = 2(|z_1|^2+|z_2|^2)$$

16. Prove that $\left|\dfrac{az+b}{\bar{b}z+\bar{a}}\right| = 1$ for $|z| = 1$.

17. Find the locus of points in the plane satisfying each of the given conditions:

 (i) $|z-5| = 6$ (ii) $|z-2i| \ge 1$

 (iii) $\text{Re}\,(z+2) = -1$ (iv) $\text{Re}\,(i\,\bar{z}) = 3$

 (v) $|z+i| = |z-i|$ (vi) $|z+3|+|z+1| = 4$

 (vii) $-1 \le \text{Re}\,z < 1$ (viii) $\text{Im}\,z < 0$

 (ix) $\text{Arg}\,z = \dfrac{\pi}{3}$ (x) $\text{Arg}\,(z-1) = -\dfrac{3\pi}{4}$

DE MOIVRE'S THEOREM

(1.11) Definition. As in the case of real numbers, integral powers of a complex number z are defined as follows:

$$z^0 = 1$$

$$\left.\begin{aligned} z^{m+1} &= z^m \cdot z \\ z^{-m} &= (z^{-1})^m \end{aligned}\right\}, \quad \text{where } z \ne 0 \text{ and } m \in Z^+ \text{ (the set of positive integers)}$$

The following laws of exponents can be proved in a manner similar to that for real numbers (for $m, n \in Z$, $z, z_1, z_2 \in C$):

 (i) $z^m z^n = z^{m+n}$

 (ii) $(z^m)^n = z^{mn}$

 (iii) $(z_1 z_2)^n = z_1^n z_2^n$.

The following theorem is useful for finding integral powers of a complex number in polar form.

(1.12) (De Moivre's[1] Theorem

Let
$$z = \cos\theta + i\sin\theta. \text{ Then}$$
$$z^n = (\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta,$$

for all integers n.

Proof. When $n = 0$, we have
$$(\cos\theta + i\sin\theta)^0 = 1 = \cos 0 + i\sin 0$$
$$= \cos 0\theta + i\sin 0\theta$$

When $n = 1$,
$$(\cos\theta + i\sin\theta)^1 = \cos\theta + i\sin\theta$$
$$= \cos 1\theta + i\sin 1\theta$$

The result is true for $n = 0$ and $n = 1$.

Now suppose the result is true for $n = k$, i.e.,
$$(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta$$

Then
$$\begin{aligned}(\cos\theta + i\sin\theta)^{k+1} &= (\cos\theta + i\sin\theta)^k(\cos\theta + i\sin\theta)\\ &= (\cos k\theta + i\sin k\theta)(\cos\theta + i\sin\theta)\\ &= (\cos k\theta\cos\theta - \sin k\theta\sin\theta)\\ &\quad + i(\sin k\theta\cos\theta + \cos k\theta\sin\theta)\\ &= \cos(k\theta + \theta) + i\sin(k\theta + \theta)\\ &= \cos(k+1)\theta + i\sin(k+1)\theta.\end{aligned}$$

Thus, the truth of the statement for $n = k$ implies its truth for $n = k + 1$. Hence, by the Principle of Mathematical Induction[2],
$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta, \text{ for all } n \in Z^+.$$

Now suppose n is a negative integer, say $n = -m$, where $m \in Z^+$.

Then
$$\begin{aligned}(\cos\theta + i\sin\theta)^n &= (\cos\theta + i\sin\theta)^{-m}\\ &= \left[(\cos\theta + i\sin\theta)^{-1}\right]^m\\ &= \left[\frac{1}{\cos\theta + i\sin\theta}\right]^m\end{aligned}$$

1. Named after the French mathematician Abraham De Moivre (1667 – 1754).
2. The principle of mathematical induction states as follows:
   Let $P(n)$ be a property about natural number n. Suppose that:
   (i) $P(1)$ is true
   (ii) For a natural number $k > 1$, $P(k)$ is true implies $P(k+1)$ is true. Then $P(n)$ is true for all n.

$$\begin{aligned}&= \frac{1}{(\cos\theta + i\sin\theta)^m}\\ &= \frac{1}{\cos m\theta + i\sin m\theta}, \text{ by the result already proved}\\ &= \frac{1}{\cos m\theta + i\sin m\theta} \times \frac{\cos m\theta - i\sin m\theta}{\cos m\theta - i\sin m\theta}\\ &= \frac{\cos m\theta - i\sin m\theta}{\cos^2 m\theta + \sin^2 m\theta}\\ &= \cos m\theta - i\sin m\theta\\ &= \cos(-m)\theta + i\sin(-m)\theta\\ &= \cos n\theta + i\sin n\theta\end{aligned}$$

Thus the theorem is true for all $n \in Z$.

(1.13) Corollary.

For
$$z = \cos\theta + i\sin\theta,$$
$$\frac{1}{z} = \frac{1}{\cos\theta + i\sin\theta} = \cos\theta - i\sin\theta$$

and
$$(\cos\theta - i\sin\theta)^n = \cos n\theta - i\sin n\theta$$

Proof. Here,
$$\cos\theta - i\sin\theta = \cos(-\theta) + i\sin(-\theta)$$

So, for all $n \in Z$, and using $\cos(-\theta) = \cos\theta$, $\sin(-\theta) = -\sin\theta$, we have
$$\begin{aligned}(\cos\theta - i\sin\theta)^n &= \left[\cos(-\theta) + i\sin(-\theta)\right]^n\\ &= \cos(-n\theta) + i\sin(-n\theta), \text{ by } \mathbf{(1.12)},\\ &= \cos n\theta - i\sin n\theta,\end{aligned}$$

as required.

Example 8. Evaluate
$$\left(\frac{\sqrt{3} - i}{\sqrt{3} + i}\right)^6$$

Solution. Here
$$z_1 = \sqrt{3} - i$$

So
$$\text{Arg } z_1 = \tan^{-1}\left(-\frac{1}{\sqrt{3}}\right)$$
$$= -\frac{\pi}{6}, \quad \text{as } \tan\frac{1}{\sqrt{3}} = \frac{\pi}{6}$$

Also $\qquad |z_1| = \sqrt{3+1} = 2$

So
$$z_1 = \sqrt{3} - i$$
$$= 2\left[\cos\left(-\frac{\pi}{6}\right) + i\sin\left(-\frac{\pi}{6}\right)\right]$$

Thus
$$z_1^6 = (\sqrt{3}-1)^6$$
$$= 2^6\left[\cos\left(-\frac{\pi}{6}\right) + i\sin\left(-\frac{\pi}{6}\right)\right]^6$$
$$= 2^6\,[\cos(-\pi) + i\sin(-\pi)], \text{ by De Moivre's Theorem}$$
$$= -2^6, \qquad \text{as } \cos(-\pi) = -1,\ \sin(-\pi) = 0$$

Next
$$z_2 = \frac{1}{\sqrt{3}+i} = \frac{1}{(\sqrt{3}+i)}\frac{(\sqrt{3}-i)}{(\sqrt{3}-i)}$$
$$= \frac{1}{4}(\sqrt{3}-i)$$

So
$$z_2^6 = \frac{1}{4^6}(\sqrt{3}-i)^6 = \frac{1}{4^6}(-2^6) = \frac{1}{2^{12}}(-2^6) = -\frac{1}{2^6}$$

Hence
$$\left(\frac{\sqrt{3}-i}{\sqrt{3}+i}\right)^6 = (z_1 z_2)^6 = \frac{-2^6}{-2^6} = 1.$$

Example 9. Prove that
$$(\sin x + i\cos x)^n = \cos n\left(\frac{\pi}{2}-x\right) + i\sin n\left(\frac{\pi}{2}-x\right),\ n \in Z$$

Solution. From Trigonometry, we have
$$\sin x = \cos\left(\frac{\pi}{2}-x\right) \quad \text{and} \quad \cos x = \sin\left(\frac{\pi}{2}-x\right).$$

Hence
$$(\sin x + i\cos x)^n = \left[\cos\left(\frac{\pi}{2}-x\right) + i\sin\left(\frac{\pi}{2}-x\right)\right]^n$$
$$= \cos n\left(\frac{\pi}{2}-x\right) + i\sin n\left(\frac{\pi}{2}-x\right), \text{ by } \mathbf{(1.12)}$$

Example 10. If $x = \cos\theta + i\sin\theta$, show that
$$x^n + \frac{1}{x^n} = 2\cos n\theta \quad \text{and} \quad x^n - \frac{1}{x^n} = 2i\sin n\theta$$

Solution. Here, $\quad x = \cos\theta + i\sin\theta \quad$ Therefore,
$$x^n = (\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta, \qquad (1)$$
by **(1.12)**

Next
$$\frac{1}{x} = \frac{1}{\cos\theta + i\sin\theta} = \cos\theta - i\sin\theta$$
$$\frac{1}{x^n} = (\cos\theta - i\sin\theta)^n = [\cos(-\theta) + i\sin(-\theta)]^n$$
$$= \cos(-n\theta) + i\sin(-n\theta)$$
$$= \cos n\theta - i\sin n\theta \qquad (2)$$

From (1) and (2), we obtain
$$x^n + \frac{1}{x^n} = \cos n\theta + i\sin n\theta + \cos n\theta - i\sin n\theta = 2\cos n\theta$$
and
$$x^n - \frac{1}{x^n} = (\cos n\theta + i\sin n\theta) - (\cos n\theta - i\sin n\theta) = 2i\sin n\theta$$

(1.14) Applications of De Movre's Theorem

I. **To express $\cos n\theta$ and $\sin n\theta$ as finite sums of powers of trigonometric function of θ, where n is a positive integer.**

The Binomial Theorem holds for the set of complex numbers. So we have
$$\cos n\theta + i\sin n\theta = (\cos\theta + i\sin\theta)^n$$
$$= \cos^n\theta + \binom{n}{1}\cos^{n-1}\theta\,(i\sin\theta) + \binom{n}{2}\cos^{n-2}\theta\,(i\sin\theta)^2$$
$$+ \binom{n}{3}\cos^{n-3}\theta\,(i\sin\theta)^3 + \binom{n}{4}\cos^{n-4}\theta\,(i\sin\theta)^4$$
$$+ \binom{n}{5}\cos^{n-5}\theta\,(i\sin\theta)^5 + \cdots + \binom{n}{n-1}\cos\theta\,(i\sin\theta)^{n-1} + (i\sin\theta)^n$$

$$= \cos^n\theta + i\binom{n}{1}\cos^{n-1}\theta\sin\theta - \binom{n}{2}\cos^{n-2}\theta\sin^2\theta$$
$$-i\binom{n}{3}\cos^{n-3}\theta\sin^3\theta + \binom{n}{4}\cos^{n-4}\theta\sin^4\theta$$
$$+i\binom{n}{5}\cos^{n-5}\theta\sin^5\theta - \dots + i^{n-1}n\cos\theta\sin^{n-1}\theta + i^n\sin^n\theta \qquad (1)$$

Two cases arise according as n is even or odd.

(i) If $n = 2m$ is even, then $n-2$ is also even and $\frac{n}{2}$, $\frac{n-2}{2}$ are nonnegative integers.

Thus (1) may be written as

$$\cos n\theta + i\sin n\theta = \cos^n\theta + i\binom{n}{1}\cos^{n-1}\theta\sin\theta - \binom{n}{2}\cos^{n-2}\theta\sin^2\theta$$
$$-i\binom{n}{3}\cos^{n-3}\theta\sin^3\theta + \binom{n}{4}\cos^{n-4}\theta\sin^4\theta$$
$$+i\binom{n}{5}\cos^{n-5}\theta\sin^5\theta + \dots + i(-1)^{\frac{n-2}{2}}n\cos\theta\sin^{n-1}\theta$$
$$+(-1)^{\frac{n}{2}}\sin^n\theta.$$

Equating real and imaginary parts on both sides of the above equation, we get

$$\cos n\theta = \cos^n\theta - \binom{n}{2}\cos^{n-2}\theta\sin^2\theta + \binom{n}{4}\cos^{n-4}\theta\sin^4\theta + \dots + (-1)^{\frac{n}{2}}\sin^n\theta$$

$$\sin n\theta = \binom{n}{1}\cos^{n-1}\theta\sin\theta - \binom{n}{3}\cos^{n-3}\theta\sin^3\theta + \binom{n}{5}\cos^{n-5}\theta\sin^5\theta$$
$$+ \dots + (-1)^{\frac{n-2}{2}}n\cos\theta\sin^{n-1}\theta.$$

(ii) If $n = 2m+1$ is odd, then $n-1$ is even and so $\frac{n-1}{2}$ is a nonnegative integer.

Rewrite (1) as

$$\cos n\theta + i\sin n\theta = \cos^n\theta + i\binom{n}{1}\cos^{n-1}\theta\sin\theta$$
$$-\binom{n}{2}\cos^{n-2}\theta\sin^2\theta - i\binom{n}{3}\cos^{n-3}\theta\sin^3\theta$$
$$+\binom{n}{4}\cos^{n-4}\theta\sin^4\theta + i\binom{n}{5}\cos^{n-5}\theta\sin^5\theta$$
$$+\dots + (-1)^{\frac{n-1}{2}}n\cos\theta\sin^{n-1}\theta + i(-1)^{\frac{n-1}{2}}\sin^n\theta.$$

Equating real and imaginary parts on both sides of he above equation, we get

$$\cos n\theta = \cos^n\theta - \binom{n}{2}\cos^{n-2}\theta\sin^2\theta + \binom{n}{4}\cos^{n-4}\theta\sin^4\theta - \dots$$
$$+(-1)^{\frac{n-1}{2}}n\cos\theta\sin^{n-1}\theta$$

$$\sin n\theta = \binom{n}{1}\cos^{n-1}\theta\sin\theta - \binom{n}{3}\cos^{n-3}\theta\sin^3\theta + \binom{n}{5}\cos^{n-5}\theta\sin^5\theta -$$
$$\dots + (-1)^{\frac{n-1}{2}}\sin^n\theta.$$

II. To express powers of $\cos\theta$ (or $\sin\theta$) in a series of cosines (or sines) of multiples of θ.

(i) Let $x = \cos\theta + i\sin\theta$

Then $\frac{1}{x} = \cos\theta - i\sin\theta$

and $x + \frac{1}{x} = 2\cos\theta,\quad x - \frac{1}{x} = 2i\sin\theta$

Now $(2\cos\theta)^n = \left(x + \frac{1}{x}\right)^n$

$$= x^n + nx^{n-1}\frac{1}{x} + \frac{n(n-1)}{2!}x^{n-2}\cdot\frac{1}{x^2} + \dots + \frac{n(n-1)}{2!}x^2\cdot\frac{1}{x^{n-2}} + nx\cdot\frac{1}{x^{n-1}} + \frac{1}{x^n}$$

(by the Binomial Theorem)

$$= x^n + nx^{n-2} + \frac{n(n-1)}{2!}x^{n-4} + \dots + \frac{n(n-1)}{2!}\cdot\frac{1}{x^{n-4}} + n\cdot\frac{1}{x^{n-2}} + \frac{1}{x^n}$$

$$= \left(x^n + \frac{1}{x^n}\right) + n\left(x^{n-2} + \frac{1}{x^{n-2}}\right) + \frac{n(n-1)}{2!}\left(x^{n-4} + \frac{1}{x^{n-4}}\right) + \dots$$

sum up to $\frac{n}{2} + 1$ or $\frac{n+1}{2}$ terms according as n is even or odd.

$$= 2\cos n\theta + n\,[2\cos(n-2)\,\theta] + \frac{n(n-1)}{2!}[2\cos(n-4)\,\theta] + \dots$$

Thus

$$\cos^n\theta = \frac{1}{2^{n-1}}\left[\cos n\theta + n\cos(n-2)\,\theta + \frac{n(n-1)}{2!}\cos(n-4)\,\theta + \dots\right]$$

(ii) To find a series for $\sin^n \theta$, we proceed with

$$(2i \sin \theta)^n = \left(x - \frac{1}{x}\right)^n$$

If $n = 2m$, is even, there are $2m + 1$ terms on the R.H.S, the middle one being $\left(\frac{n}{2}+1\right)$ th term. So

$$(2i \sin \theta)^n = x^n - nx^{n-2} + \frac{n(n-1)}{2!}x^{n-4} + \dots + \frac{n(n-1)}{2!}\cdot\frac{1}{x^{n-4}} - n\cdot\frac{1}{x^{n-2}} + \frac{1}{x^n}$$

or

$$2^n (-1)^{\frac{n}{2}} \sin^n \theta = \left(x^n + \frac{1}{x^n}\right) - n\left(x^{n-2} + \frac{1}{x^{n-2}}\right) + \frac{n(n-1)}{2!}\left(x^{n-4} + \frac{1}{x^{n-4}}\right) - \cdots \text{ to } \frac{n}{2} \text{ terms}$$

Thus,

$$\sin^n \theta = (-1)^{\frac{n}{2}} \frac{1}{2^{n-1}}\left[\cos n\theta - n\cos(n-2)\theta + \frac{n(n-1)}{2!}\cos(n-4)\theta - \cdots\right]$$

If $n = 2m + 1$ is odd, there are $2m + 2$ terms, with two middle terms namely $(m + 1)$ and $(m + 2)$th terms.

These are $\binom{n}{\frac{n-1}{2}}x$ and $-\binom{n}{\frac{n+1}{2}}\frac{1}{x}$, with $\binom{n}{\frac{n-1}{2}} = \binom{n}{\frac{n+1}{2}}$

So

$$(2i \sin \theta)^n = x^n - nx^{n-2} + \frac{n(n-1)}{2}x^{n-4} + \dots + \frac{n(n-1)}{2!}\frac{1}{x^{n-4}} - n\frac{1}{x^{n-2}} - \frac{1}{x^n}$$

$$= \left(x^n - \frac{1}{x^n}\right) - n\left(x^{n-2} - \frac{1}{x^{n-2}}\right) + \frac{n(n-1)}{2!}\left(x^{n-4} - \frac{1}{x^{n-4}}\right) - \cdots \text{ to } \frac{n+1}{2} \text{ terms.}$$

Hence

$$2^n(-1)^{\frac{n-1}{2}} i \sin^n \theta = 2i \sin n\theta - n[2i \sin(n-2)\theta] + \frac{n(n-1)}{2!}[2i \sin(n-4)\theta] + \cdots$$

or

$$\sin^n \theta = (-1)^{\frac{n-1}{2}} \frac{1}{2^{n-1}}\left[\sin n\theta - n\sin(n-2)\theta + \frac{n(n-1)}{2!}\sin(n-4)\theta - \cdots\right]$$

Example 11. Prove that

(i) $\cos 5\theta = 16\cos^5\theta - 20\cos^3\theta + 5\cos\theta$

(ii) $\sin 5\theta = 16\sin^5\theta - 20\sin^3\theta + 5\sin\theta$

(iii) $\tan 5\theta = \dfrac{5\tan\theta - 10\tan^3\theta + \tan^5\theta}{1 - 10\tan^2\theta + 5\tan^4\theta}$

Solution.

$$\cos 5\theta + i\sin 5\theta = (\cos\theta + i\sin\theta)^5$$

$$= \cos^5\theta + \binom{5}{1}\cos^4\theta\,(i\sin\theta) + \binom{5}{2}\cos^3\theta\,(i\sin\theta)^2 + \binom{5}{3}\cos^2\theta\,(i\sin\theta)^3 + \binom{5}{4}\cos\theta\,(i\sin\theta)^4 + (i\sin\theta)^5$$

$$= \cos^5\theta + 5i\cos^4\theta\sin\theta + 10\cos^3\theta(-\sin^2\theta) + 10\cos^2\theta(-i\sin^3\theta) + 5\cos\theta(\sin^4\theta) + i\sin^5\theta$$

$$= \cos^5\theta - 10\cos^3\theta\sin^2\theta + 5\cos\theta\sin^4\theta + i(5\cos^4\theta\sin\theta - 10\cos^2\theta\sin^3\theta + \sin^5\theta) \quad (1)$$

Equating real and imaginary parts in (1), we have

(i)
$$\cos 5\theta = \cos^5\theta - 10\cos^3\theta\sin^2\theta + 5\cos\theta\sin^4\theta \quad (2)$$
$$= \cos^5\theta - 10\cos^3\theta(1-\cos^2\theta) + 5\cos\theta(1-\cos^2\theta)^2$$
$$= 16\cos^5\theta - 20\cos^3\theta + 5\cos\theta$$

(ii)
$$\sin 5\theta = 5\cos^4\theta\sin\theta - 10\cos^2\theta\sin^3\theta + \sin^5\theta \quad (3)$$
$$= 5(1-\sin^2\theta)^2\sin\theta - 10(1-\sin^2\theta)\sin^3\theta + \sin^5\theta$$
$$= 16\sin^5\theta - 20\sin^3\theta + 5\sin\theta$$

(iii) Dividing (3) by (2), we have

$$\tan 5\theta = \frac{\sin 5\theta}{\cos 5\theta} = \frac{5\cos^4\theta\sin\theta - 10\cos^2\theta\sin^3\theta + \sin^5\theta}{\cos^5\theta - 10\cos^3\theta\sin^2\theta + 5\cos\theta\sin^4\theta} \quad (4)$$

Dividing each term of the numerator and denominator on the right of (4) by $\cos^5\theta$, we get the result in the required form.

Example 12. If $x = \cos\theta + i\sin\theta$, express $\cos^5\theta\,\sin^3\theta$ in a series of multiples of θ.

Solution. From Example 10, we have

$$x^n + \frac{1}{x^n} = (2\cos n\theta), \qquad x^n - \frac{1}{x^n} = (2i\sin n\theta) \quad (1)$$

Since $x = \cos\theta + i\sin\theta$, $\quad \dfrac{1}{x} = \cos\theta - i\sin\theta$, we have

$$x+\frac{1}{x} = 2\cos\theta, \qquad x-\frac{1}{x} = 2i\sin\theta.$$

Now

$$(2\cos\theta)^5 = \left(x+\frac{1}{x}\right)^5$$

$$= x^5+5x^3+10x+\frac{10}{x}+\frac{5}{x^3}+\frac{1}{x^5}$$

$$= \left(x^5+\frac{1}{x^5}\right)+5\left(x^3+\frac{1}{x^3}\right)+10\left(x+\frac{1}{x}\right)$$

and $\quad (2i\sin\theta)^3 = \left(x-\frac{1}{x}\right)^3$

$$= x^3-3x+\frac{3}{x}-\frac{1}{x^3}$$

So

$$(2\cos\theta)^5 . (2i\sin\theta)^3 = \left(x+\frac{1}{x}\right)^5\left(x-\frac{1}{x}\right)^3 \qquad (2)$$

$$=\left(x^5+5x^3+10x+\frac{10}{x}+\frac{5}{x^3}+\frac{1}{x^5}\right)\left(x^3-3x+\frac{3}{x}-\frac{1}{x^3}\right)$$

$$=\left(x^8+5x^6+10x^4+10x^2+5+\frac{1}{x^2}\right)-\left(3x^6+15x^4+30x^2+30+\frac{15}{x^2}+\frac{3}{x^4}\right)$$

$$+\left(3x^4+15x^2+30+\frac{30}{x^2}+\frac{15}{x^4}+\frac{3}{x^6}\right)-\left(x^2+5+\frac{10}{x^2}+\frac{10}{x^4}+\frac{5}{x^6}+\frac{1}{x^8}\right)$$

$$=\left(x^8-\frac{1}{x^8}\right)+2\left(x^6-\frac{1}{x^6}\right)-2\left(x^4-\frac{1}{x^4}\right)-6\left(x^2-\frac{1}{x^2}\right)$$

$$= (2i\sin 8\theta)+2(2i\sin 6\theta)-2(2i\sin 4\theta)-6(2i\sin 2\theta), \text{ by (1)}$$

$$= 2i(\sin 8\theta+2\sin 6\theta-2\sin 4\theta-6\sin 2\theta) \qquad (3)$$

Now L.H.S. of (2) is:

$$(2i)^3 . 2^5 . \cos^5\theta . \sin^3\theta = -2i . 2^7\cos^5\theta . \sin^3\theta \qquad (4)$$

From (3) and (4), we have

$$\cos^2\theta . \sin^3\theta = -\frac{1}{2^7}(\sin 8\theta+2\sin 6\theta-2\sin 4\theta-6\sin 2\theta)$$

ROOTS OF A COMPLEX NUMBER

(1.15) nth Roots of a Complex Number.

Let $z = r(\cos\theta + i\sin\theta)$ and $w = \rho(\cos\alpha + i\sin\alpha)$ be two complex numbers in polar form. We call w **an nth root of** z if

$$w^n = z$$

If w is an nth root of z then (1)

$$w^n = \rho^n(\cos\alpha + i\sin\alpha)^n$$

$$= \rho^n(\cos n\alpha + i\sin n\alpha)$$

$$= r(\cos\theta + i\sin\theta) \qquad \text{from (1)}$$

Comparing the last two expressions in the equation, we have

$$\rho^n = r, \quad \cos n\alpha = \cos\theta, \quad \sin n\alpha = \sin\theta.$$

Since $\cos(\theta + 2k\pi) = \cos\theta$, $\sin(\theta + 2k\pi) = \sin\theta$, we note that the arguments θ and α are related by the equation

$$\alpha = \frac{\theta + 2k\pi}{n} \qquad (2)$$

Moreover, we take $\rho = r^{1/n}$ as the unique positive nth root of the real number r.

Thus n distinct nth roots of z are:

$$w_k = r^{1/n}\left(\cos\left(\frac{\theta+2k\pi}{n}\right)+i\sin\left(\frac{\theta+2k\pi}{n}\right)\right), \qquad (3)$$

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

Among the nth roots of z there is one for which arg $(z^{1/n})$ is the principal value. This value corresponds to the value $k = 0$. The corresponding nth root is called **the principal nth root**.

Example 13. Find the n nth roots of unity.

Solution. Here we first write 1 as:

$$1 = 1 + 0i = \cos 2k\pi + i\sin 2k\pi.$$

So

$$1^{1/n} = (\cos 2k\pi + i\sin 2k\pi)^{1/n}$$

$$= \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n} \qquad \text{by Theorem 1.12}$$

Hence the n nth roots of unity are given by

$$\cos\frac{2k\pi}{n}+i\sin\frac{2k\pi}{n},\quad k=0,1,2,\ldots,n-1$$

or by

$$\left(\cos\frac{2\pi}{n}+i\sin\frac{2\pi}{n}\right)^k,\quad k=0,1,2,\ldots,n-1$$

If we let $\omega=\cos\frac{2\pi}{n}+i\sin\frac{2\pi}{n}$, then the n, n th roots of 1 are

$$1,\ \omega,\ \omega^2,\ldots,\ \omega^{n-1}.$$

Example 14. Solve the equation $x^4-x^3+x^2-x+1=0$.

Solution. We have $(x^5+1)=(x+1)(x^4-x^3+x^2-x+1)$

The roots of $x^5+1=0$, other than $x=-1$, are the roots of the given equation.

Now $x^5=-1=\cos(2k\pi+\pi)+i\sin(2k\pi+\pi)$, $k\in Z$.

Thus the five 5th roots of -1 are given by

$$x_k=\cos\frac{2k\pi+\pi}{5}+i\sin\frac{2k\pi+\pi}{5},\ k=0,1,2,3,4$$

That is,

$$\cos\frac{\pi}{5}+i\sin\frac{\pi}{5},\quad \cos\frac{3\pi}{5}+i\sin\frac{3\pi}{5},\quad \cos\pi+i\sin\pi,$$

$$\cos\frac{7\pi}{5}+i\sin\frac{7\pi}{5}\quad\text{and}\quad\cos\frac{9\pi}{5}+i\sin\frac{9\pi}{5}$$

are the roots of $x^5=-1$.

So, except one of the roots, namely $\cos\pi+i\sin\pi=-1$, the roots of $x^4+x^3+x^2-x+1=0$ are

$$\cos\frac{\pi}{5}+i\sin\frac{\pi}{5}\qquad\cos\frac{3\pi}{5}+i\sin\frac{3\pi}{5}$$

$$\cos\frac{7\pi}{5}+i\sin\frac{7\pi}{5}\quad\text{and}\quad\cos\frac{9\pi}{5}+i\sin\frac{9\pi}{5}$$

Note that

$$\cos\frac{9\pi}{5}+i\sin\frac{9\pi}{5}=\cos\left(2\pi-\frac{\pi}{5}\right)+i\sin\left(2\pi-\frac{\pi}{5}\right)$$
$$=\cos\frac{\pi}{5}-i\sin\frac{\pi}{5}$$

and

$$\cos\frac{7\pi}{5}+i\sin\frac{7\pi}{5}=\cos\left(2\pi-\frac{3\pi}{5}\right)+i\sin\left(2\pi-\frac{3\pi}{5}\right)$$
$$=\cos\frac{3\pi}{5}-i\sin\frac{3\pi}{5}.$$

Hence the roots of $x^4-x^3+x^2-x+1=0$ are

$$\cos\frac{\pi}{5}\pm i\sin\frac{\pi}{5},\ \cos\frac{3\pi}{5}\pm\ \sin\frac{3\pi}{5}.$$

Example 15. If $n\left(=\frac{p}{q}\right)$ is a rational number, show that

$$\cos n\theta+i\sin n\theta=\left(\cos\frac{p}{q}\theta+i\sin\frac{p}{q}\theta\right)$$

is the pth power of a qth root of $\cos\theta+i\sin\theta$.

Solution. Let $n=\frac{p}{q}$, where p is an integer (positive or negative) and q is a positive integer

Now $\quad\left(\cos\frac{\theta}{q}+i\sin\frac{\theta}{q}\right)^q=\cos q\frac{\theta}{q}+i\sin q\frac{\theta}{q}=\cos\theta+i\sin\theta$

Thus $\quad\cos\frac{\theta}{q}+i\sin\frac{\theta}{q}$ is one of the qth roots of $\cos\theta+i\sin\theta$.

That is, $\left(\cos\frac{\theta}{q}+i\sin\frac{\theta}{q}\right)^p$ is the pth power of a qth root of $\cos\theta+i\sin\theta$.

i.e., $\cos\left(\frac{p}{q}\theta\right)+i\sin\left(\frac{p}{q}\theta\right)$ is the pth power of a qth root of $\cos\theta+i\sin\theta$.

Example 16. Find all the cube roots of $z=-i$.

Also find the squares of these cube roots.

Solution. Here $z=-i=\cos\left(-\frac{\pi}{2}\right)+i\sin\left(-\frac{\pi}{2}\right)$,

so that $|z|=1$, and

$$z^{1/3}=\left[\cos\left(-\frac{\pi}{2}\right)+i\sin\left(-\frac{\pi}{2}\right)\right]^{\frac{1}{3}}$$

or $\quad z^{1/3}=w_k=\cos\left(\frac{-\frac{\pi}{2}+2k\pi}{3}\right)+i\sin\left(\frac{-\frac{\pi}{2}+2k\pi}{3}\right),\quad k=0,1,2.$

Hence

$$w_0 = \cos\left(-\frac{\pi}{6}\right) + i\sin\left(-\frac{\pi}{6}\right) = \cos\frac{\pi}{6} - i\sin\frac{\pi}{6} = \frac{\sqrt{3}}{2} - i\cdot\frac{1}{2}$$

$$= \frac{1}{2}(\sqrt{3} - i)$$

$$w_1 = \cos\left(\frac{-\frac{\pi}{2} + 2\pi}{3}\right) + i\sin\left(\frac{-\frac{\pi}{2} + 2\pi}{3}\right)$$

$$= \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i$$

$$w_2 = \cos\left(\frac{-\frac{\pi}{2} + 4\pi}{3}\right) + i\sin\left(\frac{-\frac{\pi}{2} + 4\pi}{3}\right)$$

$$= \cos\left(\frac{7\pi}{6}\right) + i\sin\left(\frac{7\pi}{6}\right)$$

$$= \cos\left(\pi + \frac{\pi}{6}\right) + i\sin\left(\pi + \frac{\pi}{6}\right)$$

$$= -\left(\frac{\sqrt{3} + 1}{2}\right) \quad \text{as } \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2},\ \sin\frac{\pi}{6} = \frac{1}{2}$$

For squares of the roots of cubic roots of $-i$, we have

$$z^{\frac{2}{3}} = \left[\cos\left(-\frac{\pi}{2}\right) + i\sin\left(-\frac{\pi}{2}\right)\right]^{\frac{2}{3}}$$

$$= \left[\cos\left(2\left(-\frac{\pi}{2}\right)\right) + i\sin\left(2\left(-\frac{\pi}{2}\right)\right)\right]^{\frac{1}{3}}$$

$$= [\cos(-\pi) + i\sin(-\pi)]^{\frac{1}{3}}$$

or $$z^{1/3} = w'_k = \cos\left(\frac{-\pi + 2k\pi}{3}\right) + i\sin\left(\frac{-\pi + 2k\pi}{3}\right), \quad k = 0, 1, 2$$

For $k = 0, 1, 2$, respectively, we get

$$w'_0 = \cos\left(-\frac{\pi}{3}\right) + i\sin\left(-\frac{\pi}{3}\right)$$

$$= \frac{1}{2} - \frac{\sqrt{3}}{2}i = \frac{1}{2}(1 - \sqrt{3}\,i)$$

$$w'_1 = \cos\frac{\pi}{3} + i\sin\frac{\pi}{3} = \frac{1}{2}(1 + \sqrt{3}\,i)$$

$$w'_2 = \cos(\pi) + i\sin(\pi) = -1$$

Note. In Example 16 we see that the roots of $-i$ have been found as complex numbers with 'known' real and imaginary parts. In some cases, however, such simplification is not possible. In that case, we write the roots as $\cos\theta_k + i\sin\theta_k$ for $k = 0, 1, 2, \ldots, n-1$ (see Problem 5 EXERCISE 1.2.)

EXERCISE 1.2

Write each of the following expressions in the form $a + ib$.

(i) $(-\sqrt{3} + i)^2$ (ii) $(-3i)^4$ (iii) $\left(\frac{1 - \sqrt{3}\,i}{1 + \sqrt{3}\,i}\right)^6$

Simplify,

(i) $\dfrac{(\cos 2\theta + i\sin 2\theta)^5 (\cos 3\theta - i\sin 3\theta)^6}{(\cos 4\theta - i\sin 4\theta)^7 (\cos 5\theta + i\sin 5\theta)^8}$

(ii) $\dfrac{(\cos\alpha - i\sin\alpha)^{11}}{(\cos\beta + i\sin\beta)^9}$

(iii) $\dfrac{(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta)}{(\cos\gamma + i\sin\gamma)(\cos\delta + i\sin\delta)}$ (iv) $\dfrac{\left(3\,\text{cis}\,\frac{\pi}{6}\right)^7}{\left(4\,\text{cis}\,\frac{\pi}{3}\right)^3}$

Prove that

(i) $\left[(\cos\theta - \cos\phi) + i(\sin\theta - \sin\phi)\right]^n + \left[(\cos\theta - \cos\phi) - i(\sin\theta - \sin\phi)\right]^n$
$= 2^{n+1}\sin^n\left(\frac{\theta - \phi}{2}\right)\cos n\left(\frac{\theta + \phi + \pi}{2}\right)$

(ii) $\left(\dfrac{1 + \sin x + i\cos x}{1 + \sin x - i\cos x}\right)^n = \cos n\left(\frac{\pi}{2} - x\right) + i\sin n\left(\frac{\pi}{2} - x\right)$

If $2\cos\theta = x + \frac{1}{x}$, $2\cos\phi = y + \frac{1}{y}$, $2\cos\psi = z + \frac{1}{z}$, then prove that

(i) $2\cos(\theta + \phi + \psi) = xyz + \dfrac{1}{xyz}$

(ii) $2\cos(m\theta + n\phi) = x^m y^n + \dfrac{1}{x^m y^n}$

(i) Find the three cube roots of $8\,i$.

(ii) Find the four 4th roots of each of the complex numbers $-16\,i$, 64 and $-2\sqrt{3} + 2i$.

6. Find the six 6th roots of (i) -1 and (ii) $1+i$

7. Find the squares of all the 5th roots of $\frac{1}{2}+\frac{\sqrt{3}}{2}i$

8. Solve the following equations:

(i) $x^7+1=0$

(ii) $x^7+x^4+x^3+1=0$

(iii) $x^6+1=\sqrt{3}\,i$

9. Solve the equation $x^{12}-1=0$ and find which of its roots satisfy the equation $x^4+x^2+1=0$.

$x^2(x^2+1)=-1$

10. Express the following in series of sines or cosines of multiplies of θ

(i) $\cos^4\theta$ (ii) $\sin^4\theta$ (iii) $\sin^6\theta$

(iv) $\cos^7\theta$ (v) $\sin^9\theta$ (vi) $\sin^6\theta\cos^2\theta$

(vii) $\cos^4\theta\sin^3\theta$ (viii) $\cos^5\theta\sin^7\theta$

11. Show that $\cos^4\theta+\sin^4\theta=\frac{1}{4}(3+\cos 4\theta)$

12. Prove that:

$$64(\cos^8\theta+\sin^8\theta)=\cos 8\theta+28\cos 4\theta+35$$

13. Prove that:

(i) $\sin 3\theta = 3\sin\theta-4\sin^3\theta$

(ii) $\cos 3\theta = 4\cos^3\theta-3\cos\theta$

(iii) $\sin 4\theta = 4(\cos^3\theta\sin\theta-\cos\theta\sin^3\theta)$

(iv) $\cos 4\theta = 8\cos^4\theta-\cos^2\theta+1$

(v) $\frac{\sin 5\theta}{\sin\theta} = 16\cos^4\theta-12\cos^2\theta+1$

14. Prove that $\tan 6\theta = 2t\left(\frac{3-10t^2+3t^4}{1-15t^2+15t^4-t^6}\right)$, where $t=\tan\theta$.

15. Prove that $\tan 3\theta=\frac{3\tan\theta-4\tan^3\theta}{1-3\tan^2\theta}$ and hence solve the equation

$$1-3t^2 = 3t-t^3,\quad \text{where}\quad t=\tan\theta.$$

16. Prove that $\cos\frac{\pi}{7}-\cos\frac{2\pi}{7}+\cos\frac{3\pi}{7}=\frac{1}{2}$.

17. Prove the following relations $(m,n\in Z)$

(i) $z^m z^n = z^{m+n}$ (ii) $(z^m)^n = z^{mn}$

(iii) $(z_1z_2)^n = z_1^n z_2^n$ (iv) $\frac{z^m}{z^n} = z^{m-n},\quad z\neq 0$

(v) $\left(\frac{z_1}{z_2}\right)^n = \frac{z_1^n}{z_2^n}$, provided $z_2\neq 0$

BASIC ELEMENTARY FUNCTIONS[1]

In this section we discuss a number of important basic elementary functions. Each of these have a significance of its own.

(1.16) The Exponential Function.

The exponential function e^z is defined as follows.

For

$$z = x+iy,\quad\text{we write}$$

$$e^z = e^{x+iy} = e^x(\cos y+i\sin y) \quad (1)$$

If, in (1), we put $x=0$ then (1) becomes

$$e^{iy} = \cos y+i\sin y$$

for all real numbers y. (2)

For $y=\pi$, (2) becomes

$$e^{i\pi} = \cos\pi+i\sin\pi = -1,$$

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

which is called the **Euler's identity**.

The formula (2) is called **Euler's formula** named after the famous Swiss mathematician Leonard Euler (1707 – 1783).

Instead if we put $y=0$ in (1), we obtain the usual real exponential function e^x.

The exponential form of $z=x+iy$ is:

$$z = x+iy$$
$$= r(\cos\theta+i\sin\theta)$$
$$= re^{i\theta} \quad (3)$$

by the Euler's formula.

1. **Elementary functions** consist of (i) algebraic functions, (ii) trigonometric and inverse trigonometric functions (iii) logarithmic and exponential functions and (iv) all functions that can be constructed from these by the four fundamental arithmetical operation +, –, × and ÷.

The following laws of exponents hold.

(i) $e^{z_1} \cdot e^{z_2} = e^{z_1+z_2}$

(ii) $\dfrac{e^{z_1}}{e^{z_2}} = e^{z_1} \cdot e^{-z_2} = e^{z_1-z_2}$

Also

(iii) e^z is periodic of period $2k\pi i$ which is a pure imaginary number.

Proof.

(i) For complex numbers $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, we have

$$\begin{aligned}
e^{z_1} \cdot e^{z_2} &= e^{x_1+iy_1} \cdot e^{x_2+iy_2} \\
&= e^{x_1}(\cos y_1 + i \sin y_1) \cdot e^{x_2}(\cos y_2 + i \sin y_2) \\
&= e^{x_1} \cdot e^{x_2} [\cos y_1 \cos y_2 - \sin y_1 \sin y_2) \\
&\quad + i(\cos y_1 \sin y_2 + \sin y_1 \cos y_2)] \\
&= e^{x_1+x_2} [\cos (y_1+y_2) + i \sin (y_1+y_2)] \\
&= e^{z_1+z_2}\ ; \quad z_1 + z_2 = (x_1+x_2) + i(y_1+y_2).
\end{aligned}$$

(ii) The proof for this part is similar to that for (i).

(iii) Here

$$\begin{aligned}
e^{z+2k\pi i} &= e^z \cdot e^{2k\pi i} \\
&= e^z \cdot (\cos 2k\pi + i \sin 2k\pi), && \text{by Euler's formula} \\
&= e^z && \text{Since } \cos 2k\pi = 1, \sin 2k\pi = 0
\end{aligned}$$

So e^z has period $2k\pi i$.

(1.17) Trigonometric Functions

Using the Euler's formulas

$$e^{iy} = \cos y + i \sin y, \qquad e^{-iy} = \cos y - i \sin y$$

for real y, we have

$$\cos y = \frac{e^{iy} + e^{-iy}}{2}, \qquad \sin y = \frac{e^{iy} - e^{-iy}}{2i}$$

On the basis of analogy with these equations, we define (for a complex number z)

(i) $\sin z = \dfrac{e^{iz} - e^{-iz}}{2i}$

(ii) $\cos z = \dfrac{e^{iz} + e^{-iz}}{2}$

(These definitions coincide with the functions $\cos x$, $\sin x$ for real $z = x$)

(iii) $\tan z = \dfrac{\sin z}{\cos z} = \dfrac{e^{iz} - e^{-iz}}{i(e^{iz} + e^{-iz})}$

(iv) $\cot z = \dfrac{\cos z}{\sin z} = \dfrac{i(e^{iz} + e^{-iz})}{e^{iz} - e^{-iz}}$

(v) $\sec z = \dfrac{1}{\cos z} = \dfrac{2}{e^{iz} + e^{-iz}}$

(vi) $\operatorname{cosec} z = \dfrac{1}{\sin z} = \dfrac{2i}{e^{iz} - e^{-iz}}$.

Using the above definitions, we can prove various trigonometric identities for complex trigonometric functions.

Example 17. For all complex numbers z, prove that $\sin^2 z + \cos^2 z = 1$.

Solution.

$$\sin^2 z = \left(\frac{e^{iz} - e^{-iz}}{2i}\right)^2 = \frac{e^{2iz} + e^{-2iz} - 2}{-4}$$

$$\cos^2 z = \left(\frac{e^{iz} - e^{-iz}}{2}\right) = \frac{e^{2iz} + e^{-2iz} + 2}{4}$$

Hence,

$$\sin^2 z + \cos^2 z = -\frac{e^{2iz} + e^{-2iz} - 2}{4} + \frac{e^{2iz} + e^{-2iz} + 2}{4} = \frac{4}{4} = 1.$$

More trigonometric identities are given in EXERCISE 1.3, Problem 2.

(1.18) The Hyperbolic Functions. Still another set of complex elementary functions are the hyperbolic functions.

The hyperbolic functions of a complex number z are defined as follows:

(i) $\sinh z = \dfrac{e^z - e^{-z}}{2}$ ($\sinh z$ is read as **sine hyperbolic z**)

(ii) $\cosh z = \dfrac{e^z + e^{-z}}{2}$ ($\cosh z$ is read as **cosine hyperbolic z**)

(iii) $\tanh z = \dfrac{\sinh z}{\cosh z} = \dfrac{e^z - e^{-z}}{e^z + e^{-z}}$

(iv) $\coth z = \dfrac{\cosh z}{\sinh z} = \dfrac{e^z + e^{-z}}{e^z - e^{-z}}$

(v) $\operatorname{sech} z = \dfrac{1}{\cosh z} = \dfrac{2}{e^z + e^{-z}}$

(vi) $\operatorname{csch} z = \dfrac{1}{\sinh z} = \dfrac{2}{e^z - e^{-z}}$

(1.19) Relationships between Trigonometric and Hyperbolic Functions

We have

$$\sin iz = \frac{e^{i(iz)} - e^{-i(iz)}}{2i} = \frac{e^{-z} - e^{z}}{2i} = \frac{e^{-z} - e^{z}}{2i^2} \times i$$

$$= \frac{e^{z} - e^{-z}}{2} \times i = i \sinh z.$$

$$\cos iz = \frac{e^{i(iz)} + e^{-i(iz)}}{2} = \frac{e^{-z} + e^{z}}{2} = \cosh z$$

$$\tan iz = \frac{\sin iz}{\cos iz} = \frac{i \sinh z}{\cosh z} = i \tanh z.$$

Similarly, we can prove that

$$\sinh iz = i \sin z, \qquad \cosh iz = \cos z.$$

Example 18. Show that

$$\cos z = \cos x \cosh y - i \sin x \sinh y, \text{ where } z = x + iy.$$

Solution.

$$\cos z = \cos (x + iy)$$

$$= \cos x \cos (iy) - \sin x \sin (iy)$$

$$= \cos x \cosh y - i \sin x \sinh y.$$

Example 19. Prove that

$$\sinh (z_1 + z_2) = \sinh z_1 \cosh z_2 + \cosh z_1 \sinh z_2.$$

Solution. We know that

$$\sin (u_1 + u_2) = \sin u_1 \cos u_2 + \cos u_1 \sin u_2$$

where u_1 and u_2 are complex numbers. Let

$u_1 = iz_1, \quad u_2 = iz_2$. Then

$$\sin i(z_1 + z_2) = \sin (iz_1) \cos (iz_2) + \cos (iz_1) \sin (iz_2)$$

i.e., $$i \sinh (z_1 + z_2) = i \sinh z_1 \cosh z_2 + i \cosh z_1 \sinh z_2$$

$$= i [\sinh z_1 \cosh z_2 + \cosh z_1 \sinh z_2]$$

Hence $$\sinh (z_1 + z_2) = \sinh z_1 \cosh z_2 + \cosh z_1 \sinh z_2.$$

EXERCISE 1.3

1. Show that
   - (i) e^z is never zero
   - (ii) $|e^{ix}| = 1$
   - (iii) $e^z = 1$ if and only if z is an integral multiple of $2\pi i$
   - (iv) $e^{z_1} = e^{z_2}$ if and only if $z_1 - z_2 = 2k\pi i$, where k in an integer
   - (v) $|e^z| = e^x$, where $z = x + iy$
   - (vi) $e^{z_1} \cdot e^{z_2} \cdot \cdots \cdot e^{z_n} = e^{z_1 + z_2 + \cdots + z_n}$
   - (vii) $(e^z)^n = e^{nz}$, n being an integer.

2. Prove that (for all $z, z_1, z_2 \in C$)
   - (i) $1 + \tan^2 z = \sec^2 z$
   - (ii) $1 + \cot^2 z = \csc^2 z$
   - (iii) $\sin (z_1 - z_2) = \sin z_1 \cos z_2 - \cos z_1 \sin z_2$
   - (iv) $\cos (z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2$
   - (v) $\cos (z_1 - z_2) = \cos z_1 \cos z_2 + \sin z_1 \sin z_2$
   - (vi) $\cos 2z = \cos^2 z - \sin^2 z = 2 \cos^2 z - 1 = 1 - 2 \sin^2 z$
   - (vii) $\sin 2z = 2 \sin z \cos z$
   - (viii) $\cos z_1 - \cos z_2 = 2 \sin \frac{z_1 + z_2}{2} \sin \frac{z_2 - z_1}{2}$
   - (ix) $\sin z_1 - \sin z_2 = 2 \cos \frac{z_1 + z_2}{2} \sin \frac{z_1 - z_2}{2}$
   - (x) $\sin 3z = 3 \sin z - 4 \sin^3 z$
   - (xi) $\tan (z_1 + z_2) = \dfrac{\tan z_1 + \tan z_2}{1 - \tan z_1 \tan z_2}$
   - (xii) $\tan (z_1 - z_2) = \dfrac{\tan z_1 - \tan z_2}{1 + \tan z_1 \tan z_2}$

3. Show that
   - (i) $\overline{\sin z} = \sin \bar{z}$
   - (ii) $\overline{\cos z} = \cos \bar{z}$
   - (iii) $\overline{\tan z} = \tan \bar{z}$
   - (iv) $\sin (-z) = -\sin z$
   - (v) $\cos (-z) = \cos z$
   - (vi) $\tan (-z) = -\tan z$

(vii) $\sinh(-z) = -\sinh z$ (viii) $\cosh(-z) = \cosh z$

(ix) $\tanh(-z) = -\tanh z$ (x) $\overline{\tanh z} = \tanh \bar{z}$

4. Prove the following identities

(i) $\cosh^2 z - \sinh^2 z = 1$

(ii) $\operatorname{sech}^2 z = 1 - \tanh^2 z$

(iii) $\operatorname{csch}^2 z = \coth^2 z - 1$

(iv) $\cosh 2z = \cosh^2 z + \sinh^2 z = 2\cosh^2 z - 1 = 1 + 2\sinh^2 z$

(v) $\sinh 2z = 2 \sinh z \cosh z$

(vi) $\sinh 3z = 3 \sinh z + 4 \sinh^3 z$

(vii) $\cosh 3z = 4\cosh^3 z - 3\cosh z$

(viii) $\sinh(z_1 - z_2) = \sinh z_1 \cosh z_2 - \cosh z_1, \sinh z_2$

(ix) $\tanh(z_1 \pm z_2) = \dfrac{\tanh z_1 \pm \tanh z_2}{1 \mp \tanh z_1 \tanh z_2}$

(x) $\tanh 3z = \dfrac{3\tanh z + \tanh^3 z}{1 + 3\tanh^2 z}$

5. If $z = x + iy$, prove that

(i) $\sin z = \sin x \cosh y + i \cos x \sinh y$

(ii) $\tan z = \dfrac{\sin 2x + i \sinh 2y}{\cos 2x + \cosh 2y}$

6. If $\sin(A + iB) = x + iy$, show that

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

and $$\frac{x^2}{\cosh^2 B} + \frac{y^2}{\sinh^2 B} = 1$$

7. If $\tan(\alpha + i\beta) = x + iy$, show that

$$x^2 + y^2 + 2x \cot 2\alpha = 1 \text{ and } x^2 + y^2 - 2y \coth 2\beta = -1.$$

8. If $\sin(\theta + i\phi) = \cos\alpha + i \sin\alpha$, prove that

$$\cos^2\theta = \pm \sin\alpha$$

9. If $\tan(\theta + i\phi) = \tan\alpha + i\sec\alpha$, prove that

$$e^{2\phi} = \pm\cot\frac{\alpha}{2} \quad \text{and} \quad 2\theta = n\pi + \frac{\pi}{2} + \alpha.$$

10. Prove that

$$\sinh\left(\frac{x}{2}\right) = \sqrt{\frac{\cosh x - 1}{2}} \quad \text{if } x \ge 0$$
$$= -\sqrt{\frac{\cosh x - 1}{2}} \quad \text{if } x < 0$$

11. Show that multiplication of a vector z by $e^{i\alpha}$, where α is a real number, rotates the vector z counterclockwise through an angle of measure α.

12. Show that

(i) $2 + i = \sqrt{5}\, e^{i\tan^{-1}\left(\frac{1}{2}\right)}$ (ii) $-3 - 4i = 5e^{i\left(\pi + \tan^{-1}\left(\frac{4}{3}\right)\right)}$

LOGARITHMIC FUNCTION

Logarithmic function also is one of the basic elementary functions.

(1.20) Definition. Let z be a nonzero complex number. If, for a complex number w, $e^w = z$, then w is called the logarithm of z and is written as

$$w = \log z$$

Since e^w is never equal to 0, the complex number 0 has no logarithm. Thus log 0 is not defined.

Suppose that $w = u + iv$. Then

$$z = e^w = e^{u+iv}$$
$$= e^u(\cos v + i \sin v)$$

and

$$|z| = e^u \tag{1}$$

From (1), we have

$$u = \log|z|$$

where $\log|z|$ denotes the logarithm of $|z|$ to the base e. We also write logarithm of $|z|$ to the base e as $\ln|z|$, called the natural logarithm of $|z|$.

Also

$$\arg z = v = \operatorname{Arg} z + 2k\pi, \quad -\pi < \operatorname{Arg} z \le \pi.$$

Thus

$$\log z = u + iv$$
$$= \log|z| + i \operatorname{Arg} z + 2k\pi i$$
$$= \log|z| + i \arg z \tag{2}$$

has infinitely many values for any complex number z. For each value of k, the right hand side of (2) defines a '**branch**' of the logarithm.

The particular value, namely

$$\log|z| + i \operatorname{Arg} z$$

is called the **principal logarithm** of z and is denoted by $\operatorname{Log} z$. Thus

$$\operatorname{Log} z = \log|z| + i \operatorname{Arg} z, \qquad -\pi < \operatorname{Arg} z \leq \pi$$

is the **principal logarithm** of z.

Using (2) we can write

$$\log z = \operatorname{Log} z + 2k\pi i$$

Example 20. Find $\operatorname{Log} z$ if

(i) $z = 2i$ (ii) $z = -i$

(iii) $z = x,\ x > 0$ (iv) $z = x,\ x < 0$

(v) $z = 1 + \sqrt{3}\, i$ (vi) $z = -4 + 3i$

Solution.

(i) $$z = 2i = 0 + 2i$$

Thus $$|z| = 2, \quad \operatorname{Arg} z = \frac{\pi}{2}$$

and $$\operatorname{Log} 2i = \ln 2 + \frac{\pi}{2} i.$$

(ii) $$z = -i = 0 + (-1)\, i$$

Here $$|z| = 1, \quad \operatorname{Arg} z = -\frac{\pi}{2}$$

Here $$\operatorname{Log}(-i) = \ln 1 - \frac{\pi}{2} i = -\frac{\pi}{2} i.$$

(iii) $$z = x = x + 0i \quad \text{and} \quad x > 0$$

$$|z| = x, \quad \operatorname{Arg} z = 0$$

Hence $$\operatorname{Log} z = \ln x.$$

(iv) $$z = x = x + 0i \quad \text{and} \quad x < 0$$

Here $$|z| = -x, \quad \operatorname{Arg} z = \pi.$$

Hence $$\operatorname{Log} z = \ln(-x) + \pi i.$$

(v) $$z = 1 + \sqrt{3}\, i$$

$$|z| = \sqrt{1+3} = 2, \quad \operatorname{Arg} z = \tan^{-1}\sqrt{3} = \frac{\pi}{3}$$

Thus

$$\operatorname{Log}\left(1 + \sqrt{3}\, i\right) = \ln 2 + \frac{\pi i}{3}$$

(vi) $$z = -4 + 3i$$

$$|z| = \sqrt{4^2 + 3^2} = 5, \quad \operatorname{Arg} z = \tan^{-1}\left(-\frac{3}{4}\right) = \pi - \tan^{-1}\left(\frac{3}{4}\right)$$

Therefore

$$\operatorname{Log}(-4 + 3i) = \ln 5 + \left(\pi - \tan^{-1}\frac{3}{4}\right)$$

INVERSE HYPERBOLIC FUNCTIONS

We have already learnt about complex hyperbolic functions. The inverse hyperbolic functions are related to these by the following definition.

(1.21) Definition. Let z and w be complex numbers. If

$$z = \sinh w$$

then $w = \sinh^{-1} z$ is called the **inverse hyperbolic** sine of z.

Other hyperbolic functions like $\cosh^{-1} z$, $\tanh^{-1} z$, etc. are defined similarly.

All these inverse hyperbolic functions are multi-valued. These can be expressed in terms of natural logarithms.

(1.22) Theorem. For any complex number z,

(i) $\sinh^{-1} z = \log\left(z + \sqrt{z^2 + 1}\right)$

(ii) $\cosh^{-1} z = \log\left(z + \sqrt{z^2 - 1}\right)$

(iii) $\tan^{-1} z = \dfrac{1}{2} \log\left(\dfrac{1+z}{1-z}\right)$

Proof.

(i) Let $$w = \sinh^{-1} z. \text{ Then } z = \sinh w. \text{ So}$$

$$z = \frac{e^w + e^{-w}}{2} = \frac{e^{2w} - 1}{2e^w}$$

Hence

$$e^{2w} - 2e^w z - 1 = 0$$

which is a quadratic equation in e^w. Therefore

$$e^w = \frac{2z \pm \sqrt{4z^2+4}}{2}$$

$$= z \pm \sqrt{z^2+1}$$

Hence, taking logarithm of both the sides,

$$w = \log\left(z+\sqrt{z^2+1}\right) = \sinh^{-1} z$$

(Only the term with positive sign has been taken. This term represents one of the branches of the square root and complex logarithm).

Thus

$$\sinh^{-1} z = \log\left(z+\sqrt{z^2+1}\right)$$

(ii) Let $w = \cosh^{-1} z$ so that

$$z = \cosh w = \frac{e^w - e^{-w}}{2} = \frac{e^{2w}+1}{2e^w}$$

Hence

$$e^{2w} - 2ze^w + 1 = 0$$

which is a quadratic equation in e^w. Therefore,

$$e^w = \frac{2z \pm \sqrt{4z^2-4}}{2}$$

$$= z \pm \sqrt{z^2-1}$$

Taking logarithm of both the sides we have

$$w = \cosh^{-1} z = \log\left(z+\sqrt{z^2-1}\right)$$

(Again only the branch with the positive square root sign has been considered)

(iii) Let $w = \tanh^{-1} z$. Then

$$z = \tanh w = \frac{e^w - e^{-w}}{e^w + e^{-w}} = \frac{e^{2w}-1}{e^{2w}+1}$$

Hence

$$e^{2w}(1-z) = 1+z$$

That is

$$e^{2w} = \frac{1+z}{1-z}$$

So

$$w = \tanh^{-1} z$$

$$= \frac{1}{2}\log\left(\frac{1+z}{1-z}\right)$$

where only the principal branch of the logarithm is taken into account.

Using the definition of other inverse hyperbolic functions, one obtains

(iv) $$\coth^{-1} z = \frac{1}{2}\log\left(\frac{z+1}{z-1}\right)$$

(v) $$\operatorname{sech}^{-1} z = \log\left(\frac{1+\sqrt{1-z^2}}{2}\right)$$

(vi) $$\operatorname{cosech}^{-1} z = \log\left(\frac{1+\sqrt{z^2+1}}{2}\right)$$

Example 21. Prove that

$$\coth^{-1}\left(\frac{2}{z}\right) = \sinh^{-1}\left(\frac{z}{\sqrt{4-z^2}}\right)$$

Solution. Let $\coth^{-1}\left(\frac{2}{z}\right) = w$ so that $\frac{2}{z} = \coth w$

or $z = 2\tanh w$

Now $\sqrt{4-z^2} = \sqrt{4-4\tanh^2 w} = 2\sqrt{1-\tanh^2 w} = 2\operatorname{sech} w$

Therefore,

$$\frac{z}{\sqrt{4-z^2}} = \frac{2\tanh w}{2\operatorname{sech} w} = \tanh w \cdot \cosh w = \sinh w.$$

and so $$w = \sinh^{-1}\left(\frac{z}{\sqrt{4-z^2}}\right)$$

Thus $\coth^{-1}\left(\frac{2}{z}\right) = \sinh^{-1}\left(\frac{z}{\sqrt{4-z^2}}\right)$ by using the supposition

$$w = \coth^{-1}$$

INVERSE TRIGONOMETRIC FUNCTIONS

(1.23) Definition. If $z = \sin w$, then $w = \sin^{-1} z$ is called the inverse sine function of z.

Similarly, other inverse circular functions are defined.

These functions can be expressed in terms of natural logarithms as follows.

(i) $\sin^{-1} z = \frac{1}{i}\log\left(iz + \sqrt{1-x^2}\right)$

(ii) $\cos^{-1} z = \frac{1}{i}\log\left(z + \sqrt{z^2-1}\right)$

(iii) $\tan^{-1} z = \frac{1}{2i}\log\left(\frac{1+iz}{1-iz}\right)$

(iv) $\cot^{-1} z = \frac{1}{2i}\log\left(\frac{z+i}{z-i}\right)$

(v) $\sec^{-1} z = \frac{1}{i}\log\left(\frac{1+\sqrt{1-z^2}}{z}\right)$

(vi) $\csc^{-1} z = \frac{1}{i}\log\left(\frac{i+\sqrt{z^2-1}}{z}\right)$

We prove (i). Let $\sin^{-1} z = w$ so that

$$z = \sin w \quad \text{and} \quad \cos w = \sqrt{1-\sin^2 w} = \sqrt{1-z^2}$$

There are two branches for $\cos w$. We have taken only the branch with the positive sign.

Now $$iz + \sqrt{1-z^2} = i\sin w + \cos w = \frac{e^{iw}-e^{-iw}}{2} + \frac{e^{iw}+e^{-iw}}{2} = e^{iw}$$

Therefore,

$$iw = \log\left(iz + \sqrt{1-e^2}\right) \Rightarrow w = \frac{1}{i}\log\left(iz + \sqrt{1-z^2}\right)$$

Hence

$$\sin^{-1} z = \frac{1}{i}\log\left(iz + \sqrt{1-z^2}\right)$$

The other results can be proved in a similar manner (See EXERCISE 1.4 Problem 3).

ij

Example 22. Separate into real and imaginary parts:

(i) $\tan^{-1}(x+iy)$

(ii) $\sin^{-1}(\cos\theta + i\sin\theta)$

Solution.

(i) Let $\tan^{-1}(x+iy) = a + i\beta$ (1)

We first show that $\tan^{-1}(x-iy) = \alpha - i\beta$.

From (1), we have

$$x + iy = \tan(\alpha + i\beta)$$

$$= \frac{\tan\alpha + \tan(i\beta)}{1 - \tan\alpha\tan(i\beta)} = \frac{\tan\alpha + i\tanh\beta}{1 - i\tan\alpha\tanh\beta}$$

Taking conjugates of both the sides, we obtain

$$x - iy = \frac{\overline{\tan\alpha + i\tan\beta}}{\overline{1 - i\tan\alpha\tanh\beta}} = \frac{\tan\alpha - i\tanh\beta}{1 + i\tan\alpha\tanh\beta}$$

$$= \frac{\tan\alpha - \tan(i\beta)}{1 + \tan\alpha\tan(i\beta)} = \tan(\alpha - i\beta)$$

i.e., $$\tan^{-1}(x-iy) = \alpha - i\beta \qquad (2)$$

Adding (1) and (2), we get

$$\tan^{-1}(x+iy) + \tan^{-1}(x-iy) = 2\alpha$$

or $$\tan^{-1}\frac{(x+iy)+(x-iy)}{1-(x+iy)(x-iy)} = 2\alpha,$$

because $\tan^{-1}x + \tan^{-1}y = \tan^{-1}\left(\frac{x+y}{1-xy}\right)$

i.e., $$\tan^{-1}\frac{2x}{1-x^2-y^2} = 2\alpha$$

Therefore,

$$\alpha = \frac{1}{2}\tan^{-1}\frac{2x}{1-x^2-y^2} \qquad (3)$$

Again, from (1) and (2), we have

$$\tan^{-1}(x+iy) - \tan^{-1}(x-iy) = 2i\beta$$

i.e., $$\tan^{-1}\frac{(x+iy)-(x-iy)}{1+(x+iy)(x-iy)} = 2i\beta$$

or $$\tan^{-1}\frac{2yi}{1+x^2+y^2} = 2i\beta$$

or $$\frac{2yi}{1+x^2+y^2} = \tan(2i\beta) = i\tanh 2\beta$$

Therefore, $$2\beta = \tanh^{-1}\frac{2y}{1+x^2+y^2}$$

or $$\beta = \frac{1}{2}\tanh^{-1}\frac{2y}{1+x^2+y^2} \qquad (4)$$

So the real and imaginary parts of $\tan^{-1}(x+iy)$ are given by (3) and (4) respectively.

(ii) Let $\sin^{-1}(\cos\theta + i\sin\theta) = x+iy$

$$\cos\theta + i\sin\theta = \sin(x+iy)$$
$$= \sin x\cosh y + i\cos x\sinh y$$

Equating real and imaginary parts, we have

$$\sin x\cosh y = \cos\theta \quad \text{and} \quad \cos x\sinh y = \sin\theta$$

so that
$$\cosh y = \frac{\cos\theta}{\sin x} \quad \text{and} \quad \sinh y = \frac{\sin\theta}{\cos x}$$

Now $1 = \cosh^2 y - \sinh^2 y = \dfrac{\cos^2\theta}{\sin^2 x} - \dfrac{\sin^2\theta}{\cos^2 x}$

or $\sin^2 x \cdot \cos^2 x = \cos^2\theta\cos^2 x - \sin^2\theta\sin^2 x$

or $(1-\cos^2 x)\cos^2 x = (1-\sin^2\theta)\cos^2 x - \sin^2\theta(1-\cos^2 x)$

or $\cos^2 x - \cos^4 x = \cos^2 x - \sin^2\theta\cos^2 x - \sin^2\theta + \sin^2\theta\cos^2 x$

or $\cos^4 x = \sin^2\theta$

So $\cos x = \sqrt{\sin\theta}$ (5)

or $x = \cos^{-1}\left(\sqrt{\sin\theta}\right)$

Again, $\sinh y = \dfrac{\sin\theta}{\cos x} = \dfrac{\sin\theta}{\sqrt{\sin\theta}} = \sqrt{\sin\theta}$, using (5)

Thus
$$y = \sinh^{-1}\left(\sqrt{\sin\theta}\right)$$
$$= \log\left(\sqrt{\sin\theta} + \sqrt{\left(\sqrt{\sin\theta}\right)^2 + 1}\right)$$
$$= \log\left(\sqrt{\sin\theta} + \sqrt{\sin\theta + 1}\right)$$

COMPLEX POWERS

We have already discussed the rational powers of a complex number. Here we study the case when the exponent of z is itself a complex number.

(1.24) Definition. Let z be a nonzero complex number. For any complex number w, the complex power w of z is defined as:

$$z^w = e^{w\log z}$$
$$= \exp(w\log z) \qquad (1)$$

Since complex logarithm is a multivalued function, equation (1) is, in general, satisfied by an infinite number of values of $\log z$. In contrast to this, z^n, n a positive integer, is single valued.

Principal Value of a Complex Power

The principal value of a complex power

$$z^w = e^{w\log z}$$

is defined as:

$$z^w = e^{w\,\mathrm{Log}\,z}$$

where

$$\mathrm{Log}\,z = \log z + i\,\mathrm{Arg}\,z$$

(1.25) Theorem. For any complex numbers w_1, w_2 and z

$$z^{w_1} \cdot z^{w_2} = z^{w_1+w_2}$$

Proof. Here

$$z^{w_1} = e^{w_1\log z}, \quad e^{w_2} = e^{w_2\log z}$$

So

$$z^{w_1} \cdot z^{w_2} = e^{w_1\log z} \cdot e^{w_2\log z} = e^{w_1\log z + w_1\log z}$$
$$= e^{(w_1+w_1)\log z} = z^{w_1+w_1}$$

Example 23. Separate $(\alpha+i\beta)^{(p+iq)}$ into real and imaginary parts.

Solution. Let $(\alpha+i\beta)^{(p+iq)} = x+iy$.

Now
$$(\alpha+i\beta)^{(p+iq)} = e^{(p+iq)\log(\alpha+i\beta)}$$
$$= \exp\left[(p+iq)\left\{\ln\sqrt{\alpha^2+\beta^2} + i\left(\tan^{-1}(\beta/\alpha) + 2n\pi\right)\right\}\right]$$
$$= \exp\left[\frac{1}{2}p\ln(\alpha^2+\beta^2) - q\tan^{-1}(\beta/\alpha) - 2qn\pi + i\left\{\frac{1}{2}q\ln(\alpha^2+\beta^2) + p\tan^{-1}(\beta/\alpha) + 2np\pi\right\}\right]$$

Let
$$u = \frac{1}{2}p\ln(\alpha^2-\beta^2) - q\tan^{-1}(\beta/\alpha) - 2qn\pi \qquad (1)$$

and
$$v = \frac{1}{2}q\ln(\alpha^2+\beta^2) + p\tan^{-1}(\beta/\alpha) + 2qn\pi \qquad (2)$$

Then $(\alpha+i\beta)^{(p+iq)} = x+iy = e^{u+iv} = e^u(\cos v + i\sin y)$

Thus $x = e^u\cos v$, $y = e^u\sin v$ with u and v, as gives by (1) and (2), are the real and imaginary parts respectively.

Example 24. Show that $(-1+i)^{i+\sqrt{3}} = e^x(\cos y + i\sin y)$, where

$$x = \frac{\sqrt{3}}{2}\ln 2 - \frac{3\pi}{4} - 2n\pi \quad \text{and} \quad y = \frac{1}{2}\ln 2 + \frac{3\sqrt{3}}{4}\pi + 2\sqrt{3}\,n\pi$$

Solution.
$$(-1+i)^{i+\sqrt{3}} = e^{(i+\sqrt{3})\log(-1+i)}$$
$$= e^{x+iy} = e^x(\cos y + i\sin y).$$

But $$\log(-1+i) = \frac{1}{2}\ln 2 + \left(\frac{3\pi}{4} + 2n\pi\right)i$$

Therefore,
$$(i+\sqrt{3})\log(-1+i) = (i+\sqrt{3})\left(\frac{1}{2}\ln 2 + \frac{3\pi i}{4} + 2n\pi i\right)$$
$$= \frac{\sqrt{3}}{2}\ln 2 - \frac{3\pi}{4} - 2n\pi + i\left(\frac{1}{2}\ln 2 + \frac{3\sqrt{3}\,\pi}{4} + 2\sqrt{3}\,n\pi\right)$$

Thus
$$x = \frac{\sqrt{3}}{2}\ln 2 - \frac{3\pi}{4} - 2n\pi$$
and
$$y = \frac{1}{2}\ln 2 + \frac{3\sqrt{3}\,\pi}{4} + 2\sqrt{3}\,n\pi$$

Alternative Solution

$$e^x(\cos y + i\sin y) = (-1+i)^{i+\sqrt{3}}$$
$$e^{x+iy} = e^{(i+\sqrt{3})\log(-1+i)}$$
or
$$\exp(x+iy) = \exp\left[(i+\sqrt{3})\left\{\frac{1}{2}\ln 2 + i\left(\frac{3\pi}{4} + 2n\pi\right)\right\}\right],\ n \in Z$$
$$= \exp\left[\frac{\sqrt{3}}{2}\ln 2 - \frac{3\pi}{4} - 2n\pi + i\left(\frac{1}{2}\ln 2 + \frac{3\sqrt{3}\,\pi}{4} + 2\sqrt{3}\,n\pi\right)\right]$$

Thus
$$x = \frac{\sqrt{3}}{2}\ln 2 - \frac{3\pi}{4} - 2n\pi$$
$$y = \frac{1}{2}\ln 2 + \frac{3\sqrt{3}\,\pi}{4} + 2\sqrt{3}\,n\pi.$$

EXERCISE 1.4

1. Prove that
   (i) $\text{Log}\, i = \dfrac{\pi i}{2}$ (ii) $\text{Log}(-5) = \ln 5 + \pi i$
   (iii) $\text{Log}(-1+i) = \dfrac{1}{2}\ln 2 + \dfrac{3\pi i}{4}$ (iv) $\text{Log}(1+i) = \dfrac{1}{2}\ln 2 + \dfrac{\pi i}{4}$
   (v) $\text{Log}\left(-\dfrac{1}{2} - \dfrac{\sqrt{3}}{2}i\right) = \dfrac{-2\pi}{3}i$ (vi) $\text{Log}(1-i) = \dfrac{1}{2}\ln 2 - \dfrac{\pi i}{4}$
2. Show that
   (i) $\coth^{-1} z = \dfrac{1}{2}\log\left(\dfrac{z+1}{z-1}\right)$ (ii) $\text{sech}^{-1} z = \log\left(\dfrac{1+\sqrt{1-z^2}}{z}\right)$
   (iii) $\text{csch}^{-1} z = \log\left(\dfrac{1+\sqrt{z^2+1}}{z}\right)$
3. Prove that
   (i) $\cos^{-1} z = \dfrac{1}{i}\log\left(z + \sqrt{z^2-1}\right)$ (ii) $\tan^{-1} z = \dfrac{1}{2i}\log\left(\dfrac{1+iz}{1-iz}\right)$
   (iii) $\cot^{-1} z = \dfrac{1}{2i}\log\left(\dfrac{z+i}{z-i}\right)$ (iv) $\sec^{-1} z = \dfrac{1}{i}\log\left(\dfrac{1+\sqrt{1-z^2}}{z}\right)$
   (v) $\csc^{-1} z = \dfrac{1}{i}\log\left(\dfrac{i+\sqrt{z^2-1}}{z}\right)$
4. Prove that
   (i) $i^i = e^{-\pi/2}$ (ii) $(-1)^i = e^{-\pi}$
   (iii) $(-i)^{-i} = e^{-\pi/2}$ (iv) $a^i = \cos(\ln a) + i\sin(\ln a),\ a > 0$
5. Prove that $\tanh^{-1} z = \sinh^{-1}\left(\dfrac{z}{\sqrt{1-z^2}}\right)$
6. Show that if $z = x + iy$, then $\log\left(\dfrac{z}{\bar{z}}\right) = 2i\tan^{-1}\left(\dfrac{y}{x}\right)$

Imp

7. If $a^{\alpha+i\beta} = (x+iy)^{p+iq}$, $a > 0$, prove that
   2012 (i) $\alpha = \dfrac{1}{2}p\log_a(x^2+y^2) - q\tan^{-1}\left(\dfrac{y}{x}\right)\log_a e$
   (ii) $\log_a(x^2+y^2) = \dfrac{2(\alpha p + \beta q)}{p^2+q^2}$

8. If $\operatorname{Log}\sin(x+iy) = u+iv$, show that

(i) $\cosh 2y = \cos 2x + 2e^{2u}$ (ii) $e^{2y} = \dfrac{\cos(x-v)}{\cos(x+v)}$

9. Show that

$$\operatorname{Log}(1+\cos\theta + i\sin\theta) = \ln(2\cos\theta/2) + i\theta/2$$

10. Prove that

$$\tan^{-1}\left(\frac{x+iy}{x-iy}\right) = \frac{\pi}{4} + \frac{i}{2}\ln\frac{x+y}{x-y}, \text{ if } x>y>0$$

11. Prove that

(i) $\sec(x+iy) = 2\dfrac{\cos x\cosh y + i\sin x\sinh y}{\cos 2x + \cosh 2y}$

(ii) $\cos^{-1}(\cos\theta + i\sin\theta) = \sin^{-1}\sqrt{\sin\theta} + i\ln\left(\sqrt{1+\sin\theta} - \sqrt{\sin\theta}\right)$

11 (iii) $\tan^{-1}(\cos\theta + i\sin\theta) = \pm\dfrac{\pi}{4} + \dfrac{i}{4}\ln\dfrac{1+\sin\theta}{1-\sin\theta}$

SUMMATION OF SERIES

(1.26) We can find the sums of certain types of trigonometric series with the help of De Moivre's Theorem. The method is illustrated by means of examples. The following infinite expansions (whose proofs are beyond the scope of the book) are useful in finding the sums of trigonometric series:

(i) $e^z = 1 + z + \dfrac{z^2}{2!} + \dfrac{z^3}{3!} + \cdots + \dfrac{z^n}{n!} + \cdots$

(ii) $\sin z = z - \dfrac{z^3}{3!} + \dfrac{z^5}{5!} - \cdots + (-1)^n\dfrac{z^{2n+1}}{(2n+1)!} + \cdots$

(iii) $\cos z = 1 - \dfrac{z^2}{2!} + \dfrac{z^4}{4!} - \cdots + (-1)^n\dfrac{z^{2n}}{(2n)!} + \cdots$

(iv) $(1+z)^n = 1 + nz + \dfrac{n(n-1)}{2!}z^2 + \cdots$

$$+\frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}z^r + \cdots, \qquad |z|<1$$

(v) $\log(1+z) = z - \dfrac{z^2}{2} + \dfrac{z^3}{3} - \cdots + (-1)^n\dfrac{z^{n+1}}{n+1} + \cdots$

(vi) $\dfrac{1}{1-z} = (1-z)^{-1} = 1 + z + z^2 + \cdots + z^n + \cdots, \qquad |z|<1.$

Example 25. Find the sum of the infinite series

$$\sin\alpha + c\sin(\alpha+\beta) + c^2\sin(\alpha+2\beta) + \cdots, \qquad |c|<1.$$

Solution. Let

$$C = \cos\alpha + c\cos(\alpha+\beta) + c^2\cos(\alpha+2\beta) + \cdots$$

and

$$S = \sin\alpha + c\sin(\alpha+\beta) + c^2\sin(\alpha+2\beta) + \cdots$$

Then

$$C + iS = (\cos\alpha + i\sin\alpha) + c\left[\cos(\alpha+\beta) + i\sin(\alpha+\beta)\right] + c^2\left[\cos(\alpha+2\beta) + i\sin(\alpha+2\beta)\right] + \cdots$$

$$= e^{i\alpha} + c\,e^{i(\alpha+\beta)} + c^2 e^{i(\alpha+2\beta)} + \cdots \qquad (1)$$

Now the series on the right of (1) is an infinite geometric series whose common ratio is $c\,e^{i\beta}$ and $|c\,e^{i\beta}| = |c| < 1$.

Therefore,

$$C + iS = \frac{e^{i\alpha}}{1-c\,e^{i\beta}} = \frac{\cos\alpha + i\sin\alpha}{1 - c\cos\beta - i\,c\sin\beta}$$

$$= \frac{(\cos\alpha + i\sin\alpha)(1 - c\cos\beta + i\,c\sin\beta)}{(1-c\cos\beta)^2 + c^2\sin^2\beta}$$

$$= \frac{\cos\alpha(1-c\cos\beta) - c\sin\alpha\sin\beta + i[\sin\alpha(1-c\cos\beta) + c\cos\alpha\sin\beta]}{1-2c\cos\beta+c^2}$$

Equating imaginary parts, we obtain

$$S = \frac{\sin\alpha(1-c\cos\beta) + c\cos\alpha\sin\beta}{1-2c\cos\beta+c^2}$$

Example 26. Evaluate the sum of the infinite series

$$1 + \frac{1}{2}\cos 2\theta - \frac{1}{2\cdot4}\cos 4\theta + \frac{1\cdot3}{2\cdot4\cdot6}\cos 6\theta - \cdots$$

Solution. Let

$$C = 1 + \frac{1}{2}\cos 2\theta - \frac{1}{2\cdot4}\cos 4\theta + \frac{1\cdot3}{2\cdot4\cdot6}\cos 6\theta - \cdots$$

and

$$S = \frac{1}{2}\sin 2\theta - \frac{1}{2\cdot4}\sin 4\theta + \frac{1\cdot3}{2\cdot4\cdot6}\sin 6\theta - \cdots$$

Then $$C+iS = 1+\frac{1}{2}(\cos 2\theta + i\sin 2\theta) - \frac{1}{2\cdot4}(\cos 4\theta + i\sin 4\theta)$$

$$+\frac{1\cdot3}{2\cdot4\cdot6}(\cos 6\theta + i\sin 6\theta) - \cdots$$

$$= 1+\frac{1}{2}e^{2i\theta} - \frac{1}{2\cdot4}e^{4i\theta} + \frac{1\cdot3}{2\cdot4\cdot6}e^{6i\theta} - \cdots$$

The series on the right is a binomial series and it is the expansion of

$$\left[1+e^{2i\theta}\right]^{\frac{1}{2}}$$

Therefore,

$$C+iS = \left[1+e^{2i\theta}\right]^{\frac{1}{2}} = (1+\cos 2\theta + i\sin 2\theta)^{\frac{1}{2}}.$$

$$= \left[2\cos^2\theta + 2i\sin\theta\cos\theta\right]^{\frac{1}{2}}$$

$$= \sqrt{2\cos\theta}\left[\cos\theta + i\sin\theta\right]^{\frac{1}{2}}$$

$$= \sqrt{2\cos\theta}\left[\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right]$$

Hence $$C = \sqrt{2\cos\theta}\,\cos\frac{\theta}{2}$$

Example 27. Evaluate the sum of the infinite series

$$\frac{c^2}{2!}\sin 2\theta - \frac{c^4}{4!}\sin 4\theta + \frac{c^6}{6!}\sin 6\theta - \cdots$$

Solution. Let

$$S = \frac{c^2}{2!}\sin 2\theta - \frac{c^4}{4!}\sin 4\theta + \frac{c^6}{6!}\sin 6\theta - \cdots$$

$$C = 1 - \frac{c^2}{2!}\cos 2\theta + \frac{c^4}{4!}\cos 4\theta - \frac{c^6}{6!}\cos 6\theta + \cdots$$

$$C-iS = 1 - \frac{c^2}{2!}(\cos 2\theta + i\sin 2\theta) + \frac{c^4}{4!}(\cos 4\theta + i\sin 4\theta)$$

$$-\frac{c^6}{6!}(\cos 6\theta + i\sin 6\theta) + \cdots$$

$$= 1 - \frac{c^2}{2!}e^{2i\theta} + \frac{c^4}{4!}e^{4i\theta} - \frac{c^6}{6!}e^{6i\theta} + \cdots$$

$$= \cos(c\,e^{i\theta})$$

$$= \cos(c\cos\theta + i\,c\sin\theta)$$

$$= \cos(c\cos\theta)\cosh(c\sin\theta) - i\sin(c\cos\theta)\sinh(c\sin\theta)$$

Thus $$S = \sin(c\cos\theta)\sinh(c\sin\theta).$$

Example 28. Evaluate the sum of the infinite series 11

$$\cos\theta - \frac{1}{2}\cos 2\theta + \frac{1}{3}\cos 3\theta - \frac{1}{4}\cos 4\theta + \cdots$$

Solution. Let

$$C = \cos\theta - \frac{1}{2}\cos 2\theta + \frac{1}{3}\cos 3\theta - \frac{1}{4}\cos 4\theta + \cdots$$

and $$S = \sin\theta - \frac{1}{2}\sin 2\theta + \frac{1}{3}\sin 3\theta - \frac{1}{4}\sin 4\theta + \cdots$$

$$C+iS = \operatorname{cis}\theta - \frac{1}{2}\operatorname{cis}2\theta + \frac{1}{3}\operatorname{cis}3\theta - \frac{1}{4}\operatorname{cis}4\theta + \cdots$$

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

$$= \log(1+e^{i\theta})$$

$$= \log(1+\cos\theta + i\sin\theta)$$

$$= \sqrt{(1+\cos\theta)^2 + \sin^2\theta} + i\tan^{-1}\left(\frac{\sin\theta}{1+\cos\theta}\right)$$

$$= \ln\sqrt{4\cos^2\theta/2} + i\tan^{-1}\frac{2\sin\theta/2\cos\theta/2}{2\cos^2\theta/2}$$

$$= \ln 2 + \ln(\cos\theta/2) + i\,\theta/2$$

$$C = \ln 2 + \ln(\cos\theta/2)$$

Example 29. Prove that

(i) $$\cos\theta + \cos(\theta+\alpha) + \cdots + \cos(\theta+n\alpha) = \frac{\sin\frac{n-1}{2}\alpha}{\sin\frac{\alpha}{2}} \cdot \cos\left(\frac{\theta+n\alpha}{2}\right)$$

(ii) $$\sin\theta + \sin(\theta+\alpha) + \cdots + \sin(\theta+n\alpha) = \frac{\sin\frac{n-1}{2}\alpha}{\sin\frac{\alpha}{2}} \cdot \sin\left(\theta+\frac{n\alpha}{2}\right)$$

Solution. Let

$$C = \cos\theta + \cos(\theta+\alpha) + \cdots + \cos(\theta+n\alpha)$$

and $$S = \sin\theta + \sin(\theta+\alpha) + \cdots + \sin(\theta+n\alpha)$$

$$C+iS = \text{cis}\,\theta + \text{cis}(\theta+\alpha) + \cdots + \text{cis}(\theta+n\alpha)$$

$$= e^{i\theta} + e^{i(\theta+\alpha)} + \cdots + e^{i(\theta+n\alpha)}$$

$$= e^{i\theta} \cdot \frac{e^{i(n+1)\alpha/2}\left(e^{-i(n+1)\alpha/2} - e^{i(n+1)\alpha/2}\right)}{e^{i\alpha/2}\left(e^{-i\alpha/2} - e^{i\alpha/2}\right)}$$

$$= e^{i(\theta+n\alpha/2)}\left[\frac{-2i\sin(n+1)\,\alpha/2}{-2i\sin\alpha/2}\right]$$

$$= \frac{\sin(n+1)\,\alpha/2}{\sin\alpha/2}\left[\cos\left(\theta+\frac{n\alpha}{2}\right) + i\sin\left(\theta+\frac{n\alpha}{2}\right)\right]$$

Hence (i) $$C = \frac{\sin(n+1)\,\alpha/2}{\sin\alpha/2}\cos\left(\theta+\frac{n\alpha}{2}\right)$$

and (ii). $$S = \frac{\sin(n+1)\,\alpha/2}{\sin\alpha/2}\sin\left(\theta+\frac{n\alpha}{2}\right)$$

EXERCISE 1.5

In each of **Problems 1 – 5**, evaluate the indicated sum:

1. (i) $\sin A + \sin 2A + \sin 3A + \cdots + \sin nA$

   (ii) $\cos A + \cos 2A + \cos 3A + \cdots + \cos nA$

2. $\cos\theta + \cos 3\theta + \cos 5\theta + \cdots + \cos(2n-1)\theta$

3. $1 + x\cos\theta + x^2\cos 2\theta + \cdots + x^n\cos n\theta$

4. $3\sin\alpha + 5\sin 2\alpha + 7\sin 3\alpha + \cdots + (2n+1)\sin n\alpha$

5. $\cos^2\theta + \cos^2 2\theta + \cos^2 3\theta + \cdots + \cos^2 n\theta.$

2, 12

17th Ramadan
11th August
Friday

In **Problems 6 – 15**, find the sum of each infinite series:

6. $\sin\theta + \frac{1}{2}\sin 3\theta + \frac{1\cdot 3}{2\cdot 4}\sin 5\theta + \cdots$

7. $\sinh\theta + \frac{\sinh 2\theta}{2!} + \frac{\sinh 3\theta}{3!} + \cdots$

8. $\sin\alpha\sin\alpha + \sin^2\alpha\sin 2\alpha + \sin^3\alpha\sin 3\alpha + \cdots$

9. $1 - \frac{1}{2}\cos\theta + \frac{1\cdot 3}{2\cdot 4}\cos 2\theta - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cos 4\theta + \cdots$

10. $n\sin\theta + \frac{n(n+1)}{2!}\sin 2\theta + \frac{n(n+1)(n+2)}{3!}\sin 3\theta + \cdots$

11. $1 + \frac{1}{2}\cos\theta + \frac{1\cdot 3}{2\cdot 4}\cos 2\theta + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cos 3\theta + \cdots$

12. $\cos\alpha - \frac{\cos(\alpha+2\beta)}{3!} + \frac{\cos(\alpha+4\beta)}{5!} - \cdots$

13. $1 + c\cos\theta + \frac{c^2}{2!}\cos 2\theta + \frac{c^3}{3!}\cos 3\theta + \cdots$

14. $c\cos\theta + \frac{c^2}{2}\cos 2\theta + \frac{c^3}{3}\cos 3\theta + \cdots$

15. $\sin\theta - \frac{1}{2}\sin 3\theta + \frac{1}{3}\sin 5\theta - \cdots.$

As our own elders, may Allah have mercy upon them, used to say, "The study of mathematics is for the mind like soap for the clothes, which washes away from them dirt and cleans the spots and stains".

ABDAL RAHMAN IBN-E-KHALDUN
(Muslim Historian and Philosopher 1332-1406 C.E.)

Chapter 2

GROUPS

The concept of a binary operation on a nonempty set has already been explained in the previous classes. One may recall that a **binary operation** on a nonempty set A is just a function $\star : A \times A \longrightarrow A$. So, for each (a, b) in $A \times A$, $\star$ associates an element $\star(a, b)$ of A. We shall denote $\star(a, b)$ by $a \star b$. If A is a nonempty set with a binary operation $\star$, then A is said to be **closed** under $\star$.

DEFINITIONS AND EXAMPLES

We define a group as follows: binary operation

(2.1) Definition. A pair $(G, \star)$, where G is a nonempty set and $\star$ is a binary operation on G, is called a **group** if the following conditions, called **axioms of a group**, are satisfied in G:

(i) The binary operation $\star$ is **associative**. That is,

$$(a \star b) \star c = a \star (b \star c) \text{ for all } a, b, c, \in G.$$

under multiplication $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

(ii) There is an element e in G such that

$$a \star e = e \star a = a \quad \text{for all } a \in G.$$

e is called **the identity** element of G. 0 multiplication no.

(iii) For each $a \in G$, there is an $a' \in G$ such that

$$a \star a' = a' \star a = e.$$

a' is called **the inverse** of a.

Note: Use of word **the** before identity element and inverse of an element is to signify their uniqueness.

(2.2) Definition. A group $(G, \bullet)$ is said to be **abelian**[1] or **commutative** if

$$a \bullet b = b \bullet a \quad \text{for all } a, b \in G.$$

If there is a pair of elements $a, b \in G$ such that $a \bullet b \neq b \bullet a$, then G is called a **non-abelian** group.

Usually the binary operation in a group is either denoted by ., called **multiplication** or by +, called **addition**. The pair $(G, .)$ then denotes **group under multiplication** while $(G, +)$ denotes a **group under addition.** In $(G, .)$ the inverse of an element a is written a^{-1}, while in $(G, +)$ the inverse of a is written as $-a$.

In practice, the product $a \cdot b$ of two elements in a group G under multiplication is written simply as ab. Also, we shall denote a group $(G, .)$ by G only.

(2.3) Definition. An element x of a group G is said to be **idempotent** if

$$x^2 = x$$

(2.4) Theorem. The only idempotent element in a group G is the identity element.

Proof. Let $x \in G$ be an idempotent element. Then

$$
\begin{aligned}
& x^2 = x \\
\Rightarrow\quad & x^{-1} \cdot x^2 = x^{-1} \cdot x = e \\
\Rightarrow\quad & x^{-1} \cdot x \cdot x = e \\
\Rightarrow\quad & e \cdot x = e \\
\text{Thus}\quad & x = e.
\end{aligned}
$$

Example 1. Consider the set

$$G = \{1, -1\}$$

and let the binary operation defined on G be the ordinary multiplication of real numbers. Then $(G, .)$ is a group.

Example 2. The pairs $(Z, +)$, $(Q, +)$, $(R, +)$ and $(C, +)$ where Z, Q, R and C are the sets of integers, rational numbers, real numbers and complex numbers respectively and + denotes ordinary addition in them, are all groups.

To verify that a finite set is a group it is some times convenient to list the products in the form of a table called **Cayley's group table**[2]. This is illustrated by the following examples:

Example 3. (Group Tables). Let $G = \{1, \omega, \omega^2\}$, where ω is a complex cube root of unity. We write the elements of G along a row and along a column as shown in the

1. Named for the Norwegian mathematician N.H. Abel (1802 – 1829).
2. Named for the British mathematician A. Cayley (1821 – 1895).

table below indicating the binary operation in the top left corner. The column headed by a_j in the upper row is called the jth column and the row with a_i in the left column is referred to as the ith row. The blanks are then filled in by writing in the ijth position the product of an element a_i in the ith row with the element a_j in the jth column.

Thus, for $G = \{1, \omega, \omega^2\}$, we have the following table. Here the binary operation '.' is the multiplication of complex numbers

.	1	ω	ω^2
1	1	ω	ω^2
ω	ω	ω^2	1

Here we have used the fact that $\omega^3 = 1$. It is now easy to verify the conditions for the group like closure law, associative law, etc., from this table.

A group is abelian if its table is symmetric about its main diagonal.

Example 4. The set $C_4 = \{1, -1, i, -i\}$ of all the fourth roots of unity is a group under the usual multiplication of complex numbers. Here

$$(i)^{-1} = -i \quad \text{and} \quad (-i)^{-1} = i.$$

In general, the set C_n of all the nth roots of unity, for a fixed natural number n, forms a group under multiplication. The elements of C_n are

$$e^{2k\pi i/n}, \quad k = 0, 1, 2, \cdots, n-1.$$

Example 5. Let $Q \setminus \{0\}$, $R \setminus \{0\}$ and $C \setminus \{0\}$ denote the sets of nonzero rational numbers, nonzero real numbers and nonzero complex numbers respectively. Then under the usual multiplication of real and complex numbers $(Q \setminus \{0\}, .)$, $(R \setminus \{0\}, .)$ and $(C \setminus \{0\}, .)$ are all groups.

Query: Why 0 has been deleted from the respective sets?

Example 6. Let

$$G = \{I, -I, i, -i, j, -j, k, -k\}$$

where the symbols satisfy the relations

$$
\begin{aligned}
ij &= k, & jk &= i, & ki &= j \\
ji &= -k, & kj &= -i, & ik &= -j
\end{aligned}
$$

and $\quad i^2 = j^2 = k^2 = -I.$

Then, under the multiplication of the symbols defined defined above, G is a group. Since in G

$$ij = k \neq -k = ji,$$

G is non-abelian. G is called the **group of quaternions.**

Example 7. Let G be the set of all 2×2 nonsingular real matrices. Then, under the usual multiplication of matrices, G is a group. Moreover, it is a non-abelian group. For example,

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

are in G and

$$AB = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \quad BA = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

Thus $AB \neq BA$ and so G is not an abelian group.

Example 8. Let $\bar{Z}_5 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}$ be the set of residue classes modulo 5. Define $+$ in $\bar{Z}_5$ as the addition modulo 5. Thus, for $\bar{a}, \bar{b} \in \bar{Z}_5$, $\bar{a} + \bar{b} = \bar{r}$ where r is the remainder obtained after division of $a + b$ by 5. Then $(\bar{Z}_5, +)$ is a group.

Example 9. Let $S = \{\bar{1}, \bar{2}, \bar{3}, \bar{4}\}$ be the set of nonzero residue classes modulo 5 and define multiplication in S as multiplication modulo 5. Thus if $\bar{a}, \bar{b} \in S$, then $\bar{a} \, . \, \bar{b} = \bar{r}$, where r is the remainder obtained after dividing the usual product ab of a and b by 5. Then $(S, .)$ is a group.

Example 10. Let

$$G = \{I, -I, X, -X, Y, -Y, Z, -Z\}$$

be the set of 2×2 matrices where

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, X = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, Y = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, Z = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

It is easy to check that

$$X^2 = Y^2 = Z^2 = -I \quad \text{and}$$

$$YZ = X, \ ZY = -X, \ ZX = Y, \ XZ = -Y, \ XY = Z, \ YX = -Z.$$

G is a group under the usual multiplication of matrices. It is the 'same' group as the group of Example 6.

The concept of 'sameness' among groups is related with that of **isomorphism** of groups.

PROPERTIES OF GROUPS

(2.5) Theorem. (The Cancellation Laws). For any three elements a, b, c in a group G,

(i) $ab = ac$ implies $b = c$ **(Left Cancellation Law)**

(ii) $ba = ca$ implies $b = c$ **(Right Cancellation Law)**

Proof. (i) For a, b, c in G,

$$\begin{aligned} ab = ac \quad &\Rightarrow \quad a^{-1}(ab) = a^{-1}(ac) \\ &\Rightarrow \quad (a^{-1}a)\,b = (a^{-1}a)\,c, \quad \text{using the associative law} \\ &\Rightarrow \quad eb = ec, \quad e \text{ is the identity of } G. \\ &\Rightarrow \quad b = c \end{aligned}$$

Thus the left cancellation law holds.

(ii) The proof similar to (i) and is left as an exercise.

(2.6) Theorem. (Solutions of Linear Equations). For any two elements a, b in a group G, the equations

$$ax = b \quad \text{and} \quad xa = b \quad \text{have unique solutions.}$$

Proof. For a, b in G,

$$\begin{aligned} ax = b \quad &\Rightarrow \quad a^{-1}(ax) = a^{-1}b \\ &\Rightarrow \quad (a^{-1}a)\,x = a^{-1}b, \text{ by the associative law} \\ &\Rightarrow \quad ex = a^{-1}b \\ &\Rightarrow \quad x = a^{-1}b. \end{aligned}$$

So $x = a^{-1}b$ is a solution of $ax = b$.

To see that the solution is unique, suppose that, for x_1, x_2 in G,

$$ax_1 = b, \qquad ax_2 = b$$

Then $\quad ax_1 = ax_2$

so that, by the cancellation law, $x_1 = x_2$. Hence the solution is unique.

The case for the solution of $xa = b$ is similar.

(2.7) Theorem. For a, b in a group G, $(ab)^{-1} = b^{-1}a^{-1}$.

Proof. This follows from the following simplifications of the product $(ab)(b^{-1}a^{-1})$.

$$
\begin{aligned}
(ab)(b^{-1}a^{-1}) &= a(bb^{-1})a^{-1} && \textbf{(Associative Law)}\\
&= a\,e\,a^{-1} \\
&= aa^{-1} && (e \text{ is the identity})\\
&= e
\end{aligned}
$$

Similarly

$$
\begin{aligned}
(b^{-1}a^{-1})(ab) &= b^{-1}(a^{-1}a)\,b && \textbf{(Associative Law)}\\
&= b^{-1}e\,b && e \text{ is the identity}\\
&= b^{-1}b \\
&= e
\end{aligned}
$$

Hence $\qquad (ab)^{-1} = b^{-1}a^{-1}$

Remark: In general, for $a_1, a_2, \cdots, a_k$ in G, we have

$$(a_1 a_2 \cdots a_k)^{-1} = a_k^{-1}\, a_{k-1}^{-1} \cdots a_2^{-1}\, a_1^{-1}$$

(2.8) Theorem For any element a of a group G, the following exponentiation rules hold $(m, n \in Z^+)$

(i) $a^m = a \cdot a \cdot \cdots a$ (m factors)

(ii) $(a^{-1})^m = a^{-m} = (a^m)^{-1}$

(iii) $a^m \cdot a^n = a^{m+n}$

(iv) $(a^m)^n = a^{mn}$

The proofs follow by induction on m and n and are left as an exercise.

(2.9) Definition (Order of a group). The number of elements in a group G is called the **order** of G and is denoted by $|G|$. A group G is said to be **finite** if G consists of only a finite number of elements. Otherwise G is said to be an **infinite** group.

The groups in Examples 1, 3, 4, 6, 8, 9 and 10 are finite groups.

The orders of these groups are 2, 3, 4, 8, 5, 4 and 8 respectively.

The groups in Examples 2, 5 and 7 are infinite groups.

(2.10) Definition (Order of an Element). Let a be an element of a group G. A positive integer n is said to be the **order** of a if $a^n = e$ and n is the least such positive integer. (e is the identity element of G)

For any element x of G we always take $x^0 = e$. If $n = 0$ is the only integer for which $a^n = e$ then a is said to be of **infinite order**.

The order of an element $a \in G$ is denoted by $|a|$

(2.11) Theorem. Let G be any group and let $a \in G$ have order n. Then, for any integer k, $a^k = e$ if and only if $k = nq$, where q is an integer.

Proof. Suppose that n is the order of a and, for some integer k, $a^k = e$. By the division algorithm[1], there are unique integers q and r such that

$$k = nq + r, \quad 0 \le r < n$$

so that

$$e = a^k = (a)^{nq+r} = (a^n)^q \cdot a^r = e \cdot a^r = a^r \qquad (\text{as } a^n = e)$$

Since a has order n, therefore, n is the smallest integer for which $a^n = e$ and so $r = 0$

Thus $\quad k = nq$

Conversely, suppose that $\quad k = nq$. Then

$$a^k = e^{nq} = (a^n)^q = e^q = e.$$

Example 11. Show that the set

$$S = \{\bar{1}, \bar{3}, \bar{5}, \bar{7}\}$$

is a group under multiplication modulo 8. Find the order of each element of S.

Solution. The Cayley's table for S under multiplication modulo 8 is as given below:

$\cdot$	$\bar{1}$	$\bar{3}$	$\bar{5}$	$\bar{7}$
$\bar{1}$	$\bar{1}$	$\bar{3}$	$\bar{5}$	$\bar{7}$
$\bar{3}$	$\bar{3}$	$\bar{1}$	$\bar{7}$	$\bar{5}$
$\bar{5}$	$\bar{5}$	$\bar{7}$	$\bar{1}$	$\bar{3}$
$\bar{7}$	$\bar{7}$	$\bar{5}$	$\bar{3}$	$\bar{1}$

For example, here $5 \times 3 = 15$ which gives 7 as remainder after division by 8.

So $\qquad \bar{5} \times \bar{3} = \bar{7}$

Similarly, $\qquad \bar{3} \times \bar{7} = \bar{5}$

One may observe from the multiplication table that $\bar{1}$ is the identity element in S. The closure law and the existence of inverse of each element can easily be verified from the table.

[1] The division algorithm states that, for any two integers a and b with $a > 0$, there are integers q and r such that

$$b = aq + r, \quad 0 \le r < a$$

For the orders of elements, we see that the order of $\bar{1}$ is 1.

The order of $\bar{3}$ is 2 because $\bar{3} \times \bar{3} = \bar{1}$.

Similarly, the order of each of $\bar{5}$ and $\bar{7}$ is 2.

Example 12. Let G be a group and $a, b \in G$. Show that

(i) The orders of a and a^{-1} are equal

(ii) The orders of a and $b^{-1}ab$ are equal

(iii) The orders of ab and ba are equal

Solution. (i) Let $a \in G$. Suppose the order of a is m so that $a^m = e$.

$$\text{Now, } a^m = e \quad \Leftrightarrow \quad a^{-m} . a^m = a^{-m} . e$$
$$\Leftrightarrow \quad e = (a^{-1})^m$$

So the order of a^{-1} is a divisor k of m.

But $(a^{-1})^k = e$ implies $a^{-k} . a^k = e . a^k = a^k$

Thus $e = a^k$.

But then m divides k. Thus $k = m$. So the order of a^{-1} is also m.

(ii) Let $a, b \in G$. Suppose, the order of a is m then $a^m = e$. Therefore,

$$a^m = e \quad \Leftrightarrow \quad b^{-1} a^m b = b^{-1} eb$$
$$\Leftrightarrow \quad (b^{-1} ab)^m = e$$

Hence the orders of a and $b^{-1} ab$ are equal.

[Here we have used the fact that $(b^{-1} ab)^m = b^{-1} a^m b$ which is proved by induction on m as follows.

The result is true for $m = 1$. Suppose it is true for $m = k$, i.e., $(b^{-1} ab)^k = b^{-1} a^k b$.

Now $(b^{-1} ab)^{k+1} = (b^{-1} ab)^k (b^{-1} ab) = (b^{-1} a^k b)(b^{-1} ab) = b^{-1} a^k e ab = b^{-1} a^{k+1} b$

Hence $(b^{-1} ab)^m = b^{-1} a^m b$ for all $m \in N$]

(iii) Suppose $|a b| = m$.

$$\text{Now, } a b = b^{-1} b a b = b^{-1} (b a) b, \qquad \textbf{(Associative Law)}$$
$$\text{or} \quad (a b)^m = e = \left[b^{-1} (b a) b \right]^m$$
$$= b^{-1} (b a)^m b, \qquad \text{as in } \textbf{(ii)}$$
$$\text{or} \quad b e b^{-1} = b b^{-1} (b a)^m b b^{-1}$$
$$\text{or} \quad e = (b a)^m$$
$$\text{Thus} \quad |b a| = m.$$

Example 13 Let G be a group of even order. Prove that there is at least one element of order 2 in G.

Solution. Let G be a group of even order. Then the non-identity elements in G will be odd in number. Also the inverse of each element of G belongs to G and that $e^{-1} = e$.

There occur pairs each consisting of some non-identity element x and x^{-1} in G such that $x \neq x^{-1}$. As there are odd number of non-identity elements in G, after pairing off such non-identity elements for which $x \neq x^{-1}$, we must have at least one element $a\ (\neq e) \in G$ such that

$$a = a^{-1}$$
$$\text{But then} \quad aa = aa^{-1}$$
$$\text{or} \quad a^2 = e$$
$$\text{Hence} \quad |a| = 2.$$

Example 14. Let G be a group and x be an element of odd order in G. Then there exists an element y in G such that $y^2 = x$.

Solution. For some nonnegative integer m and $x \in G$, let $|x| = 2m + 1$,

so that we have $x^{2m+1} = e$ (1)

Clearly $x, x^2, \cdots, x^m, x^{m+1}, \cdots, x^{2m} \in G$

Let $y = x^{m+1}$. Then

$$y^2 = x^{2m+2} = x^{2m+1} x = ex = x, \text{ by (1)}$$

EXERCISE 2.1

1. Answer true or false. Justify your answer.

(i) A group can have more that one identity element.

(ii) The null set can be considered to be a group.

(iii) There may be groups in which the cancellation law fails.

(iv) Every set of numbers which is group under addition is also a group under multiplication and vice versa.

(v) The set R of all real numbers is a group with respect to subtraction.

(vi) The set of all nonzero integers is a group with respect to division.

(vii) To each element of a group, there does not correspond an inverse element.

(viii) To each element of a group, there corresponds only one inverse element.

(ix) To each element of a group, there correspond more that one inverse elements.

2. Show that in a group G
   (i) the identity element is unique
   (ii) the inverse of each element is unique

3. Which of the following sets are groups and why?
   (i) The set of all positive rational numbers under multiplication
   (ii) The set of all complex numbers z such that $|z| = 1$, under multiplication defined for complex numbers
   (iii) The set Z of all integers under binary operation o defined by

   $$a \, o \, b = a - b \qquad \text{for all } a, b \in Z$$

   (iv) The set Q' of all irrational numbers under multiplication
   (v) $R^{+} = \{ x \in R : x > 0 \}$ under multiplication
   (vi) $R^{-} = \{ x \in R : x < 0 \}$ under multiplication
   (vii) $E = \{ e^{x} : x \in R \}$ under multiplication

4. Show that the set $\{ \bar{1}, \bar{2}, \bar{4}, \bar{5}, \bar{7}, \bar{8} \}$ under multiplication modulo 9 is a group

5. Is (Z, o) a group? where o is defined by $a \, o \, b = 0$ for all $a, b \in Z$

6. Show that the matrices:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

form a group under matrix multiplication

7. Prove that the set of complex-valued functions I, f, g and h defined on the set $C \setminus \{ 0 \}$ of nonzero complex numbers by:

$$I(z) = z, \; f(z) = -z, \; g(z) = 1/z, \; h(z) = -1/z, \; z \in C \setminus \{ 0 \}$$

forms a group under composition of functions $\quad (g \, o \, f)(z) = g(f(z))$

8. Show that the set

$$G = \{ 2^{k} : k = 0, \pm 1, \pm 2, \cdots \}$$

is a group under multiplication

9. Show that if a group G is such that $x \cdot x = e$, for all $x \in G$, where e is the identity element of G, then G is an abelian group

10. If a group G has three elements, show that it is abelian

11. If every element of a group G is its own inverse, show that G is abelian

12. Prove that if every non-identity element of a group G is of order 2, then G is abelian

13. In a group G, let a, b and ab all have order 2. Show that $ab = ba$

14. Show that a group G is abelian if and only if $(ab)^{2} = a^{2}b^{2}$ for all $a, b \in G$

15. Suppose that a group G has only one element a of order 2. Show that, for all $x \in G$, $ax = xa$

16. Let G be a group such that $(ab)^{n} = a^{n} b^{n}$ for three consecutive natural numbers n and all a, b in G. Show that G is abelian

17. If G is an abelian group, show that

$$(ab)^{n} = a^{n} b^{n} \qquad \text{for all } a, b \in G$$

18. Show that the set $GL_{2}(R)$ of all 2×2 nonsingular matrices over R is a group under the usual multiplication of matrices.

(This group is called the **general linear group of degree 2**)

SUBGROUPS

(2.12) Definition. Let $(G, .)$ be a group and H be a nonempty subset of G. If H is itself a group with the binary operation of G restricted to H, then H is called a **subgroup** of G

Example 15. $(Z, +)$ is a subgroup of $(Q, +)$ and $(Q, +)$ is a subgroup of $(R, +)$

Example 16. The set of cube roots of unity forms a subgroup of $C \setminus \{ 0 \}$, where $C \setminus \{ 0 \}$ is the group of nonzero complex numbers under multiplication of complex numbers

Example 17. Every group G has at least two subgroups namely G itself and the identity group $\{ e \}$. These are called **trivial subgroups**. Any other subgroup of G is called a **nontrivial** subgroup of G

The following theorem establishes an easy criterion for determining whether or not a subset H of a group G is a subgroup of G.

(2.13) Theorem. Let $(G, .)$ be a group. Then a nonempty subset H of G is a subgroup if and only if, for $a, b \in H$, the element $ab^{-1} \in H$.

Proof. Suppose that H is a subgroup of G. Then for all $a, b \in H$, $a, b^{-1} \in H$. Hence $ab^{-1} \in H$ by the closure law in H.

Conversely, suppose that for all $a, b \in H$, $ab^{-1} \in H$.

Now $a, a^{-1} \in H$. Since $ab^{-1} \in H$ for $a, b \in H$, so $e = aa^{-1} \in H$. Since e is the identity element of G, it is the identity element of H as well. Let $a \in H$. Then [illegible] So each element of H has an inverse in H.

Next let $a, b \in H$. Then $a, b^{-1} \in H$. Hence

$$ab = a(b^{-1})^{-1} \in H$$

Thus H is closed with respect to the binary operation in G restricted to H. The associative law holds for elements of H as it holds, in general, for the elements of G. Hence all the axioms of a group are satisfied by the elements of H. So H is a group under the binary operation of G restricted to H. Thus it is a subgroup of G.

(2.14) Theorem. The intersection of any collection of subgroups of a group G is a subgroup of G.

Proof. Let $\{H_i : i \in I\}$ be a family of subgroups of a group G. Let $H = \cap \{H_i : i \in I\}$. Let $a, b \in H$ then $a, b \in H_i$ for each $i \in I$. Since each H_i is a subgroup of G, $ab^{-1} \in H_i$ for each $i \in I$. Therefore,

$$ab^{-1} \in \cap \{H_i : i \in I\} = H$$ Hence H is a subgroup of G.

(2.15) Theorem. Let G be a group and H a subgroup of G. Then the set

$$aHa^{-1} = \{aha^{-1} : h \in H\}$$ is a subgroup of G.

Proof. Here, for $x, y \in aHa^{-1}$. Then

$$x = ah_1a^{-1},\ y = ah_2a^{-1},\ h_1, h_2 \in H.$$

So
$$xy^{-1} = (ah_1a^{-1})(ah_2a^{-1})^{-1}$$
$$= ah_1a^{-1} \cdot ah_2^{-1}a^{-1} = ah_1 e h_2^{-1}a^{-1} = ah_1h_2^{-1}a^{-1}$$

Since H is a subgroup and $h_1, h_2 \in H$, $h_1h_2^{-1} \in H$. So

$$xy^{-1} = ah_1h_2^{-1}a^{-1} \in aHa^{-1}$$

Hence aHa^{-1} is a subgroup

(2.16) Remark. For two subgroups H and K of a group G, their union $H \cup K$ need not be a subgroup of G. For example, let

$$G = \{\bar{1}, \bar{3}, \bar{5}, \bar{7}\}$$

be the group of residue classes under multiplication modulo 8. Then

$$H = \{\bar{1}, \bar{3}\},\quad K = \{\bar{1}, \bar{5}\}$$

are both subgroups of G. However, their union

$$H \cup K = \{\bar{1}, \bar{3}, \bar{5}\}$$

is not a subgroup of G because

$\bar{3} \cdot \bar{5} = \bar{7}$ and $\bar{7}$ does not belong to $H \cup K$

The following theorem gives a necessary and sufficient condition for the union of two subgroups H and K of a group G to be subgroup of G

(2.17) Theorem. The union $H \cup K$ of two subgroups H and K of a group G is a subgroup of G if and only if either $H \subset K$ or $K \subset H$

Proof. Let H and K be subgroups of a group G. Suppose that $H \subset K$ or $K \subset H$. Then

$$H \cup K = K \quad \text{or} \quad H \cup K = H$$

Since both K and H are subgroups of G, so $H \cup K$ is a subgroup of G

Conversely, suppose that $H \cup K$ is a subgroup of G and also that $H \not\subset K$ and $K \not\subset H$. Then there are elements $a \in H \setminus K$, $b \in K \setminus H$ and both $a, b \in H \cup K$. As $H \cup K$ is a subgroup, $ab \in H \cup K$.

But then $ab \in H$ or $ab \in K$. If $ab \in H$ then

$$b = a^{-1}(ab) \in H.$$

If $ab \in K$ then

$$a = (ab)b^{-1} \in K.$$

It is contradiction in both cases by the choice of b and a

Hence either $H \setminus K = \varnothing$ or $K \setminus H = \varnothing$.

That is, either $K \subset H$ or $H \subset K$, as required.

Example 18. Find the subgroups of the group $G = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$ with the following multiplication table:

+	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

Solution. Here the binary operation is addition modulo 4. Consider all nonempty subsets of the group $G = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$ with respect to addition modulo 4. It is found that the only nontrivial subgroup is $\{\bar{0}, \bar{2}\}$ so that $\{\bar{0}\}$, $\{\bar{0}, \bar{2}\}$ and G are the only subgroups of G.

Example 19. Find the subgroups of the group $G = \{e, a, b, c\}$ with the multiplication table as below.

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Solution. Here the nontrivial subgroups are

$$\{e, a\}, \{e, b\} \text{ and } \{e, c\}$$

Note that here $c = ab$ and $a^2 = b^2 = (ab)^2 = e$. The group G is called the **Klein's four group**[1]

Example 20. Let $C \setminus \{0\}$ be the group of all nonzero complex numbers under multiplication of complex numbers. Prove that the set

$$H = \{a + ib \in C \setminus \{0\} : a^2 + b^2 = 1\}$$

is a subgroup of $C \setminus \{0\}$

Solution. Let $a_1 + ib_1, a_2 + ib_2 \in H$, so that $a_1^2 + b_1^2 = 1$, $a_2^2 + b_2^2 = 1$. We have

$$\begin{aligned}(a_1 + ib_1)(a_2 + ib_2)^{-1} &= \frac{a_1 + ib_1}{a_2 + ib_2} \\ &= \frac{(a_1 + ib_1)(a_2 - ib_2)}{a_2^2 + b_2^2} \\ &= (a_1a_2 + b_1b_2) + i(a_2b_1 - a_1b_2), \text{ since } a_2^2 + b_2^2 = 1\end{aligned}$$

But

$$\begin{aligned}(a_1a_2 + b_1b_2)^2 + (a_2b_1 - a_1b_2)^2 &= a_1^2a_2^2 + b_1^2b_2^2 + a_2^2b_1^2 + a_1^2b_2^2 \\ &= a_1^2(a_2^2 + b_2^2) + b_1^2(a_2^2 + b_2^2) \\ &= (a_1^2 + b_1^2)(a_2^2 + b_2^2) = 1\end{aligned}$$

Thus $(a_1 + ib_1)(a_2 + ib_2)^{-1} \in H$. This shows that H is a subgroup of $C \setminus \{0\}$.

1 Named for the German mathematician Felix Klein (1849 – 1925).

CYCLIC GROUPS

(2.18) Definition. A group G is said to be **cyclic** if every element of G is a power of one and the same element a (say), of G.

Such an element a of G is called a **generator** of G.

If the order of a is finite, that is, if there is the least positive integer n such that $a^n = e$, then G is said to be a **finite** cyclic group of order n and is written as

$$G = \langle a : a^n = e \rangle$$

(read as G is a cyclic group of order n generated by a)

If the order of a is infinite then, for no positive integer n, $a^n = e$.

Also, in this case, if for some positive integer k, $(a^{-1})^k = e$, then

$$a^k = a^k e = a^k (a^{-1})^k = e$$

so that a has finite order, a contradiction. Thus a^{-1} also has infinite order. So if a has infinite order, then a^{-1} also has infinite order.

If G is a finite cyclic group of order n with generator a, then n is the smallest positive integer such that $a^n = e$. Thus, for no positive integer r, $0 < r < n$, $a^r = e$. Therefore, $a^n = e, a^1, a^2, \cdots, a^{n-1}$ are all the distinct elements of G.

The following theorem describes the nature of subgroups of a cyclic group.

(2.19) Theorem. Every subgroup of a cyclic group is cyclic.

Proof. Let G be a cyclic group generated by a. Then every element of G is a power of a. Let H be a subgroup of G. Let k be the least positive integer such that $a^k \in H$. Then $\langle a^k \rangle \subset H$. We show that every element of H is a power of a^k. For this, let $a^m \in H$. Then, by the division algorithm, there are unique integers q and r such that

$$m = qk + r, \qquad 0 \le r < k$$

So $\quad a^m = a^{qk+r} = (a^k)^q a^r \quad$ and $\quad (a^k)^q \in H$.

That is, $\quad a^r = a^m (a^k)^{-q} \in H, r < k$

which is a contradiction to the choice of k unless $r = 0$. But then $m = qk$ and $a^m = (a^k)^q$. Hence H is cyclic.

(2.20) Theorem. Let G be a cyclic group of order n generated by a. Then, for each positive divisor d of n, there is a unique subgroup (of G) of order d.

Proof. Let

$$G = \langle a : a^n = e \rangle$$

Let d be a positive divisor of n. Then there is an integer q such that

$$n = qd$$

Take $b = a^q$. Then

$$b^d = a^{qd} = a^n = e.$$

So $H = \langle b : b^d = e \rangle$ is the required subgroup.

To see that H is unique, suppose that K is another subgroup of G of order d. Then K is generated by an element $c = a^k$, where k is the least such positive integer. As K has order d,

$$a^{kd} = c^d = e,$$

where $kd = n$

so that $k = \frac{n}{d} = q$

Thus $b = a^q = a^k = c,$

so that $K = H$ and hence H is unique.

(2.21) Theorem. Every cyclic group is abelian.

Proof. Let G be a cyclic group generated by a and $x, y \in G$. Then there are positive integers k and m such that

$$x = a^k, \quad y = a^m.$$

So $xy = a^k a^m = a^{k+m} = a^{m+k} = a^m a^k = yx$

Thus G is abelian.

Example 21. If G is a cyclic group of even order, then prove that there is only one subgroup of order 2 in G.

Solution. Let $G = \langle a : a^{2n} = e \rangle$ be a cyclic group of order $2n$, where n is a positive integer. By Theorem 2.20, if a positive integer d divides $|G|$, then G has exactly one subgroup of order d.

Now $|G| = 2n$ and 2 divides $2n$, so G has only one subgroup of order 2.

Example 22. Find all the subgroups of a cyclic group of order 12.

Solution. Let G be a cyclic group of order 12 and a be its generator. Then the elements of G are

$$e = a^{12}, a^1, a^2, \cdots, a^{11}.$$

By Theorem 2.20, for each divisor d of 12, G has a subgroup of order d. Now the divisors of 12 are

1, 2, 3, 4, 6 and 12.

Therefore, G has

the subgroup of order 1	which is	$\{e\}$
the subgroup of order 2	which is	$\{e, a^6\}$
the subgroup of order 3	which is	$\{e, a^4, a^8\}$
the subgroup of order 4	which is	$\{e, a^3, a^6, a^9\}$
the subgroup of order 6	which is	$\{e, a^2, a^4, a^6, a^8, a^{10}\}$
the subgroup of order 12	which is	G itself

COSETS — LAGRANGE'S THEOREM

In this section we prove one of most important theorems of group theory called the **Lagrange's Theorem.** For this we first explain the concept of **cosets** of a subgroup.

(2.22) Definition. Let H be a subgroup of a group G and $a \in G$. Then the set

$$aH = \{ah : h \in H\}$$

is called a **left coset** of H in G **determined** by a (or **containing** a).

Similarly, $\quad Ha = \{ha : h \in H\}$

is called a **right coset** of H in G **determined** by a.

It is easy to see that the left (or right) coset of H in G determined by the identity element e is H itself. For

$$eH = \{eh : h \in H\} = \{h : h \in H\} = H$$

and $\quad He = \{he : h \in H\} = \{h : h \in H\} = H.$

If the binary operation in the group is denoted by +, then the left coset of H in $(G, +)$ determined by $a \in G$ is written as $a + H$ where

$$a + H = \{a + h : h \in H\}$$

Example 23. Let

$$\bar{Z}_6 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}\}$$

be the group of residue classes under addition modulo 6. Then

$$H = \{\bar{0}, \bar{2}, \bar{4}\}$$

is a subgroup of $\bar{Z}_6$. There are only two left cosets of H in G and they are:

$$\bar{0} + H = H \quad \text{and} \quad \bar{1} + H = \{\bar{1}, \bar{3}, \bar{5}\}$$

Example 24. Let $G = \{\bar{1}, \bar{3}, \bar{5}, \bar{7}\}$

be the group of residue classes under multiplication modulo 8. The multiplication table for G is as given in Example [illegible]

It is easy to see that

$$H_1 = \{\bar{1}, \bar{3}\},\ H_2 = \{\bar{1}, \bar{5}\},\ H_3 = \{\bar{1}, \bar{7}\}$$

are proper subgroups of G. (Write multiplication tables of H_1, H_2 and H_3). The left cosets of H_1 in G are only two in number. These are

$$\bar{1} \cdot H_1 = H_1, \quad \bar{5} \cdot H_1 = \bar{7} \cdot H_1 = \{\bar{5}, \bar{7}\}$$

Similarly, the left cosets of H_2 and H_3 in G are

$$\bar{1} \cdot H_2 = \{\bar{1}, \bar{5}\}, \quad \bar{3} \cdot H_2 = \{\bar{3}, \bar{7}\} \text{ and } \bar{1} \cdot H_3 = \{\bar{1}, \bar{7}\}, \bar{3} \cdot H_3 = \{\bar{3}, \bar{5}\}$$

respectively.

Notice that the union of the cosets $\bar{1} \cdot H_1$ and $\bar{5} \cdot H_1$ is G itself.

Recall that by a **partition** of a set A we mean a collection of subsets $\{A_i : i \in I\}$ of A such that

$$A = \cup\{A_i : i \in I\} \text{ and } A_i \cap A_j = \varnothing \text{ where } i, j \in I \text{ and } i \neq j$$

(2.23) Theorem. Let H be a subgroup of a group G. Then the set of all left (or right) cosets of H in G defines a partition of G.

Proof. Let $\{aH : a \in G\}$

be the collection of all the left cosets of H in G. For each $a \in G$, $a = ae \in aH$ because $e \in H$. So

$$G \subseteq \cup\{aH : a \in G\}$$

But $\cup\{aH : a \in G\}$, being the union of subsets of G is contained in G, i.e.,

$$\cup\{aH : a \in G\} \subseteq G$$

Hence $G = \cup\{aH : a \in G\}$

Next, let aH, bH be two **distinct** cosets and $x \in aH \cap bH$.

Then $x = ah_1 = bh_2$ for some $h_1, h_2 \in H$

Now $a = bh_2h_1^{-1} = bh_3$ where $h_3 = h_2h_1^{-1} \in H$, because H is a subgroup.

Thus $a \in bH$. But then, for any $h \in H$

$$ah = bh_3h \quad \text{is an element of } bH. \text{ Hence}$$

$$aH \subseteq bH$$

Likewise $bH \subseteq aH$

Hence $aH = bH$, a contradiction. Thus $aH \cap bH = \varnothing$. Therefore, $\{aH : a \in G\}$ defines a partition of G.

(2.24) Definition. The number of distinct left (or right) cosets of a subgroup H of a group G is called the **index** of H in G, and is denoted by $(G : H)$.

Example 25. Find the distinct left (or right) cosets of the subgroup

$$E = \{0, \pm 2, \pm 4, \cdots\} = \{2n : n \in Z\} \text{ in the group } (Z, +)$$

Solution. The only left cosets of E in Z are

$$0 + E \text{ and } 1 + E.$$

Here $0 + E = \{0 + 2n : n \in Z\} = E$, the set of even integers

and $1 + E = \{1 + 2n : n \in Z\}$

$= \{\pm 1, \pm 3, \pm 5, \cdots\}$ the set of odd integers

and $E \cup \{1 + E\} = Z$, $E \cap \{1 + E\} = \varnothing$.

Therefore, the index of E in Z is 2.

A relation between the order of a finite group G and the order and the index of a subgroup H of G is given by the following theorem.

(2.25) Theorem. (Lagrange[1]). Both the order and index of a subgroup of a finite group divide the order of the group.

Proof. Let H be a subgroup of order m of a finite group G of order n and let k be the index of H in G. Since $|G|$ is finite, the set

$$\{a_1H, a_2H, \cdots, a_kH\}$$

of all the distinct left cosets of H in G is also finite. By Theorem 2.23,

$$G = \bigcup_{i=1}^{k} a_iH, \text{ where } a_iH \cap a_jH = \varnothing, i \neq j, i, j = 1, 2, \cdots, k.$$

Now the mapping $\phi : H \longrightarrow a_iH$ defined by

$$\phi(h) = a_ih, \ h \in H$$

1. Joseph Louis Lagrange (1736 – 1813), a French mathematician.

is obviously onto. Also,

$$\phi(h_1) = \phi(h_2),\ h_1, h_2 \in H$$

implies

$$a_i h_1 = a_i h_2$$

which, by the cancellation law, implies $h_1 = h_2$. So ϕ is one-to-one. Hence the number of elements in H and in a_iH is the same, for $i = 1, 2, \cdots, k$. As H has m elements, each a_iH also has m elements for $i = 1, 2, \cdots, k$.

Now from $G = \bigcup_{i=1}^{k} a_iH$, we have

$$n = m + m + \cdots + m \ (k \text{ times})$$

$$= km.$$

Thus both m and k, the order and index of H in G, divide n (the order of G).

(2.26) Corollary. The order of an element of a finite group divides the order of the group.

Proof. Let G be a finite group of order n and $a \in G$. Then the order m of a is also finite. But the order of a is the order of the cyclic subgroup H of G generated by a. So m, the order of the subgroup H of G, divides n the order of G.

(2.27) Corollary. A finite group G whose order is a prime number is necessarily cyclic.

Proof. Let G be a group of order p, where p is a prime number. Let a be a non-identity element of G and let H be the cyclic group generated by a. Then the order k of H divides the order p of G. But the only divisors of p are 1 and p itself. Since H is not the identity subgroup, $k = p$. So H, as a subgroup of G of the same order as that of G, must be equal to G itself. But H is cyclic, so G is cyclic.

EXERCISE 2.2

1. Give an example of an abelian group which is not cyclic.
2. Explain why a group of order 47 cannot have proper subgroups.
3. Let G be a group of order 89. Can G have a subgroup of order:
   (i) 12 (ii) 16 (iii) 24 ?
   Justify your answer.
4. Show that an infinite cyclic group has exactly two distinct generators.

5. Is $(Q, +)$ a cyclic group? Why?
6. Let G be a cyclic group of order 24 generated by a. Find the orders of the elements.
   (i) e (ii) a^9 (iii) a^{10}.
7. Find all subgroups of the cyclic group of order 60 generated by a.
8. Write all the subgroups of a cyclic group of order 18.
9. Let H be the set of real numbers $a + b\sqrt{2}$, where $a, b \in Q$ and both are not simultaneously zero. Show that H is a subgroup of the group of nonzero real numbers under multiplication.
10. Let H be the set of complex numbers of the form $a + b\sqrt{-5}$, where $a, b \in Q$ and both are not simultaneously zero. Show that H is a subgroup of the group of nonzero complex numbers under multiplication.
11. Let $G = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in R \text{ and } ad - bc \neq 0 \right\}$ be the group of all nonsingular 2×2 real matrices under multiplication. Show that the sets
    (i) $H = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} : a, b, d \in R \text{ and } ad \neq 0 \right\}$ and
    (ii) $K = \left\{ \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} : b \in R \right\}$
    are subgroups of G.
12. Let H and K be two finite subgroups (of a group G) whose orders are relatively prime. Prove that
    $$H \cap K = \{ e \}.$$
13. Let n be an integer greater than 1 and $H = \{0, \pm n, \pm 2n, \pm 3n, \cdots\}$. Find all left cosets of H in $(Z, +)$. What is their number?
14. If H is a subgroup of a group G then show that
    $$H \cdot H = \{ h_1 h_2 : h_1, h_2 \in H \} = H.$$
15. Let H, K be subgroups of an abelian group G. Show that the set
    $$HK = \{ hk : h \in H, k \in K \} \text{ is a subgroup of } G.$$
16. Let H be a subgroup of a group G and $a \in G$. If $(Ha)^{-1} = \{ (ha)^{-1} : h \in H \}$, then show that $(Ha)^{-1} = a^{-1}H$.

17. Let H, K be two subgroups of a finite group G. Prove that, for any $g \in G$,
$$g(H \cap K) = gH \cap gK.$$

18. Let $G = \{\bar{1}, \bar{2}, \bar{3}, \cdots, \bar{12}\}$ be the group of nonzero residue classes under multiplication modulo 13. Which of the following are subgroups of G?
$$H_1 = \{\bar{1}, \bar{3}, \bar{5}, \bar{7}, \bar{9}, \bar{11}\}, \quad H_2 = \{\bar{1}, \bar{2}, \bar{4}, \bar{6}, \bar{8}, \bar{10}, \bar{12}\}$$
$$H_3 = \{\bar{1}, \bar{6}, \bar{8}, \bar{10}\}, \quad H_4 = \{\bar{1}, \bar{3}, \bar{9}\}.$$

19. Show that the set $SL_2(R)$ of all 2×2 matrices with determinants 1 is a subgroup of the general linear group $GL_2(R)$ of degree 2. (Problem 18 EXERCISE 2.1) ($SL_2(R)$ is called **the special linear group**).

20. Show that the set
$$C = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} : a, b \in R \; a^2 + b^2 \neq 0 \right\}$$
is a subgroup of $GL_2(R)$ and
$$C^* = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} : a, b \in R \; a^2 + b^2 = 1 \right\}$$
is a subgroup of $SL_2(R)$.

PERMUTATIONS

In this section we discuss a class of groups called **permutation groups.** In the early days of development of groups, properties of only such groups were used to be considered. These groups also provide examples of non-abelian groups.

(2.28) Definition. Let X be a nonempty set. A bijective mapping $f : X \longrightarrow X$ is called a **permutation** on X.

The set of all permutations on X is denoted by S_X. If X consists of n elements then we write S_n for S_X.

From now onwards we shall always take X as a finite set consisting of n elements.

Let $f : X \longrightarrow X$ be a permutation. For each $x \in X$, the image of x under f will be denoted by $(x)f$ instead of $f(x)$. This notation is used just for the sake of convenience and is in conformity with the current literature on permutation groups.

(2.29) Theorem. The set S_n of all permutations on a set X with n elements is a group under the operation of **composition (or product)** of permutations.

Proof. For permutations $f : X \longrightarrow X$, $g : X \longrightarrow X$, we define $f o g$ by $(x)(f o g) = ((x)f)g$.

Since the composition of two bijective mappings is bijective, fog is also a permutation on X. Thus S_n is closed under multiplication.

For associative law in S_n, let f, g and h be three permutations on X. Then
$$\begin{aligned} (x)((fog)oh) &= ((x)(fog))oh \\ &= (((x)f)g)h \\ &= ((x)f)(goh) \\ &= (x)(fo(goh)) \end{aligned}$$
for all $x \in X$. Hence
$$(fog)oh = fo(goh)$$
so that the associative law is satisfied in S_n.

The mapping $I : X \longrightarrow X$ defined by
$$(x)I = x \quad \text{for all } x \in X$$
is the identity element in S_n because, for any $f \in S_n$,
$$(x)(foI) = ((x)f)I = (x)f.$$
So
$$foI = f \quad \text{for all } f \in S_n.$$
Hence I is the identity element of S_n.

For any permutation f in S_n, its inverse $f^{-1} : X \longrightarrow X$ is also bijective and is in S_n.

Moreover, by definition of f^{-1}, for any $x \in X$
$$(x)f = y \Rightarrow (y)f^{-1} = x.$$
So $\quad (x)(fof^{-1}) = ((x)f)f^{-1} = (y)f^{-1} = x$

and $\quad (y)(f^{-1}of) = ((y)f^{-1})f = (x)f = y$

for all $x, y \in X$.

Hence $\quad fof^{-1} = I = f^{-1}of.$

Therefore, S_n is a group.

(2.30) Definition. The group S_n is called the **symmetric group of degree** n.

Since every element in S_n is a permutation of n objects (namely elements of X), taken n at a time, there are $n!$ such permutations. Hence the order of S_n is $n!$.

Example 26. Let $X = \{1, 2, 3\}$. Then the mappings

$I, f_1, f_2, f_3, f_4, f_5$ from X to X defined by

$(1)I = 1,\ (2)I = 2,\ (3)I = 3;\quad (1)f_1 = 2,\ (2)f_1 = 3,\ (3)f_1 = 1;$

$(1)f_2 = 3,\ (2)f_2 = 1,\ (3)f_2 = 2;\quad (1)f_3 = 2,\ (2)f_3 = 1,\ (3)f_3 = 3;$

$(1)f_4 = 3,\ (2)f_4 = 2,\ (3)f_4 = 1;\quad (1)f_5 = 1,\ (2)f_5 = 3,\ (3)f_5 = 2$

are all the permutations that can be defined on X. Hence

$$S_3 = \{I, f_1, f_2, f_3, f_4, f_5\}.$$

We shall use permutation notation to express the elements of S_3 by placing 1, 2, 3 in a row in that order and then write their corresponding images below them and enclose these within parentheses. Thus we write

$$I = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix},\quad f_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},\quad f_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

$$f_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},\quad f_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},\quad f_5 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Let us calculate product of permutations f_1 and f_3. Here

$(1)(f_1 o f_3) = ((1)f_1)f_3 = (2)f_3 = 1$

$(2)(f_1 o f_3) = ((2)f_1)f_3 = (3)f_3 = 3$

$(3)(f_1 o f_3) = ((3)f_1)f_3 = (1)f_3 = 2.$

So $$f_1 o f_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = f_5 \qquad (1)$$

Similarly,

$$f_3 o f_1 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = f_4. \qquad (2)$$

From (1) and (2) we note that in S_3

$$f_1 o f_3 \neq f_3 o f_1$$

so that S_3 is a non-abelian group.

If we write $f_1 = a,\ f_3 = b$

then, one can verify that

$$f_2 = a^2,\quad f_4 = a^2b,\quad f_5 = ab$$

Also $a^3 = b^2 = (ab)^2 = I.$

Note that $ba = f_4 = a^2b.$ (3)

(4)

Using the relations (3) and (4), the multiplication table for S_3 is as follows:

$\cdot$	I	a	a^2	b	ab	a^2b
I	I	a	a^2	b	ab	a^2b
a	a	a^2	I	ab	a^2b	b
a^2	a^2	I	a	a^2b	b	ab
b	b	a^2b	ab	I	a^2	a
ab	ab	b	a^2b	a	I	a^2
a^2b	a^2b	ab	b	a^2	a	I

We now describe a method of multiplying two permutations. This is illustrated by an example.

Example 27. Let

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix},\quad g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

be two permutations of degree 4. Then

$$fg = f o g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}\begin{pmatrix} 2 & 3 & 4 & 1 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

represents one and the same permutation because, in both the permutations, the images of corresponding elements are the same.

Now we ignore the second row of the permutation f and the first row of the permutation g and obtain the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

which represents $f \circ g$.

By this method $g \circ f$ is represented by

$$gf = g\circ f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}.$$

From the above we note that

$$f\circ g \neq g \circ f.$$

This method can be used to find the product of any two permutations.

Example 28. Find the product of the permutations √

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 3 & 2 & 6 & 4 & 1 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 2 & 6 & 5 \end{pmatrix}$$

Solution. Here

$$\alpha\circ\beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 3 & 2 & 6 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 2 & 6 & 5 \end{pmatrix}$$

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

$$= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 1 & 4 & 5 & 2 & 3 \end{pmatrix} \qquad (1)$$

Likewise,

$$\beta\circ\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 2 & 6 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 3 & 2 & 6 & 4 & 1 \end{pmatrix}$$

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

$$= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 6 & 5 & 3 & 4 & 1 \end{pmatrix} \qquad (2)$$

From (1) and (2) we note that $\alpha\circ\beta \neq \beta\circ\alpha$.

A similar procedure for multiplication of two permutations, say α and β given above, is as follows.

We look at the effect of α on the first element namely 1 of the first row of α. Under α, 1 is mapped onto 5. Then we look at the effect of β on 5 which occurs in the first row of β. Under β, 5 is mapped onto 6. So under the combined effect of α, β, 1 is mapped onto 6. Similarly, under $\alpha\circ\beta$, 2 is mapped onto, 1, 3 onto 4, 4 onto 5, 5 onto 2 and finally, 6 is mapped onto 3. Thus $\alpha\circ\beta$ is as obtained in (1). Similarly for $\beta\circ\alpha$.

CYCLES ✓

(2.31) Definition. A permutation of the form

$$\alpha = \begin{pmatrix} a_1 & a_2 & \cdots & a_k \\ a_2 & a_3 & \cdots & a_1 \end{pmatrix}$$

in which $(a_1)\alpha = a_2,\ (a_2)\alpha = a_3, \cdots, (a_k)\alpha = a_1$

is called a **cyclic permutation** or **cycle** of length k and is written as

$$\alpha = (a_1\ a_2\ \cdots\ a_k).$$

Here $a_1, a_2, \cdots, a_k$ are elements of the set X which consists of n elements so that $k \le n$. Elements of X other than $a_1, a_2, \cdots, a_k$ remain unchanged under the action of α.

For example, the permutation

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 1 \end{pmatrix}$$

is a cyclic permutation of length 6.

The product of two cyclic permutations need not be a cyclic permutation. For example consider the product

$$(1\ 2\ 5)(2\ 1\ 4\ 5\ 6)$$

of cycles $\alpha = (1\ 2\ 5)$, $\beta = (2\ 1\ 4\ 5\ 6)$

To express this as a permutation, we first note that if an element of X is not written in a cycle then that element is fixed under the effect of that permutation. Thus, 6 does not occur in the cycle $(1\ 2\ 5)$. So $(6)\,\alpha = 6$ and so on. The above product is then obtained by applying the permutations successively. Thus

$$1 \xrightarrow{\alpha} 2 \xrightarrow{\beta} 1. \quad \text{So} \quad 1 \xrightarrow{\alpha o \beta} 1$$

$$2 \xrightarrow{\alpha} 5 \xrightarrow{\beta} 6. \quad \text{So} \quad 2 \xrightarrow{\alpha o \beta} 6$$

$$5 \xrightarrow{\alpha} 1 \xrightarrow{\beta} 4 \quad \text{So} \quad 5 \xrightarrow{\alpha o \beta} 4$$

$$4 \xrightarrow{\alpha} 4 \xrightarrow{\beta} 5 \quad \text{So} \quad 4 \xrightarrow{\alpha o \beta} 5$$

$$6 \xrightarrow{\alpha} 6 \xrightarrow{\beta} 2 \quad \text{So} \quad 6 \xrightarrow{\alpha o \beta} 2$$

Hence $\quad \alpha o \beta = \alpha\beta = (1\ 2\ 5)(2\ 1\ 4\ 5\ 6)$

$$= \begin{pmatrix} 1 & 2 & 4 & 5 & 6 \\ 1 & 6 & 5 & 4 & 2 \end{pmatrix}$$

which is not a cyclic permutation.

If two cycles act on mutually disjoint sets then they commute.

Example 29. The cycles

$$\alpha = (1\ 2\ 3), \qquad \beta = (4\ 5\ 6)$$

are such that

α acts on $\{1, 2, 3\}$, β acts on $\{4, 5, 6\}$ and these sets are disjoint.

$$\alpha\beta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}\begin{pmatrix} 4 & 5 & 6 \\ 5 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 6 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 4 & 5 & 6 \\ 5 & 6 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \beta\alpha$$

The following theorem shows that every permutation can be written as a product of cycles acting on mutually disjoint sets.

(2.32) Theorem. Every permutation of degree n can be written as a product of cyclic permutations acting on mutually disjoint sets.

Proof. Let α be a permutation of degree n (i.e., an element of S_n). Let a_1 be one of the elements on which α acts. Suppose that, under the action of α,

$$a_1 \xrightarrow{\alpha} a_2,\ a_2 \xrightarrow{\alpha} a_3, \cdots$$

Since n is finite, there is a natural number k such that $a_k \xrightarrow{\alpha} a_1$

Thus a part of the effect of α is the cyclic permutation

$$\alpha_1 = (a_1\ a_2\ \cdots\ a_k)$$

If $k = n$ then $\alpha = \alpha_1$ is the required cyclic decomposition of α as cyclic permutation. However, if $k < n$ then there is a b_1, different from $a_1, a_2, \cdots, a_k$, on which α acts. Suppose that, under the action of α,

$$b_1 \xrightarrow{\alpha} b_2,\ b_2 \xrightarrow{\beta} b_3, \cdots, b_p \xrightarrow{\alpha} b_1$$

Then no b_i, $1 \le i \le p$ occurs among $a_1, a_2, \cdots, a_k$ because α, as permutation, is an injective mapping. Also then a part of the effect of α is the cyclic permutation

$$\alpha_2 = (b_1\ b_2\ \cdots\ b_p)$$

So we have extracted two cyclic permutations α_1 and α_2 from α.

If $k + p = n$ then α is the product of cycles α_1 and α_2. If, however, $k + p < n$ then the process is continued. This process of extracting a cycle each time must end after a finite number of steps because n is finite. Thus there is a natural number q such that

$$k + p + \cdots + q = n$$

and a part of the effect of α is a cycle

$$\alpha_r = (c_1\ c_2\ \cdots\ c_q)$$

where no c_i occurs among the a's, b's, etc. Therefore

$$\alpha = \alpha_1\ \alpha_2\ \cdots\ \alpha_r \qquad (1)$$

where each of $\alpha_1, \alpha_2, \cdots, \alpha_r$ acts on mutually disjoint subsets of X and are uniquely determined. Since any two permutations acting on mutually disjoint sets commute, so, apart from the order in which α_i's are taken, the expression (1) for α is unique.

TRANSPOSITIONS

(2.33) Definition. A cycle of length 2 is called a **transposition.**

Thus a permutation of the type

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

where a and b are interchanged by this permutation and other elements remain fixed, is a transposition.

(2.34) Theorem. Every cyclic permutation can be expressed as a product of transpositions.

Proof. Let

$$\alpha = (a_1\ a_2\ a_3\ \cdots\ a_k)$$

be a cyclic permutation. Consider the permutation

$$\alpha' = (a_1\ a_2)\,(a_1\ a_3)\cdots(a_1\ a_k)$$

Then, under the effect of α',

$$a_1 \longrightarrow a_2,\ a_2 \longrightarrow a_1 \longrightarrow a_3,\ a_3 \longrightarrow a_4, \cdots, a_k \longrightarrow a_1.$$

So the effect of α and α' on $a_1, a_2, a_3, \cdots, a_k$ is the same. Hence $\alpha = \alpha'$ is a product of transpositions.

Since $(a\ \ b)\,(b\ \ a) = I$, any number of such pairs of transpositions can be inserted between the pairs $(a_1\ \ a_i)$, $(a_1\ \ a_{i+1})$ involved in α.

Hence α can be expressed as a product of transpositions, possibly in infinitely many ways.

(2.35) Theorem. Every permutation of degree n can be expressed as a product of transpositions.

Proof. Let α be a permutation of degree n. Then, by Theorem 2.32, α is expressible as a product of cyclic permutations acting on mutually disjoint sets. Also, by Theorem 2.34, every cyclic permutations can be expressed as a product of transpositions in infinitely many ways. Hence α can be expressed as a product of transpositions in infinitely many ways.

Although representation of a permutation as product of transpositions is not unique, yet the following theorem proves that, for a given permutation, the number of transpositions is always either even or odd.

(2.36) Theorem. Let a permutation α in S_n be written as a product of m transpositions and as a product of p transpositions. Then $m - p$ is a multiple of 2, i.e., $m - p$ is divisible by 2. This implies that either both m and p are even or both of them are odd.

Proof. Let

$$\alpha = \lambda_1\ \lambda_2\ \cdots\ \lambda_m = \phi_1\ \phi_2\ \cdots\ \phi_p,$$

where α is a permutation on a set

$$X = \{\, x_1, x_2, \cdots, x_n \,\}$$

consisting of n elements and λ_i, φ_j are transpositions, $1 \le i \le m$, $1 \le j \le p$.

Consider the product

$$\begin{aligned} P &= \prod_{i<j}^{n} (x_i - x_j) \\ &= (x_1 - x_2)\,(x_1 - x_3)\cdots(x_1 - x_n) \\ &\qquad (x_2 - x_3)\cdots(x_2 - x_4) \\ &\qquad\quad \vdots \qquad\qquad \vdots \\ &\qquad\qquad (x_{n-1} - x_n) \end{aligned}$$

Then $$P\alpha = \prod_{i<j}^{n} (x_{(i)\alpha} - x_{(j)\alpha})$$

Now, for any transposition $\tau = (k\ \ l)$ in S_n, $k \ne l$

$$(P)\tau = \prod_{i<j}^{n} (x_{(i)\tau} - x_{(j)\tau})$$

Every factor of P that contains neither x_k nor x_l remains unchanged in $(P)\tau$. But the factor $(x_k - x_l)$ of P becomes $(x_l - x_k)$. i.e., $-(x_k - x_l)$ in $(P)\tau$.

Also those factors of P which contain either x_k or x_l, but not both x_k and x_l can be grouped into pair of products

$$\pm (x_m - x_k)(x_r - x_l)$$

where $(m, r) \neq (k, l)$. Such a product remains unchanged in $(P)\tau$. Thus it follows that $(P)\tau = -P$.

So, by successive application of the transpositions in α, in both cases

$$(P)\alpha = (P)\lambda_1 \lambda_2 \cdots \lambda_m = (-1)^m P$$

and
$$(P)\alpha = (P)\phi_1 \phi_2 \cdots \phi_p = (-1)^p P.$$

So
$$(-1)^m P = (-1)^p P,$$

that is
$$P = (-1)^{m-p} P$$

which shows that $m - p$ must be a multiple of 2, as required.

(2.37) Definition. A permutation α in S_n is said to be an **even permutation** if it can be written as a product of an even number of transpositions. Otherwise it is said to be an **odd permutation**. Thus α in S_n is an odd permutation if it can be expressed as a product of an odd number of transpositions.

For example, the number of transpositions in the decomposition of the identity permutation is zero which is an even integer. So the identity permutation is an even permutation. Similarly,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

is an even permutation.

Every transposition is an odd permutation.

Similarly,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} = (1\ 2\ 3\ 4) = (1\ 2)(1\ 3)(1\ 4)$$

is an a odd permutation.

With each permutation α of finite degree we can associate a positive integer m_α which is the number of transpositions into which α has been expressed as a product. Thus a permutation α is even or odd according as m_α is even or odd.

(2.38) Theorem. (i) The product of two even or odd permutations is an even permutation.

(ii) The product of an even permutation and an odd permutation is an odd permutation.

Proof. Let α_1 and α_2 be any two permutations of degree n. Then α_1 and α_2 can be expressed as a product of m_1 and m_2 transpositions respectively.

So the product $\alpha_1 \alpha_2$ contains $m_1 + m_2 - 2k$ transpositions where k is 0 or a natural number. The term $2k$ occurs because of the possible cancellation (simplification) of pairs of transpositions.

So

(i) If both α_1 and α_2 are even permutations, in which case, both m_1 and m_2 are even, or both α_1 and α_2 are odd permutations, in which case, both m_1 and m_2 are odd, then $m_1 + m_2 - 2k$ is an even integer so that $\alpha_1\alpha_2$ is an even permutation.

(ii) If one of the permutations, say α_1, is even and α_2 is odd then $m_1 + m_2 - 2k$ is an odd integer so that $\alpha_1\alpha_2$ is an odd permutation.

(2.39) Corollary. Let α be any permutation of degree n and τ a transposition. Then $\alpha\tau$ or $\tau\alpha$ is an even or odd according as α is odd or even respectively.

Proof. If α is an even permutation then $\alpha\tau$ or $\tau\alpha$ will have an odd number of transpositions so that $\alpha\tau$ or $\tau\alpha$ are odd permutations. Similar is the case of α being an odd permutation.

(2.40) Theorem. Let $n \geq 2$. The number of even permutations in S_n is equal to the number of odd permutations is S_n.

Proof. Let

$$\alpha_1, \alpha_2, \cdots, \alpha_k \quad (1)$$

be all even permutations and

$$\beta_1, \beta_2, \cdots, \beta_l \quad (2)$$

be all odd permutations in S_n so that $k + l = n!$.

Let τ be a transposition. Then

$$\alpha_1\tau, \alpha_2\tau, \cdots, \alpha_k\tau \quad (3)$$

are all odd, while

$$\beta_1\tau, \beta_2\tau, \cdots, \beta_l\tau \quad (4)$$

are all even permutations. Hence, from (2) and (3), we have

$$k \le l$$

and from (1) and (4), we have

$$l \le k.$$

Thus $k = l = \dfrac{n!}{2}$, as required.

Let us denote the set of all even permutations in S_n by A_n. From Theorem 2.40, it follows that the number of elements in A_n is $\dfrac{n!}{2}$.

(2.41) Theorem. The set A_n of all even permutations in S_n forms a subgroup of S_n.

Proof. Let $\alpha_1, \alpha_2 \in A_n$. Then, by Theorem 2.38, $\alpha_1 \alpha_2$ is an even permutation.

Also, since the inverse of a transposition is a transposition, the inverse of an even permutation is an even permutation.

Hence A_n is a subgroup of S_n.

Example 30. Find all the subgroups of S_3.

Solution. As in Example 26, the elements of S_3 are

$$I = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad a^2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

$$b = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad ab = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad a^2 b = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

Hence $a^3 = b^2 = (ab)^2 = I$.

So the subgroups of S_3 are

$$\{I\}, \{I, a, a^2\}, \{I, b\}, \{I, ab\}, \{I, a^2 b\} \text{ and } S_3.$$

ORDER OF A PERMUTATION

(2.42) Definition. Let α be a permutation in S_n. The **order** of α is the least positive integer m such that $\alpha^m = I$, the identity permutation.

For example, if

$$\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

then

$$\alpha^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

and

$$\alpha^3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = I$$

Hence the order of α is 3.

The order of a transposition is 2 because if

$$\tau = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

then

$$\tau^2 = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \begin{pmatrix} a & b \\ b & a \end{pmatrix} = \begin{pmatrix} a & b \\ a & b \end{pmatrix} = I.$$

(2.43) Theorem. The order of a cyclic permutation of length m is m.

Proof. Let

$$\alpha = (a_1 \; a_2 \; \cdots \; a_m)$$

be a cyclic permutation of length m. Then, under the action of

$$\alpha^2 : a_1 \longrightarrow a_3,\; a_2 \longrightarrow a_4,\; \cdots,\; a_m \longrightarrow a_2$$

$$\alpha^3 : a_1 \longrightarrow a_4,\; a_2 \longrightarrow a_5,\; \cdots,\; a_m \longrightarrow a_3$$

$$\vdots \quad \vdots \qquad \vdots \quad \vdots \qquad \vdots \quad \vdots \qquad \vdots$$

$$\alpha^m : a_1 \longrightarrow a_1,\; a_2 \longrightarrow a_2,\; \cdots,\; a_m \longrightarrow a_m$$

so that $\alpha^m = I$. Also for, no smaller power k of α, $\alpha^k = I$. Hence the order of α is m. So a cycle of length m has order m.

To find the order of a permutation α, first decompose α as a product of cyclic permutations

$$\alpha_1, \alpha_2, \cdots, \alpha_k$$

of length $m_1, m_2, \cdots, m_s$ respectively, ignoring cycles of length 1 which represent the identity permutation. The order of α is then obtained by taking the least common multiple of $m_1, m_2, \cdots, m_r$.

Example 31. Let

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 2 & 3 & 4 & 1 & 7 & 9 & 6 & 5 & 8 & 12 & 11 & 10 \end{pmatrix}$$

$$= (1\ 2\ 3\ 4)(5\ 7\ 6\ 9\ 8)(10\ 12)(11)$$

$$= (1\ 2\ 3\ 4)(5\ 7\ 6\ 9\ 8)(10\ 12)$$

The cycles in the decomposition of α are

$$(1\ 2\ 3\ 4), (5\ 7\ 6\ 9\ 8) \text{ and } (10\ 12)$$

with lengths 4, 5 and 2 respectively. The L.C.M. of 4, 5, 2 is 20. So the order of α is 20.

EXERCISE 2.3 Complete

1. Multiply the following permutations:

(i) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & 4 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 4 & 2 & 5 \end{pmatrix}$

(ii) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 5 & 6 & 4 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 3 & 2 & 5 & 1 & 4 \end{pmatrix}$

(iii) $\begin{pmatrix} 2 & 3 & 6 & 7 & 8 & 9 \\ 8 & 6 & 7 & 9 & 2 & 3 \end{pmatrix}$ and $\begin{pmatrix} 2 & 3 & 6 & 7 & 8 & 9 \\ 8 & 6 & 7 & 3 & 9 & 2 \end{pmatrix}$

2. Find the inverse of the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$$

3. Let $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 4 & 1 & 7 & 2 & 6 \end{pmatrix}$. Find the inverse of α.

(iii) $\cdot$ is distributive with respect to $+$, i.e., for all $a, b, c \in R$

$a \cdot (b + c) = a \cdot b + a \cdot c$ (Left Distributive Law)

and $(a + b) \cdot c = a \cdot c + b \cdot c$ (Right Distributive Law)

The dot notation for products is usually omitted and it is customary to write $a \cdot b$ just as ab.

The binary operation $+$ is called **addition** and the binary operation $\cdot$ is called **multiplication**. The ring R is denoted by $(R, +, \cdot)$.

Example 32. $(Z, +, \cdot)$ is a ring, where $+$ and $\cdot$ denote respectively the ordinary addition and multiplication in Z. The set E of even integers is also a ring[1].

Solution.

(i) To show that $(Z, +, \cdot)$ is a ring we verify, one by one, the properties mentioned in the definition of a ring. $(Z, +)$ is a group under addition $+$. This follows from the following known results.

1. For all $m, n \in Z$, then sum $m + n$ also is a an integer and so belongs to Z.
2. For all $m, n, p \in Z$,

$$(m + n) + p = m + (n + p)$$

3. The number $0 \in Z$ and

$$n + 0 = n \quad \text{for all } n \in Z$$

4. For each $n \in Z$, $-n \in Z$ and $n + (-n) = n - n = 0$
5. For all $m, n \in Z$,

$$m + n = n + m \qquad \textbf{(Commutative Law)}$$

From 1 – 5 we see that $(Z, +)$ is an abelian group.

Next, under multiplication

(ii) For all $m, n \in Z$, $m \cdot n \in Z$

For all $m, n, p \in Z$,

$$(m \cdot n) \cdot p = m \cdot (n \cdot p)$$

1. Some authors assume a ring to have a multiplicative identity. Under that hypothesis, E is not regarded as a ring.

(iii) The distributive laws hold

For all $m, n, p \in Z$,

$(m + n) \cdot p = m \cdot p + n \cdot p$ **(Right Distributive Law)**

$m \cdot (n + p) = m \cdot n + m \cdot p$ **(Left Distributive Law)**

Hence $(Z, +, \cdot)$ is a ring with identity 1.

The statement given above for $(Z, +, \cdot)$ also hold for $(E, +, \cdot)$ where E is the set of even integers. Note that E has no multiplicative identity.

The proofs for Examples 33 – 37 are on similar lines as given for $(Z, +, \cdot)$

Example 33. $(Q, +, \cdot)$ is a ring, where + and $\cdot$ denote respectively ordinary addition and multiplication in Q

Example 34. $(R, +, \cdot)$ where R is the set of real numbers is a ring

Example 35. $(C, +, \cdot)$ is a ring, where + and $\cdot$ denote respectively the ordinary addition and multiplication of complex numbers in C

Example 36. $\left(\{r : r \in R, r = x + y\sqrt{2}, x, y \in Z\}, +, \cdot\right)$ is a ring, where + and $\cdot$ denote respectively ordinary addition and multiplication in R

Example 37. Let $\bar{Z}_n$ denote the set of all residue classes of integers modulo n i.e.,

$$\bar{Z}_n = \{\bar{0}, \bar{1}, \bar{2}, \cdots, \overline{n-1}\}$$

Then $\bar{Z}_n$ is a ring under the binary operations of addition and multiplication of residue classes modulo n

Recall that if $\bar{a} \in \bar{Z}_n$, $\bar{b} \in \bar{Z}_n$, then $\bar{a} + \bar{b}$ is the remainder obtained by dividing the sum $a + b$ of the integers a and b by n. Similarly, $\overline{ab}$ is the remainder obtained by dividing ab by n

Example 38. $(S, +, \cdot)$ where $S = \{a, b\}$ with + and $\cdot$ defined as in the table below

+	a	b
a	a	b
b	b	a

and

.	a	b
a	a	a
b	a	b

is a ring. This is a very important ring. Here if we take $a = 0$, $b = 1$ then $S = \{0, 1\}$ with addition and multiplication modulo 2 is a ring

(2.45) Properties of Rings. Let $(R, +, \cdot)$ be a ring. Since $(R, +)$ is a group, all the group properties for addition hold in R. The unique identity of the additive group is called the **zero element** of the ring, and we denote it by 0. Thus $a + 0 = 0 + a = a$ for every $a \in R$. We denote the unique additive inverse of an element a by $-a$, and write $a + (-a) = a - a = 0$ and $a + (-b) = a - b$. We note that the properties of the additive inverse and the zero element give us the cancellation law, if $a + b = a + c$, then $b = c$, a unique solution $x = b - a$ of the equation $a + x = b$, and the rule $-(-a) = a$. If $n \in Z$ and $n > 0$, we define $na = a + a + a + \cdots + a$ (n terms). If $n < 0$, we define $na = (-a) + (-a) + (-a) + \cdots + (-a)$ (n terms). If $n = 0$, we define $0a = 0$, where $0 \in Z$ on the left side of the equation and $0 \in R$ on the right side. The reader may note that na is not to be considered as a product of n and a in the ring, for the integer n may not be in the ring at all. The equation $0a = 0$ holds also for $0 \in R$ on both sides

The following theorem proves this and various other easy but important facts

Note the strong use of the distributive laws in the proof of this theorem. These distributive laws are the only tools of relating additive concepts to multiplicative concepts in a ring.

(2.46) Theorem. If $(R, +, \cdot)$ is a ring with additive identity 0, then, for all $a, b \in R$, we have

(i) $a0 = 0a = 0$

(ii) $a(-b) = (-a)b = -ab$

(iii) $(-a)(-b) = ab$

Proof. (i) $a0 = a(0 + 0) = a0 + a0$

Hence $0 + a0 = a0 + a0$

By the cancellation law for the additive group $(R, +)$, we have

$$0 = a0$$

Likewise, $0a = (0 + 0)a = 0a + 0a$ implies $0a = 0$

(ii) By definition, $-(ab)$ is the element which added to ab gives 0. Thus in order to show $a(-b) = -(ab)$, we must show $a(-b) + ab = 0$. By the left distributive law

$$a(-b) + ab = a(-b + b) = a0 = 0$$

Similarly, $(-a)b + ab = (-a + a)b = 0b = 0$

The result follows from the above equations.

(iii) Using (ii) repeatedly, we have

$$(-a)(-b) = -(a(-b)) = -(-(ab)) = ab$$

(2.47) Definition. A ring in which multiplication is commutative is called a **commutative ring.** A ring $(R, +, \cdot)$ with **multiplicative identity** 1 such that $1a = a1 = a$ for all $a \in R$ is called a **ring with unity**. The multiplicative identity in a ring is called **unity.**

In a ring with unity, the nonzero elements satisfy all the axioms for a group under multiplication except possibly the existence of multiplicative inverses. A multiplicative inverse of an element a in a ring $(R, +, \cdot)$ with unity 1 is an element $a^{-1} \in R$ such that $aa^{-1} = a^{-1}a = 1$. Since $0a = a0 = 0$, 0 cannot have a multiplicative inverse unless we regard the set $\{0\}$, where $0 + 0 = 0$ and $(0)\cdot(0) = 0$, as a ring with 0 as both the additive and multiplicative identity. We agree to exclude this trivial case whenever we speak of a ring with unity i.e., we shall regard a ring with unity to be a nonzero ring.

(2.48) Definition. Let $(R, +, \cdot)$ be a ring with unity. An element $a \in R$ is called a **unit** of $(R, +, \cdot)$ if it has a multiplicative inverse in R. If every nonzero element of R is a unit, then $(R, +, \cdot)$ is called a **skew field** (or **division ring**).

The term **unit** and **unity** should not be confused. Unity is the multiplicative identity, while a unit is any element having a multiplicative inverse. Thus the multiplicative identity or unity is a unit, but every unit is not a unity. For example, -1 is a unit in the ring $(Z, +, \cdot)$ of all integers, but -1 is not a unity, that is, $-1 \neq 1$.

(2.49) Definition. A commutative division ring is called **a field.**

Of interest is a second definition of a field which is easily proved to be equivalent to the first definition given above.

(2.50) Definition. A field $(F, +, \cdot)$ is a nonempty set F having at least two elements and two binary operations $+$ and $\cdot$ defined on F such that the following axioms are satisfied:

(i) $(F, +)$ *is an abelian group under addition*

(ii) $(F - \{0\}, \cdot)$ *is an abelian group under multiplication*

(iii) For all $a, b, c \in F$, the right distributive law holds: i.e.,

$$(a + b)c = ac + bc$$

Example 39.

(i) $(Q, +, \cdot)$ is a field, called the field of rational numbers.

(ii) $(R, +, \cdot)$ is a field, called the field of real numbers.

(iii) $(C, +, \cdot)$ is a field, called the field of complex numbers.

(iv) Let $S = \{a + b\sqrt{3} : a, b \in Q\}$. Then $(S, +, \cdot)$, where $+$ and $\cdot$ denote ordinary addition and multiplication in R, is a field.

(v) Let p be a prime number. Then $\bar{Z}_p = \{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{p-1}\}$ is a field under binary operations of addition and multiplication of residue classes modulo p. $\bar{Z}_p$ is a field having only a finite number of elements.

In particular, for $p = 2, 3$ the rings

$$\bar{Z}_2 = \{\bar{0}, \bar{1}\}, \bar{Z}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$$

form finite fields with addition and multiplication tables as given below.

$\bar{Z}_2$:

+	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$

·	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$

and

$\bar{Z}_3$:

+	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

·	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

The conditions for $\bar{Z}_2$ and $\bar{Z}_3$ to be field can be verified from the addition and multiplication tables

Note that the inverse of $\bar{2}$ in $\bar{Z}_3$ is $\bar{2}$ itself

Solution:

(i) It is easy to verify that

(a) $(Q, +)$ is an abedian group with addition '+', defined as follows

For $\frac{p}{q}, \frac{p'}{q'} \in Q$, $\frac{p}{q} + \frac{p'}{q'} = \frac{pq' + qp'}{qq'}$

(b) $(Q \setminus \{0\}, \cdot)$ is a commutative group under multiplication. Here

$$\frac{p}{q} \cdot \frac{p'}{q'} = \frac{pq' + qp'}{qq'}, \frac{p}{q}, \frac{p'}{q'} \in Q \setminus \{0\}$$

The identity element in $Q \setminus \{0\}$ and hence in Q is 1, under the multiplication given above

For each $\frac{p}{q} \neq 0$ in Q, $\frac{q}{p}$ is its inverse

All other requirements of a field can now be easily verified

So $(Q, +, \cdot)$ is a field

These remarks hold good for $(R, +, \cdot)$ and $(C, +, \cdot)$ as well

(iv) $S = \{a + b\sqrt{3} : a, b \in Q\}$, under addition and multiplication defined by

$$(a + b\sqrt{3}) + (a' + b'\sqrt{3}) = (a + a') + (b + b')\sqrt{3}$$

and $(a + b\sqrt{3})(a' + b'\sqrt{3}) = aa' + 3bb' + (ab' + ba')\sqrt{3}$

satisfies the requirements of a ring with $1 = 1 + 0\sqrt{3}$ as the identity.

For S to be a field we have only to note that the inverse of a nonzero element

$$a + b\sqrt{3}, \quad (a \neq 0 \text{ or } b \neq 0)$$

is $\frac{a - \sqrt{3}b}{a^2 - 3b^2}$.

EXERCISE 2.4

1. Answer true or false

(i) Every field is also a ring

(ii) Every ring has a multiplicative identity

(iii) The distributive laws for a ring are not very important

(iv) Multiplication in a field is commutative

(v) The nonzero elements of a field form a group under multiplication

(vi) Addition in every ring is commutative

(vii) Every element in a ring has an additive inverse

2. Determine whether the following sets with respect to ordinary addition and multiplication are rings. If they are rings, state are they also fields?

(i) The positive integers

(ii) Numbers of the form $b\sqrt{2}$, $b \in Q$

(iii) Numbers of the form $3m$, m is an integer

(iv) The four 4th roots of unity

(v) Numbers of the form $a + b\sqrt{2}$; a and b are integers

(vi) Numbers of the form $a + b\sqrt{2}$; a and b are rational numbers

(vii) Numbers of the form $a + ib$, a and b are integers and $i = \sqrt{-1}$

(viii) Numbers of the form $a + ib$, a and b are rational numbers and $i = \sqrt{-1}$

(ix) The ordered pairs of rational numbers (a, b) with equality, addition and multiplication defined as follows:

$$(a, b) = (c, d) \text{ if and only if } a = c \text{ and } b = d$$

$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b)(c, d) = (ac, bd).$$

3. Let $R = \{a, b, c, d\}$ Define $+$ and $\cdot$ in R as follows

+	a	b	c	d
a	a	b	c	d
b	b	a	d	[illegible]
c	c	d	a	b
d	d	c	b	a

and

·	a	b	c	d
a	a	a	a	a
b	a	b	a	b
c	a	a	c	c
d	a	b	c	d

Show that $(R, +, \cdot)$ is a ring

4. Prove that in a ring $(R, +, \cdot)$
$(a + b + c)\,d = ad + bd + cd$ for all a, b, c, d belonging to R

5. Let $(R, +, \cdot)$ be a ring such that $a^2 = a$ for all $a \in R$. Prove that
   (i) $2a = 0$ for all $a \in R$
   (ii) $ab = ba$ for all $a, b \in R$

6. Verify that the following definition of a number field is equivalent to the axioms of a field
A set consisting of at least two numbers forms a field if the sum, difference, product, and quotient of any two numbers are again members of the set, division by zero being excluded

Chapter 3

MATRICES

INTRODUCITON

The importance of study of matrices lies in the fact that many situations in both pure and applied mathematics involve rectangular arrays of numbers. In many branches of business, biological and social sciences, it is necessary to express and use a set of numbers arranged in a rectangular array. For example, suppose a firm produces three types of good G_1, G_2 and G_3 which it sells to two customers C_1 and C_2. The monthly sales of these goods (in hundreds) are given in the following table

		Monthly sales of goods		
		G_1	G_2	G_3
Customers	C_1	9	4	5
	C_2	2	6	7

Ignoring table headings, we usually write this information more concisely as

$$\begin{bmatrix} 9 & 4 & 5 \\ 2 & 6 & 7 \end{bmatrix}$$

which is an example of a matrix. It has two rows and three columns. A formal definition of a matrix as a rectangular array of numbers follows.

(3.1) Definition. Let F denote the field of real or complex numbers. A rectangular array of mn elements $a_{ij} \in F$, $(1 \le i \le m,\ 1 \le j \le n)$ arranged in m **rows** and n **columns** and enclosed by square brackets such as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

is called an $m \times n$ **matrix** over the field F (or with entries from the field F). Matrices are usually designated by capital letters. Instead of writing all the entries of a matrix as above, it is convenient to write the same in an abbreviated notation as $A = [\,a_{ij}\,]_{m \times n}$. Here a_{ij} denotes the **element** (or **entry**) in the ith row and the jth column of the matrix A which has m **rows** and n **columns**. Thus the subscripts i and j of he element a_{ij} of a matrix indicate respectively the row and the column in which a_{ij} is located.

If a matrix A has m rows and n columns then A is said to be a **rectangular** matrix **of order (or size)** $m \times n$. In case $m = n$, A is said to be a **square** matrix of order n.

If $A = [\,a_{ij}\,]_{m \times n}$, a single row $[\,a_{i1}\ a_{i2}\ \cdots\ a_{in}\,]$ is a $1 \times n$ matrix and is called **a row vector**. Similarly, a single column

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

is an $m \times 1$ matrix and is called a **column vector**.

(3.2) Definition. An **upper (lower) triangular matrix** (or **a triangular matrix**) is a square matrix all of whose elements below (above) the **main diagonal** (running from upper left to lower right corner) are zero. Thus

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

is an upper triangular matrix.

$a_{11}, a_{22}, \cdots, a_{nn}$ are called elements of the **main diagonal**.

(3.3) Definition. A square matrix, all of whose elements are zero except those in the main diagonal, is called **a diagonal matrix**.

Thus, a square matrix $A = [\,a_{ij}\,]$ is a diagonal matrix if $a_{ij} = 0$ whenever $i \ne j$.

(Note that this does not say $a_{ij} \ne 0$ whenever $i = j$.)

An example of a diagonal matrix is

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

Here $a, b, c \in F$ which may or may not be zero.

(3.4) Definition. A square matrix of size n in which each element of the main diagonal is one and the same nonzero number $k \in F$ and all other elements are zero is called **a scalar matrix** of order n. Thus

$$\begin{bmatrix} k & 0 & 0 & 0 \\ 0 & k & 0 & 0 \\ 0 & 0 & k & 0 \\ 0 & 0 & 0 & k \end{bmatrix}$$ is a scalar matrix of order 4.

Every scalar matrix is a diagonal matrix.

When each $k = 1$ (the identity element of F with respect to multiplication), the scalar matrix of order n is called the **identity matrix** and is denoted by I_n. Thus

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

are identity matrices of order 3 and 4 respectively.

ALGEBRA OF MATRICES

(3.5) Definition. (Equality of Matrices). Two $m \times n$ matrices $A = [\,a_{ij}\,]$ and $B = [\,b_{ij}\,]$ are equal if and only if $a_{ij} = b_{ij}$ for each i and for each j.

When A and B are equal, we write $A = B$.

(3.6) Definition. (Zero Matrix) A matrix, whose every element is zero, is called **a zero matrix**. If it has m rows and n columns, we denote it by $0_{m \times n}$ or simply by 0 if there is no danger of ambiguity about the number of its rows and columns.

(3.7) Definition. (Addition of Matrices). Two matrices are said to be conformable for addition when they have the same number of rows and the same number of columns. Thus if $A = [\,a_{ij}\,]$ and $B = [\,b_{ij}\,]$ are $m \times n$ matrices over the same field F, then they can be added and their sum is the matrix

$$A + B = [a_{ij} + b_{ij}]$$

of order $m \times n$. That is, to find the sum $A + B$ of two matrices A and B of the same order we add their corresponding elements.

(3.8) Definition. (Additive Inverse of a Matrix). Given an $m \times n$ matrix $A = [a_{ij}]$, we define $-A = [\,-a_{ij}\,]$. Thus $-A$ is an $m \times n$ matrix and by definition

$$A + (-A) = [a_{ij} + (-a_{ij})\,] = 0 = [(-a_{ij}) + a_{ij}\,]$$
$$= (-A) + A.$$

Thus $-A$ is the additive inverse of A.

Also, $0 + A = A + 0 = A$, i.e., 0 is the identity with respect to $+$.

(3.9) Theorem. For any three $m \times n$ matrices A, B and C, we have

$A + (B + C) = (A + B) + C$ **(Associative Law for Matrix Addition)**

Proof. Let $A = [\,a_{ij}\,]$, $B = [\,b_{ij}\,]$, $C = [\,c_{ij}\,]$. Then,

$$\begin{aligned} A + (B + C) &= [\,a_{ij}\,] + [\,b_{ij} + c_{ij}\,] \\ &= [\,a_{ij} + (b_{ij} + c_{ij})], && \textbf{(Definition 3.7)} \\ &= [(a_{ij} + b_{ij}) + c_{ij}], && \textbf{(Associative Law in } F\textbf{)} \\ &= [\,a_{ij} + b_{ij}\,] + [\,c_{ij}\,], && \textbf{(Definition 3.7)} \end{aligned}$$

In fact we have

(3.10) Theorem. The set of all $m \times n$ matrices (over a field F) is an abelian group with respect to $+$ defined for matrices in (3.7).

Proof. Left as an exercise.

(3.11) Definition. (Subtraction of Matrices). If A and B are two matrices of the same order, we define $A - B$ as $A - B = A + (-B)$. For example, let

$$A = \begin{bmatrix} 2 & 1 & -3 \\ -4 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\begin{aligned} \text{Then } A - B &= A + (-B) \\ &= \begin{bmatrix} 2 & 1 & -3 \\ -4 & 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 1 & -2 \\ 0 & -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2-1 & 1+1 & -3-2 \\ -4+0 & 0-1 & 1+0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -5 \\ -4 & -1 & 1 \end{bmatrix} \end{aligned}$$

(3.12) Definition. (Multiplication of a Matrix by a Scalar). Let $A = [\,a_{ij}\,]$ and $k \in F$. Then we define the scalar multiple kA of A by k as $kA = [\,ka_{ij}\,]$ i.e., kA is the matrix obtained by multiplying each element of A by k.

The additive inverse of a matrix A can also be defined as $-A = (-1)A$, where 1 is the multiplicative identity of F.

Example 1. Let

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{bmatrix}, \text{ then } 2A = \begin{bmatrix} 2 & -4 & 6 \\ 8 & 10 & -12 \end{bmatrix}.$$

(3.13) Theorem. Let A and B be $m \times n$ matrices over a field F and let $a, b \in F$. Then

(i) $(a + b)A = aA + bA$

(ii) $a(A + B) = aA + aB$

(iii) $a(bA) = (ab)A$

(iv) $1A = A$

Proof. (i) Let $A = [\,a_{ij}\,]$, $B = [\,b_{ij}\,]$. Then

$$\begin{aligned} (a + b)A &= [\,(a + b)\,a_{ij}\,] \\ &= [\,aa_{ij} + ba_{ij}\,] = [\,aa_{ij}\,] + [\,ba_{ij}\,] \\ &= a\,[\,a_{ij}\,] + b\,[\,a_{ij}\,] = aA + bA \end{aligned}$$

(ii)
$$\alpha(A+B) = \alpha[a_{ij}+b_{ij}] = [\alpha a_{ij}+\alpha b_{ij}] = [\alpha a_{ij}] + [\alpha b_{ij}]$$
$$= \alpha[a_{ij}] + \alpha[b_{ij}] = \alpha A + \alpha B$$

The other results follow similarly and are left as an exercise.

(3.14) Definition. (Multiplication of Matrices). Two matrices A and B are said to be conformable for the **product** AB if the number of columns in A is equal to the number of rows in B. Thus if A is an $m \times n$ matrix and B is an $n \times p$ matrix, then the product $C = AB$ of the matrices A and B is an $m \times p$ matrix C whose entries are determined as follows.

Let $A = [a_{ij}]_{m\times n}$ and $B = [b_{ij}]_{n\times p}$ then $AB = C$, where

$C = [c_{ij}]_{m\times p}$ and

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

In other words, c_{ij} is the sum of the products of the elements of the ith row of A with the corresponding elements of the jth column of B i.e., c_{ij} is the dot product of the ith row vector of A with the jth column vector of B.

Example 2. Let

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 9 & 7 \\ 5 & 3 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & 1 \\ 0 & 4 \\ -2 & 3 \end{bmatrix}, \text{ then}$$

$$AB = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 9 & 7 \\ 5 & 3 & 0 \end{bmatrix}\begin{bmatrix} 6 & 1 \\ 0 & 4 \\ -2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1(6)+0(0)-1(-2) & 1(1)+0(4)-1(3) \\ 2(6)+9(0)+7(-2) & 2(1)+9(4)+7(3) \\ 5(6)+3(0)+0(-2) & 5(1)+3(4)+0(3) \end{bmatrix} = \begin{bmatrix} 8 & -2 \\ -2 & 59 \\ 30 & 17 \end{bmatrix}$$

Note that here A is a 3×3 matrix, B is a 3×2 matrix and AB is a 3×2 matrix. Moreover, the product BA is not defined because the number of columns of B is not equal to the number of rows of A.

The commutative law for multiplication of matrices does not hold in general. That is, for matrices A and B, $AB \neq BA$ even when both AB and BA are defined.

Example 3. Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}. \text{ Then}$$

$$AB = \begin{bmatrix} 1(1)+2(0) & 1(1)+2(2) \\ 3(1)+4(0) & 3(1)+4(2) \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 3 & 11 \end{bmatrix}$$

$$BA = \begin{bmatrix} 1(1)+1(3) & 1(2)+1(4) \\ 0(1)+2(3) & 0(2)+2(4) \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 8 \end{bmatrix}$$

Thus $AB \neq BA$.

We may have $AB = \mathbf{0}$ when neither $A = \mathbf{0}$ nor $B = \mathbf{0}$.

Example 4. Let

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 4 & 9 \end{bmatrix}. \text{ Then}$$

$$AB = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ but neither } A = \mathbf{0} \text{ nor } B = \mathbf{0}$$

For three matrices A, B, C, we may have $AB = AC$ (or $BA = CA$) with $B \neq C$. Thus cancellation law does not hold for multiplication of matrices.

Example 5. Let

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 2 & 2 & 2 \end{bmatrix}$$

$$\text{and } C = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}. \text{ Then}$$

$$AB = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \\ 3 & 2 & -7 \end{bmatrix} \text{ and } AC = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \\ 3 & 2 & -7 \end{bmatrix}$$

Here $AB = AC$ but $B \neq C$

(3.15) Theorem. If the matrices A, B and C are conformable for the indicated sums and products then

(i) $A(BC) = (AB)C$ **(Associative Law for Matrix Multiplication)**

(ii) $A(B+C) = AB + AC$ **(Left Distributive Law for Matrix Multiplication)**

$(A+B)C = AC + BC$ **(Right Distributive Law for Matrix Multiplication)**

(iii) $k(AB) = (kA)B = A(kB)$, where $k \in F$

Proof. (i) Let $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{n\times p}$ and $C = [c_{ij}]_{p\times q}$

Then $A(BC)$ is an $m \times q$ matrix and (i, j)th element of $A(BC)$ is the sum of the product of elements of the ith row of A and the corresponding elements of the jth column of BC. Now jth column of BC consists of elements

$$\sum_{\mu=1}^{p} b_{1\mu} c_{\mu j},\ \sum_{\mu=1}^{p} b_{2\mu} c_{\mu j},\ \cdots,\ \sum_{\mu=1}^{p} b_{n\mu} c_{\mu j}$$

Hence (i, j) element of $A(BC)$ is

$$a_{i1}\left(\sum_{\mu=1}^{p} b_{1\mu} c_{\mu j}\right) + a_{i2}\left(\sum_{\mu=1}^{p} b_{2\mu} c_{\mu j}\right) + \cdots + a_{in}\left(\sum_{\mu=1}^{p} b_{n\mu} c_{\mu j}\right)$$

$$= \sum_{\lambda=1}^{n} a_{i\lambda}\left(\sum_{\mu=1}^{p} b_{\lambda\mu} c_{\mu j}\right)$$

$$= \sum_{\lambda=1}^{n}\sum_{\mu=1}^{p} a_{i\lambda}(b_{\lambda\mu} c_{\mu j})$$

$$= \sum_{\lambda=1}^{n}\sum_{\mu=1}^{p} (a_{i\lambda} b_{\lambda\mu}) c_{\mu j} \qquad \text{(Associative Law in } F\text{)}$$

$$= \sum_{\mu=1}^{p}\left(\sum_{\lambda=1}^{n} a_{i\lambda} b_{\lambda\mu}\right) c_{\mu j}, \qquad \text{(Intercanging the summation symbols.)}$$

Hence, (i, j)th element of $A(BC) = (i, j)$th element of $(AB)C$ which is true for each i and each j.

Hence $A(BC) = (AB)C$.

(ii) Here we assume that $A = [a_{ij}]$ is an $m \times n$ matrix and each of $B = [b_{ij}]$ and $C = [c_{ij}]$ is an $n \times p$ matrix, so that $B + C$ is defined and is an $n \times p$ matrix. Moreover, $A(B + C)$, AB and AC are also defined and each of them is an $m \times p$ matrix. $AB + AC$ is also an $m \times p$ matrix.

Now $B + C = [(b_{\lambda j} + c_{\lambda j})]$

So (i, j)th element of $A(B + C) = \sum_{\lambda=1}^{n} a_{i\lambda}(b_{\lambda j} + c_{\lambda j})$

$$= \sum_{\lambda=1}^{n} (a_{i\lambda} b_{\lambda j} + c_{\lambda j})$$

$$= \sum_{\lambda=1}^{n} a_{i\lambda} b_{\lambda j} + \sum_{\lambda=1}^{n} a_{i\lambda} c_{\lambda j}$$

$= (i, j)$th element of AB + (i, j)th element of AC

$= (i, j)$th element of $(AB + AC)$

Since the above equation holds for all i, j, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, p$, we have
$A(B + C) = AB + AC$

We can prove similarly that $(A + B)C = AC + BC$

(iii) Suppose $A = [a_{ij}]$ is of order $m \times n$ and $B = [b_{ij}]$ is of order $n \times p$. Hence

$$k\left(\sum_{\lambda=1}^{n} a_{i\lambda} b_{\lambda j}\right) = \sum_{\lambda=1}^{n} (ka_{i\lambda}) b_{\lambda j} = \sum_{\lambda=1}^{n} a_{i\lambda}(kb_{\lambda j}).$$

Because the terms involved in the summations are all elements of the field in which commutative and associative laws hold. Hence

(i, j)th element of $k(AB)$ = (i, j)th element of $(kA)B$

= (i, j)th element of $A(kB)$

So $k(AB) = (kA)B = A(kB)$

(3.16) Definition. (Transpose of a Matrix). Let $A = [a_{ij}]$ be an $m \times n$ matrix over a field F. The **transpose** of A, denoted by A^T, is an $n \times m$ matrix obtained by interchanging rows and columns of A. Thus $A^T = [b_{ij}]$, where $b_{ij} = a_{ji}$. In other words, (i, j)th element of $A^T = (j, i)$th element of A.

Thus the first row of A^T is the first column of A, second row of A^T is the second column of A and so on.

(3.17) Theorem. If the matrices A and B are conformable for the sum $A + B$ and the product AB, then

(i) $(A \pm B)^T = A^T \pm B^T$

(ii) $(A^T)^T = A$

(iii) $(kA)^T = kA^T$, where $k \in F$

(iv) $(AB)^T = B^T A^T$.

Proof. (i) Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $m \times n$ matrices. Then $A + B$ is an $m \times n$ matrix and $(A + B)^T$ is an $n \times m$ matrix. Now A^T is an $n \times m$ matrix and B^T also is an $n \times m$ matrix. Therefore, $A^T + B^T$ is an $n \times m$ matrix.

Thus each of $(A + B)^T$ and $A^T + B^T$ is an $n \times m$ matrix. Now

$$\begin{aligned}(i, j)\text{th element of } (A + B)^T &= (j, i)\text{th element of } A + B \\ &= a_{ji} + b_{ji}\end{aligned}$$

And
$$\begin{aligned}(i, j)\text{th element of } A^T + B^T &= (i, j)\text{th element of } A^T + (i, j)\text{th element of } B^T \\ &= (j, i)\text{th element of } A + (j, i)\text{th element of } B \\ &= a_{ji} + b_{ji}\end{aligned}$$

Thus, for all i and j,

$$(i, j)\text{th element of } (A + B)^T = (i, j)\text{th element of } A^T + B^T$$

Hence $\quad (A + B)^T = A^T + B^T$

Similarly $\quad (A - B)^T = (A + (-B))^T = A^T + (-B)^T = A^T - B^T$

We can similarly prove that

$$(A_1 + A_2 + \cdots + A_n)^T = A^T_1 + A^T_2 + \cdots + A^T_n.$$

(ii)
$$\begin{aligned}(i, j)\text{th element of } (A^T)^T &= (j, i)\text{th element of } A^T \\ &= (i, j)\text{th element of } A.\end{aligned}$$

Thus $\quad (A^T)^T = A.$

(iii)
$$\begin{aligned}(i, j)\text{th element of } (kA)^T &= (j, i)\text{th element of } kA \\ &= ka_{ji} \\ &= (i, j)\text{th element of } kA^T\end{aligned}$$

Hence $\quad (kA)^T = kA^T.$

(iv) Let $A = [a_{ij}]$ be an $m \times n$ matrix and $B = [b_{ij}]$ be an $n \times p$ matrix.

Then AB is of order $m \times p$ and so $(AB)^T$ is of order $p \times m$.

Now B^T is of order $p \times n$ and A^T is of order $n \times m$. Hence $B^T A^T$ is of order $p \times m$.

So

$$\begin{aligned}(i, j)\text{th element of } (AB)^T &= (j, i)\text{th element of } AB \\ &= \sum_{\lambda=1}^{n} a_{j\lambda} b_{\lambda i}\end{aligned}$$

Hence $\quad (AB)^T = B^T A^T.$

This result can be generalized for the product of a finite number of matrices $A_1, A_2, \cdots, A_n$. Thus

$$(A_1 A_2 \cdots A_n)^T = A^T_n \cdots A^T_2 A^T_1$$

(3.18) Definition. Let A be a square matrix of order n. We define

$$A^n = A \cdot A \cdot A \cdots A \text{ } (n \text{ factors})$$

It can be proved by induction that, for all $m, n \in N$

$$A^m \cdot A^n = A^{m+n},$$

$$(A^m)^n = A^{mn}$$

We also define $A^0 = I_n$.

Also, if A and B are conformable for multiplication then $(AB)^n \neq A^n B^n$ in general. However, $(AB)^n = A^n B^n$ holds if $AB = BA$.

(3.19) Definition. A square matrix A for which $A^{k+1} = A$, (k being a positive integer), is called a **periodic matrix**.

If k is the least positive integer for which $A^{k+1} = A$, then A is said to be of period k. If $k = 1$, so that $A^2 = A$, then A is called an **idempotent matrix**.

Example 6. Consider

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

$$A^2 = \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}\begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix} = A.$$

Hence A is idempotent.

(3.20) Definition. A square matrix A for which $A^p = 0$, (p being a positive integer), is called **nilpotent**. If p is the least positive integer for which $A^p = 0$, then A is said to be **nilpotent of nilpotency index p**.

Example 7. Let

$$A = \begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix} \text{ Then}$$

$$A^2 = \begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}\begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}$$

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

$$A^3 = A^2A = \begin{bmatrix} 0 & 0 & 0 \\ 3 & 3 & 9 \\ -1 & -1 & -3 \end{bmatrix}\begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0$$

Hence A is nilpotent of nilpotency index 3

(3.21) Definition. A square matrix A such that $A^2 = I$ is called an **involutory matrix**

Example 8. Let

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ Then}$$

$$A^2 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

so that A is involuntary

(3.22) Definition. A square matrix A for which $A^T = A$ is called a **symmetric matrix**

A square matrix A for which $A^T = -A$ is called a **skew symmetric matrix**

Note that the product of two symmetric matrices may not be symmetric. That is, if $A^T = A,\ B^T = B$ then $(AB)^T = B^T A^T = BA \neq AB$ is general.

Example 9. For

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -3 \\ 3 & -3 & 6 \end{bmatrix}, \text{ we have } A^T = A. \text{ Hence } A \text{ is symmetric}$$

Also for the matrix

$$A = \begin{bmatrix} 0 & -2 & 3 \\ 2 & 0 & 4 \\ -3 & -4 & 0 \end{bmatrix}, A^T = -A. \text{ Hence } A \text{ is skew symmetric}$$

(3.23) Definition. Let $A = [\, a_{ij}\,]$ be a matrix with complex entries. Then the matrix $[\, \bar{a}_{ij}\,]$, obtained from A by replacing each a_{ij} by its complex conjugate, is called the **conjugate matrix of A** and is denoted by $\bar{A}$ (read A conjugate). $(\bar{A})^T$ is called **Hermitian transpose of A.**

A square matrix A such that $(\bar{A})^T = A$ is called **Hermitian** (after the name of French mathematician Charles Hermit, 1822–1901). $(\bar{A})^T$ is also denoted by A^H. Thus A is Hermitian if $A^H = A$.

A square matrix $A = [\, a_{ij}\,]$ over C for which $(\bar{A})^T = -A$ is called **skew Hermitian**.

As for the product of two symmetric matrices, the product of two Hermitian matrices A and B need not be Hermitian. That

$$(AB)^H = B^H A^H = BA \neq AB \text{ in general.}$$

However, if A and B are conformable for multiplication then

$$(B^T AB)^T = B^T A^T B \text{ and } (B^H AB)^H = B^H A^H B$$

So, for symmetric (Hermitian) A, $B^T AB$, $(B^H AB)$ is symmetric (Hermitian).

Example 10. Let

$$A = \begin{bmatrix} 1 & 1-i & 2 \\ 1+i & 3 & i \\ 2 & -i & 0 \end{bmatrix} \text{ Then } \bar{A} = \begin{bmatrix} 1 & 1+i & 2 \\ 1-i & 3 & -i \\ 2 & i & 0 \end{bmatrix}$$

$$\text{and } (\bar{A})^T = \begin{bmatrix} 1 & 1-i & 2 \\ 1+i & 3 & i \\ 2 & -i & 0 \end{bmatrix} = A. \text{ Thus } A \text{ is Hermitian.}$$

Example 11. Let

$$A = \begin{bmatrix} i & 1-i & 2 \\ -1-i & 3i & i \\ -2 & i & 0 \end{bmatrix}$$ It is easy to verify that $(\bar{A})^T = -A$

Hence A is skew Hermitian.

Example 12. Show that for any square matrix A, $A + A^T$ is symmetric and $A - A^T$ is skew symmetric.

Solution. Here we use the results $(A \pm B)^T = A^T \pm B^T$ and $(A^T)^T = A$. Thus

$$(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$$

Hence $A + A^T$ is symmetric.

Again, $(A - A^T)^T = A^T - (A^T)^T = A^T - A = -(A - A^T)$

Hence $A - A^T$ is skew symmetric.

Analogous to **Theorem 3.17**, we have the following relations between two matrices A and B and their Hermitian forms.

(3.24) Theorem. Let A and B be square matrices of the same order. Then

(i) $(A^H)^H = A$, (ii) $(A \pm B)^H = A^H \pm B^H$

(iii) $(cA)^H = (\bar{c})A^H$, (iv) $(AB)^H = B^H A^H$

Proof. Left to the reader.

PARTITIONING OF MATRICES

(3.25) Definition. Some time we partition a matrix into blocks of elements and consider such blocks as elements of the matrix. These blocks are called **submatrices** of the original matrix and are labeled in the same manner as elements of a matrix. A matrix can be partitioned into submatrices in many different ways and the manner in which it is to be partitioned is often indicated by drawing horizontal and vertical lines between selected rows and columns.

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

be an $m \times n$ matrix. We can write a partitioning of A as

$$A = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \text{ where}$$

$$B_{11} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pk} \end{bmatrix}$$

$$B_{12} = \begin{bmatrix} a_{1\,k+1} & a_{1\,k+2} & \cdots & a_{1n} \\ a_{2\,k+1} & a_{2\,k+2} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{p\,k+1} & a_{p\,k+2} & \cdots & a_{pn} \end{bmatrix}$$

$$B_{21} = \begin{bmatrix} a_{p+1\,1} & a_{p+1\,2} & \cdots & a_{p+1\,k} \\ a_{p+2\,1} & a_{p+2\,2} & \cdots & a_{p+2\,k} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix}$$

$$B_{22} = \begin{bmatrix} a_{p+1\,k+1} & a_{p+1\,k+2} & \cdots & a_{p+1\,n} \\ a_{p+2\,k+1} & a_{p+2\,k+2} & \cdots & a_{p+2\,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m\,k+1} & a_{m\,k+2} & \cdots & a_{mn} \end{bmatrix}$$

Here

B_{11} is a $p \times k$ matrix.

B_{12} is a $p \times (n-k)$ matrix.

B_{21} is an $(m-p) \times k$ matrix.

B_{22} is an $(m-p) \times (n-k)$ matrix.

Let A, B be two $m \times n$ matrices. A and B are said to be **identically partitioned** if the corresponding submatrices C_{ij} and D_{ij} of A and B are of the same order (i.e. have the same size).

In such a case $C_{ij} \pm D_{ij}$ are both defined.

Example 13. Consider

$$A = \begin{bmatrix} 1 & -2 & 5 \\ 3 & 0 & 2 \\ 4 & 2 & -6 \end{bmatrix}$$

A can be partitioned as shown below:

$$A = \left[\begin{array}{cc|c} 1 & -2 & 5 \\ 3 & 0 & 2 \\ \hline 4 & 2 & -6 \end{array}\right]$$

$$= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \text{ where}$$

$A_{11} = \begin{bmatrix} 1 & -2 \\ 3 & 0 \end{bmatrix}$ is a 2×2 submatrix

$A_{12} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ is a 2×1 submatrix

$A_{21} = [4 \;\; 2]$ is a 1×2 submatrix

$A_{22} = [-6]$ is a 1×1 submatrix

Example 14. Partitioning of the matrices A and B are given as under:

$$A = \left[\begin{array}{cc|c} 1 & 2 & 5 \\ \hline -4 & 3 & 0 \\ 2 & 1 & 5 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

and
$$B = \left[\begin{array}{cc|c} a & b & c \\ \hline d & e & f \\ g & h & i \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

Here A and B are identically partitioned.

A_{11} and B_{11} are 1×2 submatrices

A_{12} and B_{12} are 1×1 submatrices

A_{21} and B_{21} are 2×2 submatrices

A_{22} and B_{22} are 2×1 submatrices

Moreover, $A_{11} + B_{11}$, $A_{12} + B_{12}$, $A_{21} + B_{21}$ and $A_{22} + B_{22}$ are all defined. We have

$$A + B = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix}$$

$$= \left[\begin{array}{cc|c} 1+a & 2+b & 5+c \\ \hline -4+d & 3+e & 0+f \\ 2+g & 1+h & 5+i \end{array}\right]$$

Partitioning can be used in matrix multiplication provided that the matrices are partitioned in such a way that the corresponding submatrices to be multiplied, are conformable for multiplication.

Example 15. Let

$$A = \left[\begin{array}{cc|c} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \hline a_{31} & a_{32} & a_{33} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$B = \left[\begin{array}{cc} b_{11} & b_{12} \\ b_{21} & b_{22} \\ \hline b_{31} & b_{32} \end{array}\right] = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}, \text{ where}$$

$$A_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, A_{12} = \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}, A_{21} = [a_{31} \;\; a_{32}]$$

$$A_{22} = [a_{33}], \quad B_{11} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, B_{21} = [b_{31} \;\; b_{32}]$$

Then
$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix}$$

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

$$A_{12}B_{21} = \begin{bmatrix} a_{13}b_{31} & a_{13}b_{32} \\ a_{23}b_{31} & a_{33}b_{23} \end{bmatrix}$$

$$A_{21}B_{11} = [a_{31}b_{11} + a_{32}b_{21} \;\; a_{31}b_{12} + a_{32}b_{22}]$$

$$A_{22}B_{21} = [a_{33}a_{13} + a_{33}b_{32}]$$

Thus

$$\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} \end{bmatrix}$$

$$= AB$$

This method can be generalized for the product of any pair of matrices which are conformable for multiplication. It is called **block multiplication** of matrices.

Example 16. Let

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

We partition A as shown below:

$$A = \left[\begin{array}{ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$= \begin{bmatrix} I_3 & 0_{3\times2} & A_1 \\ 0_{2\times3} & I_2 & 0_{2\times1} \\ A^T_1 & 0_{1\times2} & [1] \end{bmatrix}, \text{ where } A_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\text{Now } A^2 = \begin{bmatrix} I_3 & 0_{3\times2} & A_1 \\ 0_{2\times3} & I_2 & 0_{2\times1} \\ A^T_1 & 0_{1\times2} & [1] \end{bmatrix} \begin{bmatrix} I_3 & 0_{3\times2} & A_1 \\ 0_{2\times3} & I_2 & 0_{2\times1} \\ A^T_1 & 0_{1\times2} & [1] \end{bmatrix}$$

$$= \begin{bmatrix} I_3 + A_1A^T_1 & 0_{3\times2} & A_1 + A_1 \\ 0_{2\times3} & I_2 & 0_{2\times1} \\ A^T_1 + A^T_1 & 0_{1\times2} & A^T_1A_1 + [1] \end{bmatrix}$$

$$I_3 + A_1A^T_1 = I_3 + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} [1 \;\; 1 \;\; 1]$$

$$= I_3 + \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

$$A_1 + A_1 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

$A^T_1 + A^T_1 = [2 \;\; 2 \;\; 2]$, $A^T_1A_1 + [1] = [3] + [1] = [4]$[1] $= 4$

$$\text{Hence } A^2 = \begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 2 \\ 1 & 2 & 1 & 0 & 0 & 2 \\ 1 & 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 2 & 2 & 2 & 0 & 0 & 4 \end{bmatrix}$$

Note. In a number of cases dealing with matrices having bigger orders, the operations on matrices are done by partitioning them into matrices of smaller orders. This is done specially in the situation when the memory capacity of computer is limited. The sizes of the submatrices must conform to the matrix operations to be done on the submatrices.

Since each matrix always indicates some information (see the Example in **Introduction**), partitioning of a matrix leads to the division of the information into separate blocks or parts.

The same process can be reversed by taking into account the given matrices and adjoin them together and get a matrix of a bigger size which gives the information in consolidated form.

The partition of matrices is useful in a number of ways. Since matrices present information in a compact form, the same information could be divided into a number of inter connected sub-systems to make the computation and calculation easier.

1. A 1×1 matrix is usually written without brackets.

EXERCISE 3.1

1. Let

$$A = \begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 3 & 5 \\ 1 & -3 & -5 \\ -1 & 3 & 3 \end{bmatrix}$$

Find (i) $A + B$ (ii) $A - B$

(iii) $2A + 3B$ (iv) $3A - 5B$

(v) AB (vi) BA

Is $AB = BA$?

2. Evaluate

(i) $\begin{bmatrix} 1 & 2 & 1 \\ 4 & 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & -4 \\ 1 & 5 \\ -2 & 2 \end{bmatrix}$ (ii) $[x \; y \; z]\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 2 \\ 4 & -3 \end{bmatrix}^2$ (iv) $\begin{bmatrix} 1 & 2 \\ 4 & -3 \end{bmatrix}^3$

(v) $\begin{bmatrix} -1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^2$ (vi) $\begin{bmatrix} -1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^3$

3. Prove that the product of matrices

$$A = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix} \text{ and } B = \begin{bmatrix} \cos^2\phi & \cos\phi\sin\phi \\ \cos\phi\sin\phi & \sin^2\phi \end{bmatrix}$$

is the zero matrix when θ and ϕ differ by an odd multiple of $\frac{\pi}{2}$.

4. The direction cosines of two lines are $\lambda_1, \lambda_2, \lambda_3$ and μ_1, μ_2, μ_3. Prove that the product

$$\begin{bmatrix} \lambda_1^2 & \lambda_1\lambda_2 & \lambda_1\lambda_3 \\ \lambda_1\lambda_2 & \lambda_2^2 & \lambda_2\lambda_3 \\ \lambda_1\lambda_3 & \lambda_2\lambda_3 & \lambda_3^2 \end{bmatrix}\begin{bmatrix} \mu_1^2 & \mu_1\mu_2 & \mu_1\mu_3 \\ \mu_1\mu_2 & \mu_2^2 & \mu_2\mu_3 \\ \mu_1\mu_3 & \mu_2\mu_3 & \mu_3^2 \end{bmatrix}$$

is zero if and only if the lines are perpendicular to each other.

5. Show that, in general, for any two matrices A and B which are conformable for addition and multiplication,

$$(A + B)^2 \neq A^2 + 2AB + B^2$$

and $$A^2 - B^2 \neq (A - B)(A + B)$$

Under what conditions equality holds in each case?

6. If

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}, \text{ show that}$$

$$A^2 - 4A - 5I = \begin{bmatrix} 2 & -2 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & -2 \end{bmatrix} \text{ and}$$

$$A^3 - 3A^2 - 7A - 3I = 0$$

7. Show that the matrix

$$A = \begin{bmatrix} 1 & -2 & -6 \\ -3 & 2 & 9 \\ 2 & 0 & -3 \end{bmatrix} \text{ is periodic having period 2}$$

8. Show that

$$\begin{bmatrix} 1 & -3 & -4 \\ -1 & 3 & 4 \\ 1 & -3 & -4 \end{bmatrix} \text{ is nilpotent. What is its nilpotency index?}$$

9. Show that

$$\begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix} \text{ is involutory.}$$

10. Show that every square matrix A with entries from R, can be written as a symmetric matrix $B = \frac{1}{2}(A + A^T)$ and a skew symmetric matrix $D = \frac{1}{2}(A - A^T)$.

11. Show that every square matrix A with entries from C, can be written as a Hermitian matrix $B = \frac{1}{2}(A + (\bar{A})^T)$, and as a skew Hermitian matrix $D = \frac{1}{2}(A + (-\bar{A})^T)$

12. If A and B are symmetric matrices, then prove that AB is symmetric if and only if A and B commute.

13. If A is an $m \times m$ symmetric (skew symmetric) matrix and P is an $m \times n$ matrix, then prove that $B = P^T AP$ is symmetric (skew symmetric).

14. Show that AA^T and A^TA are symmetric for any square matrix A.

15. If A is a square matrix over C, then show that $A + (\bar{A})^T$, $A(\bar{A})^T$ and $(\bar{A})^T A$ are all Hermitian.

16. Show that every square matrix over C can be expressed in a unique way as $P + iQ$, where P and Q are Hermitian.

17. Show that every Hermitian matrix can be written as $A + iB$, where A is real and symmetric and B is real and skew symmetric.

18. Show that, for any matrices A and B over C and $k \in C$,

(i) $\bar{\bar{A}} = A$ (ii) $\overline{kA} = \bar{k}\bar{A}$

(iii) $\overline{A+B} = \bar{A} + \bar{B}$ (iv) $\overline{AB} = \bar{A}\,\bar{B}$

(v) $(\bar{A})^T = \overline{(A)^T}$

19. If A is a matrix over R and $AA^T = 0$, show that $A = 0$.

20. If A is a matrix over C and $A(\bar{A})^T = 0$, then show that $A = 0 = A$.

21. Show that for any matrix A, $(A^p)^T = (A^T)^p$, where p is a positive integer.

22. Let $x = [x_1 \;\; x_2 \;\; x_3]$, $y = [y_1 \;\; y_2 \;\; y_3]$

be two 1×3 matrices. Define a product $x \times y$ (called a **vector product** of x and y), as follows:

$$x \times y = [x_2y_3 - y_2x_3 \quad x_3y_1 - y_3x_1 \quad x_1y_2 - y_1x_2]$$

Show that

(i) $y \times x = -(x \times y)$ (ii) $x \times x = 0$ for all x.

(iii) $(x \times y) \times z \neq x \times (y \times z)$, where $z = [z_1 \;\; z_2 \;\; z_3]$

23. Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Define **inner product** of A and B by

$A \cdot B = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}$

Show that: (i) $A \cdot B = 0$ but neither $A = 0$ nor $B = 0$

(ii) $A \cdot B = B \cdot A$

(iii) $(\alpha A + \beta B) \cdot C = \alpha A \cdot C + \beta B \cdot C$, $C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$

(iv) $A \cdot A \geq 0$ and $A \cdot A = 0 \Leftrightarrow A = 0$.

24. Let

$$A = \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 3 & 4 & 0 \\ \hline 0 & 0 & 2 \end{array}\right]$$

and

$$B = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 3 & 6 & 1 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

Compute AB using the indicated partitionings.

25. Let

$$A = \left[\begin{array}{cc|ccc} 1 & 2 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$B = \left[\begin{array}{ccc|cc} 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & -2 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{array}\right]$$

Find AB using the indicated partitionings.

INVERSE OF A MATRIX

(3.26) Definition. (Inverse of a Matrix). Let A be a square matrix of order n. A matrix B of order n is said to be the inverse of A if $AB = BA = I_n$

Note that the inverse of an arbitrary square matrix may not exist.

For example the matrix $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ has no inverse.

An inverse of a matrix, if it exits, is unique. For, suppose that A has two inverses, say B and C. Then

$$AB = BA = I \quad \text{and} \quad AC = CA = I$$

Therefore, by the associative law for multiplication, we have

$$B(AC) = BI = B$$

and $\quad (BA)C = IC = C$

and so $\quad B = C.$

The unique inverse of a matrix A, if it exists, is denoted by A^{-1}.

A square matrix A, whose inverse exists, is called a **nonsingular** (or **invertible**) matrix. Square matrices which do not have inverses are called **singular** matrices.

(3.27) Properties of Inverse of a Matrix. It is easy to verify the following properties of inverses of matrices.

(i) For any invertible matrix A, $(A^{-1})^{-1} = A$.

(ii) For nonsingular matrices A and B, $(AB)^{-1} = B^{-1}A^{-1}$.

(iii) For any invertible matrix A, $(A^T)^{-1} = (A^{-1})^T$.

Proof. (ii) Since A and B are nonsingular, A^{-1} and B^{-1} exist. Also, since A and B are square matrices AB and $B^{-1}A^{-1}$ are defined.

To prove that AB is nonsingular and $B^{-1}A^{-1}$ is the inverse of AB, we show that

$$(AB)(B^{-1}A^{-1}) = I = (B^{-1}A^{-1})(AB)$$

Now $\quad (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$

and $\quad (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I.$

Thus the product AB of two nonsingular matrices A and B is nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$

We can generalize the above result for the product of a finite number of nonsingular matrices of the same order. Thus, if $A_1, A_2, \cdots, A_m$ are all nonsingular matrices of the same order, then

$$(A_1 A_2 \cdots A_m)^{-1} = A_m^{-1} \cdots A_2^{-1} A_1^{-1}$$

So the product of any finite number of nonsingular matrices is nonsingular.

Proofs of **(i)** and **(iii)** are left as an exercise.

Example 17. Let

$$A = \begin{bmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1/4 & -1/4 & 3/4 \\ 1/4 & 3/4 & 3/4 \\ 0 & -1 & 1 \end{bmatrix}$$

Now $AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$ and $BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$

Hence $B = A^{-1}$. Also $B^{-1} = A$.

ELEMENTARY ROW OPERATIONS

Consider the identity matrix

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The matrices

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \; E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{bmatrix}$$

have been obtained from I_3 by the following operations.

E_1: Interchange of two rows of I_3. (Second and third rows interchanged)

E_2: Multiply a row of I_3 by a nonzero number. (Here row two is multiplied by 3)

E_3: Add a multiple of a row of I_3 to another row. (-5 times row two is added to row three)

These operations on I_3 are called **elementary row operations** and corresponding matrices E_1, E_2, E_3 are called **elementary matrices** of order 3

Now let

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

We examine the effect of premultiplying A by E_1, E_2 and E_3 respectively

$$E_1A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

= A matrix obtained from A by interchange of rows two and three

$$E_2A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{21} & 3a_{22} & 3a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

= A matrix obtained from A by multiplying row two by 3

$$E_3A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31}-5a_{21} & a_{32}-5a_{22} & a_{33}-5a_{23} \end{bmatrix}$$

= A matrix obtained from A by subtracting 5 times the second row from the third row

If we postmultiply the matrix A by

$$E_1' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, E_2' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_3' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{bmatrix}$$

Then

$$AE_1', AE_2' \text{ and } AE_3'$$

are the matrices obtained from A by doing similar operations on columns of A instead of rows of A

This fact can be generalized to any $m \times n$ matrix

That is the same operations can be performed on any arbitrary matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

by premultiplying A by the $m \times m$ elementary matrices E_1, E_2 and E_3 obtained from I_m by the row operations mentioned above respectively

This is later proved in **Theorem 3.31**

The above illustrations lead to

(3.28) Definition. The following operations on a matrix A are called **elementary row operations**

(i) Interchange of any two rows (R_{ij} denotes the interchange of the ith row with the jth row)

(ii) Multiplication of a row of A by any nonzero real or complex number (kR_i denote that ith row is multiplied by $k \neq 0$)

(iii) Addition of a scalar multiple of one row to another row ($R_j + kR_i$) denotes that k times row i is added to the jth row

Note that each of these row operations on an $m \times n$ matrix can be effected by premultiplying A by an $m \times m$ elementary matrix.

Let A be an $m \times n$ matrix. An $m \times n$ matrix B is called **row equivalent** to A if B is obtained from A by performing a finite sequence of elementary row operations on A. We write $B \overset{R}{\sim} A$ to denote B is row equivalent to A.

(3.29) Definition. An $m \times n$ matrix A is said to be in **(row) echelon**[1] **form** (or **an echelon matrix**) if it has the following properties

(i) All nonzero rows are above any zero rows (consisting of all zeros)

(ii) The first nonzero entry in each nonzero row is to the right of the first nonzero entry of each preceding row That is, the number of zeros occurring before the first nonzero entry in each nonzero row is greater than the number of zeros that appear before the first nonzero element in any preceding row

In an echelon matrix, the first nonzero entry of a row is called a **pivot** (or **a row leader**) There is at the most one pivot in each row and in each column of an echelon matrix A column containing a pivot is called a **pivot column.**

An echelon matrix in which each pivot is 1 and every other entry of the pivot column is zero, is said to be in **row reduced echelon form.**

The matrices

$$\begin{bmatrix} 0 & 5 & 2 & 1 \\ 0 & 0 & 9 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

are in echelon form The second matrix is in row reduced echelon form

(3.30) Theorem. Every matrix is row equivalent to a matrix in echelon form (reduced echelon form)

Proof. Let A be a given $m \times n$ matrix Either every element in the first column of A is zero or there exists a nonzero element a (say) in the kth row, of this column In the second case, interchanging the first and kth rows of A, we obtain a matrix B whose first element $b_{11} = a \neq 0$. Multiplying the elements of the first row of B by a^{-1}, we obtain a matrix C whose first entry $c_{11} = 1$ By adding proper multiples of the first row of C to other rows, the remaining elements of the first column can be made zeros In the first case, we consider the second column and so on till we find a column with a nonzero entry We then repeat the process as for the second case Thus A is row equivalent to a matrix in either of the forms

$$[\,\boldsymbol{0} \;\; C\,] \quad \text{or} \quad \begin{bmatrix} 1 & D \\ \boldsymbol{0} & E \end{bmatrix}$$

where, in the first case, $\boldsymbol{0}$ represents the $m \times 1$ zero matrix and C an $m \times (n-1)$ matrix.

1 Staircase

In the second case $\boldsymbol{0}$ represents the $(m-1) \times 1$ zero matrix, D a $1 \times (n-1)$ matrix and E an $(m-1) \times (n-1)$ matrix In the first case repeat the above process with the matrix C, whereas in the second case repeat the process with the matrix E. A continuation of this process leads to the desired form

When a matrix is in echelon form, then by adding suitable multiplies of second, third, ... rows to the first, second, ... rows, in succession, we get the reduced echelon form

Example 18. Reduce the matrix

$$A = \begin{bmatrix} 6 & 3 & -4 \\ -4 & 1 & -6 \\ 1 & 2 & -5 \end{bmatrix} \quad \text{into the echelon form}$$

Solution. Here

$$A \underset{\sim}{R} \begin{bmatrix} 1 & 2 & -5 \\ -4 & 1 & -6 \\ 6 & 3 & -4 \end{bmatrix} \quad \text{by } R_{13}$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & -9 & 26 \end{bmatrix} \quad \text{by } R_2 + 4R_1 \text{ and } R_3 - 6R_1$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{by } R_3 + R_2$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & 2 & -5 \\ 0 & 1 & -\frac{26}{9} \\ 0 & 0 & 0 \end{bmatrix} \quad \text{by } \frac{1}{9} R_2$$

which is the desired echelon form.

By applying $R_1 - 2R_2$ further, we get the reduced echelon form

$$\begin{bmatrix} 1 & 0 & \frac{7}{9} \\ 0 & 1 & -\frac{26}{9} \\ 0 & 0 & 0 \end{bmatrix}$$

(3.31) Notation.

Let I_{R_k} denote the kth row of the identity matrix I of order n. When we apply the operation R_{ij} to I we obtain an elementary matrix H whose rows H_{R_α} are

$$H_{R_i} = I_{R_j}, \quad H_{R_j} = I_{R_i}$$

and $$H_{R_k} = I_{R_k}, \quad (k \neq i, j)$$

(that is ith row of H is the same as jth row of I and jth row of H is the same as ith row of I and rows other than ith and jth remain unaltered).

When we apply the operation kR_i to I, we obtain the elementary matrix M whose rows M_{R_α} are given by

$$M_{R_i} = kI_{R_i}, \quad M_{R_j} = I_{R_j}, \quad (j \neq i)$$

Again the operation $R_i + kR_j$ applied to I produces the elementary matrix K whose rows K_{R_α} are given by

$$K_{R_i} = I_{R_i} + kI_{R_j}$$
$$K_{R_k} = I_{R_k}, \quad (k \neq i)$$

(3.32) Theorem. An elementary row operation on an $m \times n$ matrix A amounts to pre-multiplying A by the corresponding elementary $m \times m$ matrix E.

Proof. Suppose we apply the operation $R_i + kR_j$ to the matrix A and obtain the matrix B. Let E be the elementary matrix obtained from I_m when $R_i + kR_j$ is applied to I_m. Then we have to prove that $B = EA$.

$$\begin{aligned}(EA)_{R_i} &= (E)_{R_i} A \\ &= (I_{R_i} + kI_{R_j})A \\ &= I_{R_i}A + kI_{R_j}A \\ &= (IA)_{R_i} + k(IA)_{R_j} \\ &= A_{R_i} + kA_{R_j} = B_{R_i}\end{aligned}$$

For $l \neq i$, $(EA)_{R_l} = E_{R_l}A = I_{R_l}A = (IA)_{R_l} = A_{R_l} = B_{R_l}$

Hence $$B = EA$$

The proof for other row operations is similar.

(3.33) Corollary. Every elementary matrix E is nonsingular.

Proof. E is obtained from I by some elementary row operation. The reverse elementary operation corresponds to some elementary matrix E' and transforms E into I. By Theorem 3.32 this reverse operation takes E into $E'E$. Hence $E'E = I$ and so E is nonsingular.

(3.34) Corollary. If $B \overset{R}{\sim} A$ then $B = PA$, where P is nonsingular.

Proof. By Theorem 3.32, $B = E_n E_{n-1} \cdots E_1 A$, where E_i are elementary matrices and so are nonsingular. But the product of nonsingular matrices is nonsingular. Hence $P = E_n E_{n-1} \cdots E_1$ is nonsingular and $B = PA$.

(3.35) Theorem. A square matrix A of order n is nonsingular if and only if it is row equivalent to I_n.

Proof. Suppose A is nonsingular and let B be the matrix in reduced echelon form row equivalent to A. Then $B = PA$, where P is nonsingular. Since the product of nonsingular matrices is nonsingular, B is nonsingular and so B^{-1} exists. Let $B = [b_{ij}]$. Since B is in echelon form, at least $j - 1$ zeros precede the first nonzero element 1 in the jth row of B. Hence if $b_{kk} = 0$, the $(k+1)$st row has at least $k + 1$ zeros preceding the first nonzero element 1. If $k = n$, the nth row of B consists of zeros, and if $k < n$, at least one row of B after the kth consists of zeros. Consequently, if $b_{kk} = 0$, B has a row of zeros. Thus $BB^{-1} = I_n$ has a row of zeros contrary to the definition of I_n. Thus $b_{kk} \neq 0$ for any k. Hence $b_{ii} = 1$ for all i and $b_{ij} = 0$ when $i \neq j$. Thus $B = I_n$ and so $A \overset{R}{\sim} I_n$.

Conversely, let $A \overset{R}{\sim} I_n$. Then $I_n = PA$ where P is nonsingular. Hence $A = P^{-1} I_n$ is nonsingular.

(3.36) Theorem. If a square matrix A is reduced to the identity matrix by a sequence of elementary row operations, the same sequence of operations performed on the identity matrix produces the inverse of A.

Proof. We are given

$$I \overset{R}{\sim} A$$

Hence $I = (E_r E_{r-1} \cdots E_2 E_1) A$, where E_i's are suitable elementary matrices

Thus $$\begin{aligned}A^{-1} &= A^{-1} I = IA^{-1} = (E_r E_{r-1} \cdots E_2 E_1) A A^{-1} \\ &= (E_r E_{r-1} \cdots E_2 E_1) I\end{aligned}$$

(3.37) Corollary. A square matrix A is nonsingular if and only if it can be written as a product of elementary matrices i.e.,

$$A = E_r E_{r-1} \cdots E_2 E_1.$$

Proof. Here $A = (A^{-1})^{-1}$ and by Theorem **3.36**, we have

$$(A^{-1}) = E_r E_{r-1} \cdots E_2 E_1 I$$
$$= E_r E_{r-1} \cdots E_2 E_1.$$

Hence
$$A = (E_r E_{r-1} \cdots E_2 E_1)^{-1}$$
$$= E_1^{-1} E_2^{-1} \cdots E_{r-1}^{-1} E_r^{-1}.$$

Since the inverse of an elementary matrix is an elementary matrix, A is the product of elementary matrices.

Conversely, suppose that A is the product of elementary matrices. Since each elementary matrix is nonsingular, product of any finite number of nonsingular matrices is nonsingular. Hence A is nonsingular.

(3.38) Corollary. $B \overset{R}{\sim} A$ if and only if $B = PA$, where P is nonsingular.

Example 19. By using elementary row operations, find the inverse of the matrix

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

Solution. To find the inverse of A we perform row operations on A and simultaneously the same row operations on I_3 in such a way that A reduces to the identity matrix while I_3 is transformed into a matrix B. This matrix B is then the inverse of A.

(Note that we can perform two row operations in one and the same step as explained in the solution).

Here we write the matrix A and he identity matrix I_3 separately as follows, indicating the row operations on A and I_3.

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 4 & 1 \\ 1 & 3 & 0 \end{bmatrix} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A \overset{R}{\sim} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 4 & -5 \\ 0 & 3 & -3 \end{bmatrix} \text{ by } R_2 - 2R_1,\ R_3 - R_1 \qquad I_3 \overset{R}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \text{ by } R_2 - 2R_1,\ R_3 - R_1$$

$$\overset{R}{\sim} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 3 & -3 \end{bmatrix} \text{ by } R_2 - R_3 \qquad \overset{R}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \text{ by } R_2 - R_3$$

$$\overset{R}{\sim} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \end{bmatrix} \text{ by } R_3 - 3R_2 \qquad \overset{R}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ 2 & -3 & 4 \end{bmatrix} \text{ by } R_3 - 3R_2$$

$$\overset{R}{\sim} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \text{ by } \frac{1}{3}R_3 \qquad \overset{R}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ \frac{2}{3} & -1 & \frac{4}{3} \end{bmatrix} \text{ by } \frac{1}{3}R_3$$

$$\overset{R}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ by } R_1 - 3R_3,\ R_2 + 2R_3 \qquad \overset{R}{\sim} \begin{bmatrix} -1 & 3 & -4 \\ \frac{1}{3} & -1 & \frac{5}{3} \\ \frac{2}{3} & -1 & \frac{4}{3} \end{bmatrix} \text{ by } R_1 - 3R_3,\ R_2 + 2R_3$$

Thus

$$B = A^{-1} = \begin{bmatrix} -1 & 3 & -4 \\ \frac{1}{3} & -1 & \frac{5}{3} \\ \frac{2}{3} & -1 & \frac{4}{3} \end{bmatrix}$$

Verification:

$$AB = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 4 & 1 \\ 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} -1 & 3 & -4 \\ \frac{1}{3} & -1 & \frac{5}{3} \\ \frac{2}{3} & -1 & \frac{4}{3} \end{bmatrix}$$

$$= \begin{bmatrix} -1+2 & 3-3 & -4+4 \\ -2+2 & 6-5 & -8+8 \\ -1+1 & 3-3 & -4+5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

NOTE: If, in the above process, a row in the matrix on the left hand side consists of just zeros then the given matrix has no inverse.

ELEMENTARY COLUMN OPERATIONS

So far we have been considering elementary row operations. We can consider elementary column operations in an analogous manner.

(3.39) Definition. The following operations on a matrix A are called **elementary column operations.**

(i) Interchange of any two columns of A

(ii) Multiplication of a column of A by a nonzero scalar

(iii) Addition of a scalar multiple of one column of A to another column

On applying an elementary column operation to the identity matrix, we obtain a corresponding elementary matrix, which is nonsingular.

(3.40) Definition. An $m \times n$ matrix B is called **column equivalent to** an $m \times n$ matrix A if B is obtained by performing a sequence of a finite number of elementary column operations on A. We write $B \overset{C}{\sim} A$ to denote B is column equivalent to A.

An elementary column operation on a matrix A becomes an elementary row operation on the transpose A^T of A. Thus if $B \overset{C}{\sim} A$, then $B^T \overset{R}{\sim} A^T$. Hence $B^T = EA^T$ where E is an elementary matrix. Thus $B = AE^T$ and we have proved

(3.41) Theorem. An elementary column operation on an $m \times n$ matrix A amounts to post-multiplying A by an $n \times n$ elementary matrix corresponding to the elementary column operation.

(3.42) Theorem. An $m \times n$ matrix B is column equivalent to an $m \times n$ matrix A if and only if $B = AQ$, where Q is an $n \times n$ nonsingular matrix.

Proof. Let $B \overset{C}{\sim} A$. Then $B^T \overset{R}{\sim} A^T$. Hence $B^T = PA^T$, where P is an $n \times n$ nonsingular matrix. Now

$$B = (B^T)^T = (PA^T)^T = (A^T)^T P^T = AP^T = AQ,$$ where $Q = P^T$ is an $n \times n$ nonsingular matrix.

So $B \overset{C}{\sim} A \;\Rightarrow\; B = AQ$ for a nonsingular matrix Q.

Conversely, if $B = AQ$, where Q is nonsingular, then

$$B^T = (AQ)^T = Q^T A^T \quad \text{so that} \quad B^T \overset{R}{\sim} A^T. \text{ That is } B \overset{C}{\sim} A.$$

(3.43) Definition. If an $m \times n$ matrix B can be obtained from an $m \times n$ matrix A by a finite number of elementary row and column operations, then B is said to be **equivalent** to A. We write $B \sim A$ to denote B is equivalent to A.

The row and column equivalences are special cases of the general concepts of equivalence of matrices. Combining our previous results, we have

(3.44) Theorem. An $m \times n$ matrix B is equivalent to an $m \times n$ matrix A if and only if $B = PAQ$ where P and Q are nonsingular and of orders $m \times m$ and $n \times n$ respectively.

(3.45) Theorem. Every nonzero $m \times n$ matrix is equivalent to an $m \times n$ matrix D where

$$D = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

(D is called the **canonical** (or **normal form**) of the matrix A)

Proof. Since $A \neq 0$, there exists a nonzero element a of A. By performing elementary row and column operations on A, we obtain an equivalent matrix with element a in the (1, 1) position. Applying the operation $a^{-1} R_1$, we get 1 in the (1, 1) position. By adding suitable multiples of the first row to the remaining rows, every other element of the first column can be made zero. Having done that we can reduce every element of the first row, except that in the (1, 1) position, to a zero. We thus obtain an equivalent matrix of the form

$$B = \begin{bmatrix} I & 0 \\ 0 & C \end{bmatrix}$$

Here 0 in the first row is a $1 \times (n-1)$ zero matrix, 0 in the first column is an $(m-1) \times 1$ zero matrix and C is an $(m-1) \times (n-1)$ matrix. Repeating the process with C, we get the required form.

The procedure is explained in the following example.

Example 20. Reduce the matrix

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix}$$ to the normal form. Also find the nonsingular matrices P and Q such that PAQ is in the normal form.

Solution. Here we write the given matrix A and the identity matrix I_3 (twice) in three columns as below. The same row operations as on A are performed on the first identity matrix and the same column operations as on A are performed on the second identity matrix. Thus

A	I_3 Row operations	I_3 Column operations
$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

By $R_2 - R_1$, we have

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

By $C_2 - C_1$ and $C_3 - 2C_1$, we get

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

By $R_3 + R_2$, we obtain

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

By $C_3 - C_2$, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} I_2 & 0_{2\times 1} \\ 0_{1\times 2} & 0_{1\times 1} \end{bmatrix} \qquad P \qquad Q$$

Thus the required normal form is $\begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}$, with

$$P = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

Verification:

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

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

$$= \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}$$

(3.46) Definition. (Rank of a Matrix). The **rank** of a matrix A is equal to the number of nonzero rows in its echelon (or reduced echelon) form or the order of I_r in the canonical form of A.

A more formal definition of the rank of a matrix is given in **(6.37)**

To find the rank of A we just reduce A to its echelon (or reduced echelon) form or canonical form and count its nonzero rows.

Example 21. Find the rank of the matrix

$$A = \begin{bmatrix} 5 & 9 & 3 \\ -3 & 5 & 6 \\ -1 & -5 & -3 \end{bmatrix}$$

Solution. To find the rank of A, we reduce A to an echelon form. Thus:

$$A \underset{\sim}{R} \begin{bmatrix} -1 & -5 & -3 \\ -3 & 5 & 6 \\ 5 & 9 & 3 \end{bmatrix} \quad \text{by } R_{13}$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & 5 & 3 \\ -3 & 5 & 6 \\ 5 & 9 & 3 \end{bmatrix} \quad \text{by } (-1)R_1$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & 5 & 3 \\ 0 & 20 & 15 \\ 0 & -16 & -12 \end{bmatrix} \quad \text{by } R_2 + 3R_1 \text{ and } R_3 - 5R_1$$

$$\underset{\sim}{R}\begin{bmatrix} 1 & 5 & 3 \\ 0 & 1 & 3/4 \\ 0 & -16 & -12 \end{bmatrix} \quad \text{by } \frac{1}{20}R_2$$

$$\underset{\sim}{R}\begin{bmatrix} 1 & 5 & 3 \\ 0 & 1 & 3/4 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{by } R_3 + 16R_2.$$

This is an echelon form of A and the number of its nonzero rows is 2. Hence the rank of A is 2.

For an **Alternative Method** of finding the rank of a matrix, see **(6.42)**

EXERCISE 3.2

1. (i) Show that the inverse of a diagonal matrix, with all diagonal elements nonzero, is a diagonal matrix.

   (ii) Show that the inverse of a scalar matrix is a scalar matrix.

2. For a nonsingular matrix A, show that

   (i) $(A^n)^{-1} = (A^{-1})^n$, here n is a positive integer.

   (ii) $(kA)^{-1} = k^{-1}A^{-1}$, k is any nonzero scalar.

   (iii) $(A^{-1})^T = (A^T)^{-1}$

   (iv) $(\bar{A})^{-1} = (\overline{A^{-1}})$

   (v) $\left(\overline{A^T}\right)^{-1} = (\overline{A^{-1}})^T$.

3. If A is invertible and $AB = 0$, then show that $B = 0$.

4. Let A and B be distinct $n \times n$ matrices with real entries. If $AB^2 = BA^2$ and $A^3 = B^3$, show that $A^2 + B^2$ is not invertible.

5. Find the inverse of each of the following matrices:

(i) $\begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}$

(iii) $\begin{bmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{bmatrix}$ (iv) $\begin{bmatrix} 2 & 1 & -1 \\ 0 & 2 & 1 \\ 5 & 2 & -3 \end{bmatrix}$

(v) $\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}, \left(i = \sqrt{1}\right)$ (vi) $\begin{bmatrix} i & -1 & 2i \\ 2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \left(i = \sqrt{-1}\right)$

(vii) $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (viii) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 & 7 \end{bmatrix}$

6. Reduce each of the following matrices into the indicated form:

(i) $\begin{bmatrix} 1 & -2 & 3 & -1 \\ 2 & -1 & 2 & 2 \\ 3 & 1 & 2 & 3 \end{bmatrix}$ reduced echelon form

(ii) $\begin{bmatrix} 0 & 1 & 3 & -2 \\ 2 & 1 & -4 & 3 \\ 2 & 3 & 2 & -1 \end{bmatrix}$ reduced echelon form

(iii) $\begin{bmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & -2 & 3 \\ 3 & 6 & 2 & -6 & 3 \end{bmatrix}$ echelon form

(iv) $\begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{bmatrix}$ reduced echelon form.

7. Show that

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \underset{\sim}{R} I_3.$$

8. Find the rank of each of the following matrices:

(i) $\begin{bmatrix} 1 & 3 \\ 0 & -2 \\ 5 & -1 \\ -2 & 3 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 1 & 2 & -3 \\ 2 & 1 & 0 \\ -2 & -1 & 3 \\ -1 & 4 & -2 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{bmatrix}$ **(iv)** $\begin{bmatrix} 1 & 3 & -2 & 5 & 4 \\ 1 & 4 & 1 & 3 & 5 \\ 1 & 4 & 2 & 4 & 3 \\ 2 & 7 & -3 & 6 & 13 \end{bmatrix}$

(v) $\begin{bmatrix} ar^{(1-1)n} & ar^{(1-1)n+1} & \cdots & ar^{(1-1)n+(n-1)} \\ ar^{(2-1)n} & ar^{(2-1)n+1} & \cdots & ar^{(2-1)n+(n-1)} \\ \vdots & \vdots & \cdots & \vdots \\ ar^{(n-1)n} & ar^{(n-1)n+1} & \cdots & ar^{(n-1)n+(n-1)} \end{bmatrix}$, ($a$ and r are nonzero)

9. Reduce each of the following matrices to the canonical form. In each case also find nonsingular matrices P and Q.

(i) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & -1 & 3 \\ 2 & -4 & 1 \\ 0 & 3 & 2 \end{bmatrix}$ **(iv)** $\begin{bmatrix} 1 & 2 & 0 & -1 \\ 3 & 4 & 1 & 2 \\ -2 & 3 & 2 & 5 \end{bmatrix}$

10. **(i)** Using the row operations, show that the matrix

$\begin{bmatrix} 1 & 2 & -3 \\ 1 & -2 & 1 \\ 5 & -2 & -3 \end{bmatrix}$ has no inverse.

(ii) If $A^{-1} = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 4 \end{bmatrix}$ and $B^{-1} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 0 & -2 \\ 1 & 1 & -1 \end{bmatrix}$, compute $(AB)^{-1}$.

11. Let $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ be $n \times 1$ matrices.

Then the products $x^T y$ and xy^T are 1×1 and $n \times n$ matrices respectively. The products $x^T y$ and xy^T are respectively called the **inner (or scalar)** and **outer products** of x and y.

(i) Write the inner and outer products of x and y.

(ii) Show that the inner product of x and y is equal to the inner product of y and x (i.e., $x^T y = y^T x$).

(iii) Prove that rank $(x y^T) = 1$.

12. Let A and B be idempotent matrices i.e. $A^2 = A$, $B^2 = B$.

Show that:

(i) if $AB = BA$ then AB is idempotent.

(ii) if A^T is idempotent so is A.

Is the sum of two idempotent matrices idempotent? Justify your answer.

13. If A is an $n \times n$ nilpotent matrix, show that $I_n - A$ is nonsingular.

In most sciences one generation tears down what another has built and what one has established another undoes. In mathematics alone each generation builds a new story to the old structure.

HERMANN HANKEL

(German Mathematician 1839-1873 C.E.)

Chapter 4

SYSTEMS OF LINEAR EQUATIONS

PRELIMINARIES

An equation of the form

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b$$

is called a **linear equation** in the variables $x_1, x_2, x_3, \cdots, x_n$.

Here $a_1, a_2, \cdots, a_n, b$ are constants, usually real numbers.

A collection of several such equations is called a **system of linear equations.**

The theory of matrices has been usefully employed in various branches of pure as well as applied mathematics. In what follows, we shall apply this theory to the solution of m linear equations in n unknowns.

Consider the m linear equations:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \quad\quad \vdots \quad\quad \cdots \quad\quad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}\right\} \qquad (1)$$

in n unknowns $x_1, x_2, \cdots, x_n$, where a_{ij}, b_i are scalars, $i = 1, 2, \cdots, m; j = 1, 2, \cdots, n$.

Using the matrix notation, the system (1) can be written as

$$Ax = b$$

where A, x, b are the matrices as given below:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

The matrix A is called the **matrix of coefficients** of the system of equations (1), the column of constants b_i forms a column vector b and the unknowns x_j form the column vector x.

(4.1) Definition. The equations of the type given in (1) above with $b \neq 0$, are called a system of **nonhomogeneous linear equations**. If $b = 0$, then the system of equations (1) is known as a **homogeneous linear system**.

(4.2) Definition. An ordered n-tuple $y = (y_1, y_2, \cdots, y_n)$ is called a **solution** of the system (1) if, whenever $x_1, x_2, \cdots, x_n$ are replaced by $y_1, y_2, \cdots, y_n$ respectively in each equation of (1), the resulting statements are true statements. The vector y is then also said to satisfy the system of equations (1).

For example, the system of linear equations

$$x_1 + 2x_2 = 1$$
$$2x_1 + x_2 = 2$$

has the ordered pair (1, 0) as a solution. Here $y = (1, 0)$.

(4.3) Definition. The set S of all ordered n-tuples $y = (y_1, y_2, \cdots, y_n)$ satisfying the equations (1) is called the **solution set** of the given equations.

EQUIVALENT EQUATIONS

(4.4) Definition. Let $Ax = b$,

$$\text{where } A = [a_{ij}]_{m \times n}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad (1)$$

be a given system of m linear equations in n unknowns.

A system of linear equations

$$A'x = d$$

$$\text{where} \quad A' = [a'_{ij}]_{m \times n}, \quad d = \begin{bmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_m \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad (2)$$

is said to be **equivalent** to the system (1) if the matrices

$$[A \mid b] = A_b = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

$$\text{and} \quad [A' \mid b'] = A'_b = \left[\begin{array}{cccc|c} a'_{11} & a'_{12} & \cdots & a'_{1n} & b'_1 \\ a'_{21} & a'_{22} & \cdots & a'_{2n} & b'_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a'_{m1} & a'_{m2} & \cdots & a'_{mn} & b'_m \end{array}\right]$$

are row equivalent. Since application of an elementary row operation to the matrix A_b implies application of the corresponding operation to the equations (1), we see that the system (1) is equivalent to (2) if (2) is obtained from (1) by any of the following three operations:

(i) Any two equations of (1) are interchanged.

(ii) Any equation of (1) is multiplied by a nonzero scalar.

(iii) A scalar multiple of an equation of (1) is added to another equation of (1).

The three types of operations listed above are called **elementary** or **admissible operations** for the system (1).

Correspondingly, the two matrices $[A \mid b]$ and $[A' \mid b']$, as in (1) and (2) are said to be **row equivalent** if $[A' \mid b']$ can be obtained from $[A \mid b]$ by the following elementary row operations.

(i)′ Interchange of any two rows of $[A \mid b]$,

(ii)′ Multiplying a row of $[A \mid b]$ by a nonzero scalar.

(iii)′ Adding a scalar multiple of ith row of $[A \mid b]$ to its jth row.

The matrix

$$[A \mid b] = A_b = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

is known as the **augmented matrix** of the system of linear equations (1). It is quite important in the sense that its rank and the rank of A (the matrix of coefficients) determine whether the system of equations $Ax = b$ does or does not have a solution.

In fact the system of equations given by (1) has a unique solution if and only if rank A = rank A_b. **(Theorem 4.10)**

GAUSSIAN ELIMINATION METHOD

(4.5) One of the several methods employed to solve a system of n linear equations in n variables, is known as **Guassian Elimination Procedure** (named after its discoverer, the German mathematician **C.F. Gauss**, 1777 – 1855). Computer programmes exist to solve a system of equations by Guassian Elimination Method and the Guass-Jordan Method described in the following paragraphs. We explain these methods below.

Let

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 && (1)\\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 && (2)\\ \vdots \quad\quad \vdots \quad\quad \cdots \quad\quad \vdots \quad &\quad \vdots && \vdots\\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n && (n) \end{aligned}$$

i.e., $Ax = b$, where $A = [a_{ij}]_{n \times n}$,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

be a system of n linear equations in n unknowns $x_1, x_2, \cdots, x_n$. We can suppose, without any loss of generality, that coefficient of the first term of the first equation namely a_{11} is

not zero[1], and has the largest absolute value among the coefficients of the first terms of each equation. Such an element is called **pivot** and the corresponding equation is known as the **pivotal equation.**

Divide the first equation by a_{11} (i.e. the **pivot**) and obtain

$$x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n = \frac{b_1}{a_{11}} \qquad (1)$$

Multiply (1) by a_{i1}, $i = 2, 3, \cdots, n$ and subtract from the equations (2), (3), $\cdots$, (n) respectively, getting a new system of equivalent equations

$$\begin{aligned} x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n &= \frac{b_1}{a_{11}} \\ \left(a_{22} - a_{21}\frac{a_{12}}{a_{11}}\right)x_2 + \cdots + \left(a_{2n} - a_{21}\frac{a_{1n}}{a_{11}}\right)x_n &= b_2 - a_{21}\frac{b_1}{a_{11}} \\ \vdots \quad\quad \cdots \quad\quad \vdots \quad\quad &\quad \vdots \\ \left(a_{n2} - a_{n1}\frac{a_{12}}{a_{11}}\right)x_2 + \cdots + \left(a_{nn} - a_{n1}\frac{a_{nn}}{a_{11}}\right)x_n &= b_n - a_{n1}\frac{b_1}{a_{11}} \end{aligned}$$

or

$$\begin{aligned} x_1 + a'_{12}x_2 + \cdots + a'_{1n}x_n &= b'_1 && (1')\\ a'_{22}x_2 + \cdots + a'_{2n}x_n &= b'_2 && (2')\\ \vdots \quad \cdots \quad \vdots \quad &\quad \vdots \\ a'_{n2}x_2 + \cdots + a'_{nn}x_n &= b'_n && (n') \end{aligned}$$

Again, we can suppose that at least one of the coefficient a'_{i2}, $i = 2, \cdots, n$ is different from zero. Since we can rearrange these (n – 1) equations, it can be assumed that $a'_{22} \neq 0$ and is a pivot. Carrying out the division of (2') by a'_{22} and then eliminating x_2 from the remaining n – 2 equations, the new system takes the form

---

1. For if $a_{11} = 0$, this equation could be interchanged with one in which the first coefficient is nonzero. Of course, we assume that not all $a_{ij} = 0$, $i = 1, 2, \cdots, n$. For otherwise, the number of unknown reduces to $n - 1$.
The choice of pivot element is necessary in cases where the coefficients are very small. This is because division by small numbers may increase the coefficients leading to convergence and significant error problems. In our case we need only 1 in the (1, 1) position and then make all other entries in the first column as zero.

$$\begin{aligned}
x_1 + a'_{12}x_2 + a'_{13}x_3 + \cdots + a'_{1n}x_n &= b'_1 \\
x_2 + a''_{23}x_3 + \cdots + a''_{2n}x_n &= b''_2 \\
a''_{33}x_3 + \cdots + a''_{3n}x_n &= b''_3 \\
a''_{43}x_3 + \cdots + a''_{4n}x_n &= b''_4 \\
\vdots \quad \cdots \quad \vdots &\quad \vdots \\
a''_{n3}x_3 + \cdots + a''_{nn}x_n &= b''_n
\end{aligned}$$

Continuing this process and eliminating, step by step, the variables $x_3, x_4, \cdots, x_n$ from the remaining $(n-3), (n-4), \cdots, 2$ equations the final form of system of equations becomes

$$\left.\begin{aligned}
x_1 + p_{12}x_2 + p_{13}x_3 + \cdots + p_{1n}x_n &= q_1 \\
x_2 + p_{23}x_3 + \cdots + p_{2n}x_n &= q_2 \\
x_{n-1} + p_{(n-1)n}x_n &= q_{n-1} \\
x_n &= q_n
\end{aligned}\right\} \qquad \text{(II)}$$

A system of equations of the form (II) is said to be in **echelon form.** The system (II) is equivalent to the given system and has the same solution set. But in the system (II), x_n is explicitly known and putting its value in the last but one equation, we obtain x_{n-1}. The values of x_{n-1} and x_n are then substituted into the preceding equation giving x_{n-2} and continuing in this way, we find values all the unknowns.

The process of finding the values of the unknowns from the system of equations (II) described above, is called the **process of backward substitution** or solving the equations **recursively.**

Observe that the system of equations (II) can be written in the matrix form as:

$$\begin{bmatrix} 1 & p_{12} & p_{13} & \cdots & p_{1n} \\ 0 & 1 & p_{23} & \cdots & p_{2n} \\ 0 & 0 & 1 & \cdots & p_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_n \end{bmatrix}$$

i.e. $\qquad Px = q$

where P is the echelon (upper triangular) matrix obtained from the matrix A by elementary row operations. If E is the product of elementary matrices corresponding to these row operations then

$$P = EA \quad \text{and} \quad q = Eb.$$

The given system reduces to (II) if and only if the rank of A is n. If the rank of A is less than n, Guassian elimination method fails for one of the following two reasons:

(i) Either we are not able to determine the values of all the variables, or

(ii) the apparent inconsistency that all the p_{ij} in a certain row are zero while the corresponding q_i is not zero.

Example 1. Solve the system of equations:

$$\begin{aligned}
x_1 - x_2 + 2x_3 &= 0 && (1)\\
4x_1 + x_2 + 2x_3 &= 1 && (2)\\
x_1 + x_2 + x_3 &= -1 && (3)
\end{aligned}$$

Solution. (First Method). The coefficient of x_1 in the (1, 1) position is already 1. Following the procedure described above, multiplying equation (1) by 4 and subtracting the resulting equation from equation (2) and subtracting equation (1) from equation (3) respectively, we have

$$\begin{aligned}
x_1 - x_2 + 2x_3 &= 0 && (1)\\
5x_2 - 6x_3 &= 1 && (4)\\
2x_2 - x_3 &= -1 && (5)
\end{aligned}$$

Multiply equation (5) by 2 and subtract from equation (4). The set of equations becomes

$$\begin{aligned}
x_1 - x_2 + 2x_3 &= 0 && (1)\\
x_2 - 4x_3 &= 3 && (6)\\
2x_2 - x_3 &= -1 && (5)
\end{aligned}$$

Multiply equation (6) by 2 and subtract from equation (5). We have

$$\begin{aligned}
x_1 - x_2 + 2x_3 &= 0 && (1)\\
x_2 - 4x_3 &= 3 && (6)\\
7x_3 &= -7 && (7)
\end{aligned}$$

Now (7) gives

$$x_3 = -1.$$

By backward substitution,

$$x_2 = -1, \quad x_1 = 1.$$

Second Method. (By Gaussian Elimination Method).

Here

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 4 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}, \quad A_b = \left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 4 & 1 & 2 & 1 \\ 1 & 1 & 1 & -1 \end{array}\right]$$

So, we apply row operations on A_b to reduce it to the upper triangular form. Thus

$$A_b = \left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 4 & 1 & 2 & 1 \\ 1 & 1 & 1 & -1 \end{array}\right] \underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 0 & 5 & -6 & 1 \\ 0 & 2 & -1 & -1 \end{array}\right] \text{by } R_2 - 4R_1 \text{ and } R_3 - R_1$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 0 & 1 & -4 & 3 \\ 0 & 2 & -1 & -1 \end{array}\right] \quad \text{by } R_2 - 2R_3$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 0 & 1 & -4 & 3 \\ 0 & 0 & 7 & -7 \end{array}\right] \quad \text{by } R_3 - 2R_2$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 0 & 1 & -4 & 3 \\ 0 & 0 & 1 & -1 \end{array}\right] \quad \text{by } \frac{1}{7} R_3.$$

So $x_3 = -1$. By backward substitution, the equations

$x_2 - 4x_3 = 3$ and $x_1 - x_2 + 2x_3 = 0$ yield

$x_1 = 1, \quad x_2 = -1, \quad x_3 = -1.$

Example 2. Solve the following systems of equations by Gaussian elimination method.

$$\begin{aligned} x_1 + 5x_2 + 2x_3 &= 9 \\ x_1 + x_2 + 7x_3 &= 6 \\ -3x_2 + 4x_3 &= -2. \end{aligned}$$

Solution. Here we apply row operations on A_b indicated as under. Thus

$$A_b = \left[\begin{array}{ccc|c} 1 & 5 & 2 & 9 \\ 1 & 1 & 7 & 6 \\ 0 & -3 & 4 & -2 \end{array}\right] \underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 5 & 2 & 9 \\ 0 & -4 & 5 & -3 \\ 0 & -3 & 4 & -2 \end{array}\right] \quad \text{by } R_2 - R_1$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 5 & 2 & 9 \\ 0 & -1 & 1 & -1 \\ 0 & -3 & 4 & -2 \end{array}\right] \quad \text{by } R_2 - R_3$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 5 & 2 & 9 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right] \quad \text{by } R_3 - 3R_2.$$

So,
$$\begin{aligned} x_1 + 5x_2 + 2x_3 &= 9 \\ -x_2 + x_3 &= -1 \\ x_3 &= 1 \end{aligned}$$

By backward substitution, we obtain

$x_3 = 1, \quad x_2 = 2, \quad x_1 = -3.$

Example 3. A yarn merchant sells brands A, B, C of yarn each of which is a blend of Pakistani, Egyptian and American cotton in the ratios 1 : 2 : 1, 2 : 1 : 1 and 2 : 0 : 2 respectively. If one kilogram of A, B, C costs 40, 50 and 60 rupees respectively, find the cost of a kilogram of cotton of each country.

Solution. Let the cost of each kilogram of Pakistani, Egyptian and American cotton be x_1, x_2 and x_3 rupees respectively. Then the above mentioned problem is formulated into the following mathematical model involving equations:

$$\begin{aligned} \frac{1}{4} x_1 + \frac{2}{4} x_2 + \frac{1}{4} x_3 &= 40 \quad \text{(for brand } A) \\ \frac{2}{4} x_1 + \frac{1}{4} x_2 + \frac{1}{4} x_3 &= 50 \quad \text{(for brand } B) \\ \frac{2}{4} x_1 \qquad\quad + \frac{2}{4} x_3 &= 60 \quad \text{(for brand } C) \end{aligned}$$

which is equivalent to the system

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 160 \\ 2x_1 + x_2 + x_3 &= 200 \\ 2x_1 \qquad + 2x_3 &= 240. \end{aligned}$$

So

$$A_b = \left[\begin{array}{ccc|c} 1 & 2 & 1 & 160 \\ 2 & 1 & 1 & 200 \\ 2 & 0 & 2 & 240 \end{array}\right] \underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 160 \\ 0 & -3 & -1 & -120 \\ 0 & -4 & 0 & -80 \end{array}\right] \quad \text{by } R_2 - 2R_1 \text{ and } R_3 - 2R_1$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 160 \\ 0 & -4 & 0 & -80 \\ 0 & -3 & -1 & -120 \end{array}\right] \quad \text{by } R_{23}$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 160 \\ 0 & 1 & 0 & 20 \\ 0 & -3 & -1 & -120 \end{array}\right] \quad \text{by } -\frac{1}{4} R_2$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 160 \\ 0 & 1 & 0 & 20 \\ 0 & 0 & -1 & -60 \end{array}\right] \quad \text{by } R_3 + 3R_2.$$

Thus the given system is equivalent to

$$x_1 + 2x_2 + x_3 = 160$$
$$x_2 = 20$$
$$x_3 = 60$$

Hence, by backward substitution, we obtain,

$$x_1 = \text{Rs. } 60, \quad x_2 = \text{Rs. } 20, \quad x_3 = \text{Rs. } 60.$$

GAUSS-JORDAN ELIMINATION METHOD

(4.6) While reducing the given system of linear equations in (4.5) to the equivalent system (II), we eliminated x_k, $k = 1, 2, \cdots, n - k$ from the $(k + 1)$th, $(k + 2)$th $\cdots$ nth equations only. If we remove x_k from the first $(k - 1)$ equations also then x_k will appear only in the kth equation. With this modification we get another equivalent system from the given system expressing each x_k explicitly and we no more require any backward substitution of these variables for finding the values of other unknowns. This process is called the **Gauss-Jordan Elimination Method**. The following example explains the working of this method.

Example 4. Find the solution of the following system of linear equations by Gauss-Jordan elimination method:

$$2x_1 - x_2 - x_3 = 4$$
$$3x_1 + 4x_2 - 2x_3 = 11$$
$$3x_1 - 2x_2 + 4x_2 = 11.$$

Solution. We reduce the augmented matrix A_b into the reduced echelon form as follows:

$$A_b = \left[\begin{array}{ccc|c} 2 & -1 & -1 & 4 \\ 3 & 4 & -2 & 11 \\ 3 & -2 & 4 & 11 \end{array}\right] \underset{\sim}{R} \left[\begin{array}{ccc|c} 3 & 4 & -2 & 11 \\ 2 & -1 & -1 & 4 \\ 3 & -2 & 4 & 11 \end{array}\right] \quad \text{by } R_{12}$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 5 & -1 & 7 \\ 2 & -1 & -1 & 4 \\ 3 & -2 & 4 & 11 \end{array}\right] \quad \text{by } R_1 - R_2$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 5 & -1 & 7 \\ 0 & -11 & 1 & -10 \\ 0 & -17 & 7 & -10 \end{array}\right] \quad \text{by } R_2 - 2R_1 \text{ and } R_3 - 3R_1$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 5 & -1 & 7 \\ 0 & -11 & 1 & -10 \\ 0 & -6 & 6 & 0 \end{array}\right] \quad \text{by } R_3 - R_2$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 5 & -1 & 7 \\ 0 & -11 & 1 & -10 \\ 0 & -1 & 1 & 0 \end{array}\right] \quad \text{by } \frac{1}{6} R_3$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 5 & -1 & 7 \\ 0 & -10 & 0 & -10 \\ 0 & -1 & 1 & 0 \end{array}\right] \quad \text{by } R_2 - R_3$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 5 & -1 & 7 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{array}\right] \quad \text{by } -\frac{1}{10} R_2$$

$$\underset{\sim}{R}\left[\begin{array}{ccc|c} 1 & 5 & -1 & 7 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \quad \text{by } R_3 + R_2$$

$$\underset{\sim}{R}\left[\begin{array}{ccc|c} 1 & 5 & 0 & 8 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \quad \text{by } R_1 + R_3$$

$$\underset{\sim}{R}\left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \quad \text{by } R_1 - 5R_2.$$

So $x_1 = 3, \quad x_2 = 1, \quad x_3 = 1.$

Example 5. Solve the following system of linear equations by Gauss-Jordan elimination method:

$$\begin{aligned} 5x_1 \quad\quad + 4x_3 + 2x_4 &= 3 \\ x_1 - x_2 + 2x_3 + x_4 &= 1 \\ 4x_1 + x_2 + 2x_3 \quad\quad &= 1 \\ x_1 + x_2 + x_3 + x_4 &= 0 \end{aligned}$$

Solution. Here

$$A_b = \left[\begin{array}{cccc|c} 5 & 0 & 4 & 2 & 3 \\ 1 & -1 & 2 & 1 & 1 \\ 4 & 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{array}\right] \underset{\sim}{R} \left[\begin{array}{cccc|c} 1 & -1 & 2 & 2 & 2 \\ 1 & -1 & 2 & 1 & 1 \\ 4 & 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{array}\right] \quad \text{by } R_1 - R_3$$

$$\underset{\sim}{R}\left[\begin{array}{cccc|c} 1 & -1 & 2 & 2 & 2 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 5 & -6 & -8 & -7 \\ 0 & 2 & -1 & -1 & -2 \end{array}\right] \quad \text{by } R_2 - R_1;\ R_3 - 4R_1 \text{ and } R_4 - R_1$$

$$\underset{\sim}{R}\left[\begin{array}{cccc|c} 1 & -1 & 2 & 2 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & -4 & -6 & -3 \\ 0 & 2 & -1 & -1 & -2 \end{array}\right] \quad \text{by } (-1)R_2 \text{ and } R_3 - 2R_4$$

$$\underset{\sim}{R}\left[\begin{array}{cccc|c} 1 & -1 & 2 & 2 & 2 \\ 0 & 1 & -4 & -6 & -3 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 2 & -1 & -1 & -2 \end{array}\right] \quad \text{by } R_{23}$$

$$\underset{\sim}{R}\left[\begin{array}{cccc|c} 1 & -1 & 2 & 2 & 2 \\ 0 & 1 & -4 & -6 & -3 \\ 0 & 2 & -1 & -1 & -2 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] \quad \text{by } R_{34}$$

$$\underset{\sim}{R}\left[\begin{array}{cccc|c} 1 & 0 & -2 & -4 & -1 \\ 0 & 1 & -4 & -6 & -3 \\ 0 & 0 & 7 & 11 & 4 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] \quad \text{by } R_1 + R_2 \text{ and } R_3 - 2R_2$$

$$\underset{\sim}{R}\left[\begin{array}{cccc|c} 1 & 0 & -2 & -4 & -1 \\ 0 & 1 & -4 & -6 & -3 \\ 0 & 0 & 1 & 11/7 & 4/7 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] \quad \text{by } \frac{1}{7}R_3$$

$$\underset{\sim}{R}\left[\begin{array}{cccc|c} 1 & 0 & 0 & -6/7 & 1/7 \\ 0 & 1 & 0 & 2/7 & -5/7 \\ 0 & 0 & 1 & 11/7 & 4/7 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] \quad \text{by } R_1 + 2R_3 \text{ and } R_2 + 4R_3$$

$$\underset{\sim}{R}\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] \quad \text{by } R_1 + \left(\frac{6}{7}\right)R_4;\ R_2 - \frac{2}{7}R_4 \text{ and } R_3 - \left(\frac{11}{7}\right)R_4$$

Hence $x_1 = 1, \quad x_2 = -1, \quad x_3 = -1, \quad x_4 = 1.$

Example 6. The Housing Department of the Government plans to undertake four housing projects and lists material requirements for the houses in each of the projects as follows:

	Project 1	Project 2	Project 3	Project 4
Paint (in 100 gallons)	1	2	1	1.5
Wood (in 10,000 cu. ft.)	3	4	2.5	2.5
Bricks (in millions)	1	2	1.5	1
Labour (in 1000 hours)	10	10	9	8

If the supplier delivers 6,800 gallons of paint, 14,200,200 cubic feet of wood, 64 million bricks and 448,000 hours of labour, find the number of houses built for each project.

Solution. Let x_1, x_2, x_3 and x_4 be the number of houses built under Project 1, Project 2, Project 3 and Project 4, respectively. Then according to the conditions of the problem:

$$
\begin{aligned}
1 \times x_1 + 2 \times x_2 + 1 \times x_3 + 1.5 \times x_4 &= 68 \quad \text{(in 100 gallons)}\\
3 \times x_1 + 4 \times x_2 + 2.5 \times x_3 + 2.5 \times x_4 &= 142 \quad \text{(in 10,000 cu. ft)}\\
1 \times x_1 + 2 \times x_2 + 1.5 \times x_3 + 1 \times x_4 &= 64 \quad \text{(in millions)}\\
10 \times x_1 + 10 \times x_2 + 9 \times x_3 + 8 \times x_4 &= 448 \quad \text{(in 1000 hours)}
\end{aligned}
$$

The augmented matrix is

$$A_b = \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1.5 & 68 \\ 3 & 4 & 2.5 & 2.5 & 142 \\ 1 & 2 & 1.5 & 1 & 64 \\ 10 & 10 & 9 & 8 & 448 \end{array}\right]$$

$$\underset{\sim}{R} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1.5 & 68 \\ 0 & -2 & -0.5 & -2 & -62 \\ 0 & 0 & 0.5 & -0.5 & -4 \\ 0 & -10 & -1 & -7 & -232 \end{array}\right] \quad \text{by } R_2 - 3R_1;\ R_3 - R_1 \text{ and } R_4 - 10R_1$$

$$\underset{\sim}{R} \left[\begin{array}{cccc|c} 1 & 0 & 0.5 & -0.5 & 6 \\ 0 & -2 & -0.5 & -2 & -62 \\ 0 & 0 & 1 & -1 & -8 \\ 0 & 0 & 1.5 & 3 & 78 \end{array}\right] \quad \text{by } R_1 + R_2;\ 2R_3;\ R_4 + (-5)R_2$$

$$\underset{\sim}{R} \left[\begin{array}{cccc|c} 1 & 0 & 0.5 & -0.5 & 6 \\ 0 & 1 & 0.25 & 1 & 31 \\ 0 & 0 & 1 & -1 & -8 \\ 0 & 0 & 1.5 & 3 & 78 \end{array}\right] \quad \text{by } -\frac{1}{2}R_2$$

$$\underset{\sim}{R} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 10 \\ 0 & 1 & 0 & 1.25 & 33 \\ 0 & 0 & 1 & -1 & -8 \\ 0 & 0 & 0 & 4.5 & 90 \end{array}\right] \quad \text{by } R_1 - (0.5)R_3,\ R_2 + (-0.25)R_3 \text{ and } R_4 + (-1.5)R_3$$

$$\underset{\sim}{R} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 10 \\ 0 & 1 & 0 & 1.25 & 33 \\ 0 & 0 & 1 & -1 & -8 \\ 0 & 0 & 0 & 1 & 20 \end{array}\right] \quad \text{by } \frac{1}{4.5}R_4$$

$$\underset{\sim}{R} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 10 \\ 0 & 1 & 0 & 0 & 8 \\ 0 & 0 & 1 & 0 & 12 \\ 0 & 0 & 0 & 1 & 20 \end{array}\right] \quad \text{by } R_2 + (-1.25)R_4 \text{ and } R_3 + R_4$$

Hence $x_1 = 10$; $x_2 = 8$; $x_3 = 12$ and $x_4 = 20$ are the number of houses built for each project.

CONSISTENCY CRITERION

(ANOTHER METHOD OF SOLVING A CONSISTENT SYSTEM OF EQUATIONS)

(4.7) A general system

$$Ax = b \tag{I}$$

of m linear equations in n unknowns $x_1, x_2, \cdots, x_n$ may or may not have a solution. If the system $Ax = b$ has a solution $y = (y_1, y_2, \cdots, y_n)$ say, then it is called a **consistent system.** If the system has no solution, that is, if there does not exist a vector y which satisfies the equation $Ax = b$, then the system is termed as **inconsistent.** In the simple case when $m = n$, that is, the number of equations equals the number of unknowns and the matrix A of coefficients is nonsingular, the following theorem shows that the system (I) is consistent.

(4.8) Theorem. The system $Ax = b$, with $m = n$ and matrix A nonsingular, has a unique solution $x = A^{-1}b$.

Proof. Since A is nonsingular, its inverse A^{-1} exists. Multiplying both the sides of the equation $Ax = b$ by A^{-1}, we have

$$(A^{-1}A)x = A^{-1}b.$$

Since $A^{-1}A = I_n$, the above equation becomes $I_n x = x = A^{-1}b$, giving the solution vector $x = A^{-1}b$.

To show the uniqueness, let y be a second solution. Then $Ay = b$ and [illegible] $y = A^{-1}b = x$

Hence the theorem.

When the system of equations under consideration is homogeneous, that is, when $b = 0$, we have the following:

(4.9) Corollary. A system $Ax = 0$ of n homogeneous linear equations in n unknowns has a unique solution $x = 0$ if and only if A is a nonsingular matrix. $x = 0$ is a trivial solution.

Proof. See (6.43).

In the case when the number of equations is not equal to the number of unknowns, the system is either not consistent (so that there is no solution at all) or there may be an infinite number of solutions of the system. As to when a nonhomogeneous system of equations is consistent and when it is not, the following criterion provides [illegible] answer.

(4.10) Theorem. (Consistency Criterion). Let $Ax = b$ be a system of m linear equations in n unknowns. Then the equations have a solution if and only if the rank of A is equal to the rank of the augmented matrix A_b.

Proof. See (6.44)

Since the condition Rank A = Rank A_b is always satisfied in case of a system of homogeneous linear equations, such equations, therefore, necessarily have a solution. We have already seen in (4.9) that if Rank $A = n$, and $m = n$, then $x = 0$ is the unique solution of $Ax = 0$

(4.11) Theorem. A system of m homogeneous equation in n unknowns

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

has a nontrivial solution if $m < n$.

Proof. See (6.45).

It can be shown that, in a system of m linear equations in n unknowns, if Rank A = Rank $A_b = k < m$, then any x which satisfies k of the equations for which the corresponding rows of A in echelon form are nonzero satisfies all the equations of the set. Moreover, if $k < n$, then $n - k$ of the variables can be assigned arbitrary values and the remaining k variables can be found provided that the columns of A in echelon form associated with the k variables are nonzero. Naturally, therefore, in this case the solution set is infinite. We shall see in Chapter 6 that the solution set in this case is a vector space called the **solution space** of the system.

For a homogeneous system of linear equations, we have the following necessary and sufficient condition for a nontrivial solution.

(4.12) Theorem. A system of m homogeneous linear equations $Ax = 0$ in n variables has a nontrivial solution if and only if the rank of A is less than n.

Hence a system $Ax = 0$ of n equations in n variables has a nontrivial solution if and only if A is singular.

Proof. See (6.46).

Example 7. Examine the following homogeneous system for nontrivial solution:

$$\begin{aligned} x_1 - x_2 + 2x_3 + x_4 &= 0 \\ 3x_1 + 2x_2 \quad\quad + x_4 &= 0 \\ 4x_1 + x_2 + 2x_3 + 2x_4 &= 0 \end{aligned}$$

Solution. The matrix of coefficients is

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

We bring A into the echelon form by the indicated elementary row operations.

$$A \underset{\sim}{R} \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 5 & -6 & -2 \\ 0 & 5 & -6 & -2 \end{bmatrix} \quad \text{by } R_2 - 3R_1 \text{ and } R_3 - 4R_1$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 5 & -6 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{by } R_3 - R_2$$

$$\underset{\sim}{R}\begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -\frac{6}{5} & -\frac{2}{5} \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{by } \frac{1}{5}R_2$$

$$\underset{\sim}{R}\begin{bmatrix} 1 & 0 & \frac{4}{5} & \frac{3}{5} \\ 0 & 1 & -\frac{6}{5} & -\frac{2}{5} \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{by } R_1 + R_2$$

The rank of the matrix is $2 < 4$. Hence, by Theorem **4.12**, the equations have nontrivial solution.

The first two rows of the above matrix give the following relations:

$$x_1 + \frac{4}{5}x_3 + \frac{3}{5}x_4 = 0$$

$$x_2 - \frac{6}{5}x_3 - \frac{2}{5}x_4 = 0$$

i.e., $x_1 = -\frac{4}{5}x_3 - \frac{3}{5}x_4$

$x_2 = \frac{6}{5}x_3 + \frac{2}{5}x_4$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\frac{4}{5}x_3 - \frac{3}{5}x_4 \\ \frac{6}{5}x_3 + \frac{2}{5}x_4 \\ x_3 \\ x_4 \end{bmatrix}$$

Assigning arbitrary values to x_3, x_4, we find the corresponding values of x_1, x_2. In particular, when $x_3 = a$, $x_4 = b$,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\frac{4}{5}a - \frac{3}{5}b \\ \frac{6}{5}a + \frac{2}{5}b \\ a \\ b \end{bmatrix} = a\begin{bmatrix} -\frac{4}{5} \\ \frac{6}{5} \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} -\frac{3}{5} \\ \frac{2}{5} \\ 0 \\ 1 \end{bmatrix}$$

which give solutions for arbitrary values of a and b.

Example 8. For what value of λ the equations

$$\begin{aligned} (5-\lambda)x_1 + 4x_2 + 2x_3 &= 0 \\ 4x_1 + (5-\lambda)x_2 + 2x_3 &= 0 \\ 2x_1 + 2x_2 + (2-\lambda)x_3 &= 0 \end{aligned}$$

have nontrivial solutions. Find these solutions.

Solution. The matrix of coefficients is

$$A = \begin{bmatrix} 5-\lambda & 4 & 2 \\ 4 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{bmatrix}$$

We reduce this matrix to the echelon form by applying elementary row operations:

$$A \underset{\sim}{R} \begin{bmatrix} 2 & 2 & 2-\lambda \\ 4 & 5-\lambda & 2 \\ 5-\lambda & 4 & 2 \end{bmatrix} \quad \text{by } R_{13}$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & 1 & 1-\lambda/2 \\ 4 & 5-\lambda & 2 \\ 5-\lambda & 4 & 2 \end{bmatrix} \quad \text{by } \frac{1}{2}R_1$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & 1 & 1-\lambda/2 \\ 0 & 1-\lambda & -2+2\lambda \\ 0 & \lambda-1 & \frac{-6+7\lambda-\lambda^2}{2} \end{bmatrix} \quad \text{by } R_2 - 4R_1 \text{ and } R_3 - (5-\lambda)R_1$$

$$= \begin{bmatrix} 1 & 1 & \frac{2-\lambda}{2} \\ 0 & 1-\lambda & -2(1-\lambda) \\ 0 & -(1-\lambda) & \frac{-(1-\lambda)(6-\lambda)}{2} \end{bmatrix} \quad (1)$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & 1 & \frac{2-\lambda}{2} \\ 0 & 1 & -2 \\ 0 & -1 & \frac{(6-\lambda)}{2} \end{bmatrix} \quad \text{by } \frac{1}{1-\lambda}R_2 \text{ and } \frac{1}{1-\lambda}R_3$$

provided $1 - \lambda \neq 0$

$$\underset{\sim}{R}\begin{bmatrix} 1 & 0 & \frac{6-\lambda}{2} \\ 0 & 1 & -2 \\ 0 & 0 & \frac{(\lambda-10)}{2} \end{bmatrix} \quad \text{by } R_1 - R_2 \text{ and } R_3 + R_2 \qquad (2)$$

If $\lambda = 10$ in (2), we have

$$A \underset{\sim}{R} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{with} \quad 1 - \lambda \neq 0$$

The given system of equations reduces to

$$x_1 - 2x_3 = 0$$
$$x_2 - 2x_3 = 0.$$

So $\qquad x_1 = x_2 = 2x_3$

Take $x_3 = a$, $a \in R$. Then

$$x_1 = 2a, \quad x_2 = 2a, \quad x_3 = a$$

So $(x_1, x_2, x_3) = (2a, 2a, a) = a(2, 2, 1)$ gives all the solutions of the system for $a \in R$.

If, however $\lambda \neq 10$ with $1 - \lambda = 0$, then (1) gives

$$A \underset{\sim}{R} \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The given system reduces to the single equation

$$x_1 + x_2 + \frac{1}{2}x_3 = 0$$

Take $x_3 = 2a$, $x_2 = b$. Then $x_1 = -x_2 - \frac{1}{2}x_3 = -b - a$.

Hence the solution vector is

$$(x_1, x_2, x_3) = (-b-a, b, 2a)$$
$$= b(-1, 1, 0) + a(-1, 0, 2), \quad a, b \in \boldsymbol{R}.$$

For $\lambda \neq 1$ and $\lambda \neq 10$ the given system has only the trivial solution (0, 0, 0) because, from (2) above, we see that A is invertible.

Example 9. Three species of bacteria co-exist in a test tube and feed on three foods F_1, F_2, F_3. Suppose that a bacterium of the ith species consumes, on the average, an amount a_{ij} of the jth food per day where

$$a_{11} = 1, \quad a_{12} = 1, \quad a_{13} = 1, \quad a_{21} = 1,$$
$$a_{22} = 2, \quad a_{23} = 3, \quad a_{31} = 1, \quad a_{32} = 3, \quad a_{33} = 5.$$

Further suppose that there are 15000, 30000, 45000 units of food of types F_1, F_2, F_3 respectively. Assuming that all food is consumed, what are the populations of the three species that can co-exist in this environment?

Solution. Let the population of the three species of bacteria be x_1, x_2, x_3. Since the ith species consumes a_{ij} units of food of type j, we have

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 15000 \\ 30000 \\ 45000 \end{bmatrix}$$

That is,
$$x_1 + x_2 + x_3 = 15{,}000$$
$$x_1 + 2x_2 + 3x_3 = 30{,}000$$
$$x_1 + 3x_2 + 5x_3 = 45{,}000$$

Matrix of coefficients is $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{bmatrix}$

The augmented matrix is

$$A_b = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 15000 \\ 1 & 2 & 3 & 30000 \\ 1 & 3 & 5 & 45000 \end{array}\right]$$

$$\underset{\sim}{R}\left[\begin{array}{ccc|c} 1 & 1 & 1 & 15000 \\ 0 & 1 & 2 & 15000 \\ 0 & 2 & 4 & 30000 \end{array}\right] \quad \text{by } R_2 - R_1 \text{ and } R_3 - R_1$$

$$\underset{\sim}{R}\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 15000 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \text{by } R_1 - R_2 \text{ and } R_3 - 2R_2$$

Here we note that Rank A = Rank A_b.

Thus we may choose $x_3 = a$ arbitrarily, then

$$x_1 = x_3 = a, \qquad x_2 = 15{,}000 - 2x_3 = 15{,}000 - 2a.$$

Since x_1, x_2, x_3 are positive,

$$15{,}000 - 2x_3 \geq 0, \quad \text{that is,} \quad 0 < x_3 \leq 7{,}500.$$

From the first equation we see that the total populations that can co-exist is 15,000.

If we take $x_3 = a = 3{,}000$, then $x_1 = 3{,}000$ and $x_2 = 9{,}000$.

NETWORK FLOW PROBLEMS

Many types of real world problems involve networks along which some sort of flow is observed. A network of roads and a network of irrigation system are examples of such problems. There are points in the network at which a net flow either enters or leaves the system. Headworks of canals and junctions of roads are the points where the flow of water and volume of traffic are studied and analyzed. **The fundamental principle in the analysis of these problems is that the flow is conserved – that is the total flow entering the system at a junction must be equal to the total flow leaving the junction.** Mathematical model of the problem consists of linear equations which can be solved by the methods already discussed.

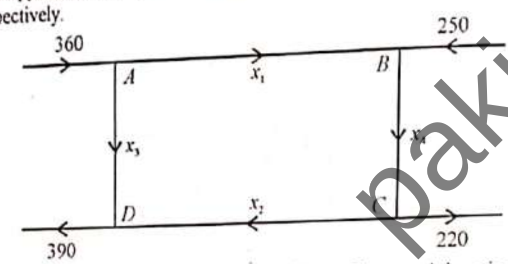
Example 10. A part of a city's road network for vehicular traffic is as shown by arrows in the following diagram:

(i) Write the equations indicating the traffic flow given in the diagram

(ii) Show that the traffic flow along AB, CD can be expressed in terms of the traffic flow along AD.

(iii) If the stretch AD or the stretch CD is closed, then show that the solution to the problem is unique.

Solution. Suppose that x_1, x_2 and x_3 denote the number of vehicles going along AB, CD and AD respectively.



Equating the incoming traffic to the outgoing traffic at each junction, we have the following mathematical model. Here we equate the incoming and outgoing vehicles. Thus:

At A: $x_1 + x_3 = 360$ (360 vehicles going along x_1 and x_3 directions)

At D: $x_2 + x_3 = 390$ (As in the case above)

At B: $x_1 + 250 = x_4$

At C: $x_2 + 220 = x_4$

i.e., $x_1 - x_2 = -30$. from equations at B and C.

Here, $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$.

The augmented matrix is

$$A_b = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 360 \\ 0 & 1 & 1 & 390 \\ 1 & -1 & 0 & -30 \end{array}\right] \underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 360 \\ 0 & 1 & 1 & 390 \\ 0 & -1 & -1 & -390 \end{array}\right] \text{ by } R_3 - R_1$$

$$\underset{\sim}{R} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 360 \\ 0 & 1 & 1 & 390 \\ 0 & 0 & 0 & 0 \end{array}\right] \text{ by } R_3 + R_2$$

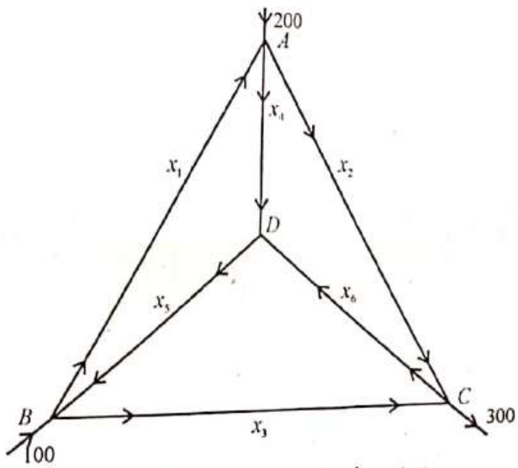
So $x_1 + x_3 = 360$ and $x_2 + x_3 = 390$

Take $x_3 = a$. Then $x_1 = 360 - a$, $x_2 = 390 - a$, $x_3 = a$ where a is an arbitrary integer such that $0 \leq a \leq 360$.

If the stretch AD is closed, then $x_3 = 0$ and $x_1 = 360$, $x_2 = 390$

Example 11. The flow through a network is as shown in the diagram.

The unknowns x_1, x_2, x_3, x_4, x_5, x_6 indicate the number of outgoing or incoming vehicles at the given junctions.



Using the fundamental principle of network flow, solve the system.

If $x_3 = 100$, $x_5 = 50$, $x_6 = 50$, find the flow.

Solution. We consider the flow of traffic at various points of the network. For this, let us consider the point A. The variables x_1, x_2, x_4 indicate the number of incoming and outgoing vehicles shown by the arrows. One can see that there are $200 + x_1$ vehicles entering at A. The number of outgoing vehicles at A is $x_2 + x_4$. By the fundamental principle of network flow, we have the following relations.

At A: $x_1 + 200 = x_2 + x_4$

Similarly, equations at other points are as follows:

At B: $x_1 + x_3 = x_5 + 100$

At C: $x_2 + x_3 = x_6 + 300$

At D: $x_4 + x_6 = x_5$

We can rewrite these equations as:

$$\begin{aligned}
x_1 - x_2 + 0x_3 - x_4 + 0x_5 + 0x_6 &= -200\\
x_1 + 0x_2 + x_3 + 0x_4 - x_5 + 0x_6 &= 100\\
0x_1 + x_2 + x_3 + 0x_4 + 0x_5 - x_6 &= 300\\
0x_1 + 0x_2 + 0x_3 + x_4 - x_5 + x_6 &= 0
\end{aligned}$$

So the augmented matrix is:

$$A_b = \left[\begin{array}{cccccc|c} 1 & -1 & 0 & -1 & 0 & 0 & -200\\ 1 & 0 & 1 & 0 & -1 & 0 & 100\\ 0 & 1 & 1 & 0 & 0 & -1 & 300\\ 0 & 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right]$$

The row operations on A_b reduces the matrix to echelon form as follows:

$$A_b \underset{\sim}{R} \left[\begin{array}{cccccc|c} 1 & -1 & 0 & -1 & 0 & 0 & -200\\ 0 & 1 & 1 & 1 & -1 & 0 & 300\\ 0 & 1 & 1 & 0 & 0 & -1 & 300\\ 0 & 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right] \text{ by } R_2 - R_1$$

$$\underset{\sim}{R} \left[\begin{array}{cccccc|c} 1 & -1 & 0 & -1 & 0 & 0 & -200\\ 0 & 1 & 1 & 1 & -1 & 0 & 300\\ 0 & 0 & 0 & -1 & 1 & -1 & 0\\ 0 & 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right] \text{ by } R_3 - R_2$$

$$\underset{\sim}{R} \left[\begin{array}{cccccc|c} 1 & 0 & 1 & 0 & -1 & 0 & 100\\ 0 & 1 & 1 & 1 & -1 & 0 & 300\\ 0 & 0 & 0 & -1 & 1 & -1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l}\text{by } R_1 + R_2\\ \text{and } R_4 + R_3\end{array}$$

$$\underset{\sim}{R} \left[\begin{array}{cccccc|c} 1 & 0 & 1 & 0 & -1 & 0 & 100\\ 0 & 1 & 1 & 0 & 0 & -1 & 300\\ 0 & 0 & 0 & -1 & 1 & -1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \text{ by } R_2 - R_3$$

So the network flow system becomes

$$\begin{aligned}
x_1 + x_3 - x_5 &= 100 && (1)\\
x_2 + x_3 - x_6 &= 300 && (2)\\
x_4 - x_5 + x_6 &= 0 && (3)
\end{aligned}$$

We can give arbitrary values to x_3, x_5, x_6 as:

$x_3 = r, \quad x_5 = s, \quad x_6 = t$

Then, from (3), we have

$x_4 = x_5 - x_6 = s - t$

So (1) becomes

$$\begin{aligned}
x_1 &= 100 - x_3 + x_5\\
&= 100 - r + s\\
x_2 &= 300 - x_3 + x_6\\
&= 300 - r + t\\
x_3 &= r\\
x_4 &= s - t\\
x_5 &= s\\
x_6 &= t
\end{aligned}$$

The values of x_i, $i = 1, 2, \cdots, 6$ given above provide the most general solution of the network system.

Obviously each arbitrary value of r, s and t gives a solution of the system.

In particular, if $x_3 = 100$, $x_5 = 50$, $x_6 = 50$, then

$$\begin{aligned}
x_1 &= 100 - x_3 + x_5 = 100 - 100 + 50 = 50\\
x_2 &= 300 - x_3 + x_6 = 300 - 100 + 50 = 250\\
x_4 &= x_5 - x_6 = 50 - 50 = 0
\end{aligned}$$

The particular solution is

$$\begin{aligned}
x_1 &= 50\\
x_2 &= 250\\
x_3 &= 100\\
x_4 &= 0\\
x_5 &= 50\\
x_6 &= 50
\end{aligned}$$

EXERCISE 4

Solve the following systems of linear equations, the field of scalars being R:

1.
$$\begin{aligned} 2x_1 + x_3 &= 1 \\ 2x_1 + 4x_2 - x_3 &= -2 \\ x_1 - 8x_2 - 3x_3 &= 2 \end{aligned}$$

2.
$$\begin{aligned} x_1 + x_2 + x_3 &= a \\ x_1 + (1+a)\,x_2 + x_3 &= 2a \\ x_1 + x_2 + (1+a)\,x_3 &= 3a \end{aligned}$$

3.
$$\begin{aligned} x_1 - x_2 + x_3 - x_4 + x_5 &= 1 \\ 2x_1 + x_2 + 3x_3 + 4x_5 &= 3 \\ 3x_1 - 2x_2 + 2x_3 + x_4 + x_5 &= 1 \\ x_2 + x_4 + x_5 &= 0 \end{aligned}$$

4.
$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ x_2 + x_3 - x_4 &= 1 \\ x_3 + x_4 - x_5 &= 1 \\ -x_3 + x_4 + x_5 &= 1 \\ -x_2 + x_3 + x_4 &= 1 \end{aligned}$$

5.
$$\begin{aligned} x_1 - 2x_2 - 7x_3 + 7x_4 &= 5 \\ -x_1 + 2x_2 + 8x_3 - 5x_4 &= -7 \\ 3x_1 - 4x_2 - 17x_3 + 13x_4 &= 14 \\ 2x_1 - 2x_2 - 11x_3 + 8x_4 &= 7 \end{aligned}$$

6.
$$\begin{aligned} x_1 + 2x_2 + x_3 &= -1 \\ 6x_1 + x_2 + x_3 &= -4 \\ 2x_1 - 3x_2 - x_3 &= 0 \\ x_1 - x_2 &= 1 \end{aligned}$$

7.
$$\begin{aligned} 2x_1 + x_2 + 5x_3 &= 4 \\ 3x_1 - 2x_2 + 2x_3 &= 2 \\ 5x_1 - 8x_2 - 4x_3 &= 1. \end{aligned}$$

8. **Solve the system of equations having the given matrices as their augmented matrices:**

(i) $\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \end{array}\right]$ (ii) $\left[\begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 3 & 0 \end{array}\right]$

(iii) $\left[\begin{array}{ccc|c} 4 & 2 & -1 & 0 \\ 3 & 3 & 6 & 3 \\ 5 & 1 & -8 & -1 \end{array}\right]$ (iv) $\left[\begin{array}{ccc|c} 2 & -1 & 3 & 4 \\ 0 & 4 & 1 & 8 \\ 0 & 0 & 2 & 0 \end{array}\right]$

For what values of λ do the following homogeneous equations have nontrivial solutions? Find these solutions: (Problems 9 – 11):

9.
$$\begin{aligned} (1-\lambda)\,x_1 + x_2 &= 0 \\ x_1 + (1-\lambda)\,x_2 &= 0 \end{aligned}$$

10.
$$\begin{aligned} (3-\lambda)\,x_1 - x_2 + x_3 &= 0 \\ x_1 - (1-\lambda)\,x_2 + x_3 &= 0 \\ x_1 - x_2 + (1-\lambda)\,x_3 &= 0 \end{aligned}$$

11.
$$\begin{aligned} (1-\lambda)\,x_1 + x_2 - x_3 &= 0 \\ x_1 - \lambda x_2 - 2x_3 &= 0 \\ x_1 + 2x_2 - \lambda x_3 &= 0 \end{aligned}$$

In each of the following use Gauss-Jordan method to reduce the given system to reduced echelon form, indicating the operations performed and determine the solution if any: (Problems 12 – 19):

12.
$$\begin{aligned} 6x_1 - 6x_2 + 6x_3 &= 6 \\ 2x_1 - 4x_2 - 6x &= 12 \\ 10x_1 - 5x_2 + 5x_3 &= 30 \end{aligned}$$

13.
$$\begin{aligned} 5x_1 + 5x_2 - x_3 &= 0 \\ 10x_1 + 5x_2 + 2x_3 &= 0 \\ 5x_1 + 15x_2 - 9x_3 &= 0 \end{aligned}$$

14.
$$\begin{aligned} 5x_1 - 2x_2 + x_3 &= 3 \\ 3x_1 + 2x_2 + 7x_3 &= 5 \\ x_1 + x_2 + 3x_3 &= 2 \end{aligned}$$

15.
$$\begin{aligned} 5x_1 - 2x_2 + x_3 &= 2 \\ 3x_1 + 2x_2 + 7x_3 &= 3 \\ x_1 + x_2 + 3x_3 &= 2 \end{aligned}$$

16.
$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 3 \\ 3x_1 + x_2 - 5x_3 &= 0 \\ 4x_1 - x_2 + x_3 &= 3 \end{aligned}$$

17.
$$\begin{aligned} x_1 + 3x_2 + 5x_3 - 4x_4 &= 1 \\ x_1 + 2x_2 + x_3 - x_4 + x_5 &= -1 \\ x_1 - 2x_2 + 3x_3 + 2x_4 - x_5 &= 3 \\ x_1 + 5x_2 + 3x_3 + x_4 + x_5 &= -11 \\ x_1 + 3x_2 - x_3 + x_4 + 2x_5 &= -3 \end{aligned}$$

18.
$$\begin{aligned} 3x_1 + 2x_2 + 4x_3 &= 7 \\ 2x_1 + x_2 + x_3 &= 4 \\ x_1 + 3x_2 + 5x_3 &= 3 \end{aligned}$$

19.
$$\begin{aligned} 5x_1 + 4x_3 + 2x_4 &= 3 \\ x_1 - x_2 + 2x_3 + x_4 &= 1 \\ 4x_1 + x_2 + 2x_3 &= 1 \\ x_1 + x_2 + x_3 + x_4 &= 0 \end{aligned}$$

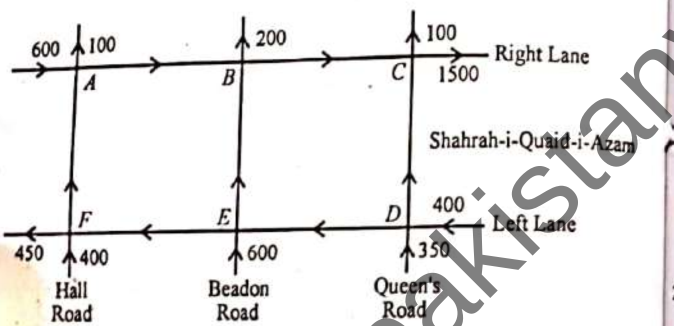
20. Show that the system
$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= a \\ 3x_1 + x_2 - 5x_3 &= b \\ -5x_1 - 5x_2 + 21x_3 &= c \end{aligned}$$
is inconsistent if $c \neq 2a - 3b$.

21. Let

$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}$. Find a matrix B such that $AB = I_2$. Does this mean A is invertible? Explain.

22. A soap manufacturer decides to spend 600,000 rupees on radio, magazine and TV advertising. If he spends as much on TV advertising as on magazines and radio together, and the amount spent on magazines and TV combined equals five times that spent on radio, what is the amount to be spent on each type of advertising?

23. Traffic counters submitted the following information for March 23 from 7 p.m. to 8 p.m. on the following roads of the city.

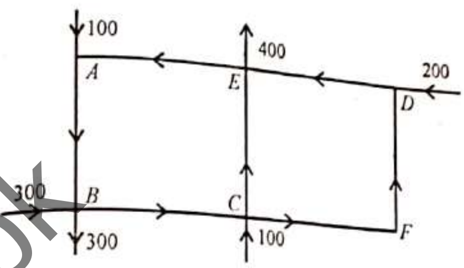


(i) Construct a mathematical model that describes this system, carefully labeling the variables you introduce.

(ii) Show that there must be at least 50 vehicles travelling to Hall Road from Beadon Road on the section of left lane during the count.

(iii) The city planners are inclined to take this traffic count as typical rush hour evening traffic in this area. In their planning of the annual closure of left lane between Queen's Road and Beadon Road for repair, how much traffic can be expected on right lane between Queen's Road and Beadon Road?

24. One part of Lahore's network of traffic is given with the number of vehicles that enter and leave during a typical rush hour as shown below. All the lanes are one-way in the direction indicated by the arrows.

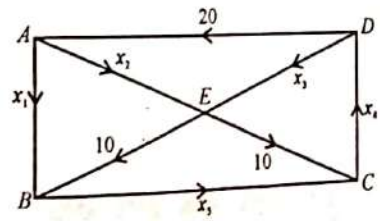


(i) Construct the linear mathematical model that describes this system.

(ii) If the stretch EA is closed for repair, what will be the traffic flow along the other stretches?

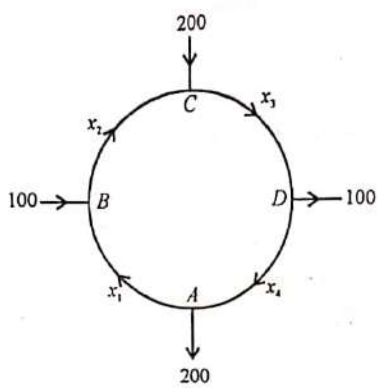
(iii) If only 100 vehicles are allowed to pass during the rush hour through EA, how will that affect the other branches?

25. Set up a system of linear equations to represent the network shown in the diagram and solve the system.



If $x_1 = x_3 = 0$, find the flow.

26. The flow of traffic at the Kalama Chowk on Ferozepur Road, Lahore is shown below:



(i) Solve the system.

(ii) Find the traffic flow when $x_4 = 300$.

Chapter 5

DETERMINANTS

The reader may already be familiar with the concept of determinants. This concept played a basic role in the solution of systems of linear equations before the use of matrices and computers. The use of determinants is becoming less and less with every passing day. Neverthless determinants still play an important but minor role in finding solutions of linear problems.

This chapter includes concept of determinant of a square matrix, minors and cofactors, properties of determinants, Laplace expansion of a determinant, adjoint and inverse of a matrix.

DETERMINANT OF A SQUARE MATRIX

(5.1) Definition. In Chapter 4, methods for solving systems of homogeneous and nonhomogeneous linear equations were given. We shall make use of those methods to give formal definition of a **determinant.** Before taking up the general case, we consider the following special case of two linear equations:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

in two unknowns x_1, x_2. Here

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad A_b = \left[\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array}\right] \qquad (1)$$

are the **coefficient matrix** and **the augmented matrix** respectively with entries as real numbers. By the Gauss-Jordan procedure, the solution of these equations is

$$x_1 = \frac{(b_1a_{22} - b_2a_{12})}{(a_{11}a_{22} - a_{21}a_{12})}, \quad x_2 = \frac{(b_2a_{11} - b_1a_{21})}{(a_{11}a_{22} - a_{21}a_{12})}$$

provided that

$$a_{11}a_{22} - a_{21}a_{12} \neq 0.$$

The scalar $a_{11}a_{22} - a_{21}a_{12}$ is uniquely determined by the matrix A. It is called the **determinant of order 2** of the square matrix A given in (1) and is denoted by det A or $|A|$. Thus

$$|A| = \det A = a_{11}a_{22} - a_{21}a_{12} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \qquad (2)$$

Note the two vertical bars instead of square brackets used for matrices.

We note that det A of order 2 is a real number associated with a matrix A of order 2. So, in this case, we may regard det as a function whose domain is the set of all square matrices of order 2 and whose range is a subset of real numbers.

The following properties of determinants of order 2 follow directly from the above definition.

(i) For any real number k,

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

(ii) If $a_{12} = b_{12} + c_{12}$, $a_{22} = b_{22} + c_{22}$, then

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

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

(iii) If the two columns of A are identical then det $A = 0$. That is

$$\begin{vmatrix} a_{11} & a_{11} \\ a_{21} & a_{21} \end{vmatrix} = 0.$$

(iv) The determinant of the unit matrix is 1. That is

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

It will be seen later that these properties are characteristics of a determinant of a square matrix of any order.

(5.2) Definition. (Determinant of Order n). Firstly we define the determinant of an $n \times n$ matrix inductively. That is, from our knowledge of a determinant of order 2, we define a determinant of order 3 and use this definition of a determinant of order 3 to describe a determinant of order 4 and so on.

Thus a determinant of order 3 is defined as follows:

Let

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

Then

$$\det A = |A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \qquad (1)$$

Note the minus sign before the second term on the right hand side of (1).

For example, if

$$A = \begin{bmatrix} 4 & 1 & 2 \\ 3 & 2 & 5 \\ 1 & 2 & 3 \end{bmatrix}, \text{ then}$$

$$\det A = \begin{vmatrix} 4 & 1 & 2 \\ 3 & 2 & 5 \\ 1 & 2 & 3 \end{vmatrix} = 4\begin{vmatrix} 2 & 5 \\ 2 & 3 \end{vmatrix} - 1\begin{vmatrix} 3 & 5 \\ 1 & 3 \end{vmatrix} + 2\begin{vmatrix} 3 & 2 \\ 1 & 2 \end{vmatrix}.$$

$$= 4(6 - 10) - 1(9 - 5) + 2(6 - 2)$$

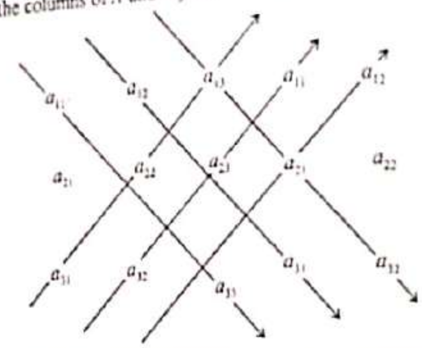
$$= -16 - 4 + 8 = -12.$$

There is a simplex method to calculate a determinant of order 3. For example, in (1) of (5.2),

$$\det A = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

We write the columns of A and adjoin to these the first two columns as below



Now calculate the six products of numbers on the directed lines taking plus sign with those products on arrows pointing downwards and minus sign with products on arrows pointing upwards. Adding these products we get the value of det A. This is known as **Sarrus's rule.**

Note: The method of arrows given above works only for $n = 2, 3$, and not for $n \geq 4$.

(5.3) Definition. Let A be a square matrix of order n. The matrix obtained from A by deleting its ith row and jth column is again a matrix M_{ij} of order $n-1$. M_{ij} is called the **ijth minor of A.**[1]

(5.4) Definition. Let M_{ij} be the ijth minor of a square matrix A of order n. Then

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

is called the ijth **cofactor** of A.

Observe that the sign on the right hand side of the above equality is positive or negative according as $i + j$ is even or odd.

Example 1. Let

$$A = \begin{bmatrix} 2 & -1 & 3 & 4 \\ 3 & 2 & 5 & -2 \\ 3 & 4 & 2 & 1 \\ -3 & 2 & 5 & 1 \end{bmatrix}$$

Then

$$M_{23} = \begin{bmatrix} 2 & -1 & 4 \\ 3 & 4 & 1 \\ -3 & 2 & 1 \end{bmatrix}, \quad M_{42} = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 5 & -2 \\ 3 & 2 & 1 \end{bmatrix}$$

1. Some authors call det M_{ij} of the matrix M_{ij} as the minor of the (i, j) th element.

and $A_{23} = (-1)^{2+3} \det M_{23} = -82$

$A_{42} = (-1)^{4+2} \det M_{42} = -45$

Consider the expansion of det A as given in (1) of (5.2). Using the fact that

$$A_{11} = (-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad A_{12} = (-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

$$A_{13} = (-1)^{1+3} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

we may write (1) of (5.2) as

$$\det A = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$$

We now define det A of a matrix A of order n as follows.

(5.5) Definition. For the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

of order n, we define det A by

$$\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \qquad (1)$$

The expression on the right hand side of (1) is an **expansion** of det A by cofactors of the ith row of det A. This is called **Laplace's Expansion** of a determinant of order n. The definition (1) given above indicates that the determinant of a matrix can be evaluated by any row of the matrix. We only have to be careful about the sign of the cofactor of the corresponding element.

Example 2. Let

$$A = \begin{bmatrix} 3 & 2 & 1 & -1 \\ 4 & 5 & 1 & 2 \\ -2 & 3 & 0 & 1 \\ 2 & 1 & 3 & 5 \end{bmatrix}. \text{ Then}$$

$$\det A = \begin{vmatrix} 3 & 2 & 1 & -1 \\ 4 & 5 & 1 & 2 \\ -2 & 3 & 0 & 1 \\ 2 & 1 & 3 & 5 \end{vmatrix}$$

$$= 3\begin{vmatrix} 5 & 1 & 2 \\ 3 & 0 & 1 \\ 1 & 3 & 5 \end{vmatrix} - 2\begin{vmatrix} 4 & 1 & 2 \\ -2 & 0 & 1 \\ 2 & 3 & 5 \end{vmatrix} + 1\begin{vmatrix} 4 & 5 & 2 \\ -2 & 3 & 1 \\ 2 & 1 & 5 \end{vmatrix}$$

$$+1\begin{vmatrix} 4 & 5 & 1 \\ -2 & 3 & 0 \\ 2 & 1 & 3 \end{vmatrix}$$

$$= 3[5(-3) - 1(14) + 2(9)] - 2[4(-3) - 1(-12) + 2(-6)]$$

$$+ 1[4(14) - 5(-12) + 2(-8)] + 1[4(9) - 5(-6) + 1(-8)]$$

$$= 3(-15 - 14 + 18) - 2(-12 + 12 - 12) + 1(56 + 60 - 16)$$

$$+ 1(36 + 30 - 8)$$

$$= -33 + 24 + 100 + 58 = 149.$$

(5.6) Remark. Although the above technique to evaluate determinant of an $n \times$ matrix seems quite straightforward, yet, in practice, it is very laborious to work w when $n > 3$ because it involves a lot of calculations. We shall later discuss methods wh simplify these calculations.

Note that, for some special type of matrices, their determinants can be eas evaluated. For instance, if a matrix A is **triangular** (upper or lower), that is, all its entr below or above the main diagonal are zero respectively, then its determinant is just t product of the elements on the main diagonal.

Example 3. Let

$$A = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

Then $\det A = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} + a_{14}A_{14}$

$$= a_{11}A_{11} + 0A_{12} + 0A_{13} + 0A_{14} = a_{11}A_{11}$$

$$= a_{11}\begin{vmatrix} a_{22} & 0 & 0 \\ a_{32} & a_{33} & 0 \\ a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{11}a_{22}\begin{vmatrix} a_{33} & 0 \\ a_{43} & a_{44} \end{vmatrix} = a_{11}a_{22}a_{33}a_{44}.$$

By induction on n, it is easy to generalize the above result to a triangular matrix of any order.

Similarly for an upper triangular matrix. Thus:

(5.7) Theorem. Let $A = [\,a_{ij}\,]$ be an $n \times n$ triangular matrix. Then $\det A$ is the product of the elements on the main diagonal; that is,

$$\det A = a_{11}\, a_{22}\, a_{33} \cdots a_{nn}.$$

Note: In general, to evaluate the determinant of a matrix A one should start with the row (or column) of A with the maximum number of zero entries.

AXIOMATIC DEFINITION OF A DETERMINANT

(5.8) Notation. Let M_n be the set of all $n \times n$ matrices with entries from F, where F is the field of real or complex numbers. An element A of M_n is an $n \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

If we denote the jth column

$$\begin{matrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{matrix}$$

of A by a_j, then we can write A as

$$A = [\,a_1 \quad a_2 \cdots a_j \cdots a_n\,]$$

The matrix

$$A^* = [\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ c\boldsymbol{a}_j \cdots \ \boldsymbol{a}_n\,]$$

denotes the matrix obtained from A by multiplying every element of its jth column by a scalar c.

Similarly, the matrix

$$A^{**} = [\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{a}_i + c\boldsymbol{a}_j \quad \cdots \ \boldsymbol{a}_n\,]$$

represents a matrix obtained from A by replacing ith column $\boldsymbol{a}_i$ of A by $\boldsymbol{a}_i + c\boldsymbol{a}_j$. Thus

$$A^{**} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1i} + ca_{1j} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i} + ca_{2j} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{ni} + ca_{nj} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}$$

We now give an axomatic definition of determinant of a square matrix of order n as follows:

(5.9) Definition. A function det: $M_n \longrightarrow F$ which associates with a matrix A in M_n a unique element det A of F, is said to be the **determinant function** if it satisfies the following conditions:

(i) det is a linear function of columns of A. That is, if

$$\boldsymbol{a}_j = \beta\boldsymbol{b}_j + \gamma\boldsymbol{c}_j, \text{ for } \beta, \gamma \in F, \quad 1 \le j \le n, \text{ then}$$

$$\begin{aligned} \det A &= \det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{a}_j \ \cdots \ \boldsymbol{a}_n\,] \\ &= \det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \beta\boldsymbol{b}_j + \gamma\boldsymbol{c}_j \ \cdots \ \boldsymbol{a}_n\,] \\ &= \beta\det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{b}_j \ \cdots \ \boldsymbol{a}_n\,] \\ &\quad + \gamma\det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{c}_j \ \cdots \ \boldsymbol{a}_n\,] \end{aligned}$$

(ii) If any two adjacent columns of A are identical then det $A = 0$.

(iii) det $\boldsymbol{I}_n = 1$. Here $\boldsymbol{I}_n$ is the $n \times n$ identity matrix and 1 is the multiplicative identity of F.

(5.10) Definition. The element det A of F associated with a matrix $A \in M_n$ is called the **determinant** of A.

Note that while a matrix A is a symbol, its determinant det A is a real or complex number according as the entries of A are from R or C.

Henceforth, for the sake of simplicity and convenience, we shall make no distinction between the properties of the determinant function and those of a determinant.

In the following theorem we give a few properties of the determinant function.

(5.11) Theorem. Let A be a square matrix of order n. Then the determinant function has the following properties:

(i) If the column vector $\boldsymbol{a}_i$ of A is of the form

$$\boldsymbol{a}_i = \alpha\boldsymbol{b}_i$$

then

$$\det A = \alpha\det B, \ (\alpha \text{ being a scalar}),$$

where B is the matrix obtained from A by replacing its ith column $\boldsymbol{a}_i$ by $\boldsymbol{b}_i$.

(ii) If a column $\boldsymbol{a}_i$ of A is $\boldsymbol{0}$, then $\det A = 0$

(iii) If a column $\boldsymbol{a}_i$ of A is of the form $\boldsymbol{a}_i = \boldsymbol{b}_i + \boldsymbol{c}_i$, then

$$\det A = \det B + \det C$$

where B and C are matrices obtained from A by replacing its ith column by $\boldsymbol{b}_i$ and $\boldsymbol{c}_i$ respectively.

(iv) If $\boldsymbol{a}_i$ is replaced by $\boldsymbol{a}_i + c\,\boldsymbol{a}_{i+1}$ in A, then det A remains unchanged. That is

$$= \det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{a}_i + c\,\boldsymbol{a}_{i+1} \quad \boldsymbol{a}_{i+1} \ \cdots \ \boldsymbol{a}_n\,], \text{ for any scalar } c.$$

$$= \det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{a}_i \quad \boldsymbol{a}_{i+1} \ \cdots \ \boldsymbol{a}_n\,]$$

(v) Interchange of any two adjacent columns of A changes the sign of det A.

(vi) Interchange of any two column of A results in the change of sign of det A.

(vii) If any two columns of A are identical then det $A = 0$.

(viii) A constant multiple of a column of A may be added to any other column without changing the value of det A, that is

$$\det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{a}_i + c\,\boldsymbol{a}_j \ \cdots \ \boldsymbol{a}_n\,], \text{ for any scalar } c.$$

$$= \det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{a}_i \ \cdots \ \boldsymbol{a}_n\,]$$

Proof. (i) Since by Definition 5.9 (i), det is a linear function of A,

$$\begin{aligned} \det A &= \det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{a}_i \ \cdots \ \boldsymbol{a}_n\,] \\ \det A &= \det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \alpha\boldsymbol{b}_i \ \cdots \ \boldsymbol{a}_n\,] \\ &= \alpha\det[\,\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \ \cdots \ \boldsymbol{b}_i \ \cdots \ \boldsymbol{a}_n\,] = \alpha\det B. \end{aligned}$$

(ii) If each entry of a column a_i, say, is zero, then

$$a_i = 0 = 0\,b_i \quad \text{for any } b_i.$$

Hence, using **(i)** of **(5.9)**, we have

$$\det A = 0.\ \det B = 0.$$

(iii) Suppose that, for a column a_i of A, $a_i = b_i + c_i$.

Then, by **(i)** of **(5.9)**, taking $\beta = \gamma = 1$, we have

$$\begin{aligned}
\det A &= \det[\,a_1\ \ a_2\ \cdots\ b_i + c_i\ \ a_{i+1}\ \cdots\ a_n\,] \\
&= \det[\,a_1\ \ a_2\ \cdots\ b_i\ \ a_{i+1}\ \cdots\ a_n\,] \\
&\quad + \det[\,a_1\ \ a_2\ \cdots\ c_i\ \ a_{i+1}\ \cdots\ a_n\,] \\
&= \det B + \det A,
\end{aligned}$$

where B and C are the matrices obtained from A by replacing its column a_i by b_i and c_i respectively.

(iv) Suppose that a constant multiple of the column a_{i+1} is added to the column a_i, that is a_i is replaced by $a_i + c\,a_{i+1}$. If B denotes the new matrix, then

$$\begin{aligned}
\det B &= \det[\,a_1\ \ a_2\ \cdots\ a_i + c_i a_{i+1}\ \ a_{i+1}\ \cdots\ a_n\,] \\
&= \det[\,a_1\ \ a_2\ \cdots\ a_i\ \ a_{i+1}\ \cdots\ a_n\,] \\
&\quad + c\det[\,a_1\ \ a_2\ \cdots\ a_{i+1}\ \ a_{i+1}\ \cdots\ a_n\,] \\
&= \det A + c\,0 = \det C,
\end{aligned}$$

(v) Let $\quad A = [\,a_1\ \ a_2\ \cdots\ a_i\ \ a_{i+1}\ \cdots\ a_n\,]$

and $\quad B = [\,a_1\ \ a_2\ \cdots\ a_{i+1}\ \ a_i\ \cdots\ a_n\,]$

be the matrix obtained from A by interchanging its ith and $(i+1)$th columns. Then

$$\begin{aligned}
\det B &= \det[\,a_1\ \ a_2\ \cdots\ a_{i+1}\ \ a_i\ \cdots\ a_n\,] \\
&= \det[\,a_1\ \ a_2\ \cdots\ a_{i+1}\ \ a_i + a_{i+1}\ \ a_{i+2}\ \cdots\ a_n\,] \\
&\qquad\qquad\downarrow \\
&= \det[\,a_1\ \ a_2\ \cdots\ -a_i + a_i + a_{i+1}\ \ a_i + a_{i+1}\ \ a_{i+2}\ \cdots\ a_n\,], && \text{by (iv)} \\
&= \det[\,a_1\ \ a_2\ \cdots\ -a_i\ \ a_i + a_{i+1}\ \ a_{i+2}\ \cdots\ a_n\,], && \text{by (iv)} \\
&= -\det[\,a_1\ \ a_2\ \cdots\ a_i\ \ a_i + a_{i+1}\ \ a_{i+2}\ \cdots\ a_n\,], && \text{by (i)} \\
&= \det[\,a_1\ \ a_2\ \cdots\ a_i\ \ a_i + a_{i+1}\ \cdots\ a_n\,], && \text{by (iv)} \\
&= -\det A.
\end{aligned}$$

(vi) Suppose that two columns $a_i, a_j, i \neq j$ of A are identical, i.e., $a_i = a_j$. Then

$$\begin{aligned}
\det A &= \det[\,a_1\ \ a_2\ \cdots\ a_i\ \cdots\ a_j\ \cdots\ a_n\,] \\
&= -\det[\,a_1\ \ a_2\ \cdots\ a_j\ \cdots\ a_i\ \cdots\ a_n\,], && \text{by (vi)} \\
&= -\det[\,a_1\ \ a_2\ \cdots\ a_i\ \cdots\ a_j\ \cdots\ a_n\,], && \text{because } a_i = a_j \\
&= -\det A.
\end{aligned}$$

Hence

$$2\det A = 0.\ \text{That is,}\ \det A = 0.$$

The other properties may be verified similarly

DETERMINANT AS SUM OF PRODUCTS OF ELEMENTS

(5.12) Definition. Let $A = [\,a_{ij}\,]$ be a square matrix of order n with entries from a field F. Then the sum

$$\sum_{\sigma} (-1)^k a_{1k_1} a_{2k_2} \cdots a_{nk_n}$$

where k is 0 or 1 according as the permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ k_1 & k_2 & k_3 & \cdots & k_n \end{pmatrix}$$

in S_n is even or odd, is called the **determinant** of A and is denoted by $\det A$ or by

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

In particular if $n = 3$, then S_3 has $3! = 6$ elements which are

$$I = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix},\ \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},\ \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},\ \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

Of the above six permutations, the first three are even and the next three are odd. So if

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

then $\det A = \sum_{\sigma}(-1)^k a_{1i_1} a_{2i_2} a_{3i_3}$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \qquad (2)$$

Here

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ i_1 & i_2 & i_3 \end{pmatrix}$$

is one of the permutations given above. Thus, for writing the products in (2), we first take the subscript i in a_{ij} as in the first row of the corresponding permutation and j as the corresponding element below i in the second row of the permutation. This process is repeated for all the six permutations.

For example, the term in sum (2) corresponding to the first permutation namely

$$I = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$$

is the product $a_{11}\, a_{22}\, a_{33}$. Likewise, for the permutation

$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

we have the product $a_{12}\, a_{21}\, a_{33}$ with minus sign because σ_3 is an odd permutation.

Similarly, if

$$A = \begin{bmatrix} 1 & 6 & 7 \\ 2 & 3 & 5 \\ 4 & 1 & 7 \end{bmatrix},$$

then, writing the corresponding values of a_{ij} in the matrix A, we have:

$$\begin{aligned} \det A &= 1\times3\times7+6\times5\times4+7\times2\times1-6\times2\times7-1\times5\times1-7\times3\times4 \\ &= 21+120+14-84-5-84 \\ &= 155-173 = -18. \end{aligned}$$

DETERMINANT OF THE TRANSPOSE

In the following theorem we shall see that all statements about columns of a matrix A in the definition of det A hold equally good for the rows of A.

(5.13) Theorem. Let A^T be the transpose of a square matrix A of order n. Then

$$\det A^T = \det A$$

Proof. Let

$A = [\, a_{ij} \,]$ be an $n \times n$ matrix and

$A^T = [\, b_{ij} \,], \quad b_{ij} = a_{ji}$

be the transpose of A. Here a_{ji} is the element in the ith row and jth column of A^T. Now

$$\begin{aligned} \det A^T &= \sum_{\sigma}(\pm)\, b_{1i_1}\, b_{2i_2} \cdots b_{ni_n} \\ &= \sum_{\sigma}(\pm)\, b_{1j_1}\, b_{2j_2} \cdots b_{nj_n}, \end{aligned}$$

where if

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{pmatrix}$$

then

$$\sigma^{-1} = \begin{pmatrix} j_1 & j_2 & j_3 & \cdots & j_n \\ 1 & 2 & 3 & \cdots & n \end{pmatrix}$$

is the inverse of σ. Since the inverse of an even or odd permutation is even or odd respectively, the signs with

$$b_{1j_1}\, b_{2j_2} \cdots b_{nj_n}$$

and

$$a_{1j_1}\, a_{2j_2} \cdots a_{nj_n} \qquad (1)$$

are the same. Moreover, a term like that in (1) also appears in det A. Hence each term in the expansion of det A^T is a term in the expansion of det A and conversely. Hence

$$\det A^T = \det A.$$

For example let

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{bmatrix}, \text{ then } A^T = \begin{bmatrix} 1 & 3 & 0 \\ -1 & 1 & -2 \\ 2 & 4 & 5 \end{bmatrix}$$

and $\det A = 16 = \det A^T$

(5.14) Remark. The expansion of det A as given in Definition 5.5 is called the **expansion of determinant by rows** or **row expansion** of det A. Similarly, one has the notion of **column expansion** of det A. Theorem 5.13 shows that these two expansions are algebraically identical.

(5.15) Remark. Theorem 5.13 also explains the fact that, in every theorem about determinants, the role of columns and rows may be interchanged to obtain another theorem. Thus, Theorem 5.11 which describes the properties of det A in terms of columns of A can be restated in terms of its rows. Thus we have

(5.16) Theorem. Let A be an $n \times n$ matrix and let det A denote the determinant function. Then

(i) If the row vector $\boldsymbol{a}_i^T = [\,a_{i1}\ a_{i2}\ \cdots\ a_{in}\,]$ of A is of the form

$$\boldsymbol{a}_i^T = \alpha \boldsymbol{b}_i^T$$

then

$$\det A = \alpha \det B$$

where B is the matrix obtained from A by replacing its ith row $\boldsymbol{a}_i^T$ by $\boldsymbol{b}_i^T$

(ii) If a row $\boldsymbol{a}_i^T$ of A is $\boldsymbol{0}$, then

$$\det A = 0$$

(iii) If a row $\boldsymbol{a}_i^T$ of A is of the form

$$\boldsymbol{a}_i^T = \boldsymbol{b}_i^T + \boldsymbol{c}_i^T$$

then

$$\det A = \det B + \det C$$

where B and C are matrices obtained from A by replacing its ith row by $\boldsymbol{b}_i^T$ and $\boldsymbol{c}_i^T$ respectively

(iv) If B is the matrix obtained from A by replacing its ith row $\boldsymbol{a}_i^T$ by $\boldsymbol{a}_i^T + c\,\boldsymbol{a}_{i+1}^T$ then

$$\det B = \det A$$

(v) Interchange of any two adjacent rows of A changes the sign of det A

(vi) Interchange of any two rows of A results in the change of sign of det A

(vii) If any two rows of A are identical then det $A = 0$

(viii) A constant multiple of a row of A may be added to any other row without changing the value of det A. That is

$$\det A = \det[\,\boldsymbol{a}_1^T\ \boldsymbol{a}_2^T\ \cdots\ \boldsymbol{a}_i^T\ \cdots\ \boldsymbol{a}_j^T\ \cdots\ \boldsymbol{a}_n^T\,]$$
$$= \det[\,\boldsymbol{a}_1^T\ \boldsymbol{a}_2^T\ \cdots\ \boldsymbol{a}_i^T + c\boldsymbol{a}_j^T\ \cdots\ \boldsymbol{a}_j^T\ \cdots\ \boldsymbol{a}_n^T\,]$$

for any scalar c

Proof. This theorem can be proved on the same lines as Theorem 5.11 and is left as an exercise.

The main use of the above theorem lies in the fact that it facilitates the evaluation of det A by its rows.

(5.17) **Recall that**, for an $n \times n$ matrix A, the following three types of operations are called **elementary row (column) operations**.

(i) Interchanging any two rows (columns) of A

(ii) Multiplying every element of a row (column) of A by a nonzero scalar

(iii) Adding a scalar multiple of a row (column) of A to any other row (column)

Also an **elementary matrix** is a matrix E obtained from the identity matrix I_n by an elementary row (column) operation

Thus

I_{pq} = the matrix obtained from I_n by interchanging the pth and qth row (column),

I_{cp} = the matrix obtained from I_n by multiplying its pth row (column) by a nonzero scalar c,

I_{p+cq} = the matrix obtained from I_n by adding to the pth row (column) c-times the qth row (column)

All of these are **examples** of elementary matrices

It is easy to see that

$$\det I_{pq} = -1, \quad \det I_{cp} = c, \quad \det I_{p+cq} = 1$$

Also for any two elementary matrices E_1 and E_2

$$\det (E_1 E_2) = \det E_1 \cdot \det E_2$$

by actual computation. For example:

$$\det (I_{pq} \; I_{cp}) = -c = \det I_{pq} \cdot \det I_{cp}$$

$$\det (I_{pq} \; I_{p+cq}) = -1 = \det I_{pq} \cdot \det I_{p+cq}$$

With the above remarks, we are now in a position to prove the following important result about the product of determinants.

(5.18) Theorem. (The Product Theorem). For any $n \times n$ matrices A and B,

$$\det (AB) = \det A \cdot \det B = \det B \cdot \det A = \det (BA)$$

Proof. We have the following cases.

(i) One of the matrices, say A, is singular, that is, $\det A = 0$. In this case, there are elementary matrices $E_1, E_2, \cdots, E_k$ such that

$$E_1 \; E_2 \cdots E_k A = \begin{bmatrix} A_r & 0 \\ 0 & 0 \end{bmatrix}$$

where A_r is an $r \times r$ invertible matrix. So

$$E_1 \; E_2 \; \cdots \; E_k AB = \begin{bmatrix} A_r & 0 \\ 0 & 0 \end{bmatrix} B.$$

Hence $\quad \text{Rank} (AB) \le \text{Rank } A < n$

So

$$\det (AB) = 0 = \det A \cdot \det B$$

Similar is the case when B is singular.

(ii) Now suppose that both A and B are nonsingular. Then there are elementary matrices

$E_1, E_2, \cdots, E_r$ and $F_1, F_2, \cdots, F_s$

such that $\quad A = E_1 E_2 \cdots E_r$ and $B = F_1 F_2 \cdots F_s$

So $\quad AB = E_1 E_2 \cdots E_r \; F_1 F_2 \cdots F_s$

and $\quad \det A = \det (E_1 E_2 \cdots E_r), \det B = \det (F_1 F_2 \cdots F_s).$

By making repeated use of the result in (5.17) about $\det (E_1 E_2)$, E_1, E_2 being elementary matrices, we have

$$\det (AB) = \det (E_1) \cdots \det (E_r) \cdot \det (F_1) \cdots \det (F_s)$$
$$= \det A \cdot \det B.$$

Similarly, $\quad \det (BA) = \det B \cdot \det A$

$$= \det A \cdot \det B$$
$$= \det (AB)$$

and the proof is complete.

We illustrate the product theorem as follows:

Let

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 3 \\ 2 & 5 \end{bmatrix}$$

Then

$$AB = \begin{bmatrix} 10 & 11 \\ 16 & 19 \end{bmatrix}$$

Now

$$\det A = 4 - 3 = 1, \quad \det B = 20 - 6 = 14$$

$$\det (AB) = 190 - 176 = 14 = \det A \cdot \det B$$

Hence

$$\det (AB) = \det A \cdot \det B.$$

AN ALGORITHM TO EVALUATE DET A

(5.19) By an **algorithm** we mean a sequence of a finite number of steps to get a desired result. The word algorithm comes from the famous Muslim mathematician Al-Khwarizmi who invented the world ALGEBRA.

A step by step evaluation of $\det A$ of order n is obtained as follows:

Step 1: By an interchange of rows of A (and taking into account the resulting sign) bring a nonzero entry to the $(1, 1)$ position (unless all the entries in the first column are zero (in which case $\det A = 0$)).

Step 2: By adding suitable multiples of the first row to all the other rows, reduce the $(n-1)$ entries, except $(1, 1)$ in the first column, to 0. Expand det A by its first column. Repeat this process.

Or continue the following steps.

Step 3: Repeat Step 1 and Step 2 with the last remaining rows concentrating on the second column.

Step 4: Repeat Step 1, Step 2 and Step 3 with the remaining $n-2$ rows, $n-3$ rows and so on, until a triangular matrix is obtained.

Step 5: Multiply all the diagonal entries of the resulting triangular matrix and then multiply it by its sign to get det A.

Example 4. Evaluate det A, where

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix}$$

Solution. Here,

$$\det A = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -7 \\ 0 & -2 & -8 & -10 \\ 0 & -7 & -10 & -13 \end{vmatrix} \quad \text{by } R_2 + (-2)R_1,\ R_3 + (-3)R_1 \text{ and } R_4 + (-4)R_1$$

$$= \begin{vmatrix} -1 & -2 & -7 \\ -2 & -8 & -10 \\ -7 & -10 & -13 \end{vmatrix}, \quad \text{expanding by the first column}$$

$$= -2 \begin{vmatrix} 1 & 2 & 7 \\ 1 & 4 & 5 \\ 7 & 10 & 13 \end{vmatrix}, \quad \text{taking } (-1), (-2) \text{ and } (-1) \text{ common from 1st, 2nd and 3rd rows respectively}$$

$$= -2 \begin{vmatrix} 1 & 2 & 7 \\ 0 & 2 & -2 \\ 0 & -4 & -36 \end{vmatrix}, \quad \text{by } R_2 + (-1)R_1,\ R_3 + (-7)R_1$$

$$= -2 \begin{vmatrix} 2 & -2 \\ -4 & -36 \end{vmatrix}, \quad \text{expanding by first column}$$

$$= -2(2)(-4) \begin{vmatrix} 1 & -1 \\ 1 & 9 \end{vmatrix}, \quad \text{taking out 2 and } -4 \text{ from 1st and 2nd rows respectively}$$

$$= 16 \begin{vmatrix} 1 & -1 \\ 0 & 10 \end{vmatrix}, \quad \text{by } R_2 + (-1)R_1$$

$$= 160.$$

Example 5. Let A and B be matrices of order 6 such that

$$\det(AB^2) = 72, \qquad \det(A^2B^2) = 144.$$

Find det (A), det $(2A)$ and det (AB^6).

Solution. Suppose $\det A = a$ and $\det B = b$.

Then, by the product rule of determinants, we have

$$\det(BA^2) = \det A \det B^2 = \det A\,(\det B)^2 = ab^2 = 72$$

$$\det(A^2B^2) = \det A^2 . \det B^2 = (\det A)^2 . (\det B)^2 = a^2b^2 = 144$$

So $\quad \det A = a = \dfrac{a^2b^2}{ab^2} = \dfrac{144}{72} = 2$

Also

$$\det(2A) = 2^6 \det A = 128.$$

Lastly,

$$\det(AB^6) = \det A \cdot \det B^6$$
$$= \det A \, (\det B)^6$$
$$= ab^6$$

Since $\quad b^2 = \dfrac{72}{a} = \dfrac{72}{2} = 36,$

$$\det(AB^6) = 2 \times (36)^3 = 93312$$

Example 6. Evaluate the determinant of the matrix

$$D = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$$

Solution. Expanding by the first row, we have

$$\det D = aA - hH + gG$$

where $\quad A = \begin{vmatrix} b & f \\ f & c \end{vmatrix} = bc - f^2,$

$$H = -\begin{vmatrix} h & f \\ g & c \end{vmatrix} = fg - ch,$$

$$G = \begin{vmatrix} h & b \\ g & f \end{vmatrix} = hf - bg$$

Hence $\quad \det D = a(bc - f^2) + h(fg - ch) + g(hf - bg)$

$$= abc - af^2 + fgh - ch^2 + fgh - bg^2$$

$$= abc + 2fgh - af^2 - bg^2 - ch^2.$$

Example 7. Evaluate the determinant of the matrix

$$A = \begin{bmatrix} 4 & 2 & 5 & 10 \\ 1 & 1 & 6 & 3 \\ 7 & 3 & 0 & 5 \\ 0 & 2 & 5 & 8 \end{bmatrix}$$

Solution.

The solution given below makes use of Theorem 5.16 and the algorithm (5.19) by reducing the determinant into a determinant of smaller size. Thus

$$\det A = \begin{vmatrix} 4 & 2 & 5 & 10 \\ 1 & 1 & 6 & 3 \\ 7 & 3 & 0 & 5 \\ 0 & 2 & 5 & 8 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 1 & 6 & 3 \\ 4 & 2 & 5 & 10 \\ 7 & 3 & 0 & 5 \\ 0 & 2 & 5 & 8 \end{vmatrix} \qquad \text{by } R_{12}\text{, changing the sign}$$

$$= -\begin{vmatrix} 1 & 1 & 6 & 3 \\ 0 & -2 & -19 & -2 \\ 0 & -4 & -42 & -16 \\ 0 & 2 & 5 & 8 \end{vmatrix} \qquad \text{by } R_2 - 4R_1,\ R_3 - 7R_1$$

(The value of det A does not change with these row operations.)

$$= -\begin{vmatrix} -2 & -19 & -2 \\ -4 & -42 & -16 \\ 2 & 5 & 8 \end{vmatrix}, \qquad \text{evaluating by the first column}$$

$$= -\begin{vmatrix} 2 & 19 & 2 \\ 4 & 42 & 16 \\ 2 & 5 & 8 \end{vmatrix} \qquad \text{by } (-1)R_1,\ (-1)R_2$$

i.e. multiplying the determinant by $(-1)^2$

$$= -2\begin{vmatrix} 2 & 19 & 2 \\ 2 & 21 & 8 \\ 2 & 5 & 8 \end{vmatrix} \qquad \text{taking out 2 from } R_2$$

$$= -2\begin{vmatrix} 2 & 19 & 2 \\ 0 & 2 & 6 \\ 0 & -14 & 6 \end{vmatrix} \quad \text{by } R_2 - R_1 \text{ and } R_3 - R_1$$

$$= -4\begin{vmatrix} 2 & 6 \\ -14 & 6 \end{vmatrix} = -4(12+84) = -384$$

Example 8. Prove the identity from definition:

$$\begin{vmatrix} \beta\gamma\delta & \alpha & \alpha^2 & \alpha^3 \\ \gamma\delta\alpha & \beta & \beta^2 & \beta^3 \\ \delta\alpha\beta & \gamma & \gamma^2 & \gamma^3 \\ \alpha\beta\gamma & \delta & \delta^2 & \delta^3 \end{vmatrix} = \begin{vmatrix} 1 & \alpha^2 & \alpha^3 & \alpha^4 \\ 1 & \beta^2 & \beta^3 & \beta^4 \\ 1 & \gamma^2 & \gamma^3 & \gamma^4 \\ 1 & \delta^2 & \delta^3 & \delta^4 \end{vmatrix}$$

Solution. Let

$$\Delta = \begin{vmatrix} \beta\gamma\delta & \alpha & \alpha^2 & \alpha^3 \\ \gamma\delta\alpha & \beta & \beta^2 & \beta^3 \\ \delta\alpha\beta & \gamma & \gamma^2 & \gamma^3 \\ \alpha\beta\gamma & \delta & \delta^2 & \delta^3 \end{vmatrix}$$

Multiply the first, second, third and fourth rows of Δ by α, β, γ and respectively. Since Δ is the multiplied by $\alpha\beta\gamma\delta$, we have

$$\Delta = \frac{1}{\alpha\beta\gamma\delta}\begin{vmatrix} \alpha\beta\gamma\delta & \alpha^2 & \alpha^3 & \alpha^4 \\ \alpha\beta\gamma\delta & \beta^2 & \beta^3 & \beta^4 \\ \alpha\beta\gamma\delta & \gamma^2 & \gamma^3 & \gamma^4 \\ \alpha\beta\gamma\delta & \delta^2 & \delta^3 & \delta^4 \end{vmatrix}$$

Take out $\alpha\beta\gamma\delta$ from the first column. Then

$$\Delta = \frac{\alpha\beta\gamma\delta}{\alpha\beta\gamma\delta}\begin{vmatrix} 1 & \alpha^2 & \alpha^3 & \alpha^4 \\ 1 & \beta^2 & \beta^3 & \beta^4 \\ 1 & \gamma^2 & \gamma^3 & \gamma^4 \\ 1 & \delta^2 & \delta^3 & \delta^4 \end{vmatrix} = \begin{vmatrix} 1 & \alpha^2 & \alpha^3 & \alpha^4 \\ 1 & \beta^2 & \beta^3 & \beta^4 \\ 1 & \gamma^2 & \gamma^3 & \gamma^4 \\ 1 & \delta^2 & \delta^3 & \delta^4 \end{vmatrix}$$

as required

Example 9. Verify that

$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (a-b)(b-c)(c-a)$$

Solution. Let

$$\Delta = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$$

Subtracting the first row from the second and third rows respectively, we have

$$\Delta = \begin{vmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{vmatrix}$$

Taking factors $(b-a)$, $(c-a)$ out of the second and third rows, we get

$$\Delta = (b-a)(c-a)\begin{vmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & 1 & c+a \end{vmatrix}$$

$$= (b-a)(c-a)\begin{vmatrix} 1 & b+a \\ 1 & c+a \end{vmatrix}, \quad \text{on expanding by the first column}$$

$$= (b-a)(c-a)(c+a-b-a)$$

$$= (a-b)(b-c)(c-a)$$

Example 10. Show that

$$\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} = -(a+b+c)(a+b\omega+c\omega^2)(a+b^2+c\omega).$$

where $\omega^3 = 1$.

Solution. Let

$$\Delta = \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$$

and consider the determinant

$$\Omega = \begin{vmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{vmatrix},$$ where ω is a complex cube root of 1.

Then $\Omega \neq 0$ because, adding the second and third columns to the first and using the identity $1 + \omega + \omega^2 = 0$, we have

$$\Omega = \begin{vmatrix} 3 & 1 & 1 \\ 0 & \omega & \omega^2 \\ 0 & \omega^2 & \omega \end{vmatrix} = 3(\omega^2 - \omega^4) = 3(\omega^2 - \omega) \neq 0$$

Now $$\Omega\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{vmatrix} \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$$

$$= \begin{vmatrix} a+b+c & b+c+a & c+a+b \\ a+b\omega+c\omega^2 & b+c\omega+a\omega^2 & c+a\omega+b\omega^2 \\ a+b\omega^2+c\omega & b+c\omega^2+a\omega & c+a\omega^2+b\omega \end{vmatrix}$$

However,

$$b + c\omega + a\omega^2 = \omega^2(a + b\omega + c\omega^2), c + a\omega + b\omega^2 = \omega(a + b\omega + c\omega^2),$$

$$b + c\omega^2 + a\omega = \omega(a + b\omega^2 + c\omega), c + a\omega^2 + b\omega = \omega^2(a + b\omega^2 + c\omega).$$

Hence, taking out factors $a + b + c$, $a + b\omega + c\omega^2$, $a + b\omega^2 + c\omega$ from the first, second and third rows respectively, we have

$$\Delta\Omega = (a+b+c)(a+b\omega+c\omega^2)(a+b\omega^2+c\omega)\begin{vmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{vmatrix}$$

$$= -(a+b+c)(a+b\omega+c\omega^2)(a+b\omega^2+c\omega)\,\Omega,$$

after interchanging the last two rows and thereby changing the sign to -ive. As $\Omega \neq 0$, we get

$$\Delta = -(a+b+c)(a+b\omega+c\omega^2)(a+b\omega^2+c\omega)$$

Alternative Solution:

We have $$\Delta = -\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}, \text{ by } R_{23}$$

$$= -\begin{vmatrix} a+c\omega+b\omega^2 & b+a\omega+c\omega^2 & c+b\omega+a\omega^2 \\ c & a & b \\ b & c & a \end{vmatrix},$$

by $R_1 + \omega R_2 + \omega^2 R_3$,

where $\omega^3 = 1$.

$$= -\begin{vmatrix} a+b\omega^2+c\omega & \omega(a+b\omega^2+c\omega) & \omega^2(a+b\omega^2+c\omega) \\ c & a & b \\ b & c & a \end{vmatrix}$$

Thus $a + b\omega^2 + c\omega$ is a factor of Δ. Since ω^2 and 1 are also cube roots of 1, $a + b(\omega^2)^2 + c\omega^2$ and $a + b + c$ are also factors of Δ

Hence $\Delta = -k(a+b+c)(a+b\omega^2+c\omega)(a+b\omega+c\omega^2)$ (1)

where k is a constant since Δ involves terms of third degree in a, b, c. Comparing coefficients of a^3 on both sides of (1), we get $k = 1$. Therefore,

$$\Delta = -(a+b+c)(a+b\omega^2+c\omega)(a+b\omega+c\omega^2)$$

Note. The determinant Δ is called a **circulant determinant** because its second and third rows are obtained by permuting the first row cyclically. By a similar procedure, a circulant determinant of any finite order can be evaluated in the factorised form.

DETERMINANTS AND INVERSES OF MATRICES

The section deals with another method of finding the inverse of a matrix. In the process we shall define the **adjugate or adjoint** of a matrix. We begin with the following simple result.

(5.20) Theorem. For an invertible matrix A, $\det A \neq 0$ and

$$\det(A^{-1}) = \frac{1}{\det A}$$

Proof. Let A be a nonsingular matrix with A^{-1} as its inverse.

Then

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

By Theorem 5.18

$$\det(A^{-1}A) = \det(A^{-1}).\det(A)$$

But $\det(A^{-1}A) = \det(I) = 1$

Hence $\det(A^{-1})\ \det A = 1$

$\Rightarrow$ $\det(A^{-1}) \neq 0$, $\det A \neq 0$

and $\det(A^{-1}) = \dfrac{1}{\det A}$.

Before we describe the other method of finding the inverse of a matrix A, we define the **adjugate or adjoint** of A as follows:

(5.21) Definition. Let $A = [a_{ij}]$ be an $n \times n$ matrix and A_{ij} be the cofactor of a_{ij} in A. Let

$$B = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \text{ be the matrix of cofactors}$$

Then **the adjugate or adjoint of A, written adj A, is the transpose of B.** Thus

$$\text{adj } A = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix} = B^T.$$

Example 11. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \text{ Find adj } A.$$

Solution. Here, the cofactors are

$$A_{11} = d, \quad A_{12} = -c, \quad A_{21} = -b, \quad A_{22} = a$$

Hence $B = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$

so that

$$\text{adj } A = B^T = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Example 12. Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}. \text{ Compute adj } A.$$

Solution. Here, the cofactors are

$$A_{11} = 5, \quad A_{12} = -1, \quad A_{13} = -7, \quad A_{21} = -1, \quad A_{22} = -7$$
$$A_{23} = 5, \quad A_{31} = -7, \quad A_{32} = 5, \quad A_{33} = -1.$$

Hence, the matrix of cofactors is

$$B = \begin{bmatrix} 5 & -1 & -7 \\ -1 & -7 & 5 \\ -7 & 5 & -1 \end{bmatrix}$$

and $\text{adj } A = B^T = \begin{bmatrix} 5 & -1 & -7 \\ -1 & -7 & 5 \\ -7 & 5 & -1 \end{bmatrix}$.

(5.22) Theorem. Let A be an $n \times n$ matrix. Then

$$A \text{ adj } A = \begin{bmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & & \det A \end{bmatrix} = (\det A)\, I_n.$$

In order to prove this theorem, we need the following

(5.23) Lemma. For any $n \times n$ matrix A, if $i \neq k$, then

$$\sum_{j=1}^{n} a_{ij} A_{kj} = a_{i1} A_{k1} + a_{i2} A_{k2} + \cdots + a_{in} A_{kn} = 0$$

Proof. Since

$$\det A = \sum_{j=1}^{n} a_{kj} A_{kj}$$

and adding the ith row to kth row of A does not change $\det A$, we have

$$\det A = \sum_{j=1}^{n} (a_{kj} + a_{ij}) A_{kj}, \quad i \neq k$$

$$= \sum_{j=1}^{n} a_{kj} A_{kj} + \sum_{j=1}^{n} a_{ij} A_{kj} = \det A + \sum_{j=1}^{n} a_{ij} A_{kj}$$

So $$\sum_{j=1}^{n} a_{ij} A_{kj} = a_{i1}A_{k1} + a_{i2} A_{k2} + \cdots a_{in} A_{kn} = 0, \quad i \neq k$$

Proof of Theorem 5.22. We have

$$A \operatorname{adj} A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

$$= [\, c_{ij} \,]_{n \times n}$$

where

$$c_{ij} = a_{i1} A_{j1} + a_{i2} A_{j2} + \cdots + a_{in} A_{jn}, \quad i, j = 1, 2, \cdots, n$$

$$= \begin{cases} \det A & \text{if } i = j \\ 0 & \text{if } i \neq k, \text{ by the Lemma 5.23} \end{cases}$$

Hence

$$A \operatorname{adj} A = \begin{bmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \det A \end{bmatrix} = (\det A)\, I_n$$

(5.24) Theorem. Let A be an $n \times n$ matrix. Then A is invertible if and only if $\det A \neq 0$. If $\det A \neq 0$, then

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A$$

Proof. If A is invertible then $\det A \neq 0$ by Theorem 5.20.

Conversely if $\det A \neq 0$, then

$$A\left(\frac{1}{\det A} \operatorname{adj} A\right) = \frac{1}{\det A} (A \operatorname{adj} A)$$

$$= \frac{\det A}{\det A} I_n = I_n$$

Hence A is invertible and

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A$$

Note. Evaluation of A^{-1} by Theorem 5.24 is known as finding the inverse of A by the **Adjoint Method.**

Example 13. Find, by adjoint method, the inverse of

$$A = \begin{bmatrix} 3 & 4 & 5 \\ 2 & -1 & 8 \\ 5 & -2 & 7 \end{bmatrix}$$

Solution. Here, the cofactors are

$$A_{11} = (-1)^{1+1} \det \begin{bmatrix} -1 & 8 \\ -2 & 7 \end{bmatrix} = 9, \quad A_{12} = (-1)^{1+2} \det \begin{bmatrix} 2 & 8 \\ 5 & 7 \end{bmatrix} = 26,$$

$$A_{13} = (-1)^{1+3} \det \begin{bmatrix} 2 & -1 \\ 5 & -2 \end{bmatrix} = 1, \quad A_{21} = (-1)^{2+1} \det \begin{bmatrix} 4 & 5 \\ -2 & 7 \end{bmatrix} = -38,$$

$$A_{22} = (-1)^{2+2} \det \begin{bmatrix} 3 & 5 \\ 5 & 7 \end{bmatrix} = -4, \quad A_{23} = (-1)^{2+3} \det \begin{bmatrix} 3 & 4 \\ 5 & -2 \end{bmatrix} = 26,$$

$$A_{31} = (-1)^{3+1} \det \begin{bmatrix} 4 & 5 \\ -1 & 8 \end{bmatrix} = 37, \quad A_{32} = (-1)^{3+2} \det \begin{bmatrix} 3 & 5 \\ 2 & 8 \end{bmatrix} = -14,$$

$$A_{33} = (-1)^{3+3} \det \begin{bmatrix} 3 & 4 \\ 2 & -1 \end{bmatrix} = -11$$

$$\therefore \det A = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} = 3 \times 9 + 4 \times 26 + 5 \times 1 = 136$$

Therefore,

$$A^{-1} = \frac{1}{\det A}\,\text{adj}\,A.$$

$$= \frac{1}{136}\begin{bmatrix} 9 & -38 & 37 \\ 26 & -4 & -14 \\ 1 & 26 & -11 \end{bmatrix}$$

Example 14. Let

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 0 \\ 3 & 1 & 5 \end{bmatrix}$$

Examine whether A is invertible and, if so, determine A^{-1}.

Solution. We have

$$\det A = \begin{vmatrix} 2 & 3 & 4 \\ 1 & 2 & 0 \\ 3 & 1 & 5 \end{vmatrix} = -\begin{vmatrix} 1 & 2 & 0 \\ 2 & 3 & 4 \\ 3 & 1 & 5 \end{vmatrix} \text{ by } R_{12}$$

$$= -\begin{vmatrix} 1 & 2 & 0 \\ 0 & -1 & 4 \\ 0 & -5 & 5 \end{vmatrix} \text{ by } R_2 - R_1 \text{ and } R_3 - R_1$$

$$= -15 \neq 0.$$

Hence A is invertible.

Now, $A_{11} = 10$, $A_{12} = -5$, $A_{13} = -5$,
$A_{21} = -11$, $A_{22} = -2$, $A_{23} = 7$,
$A_{31} = -8$, $A_{32} = 4$, $A_{33} = 1$.

Therefore,

$$A^{-1} = \frac{1}{\det A}\,\text{adj}\,A = -\frac{1}{15}\begin{bmatrix} 10 & -11 & -8 \\ -5 & -2 & 4 \\ -5 & 7 & 1 \end{bmatrix}$$

Note. If would be interesting for the reader to compare this method of finding the inverse of a matrix with the one using elementary row operations on A and the identity matrix I simultaneously discussed earlier in Example 19 of Chapter 3.

EXERCISE 5.1

1. Let M_2 be the set of all 2×2 matrices. Set up the transformation $A \longrightarrow \det A$, $A \in M_2$. What is the range of this mapping? Is the mapping one-to-one?

2. For 2×2 matrices A and B which of the following equations hold?

(i) $\det(A + B) = \det A + \det B$

(ii) $\det(A + B)^2 = (\det(A + B))^2$

(iii) $\det(A + B)^2 = \det(A^2 + B^2)$

(iv) $\det(A + B)^2 = \det(A^2 + 2AB + B^2)$

3. Evaluate the following determinants

(i) $\begin{vmatrix} 1 & 0 & 2 \\ 3 & 4 & 5 \\ 5 & 6 & 7 \end{vmatrix}$ (ii) $\begin{vmatrix} 2 & -1 & 1 \\ 3 & 2 & 4 \\ -1 & 0 & 3 \end{vmatrix}$

(iii) $\begin{vmatrix} 6 & -6 & 6 \\ 2 & 4 & -6 \\ 10 & -5 & 5 \end{vmatrix}$

4. Evaluate

(i) $\begin{vmatrix} 2 & 3 & -2 & 4 \\ 7 & 4 & -3 & 10 \\ 3 & 2 & 3 & 4 \\ -2 & 4 & 0 & 5 \end{vmatrix}$ (ii) $\begin{vmatrix} 3 & 7 & 5 & 2 \\ 2 & 4 & 1 & 1 \\ -2 & 0 & 0 & 0 \\ 1 & 1 & 3 & 4 \end{vmatrix}$

(iii) $\begin{vmatrix} 1 & -2 & 3 & -4 \\ 0 & 1 & -1 & 1 \\ 1 & 3 & 0 & -3 \\ 0 & -7 & 3 & 1 \end{vmatrix}$ (iv) $\begin{vmatrix} 9 & 93 & 12 & -6 \\ 1 & 92 & 84 & -6 \\ 2 & 185 & 100 & -12 \\ 4 & 270 & 196 & -24 \end{vmatrix}$

(v) $\begin{vmatrix} 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 & 1 \end{vmatrix}$ (vi) $\begin{vmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{vmatrix}$ in factorised form

5. Without expanding show that

(i) $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = \begin{vmatrix} e & b & h \\ d & a & g \\ f & c & k \end{vmatrix}$

(ii) $\begin{vmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{vmatrix} = 0$ (iii) $\begin{vmatrix} a+b & c & 1 \\ b+c & a & 1 \\ c+a & b & 1 \end{vmatrix} = 0$

6. Prove that each of following determinants vanishes:

(i) $\begin{vmatrix} bc & ca & ab \\ 1/a & 1/b & 1/c \\ a^2 & b^2 & c^2 \end{vmatrix}$ (ii) $\begin{vmatrix} a-b & b-c & c-a \\ b-c & c-a & a-b \\ c-a & a-b & b-c \end{vmatrix}$

(iii) $\begin{vmatrix} a & a^2 & a/bc \\ b & b^2 & b/ca \\ c & c^2 & c/ab \end{vmatrix}$ (iv) $\begin{vmatrix} \sin^2\theta & 1 & \cos^2\theta \\ \sin^2\phi & 1 & \cos^2\phi \\ \sin^2\psi & 1 & \cos^2\psi \end{vmatrix}$

(v) $\begin{vmatrix} \sin^2\alpha & \cos 2\alpha & \cos^2\alpha \\ \sin^2\beta & \cos 2\beta & \cos^2\beta \\ \sin^2\gamma & \cos 2\gamma & \cos^2\gamma \end{vmatrix}$

(vi) $\begin{vmatrix} \cos\alpha & \sin\alpha & \sin(\alpha+\delta) \\ \cos\beta & \sin\beta & \sin(\beta+\delta) \\ \cos\gamma & \sin\gamma & \sin(\gamma+\delta) \end{vmatrix}$

(vii) $\begin{vmatrix} 1 & \cos\alpha & \cos\beta \\ \cos\alpha & 1 & \cos(\alpha+\beta) \\ \cos\beta & \cos(\alpha+\beta) & 1 \end{vmatrix}$

(viii) $\begin{vmatrix} (a+b)^2 & a^2+b^2 & ab \\ (c+d)^2 & c^2+d^2 & cd \\ (g+h)^2 & g^2+h^2 & gh \end{vmatrix}$

(ix) $\begin{vmatrix} (a^m+a^{-m})^2 & (a^m-a^{-m})^2 & abc \\ (b^n+b^{-n})^2 & (b^n-b^{-n})^2 & abc \\ (c^p+c^{-p})^2 & (c^p-c^{-p})^2 & abc \end{vmatrix}$

(x) $\begin{vmatrix} \frac{1}{2!} & 1 & 0 \\ \frac{1}{3!} & \frac{1}{2!} & 1 \\ \frac{1}{4!} & \frac{1}{3!} & \frac{1}{2!} \end{vmatrix}$

(xi) $\begin{vmatrix} a^2 & b\sin A & c\sin A \\ b\sin A & 1 & \cos A \\ c\sin A & \cos A & 1 \end{vmatrix}$

where a, b, c are the magnitudes of the sides of a triangle and A is the measure of the angle opposite to the side with magnitude a

(xii) $\begin{vmatrix} a & b & c & d & 1 \\ b & c & d & a & 1 \\ c & d & a & b & 1 \\ d & a & b & c & 1 \\ b & a & d & c & 1 \end{vmatrix}$

(xiii) $\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{vmatrix}$

7. Without expansion, prove that

$$\begin{vmatrix} bc & a & a^2 \\ ca & b & b^2 \\ ab & c & c^2 \end{vmatrix} = \begin{vmatrix} 1 & a^2 & a^3 \\ 1 & b^2 & b^3 \\ 1 & c^2 & c^3 \end{vmatrix}$$

8. Show that

$$\begin{vmatrix} a & b & c \\ -c & a & b \\ -b & -c & a \end{vmatrix} = (a-b+c)(a+b\omega_1+c\omega_1^2)(a+b\omega_2+c\omega_2^2)$$

where ω_1 and ω_2 are cube roots of -1.

9. Prove that

$$\begin{vmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{vmatrix} = (a^2+b^2+c^2+d^2)^2.$$

10. Prove that

(i) $$\begin{vmatrix} (b+c)^2 & a^2 & a^2 \\ b^2 & (c+a)^2 & b^2 \\ c^2 & c^2 & (a+b)^2 \end{vmatrix} = 2abc(a+b+c)^3$$

(ii) $$\begin{vmatrix} \frac{a^2+b^2}{c} & c & c \\ a & \frac{b^2+c^2}{a} & a \\ b & b & \frac{c^2+a^2}{b} \end{vmatrix} = 4\,abc.$$

11. Prove that

$$\begin{vmatrix} x & a & a & a \\ a & x & a & a \\ a & a & x & a \\ a & a & a & x \end{vmatrix} = (x-a)^3(x+3a).$$

12. Show that

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \beta\gamma & \gamma\alpha & \alpha\beta \end{vmatrix} = (\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)$$

13. Show that

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^3 & \beta^3 & \gamma^3 \end{vmatrix} = (\beta-\gamma)(\gamma-\alpha)(\alpha-\beta)(\alpha+\beta+\gamma)$$

14. Show that

(i) $$\begin{vmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{vmatrix} = (a-1)^3(a+3)$$

(ii) $$\begin{vmatrix} 1+a & 1 & 1 & 1 \\ 1 & 1+b & 1 & 1 \\ 1 & 1 & 1+c & 1 \\ 1 & 1 & 1 & 1+d \end{vmatrix} = abcd\left(1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)$$

(iii) $$\begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1-x & 1 & 1 \\ 1 & 1 & 1+y & 1 \\ 1 & 1 & 1 & 1-y \end{vmatrix} = x^2y^2.$$

15. Prove that

$$\begin{vmatrix} a^3 & 3a^2 & 3a & 1 \\ a^2 & a^2+2a & 2a+1 & 1 \\ a & 2a+1 & a+2 & 1 \\ 1 & 3 & 3 & 1 \end{vmatrix} = (a-1)^6.$$

16. If

$$\Delta = \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}$$

and $A, B, C, \cdots$ are cofactors of $a, b, c, \cdots$ in Δ, then show that

(i)	$BC - F^2 = a\Delta$	(ii)	$CA - G^2 = b\Delta$
(iii)	$AB - H^2 = c\Delta$	(iv)	$GH - AF = f\Delta$
(v)	$HF - BG = g\Delta$	(vi)	$FG - CH = h\Delta$
(vii)	$aG + hF + gC = 0$	(viii)	$hG + bF + fC = 0$
(ix)	$gG + fF + cC = \Delta$		

17. If Δ of Problem 16 is zero, show that

(i)	$F^2 + G^2 = C(A + B)$	(ii)	$G^2 + H^2 = A(B + C)$
(iii)	$H^2 + F^2 = B(A + C)$	(iv)	$ABC = FGH$

18. Prove that

$$\begin{vmatrix} 0 & c & b \\ c & 0 & a \\ b & a & 0 \end{vmatrix}^2 = \begin{vmatrix} b^2 + c^2 & ab & ac \\ ba & c^2 + a^2 & bc \\ ca & cb & a^2 + b^2 \end{vmatrix}$$

19. Show that

$$\begin{vmatrix} 1 & \omega & \omega^2 & \omega^3 \\ 1 & \omega^2 & \omega^4 & \omega \\ 1 & \omega^3 & \omega & \omega^4 \\ 1 & \omega^4 & \omega^3 & \omega^2 \end{vmatrix}^2 = 125, \text{ where } \omega \text{ is a fifth root of } 1$$

20. Prove that the determinant

$$\begin{vmatrix} 2b_1 + c_1 & c_1 + 3a_1 & 2a_1 + 3b_1 \\ 2b_2 + c_2 & c_2 + 3a_2 & 2a_2 + 3b_2 \\ 2b_3 + c_3 & c_3 + 3a_3 & 2a_3 + 3b_3 \end{vmatrix}$$

is a multiple of the determinant

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \text{ and find the other factor}$$

21. Show that the determinant

$$\begin{vmatrix} 3 & 5 & 2 & 8 & 2 \\ 4 & 4 & 7 & 5 & 9 \\ 5 & 8 & 9 & 1 & 6 \\ 8 & 0 & 6 & 5 & 2 \\ 9 & 2 & 4 & 6 & 9 \end{vmatrix} \text{ is a multiple of } 13$$

22. Prove that

$$\begin{vmatrix} (a-x)^2 & (a-y)^2 & (a-z)^2 \\ (b-x)^2 & (b-y)^2 & (b-z)^2 \\ (c-x)^2 & (c-y)^2 & (c-z)^2 \end{vmatrix} = 2(a-b)(b-c)(c-a)(x-y)(y-z)(z-x)$$

23. Show that

$$\begin{vmatrix} 2bc - a^2 & c^2 & b^2 \\ c^2 & 2ca - b^2 & a^2 \\ b^2 & a^2 & 2ab - c^2 \end{vmatrix}$$

$$= \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} \begin{vmatrix} -a & c & b \\ -b & a & c \\ -c & b & a \end{vmatrix} = (a^3 + b^3 + c^3 - 3abc)^2.$$

24. Find, by the adjoint method, the inverse of each of the following matrices:

(i) $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ (ii) $\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & -1 \\ 3 & 2 & -2 \end{bmatrix}$

MISCELLANEOUS RESULTS

(5.25) Theorem. Let A be an $n \times n$ matrix whose elements are functions of a variable x

(i) If two rows (or columns) of A become identical when x is replaced by α then $x - \alpha$ is a factor of $\det A$

(ii) If r rows (or columns) of A become identical when x is replaced by α then $(x-\alpha)^{r-1}$ is a factor of $\det A$

Proof.(i) Suppose that, for the sake of convenience only, two rows a_1^T, a_2^T become identical when x is replaced by α in the matrix A whose elements are functions of a variable x. Then $a_2^T - a_1^T$ has $x - \alpha$ as a factor and this is then also a factor of $\det A$

(ii) If r rows

$$a_1^T, a_2^T, \cdots, a_r^T$$

of A, say, become identical when x is replaced by α, then $x - \alpha$ is a factor of each of

$$a_{i+1}^T - a_1^T, \qquad i = 1, 2, \cdots, r-1$$

Thus $(x-\alpha)^{r-1}$ is a factor of $\det A$

Example 15. Evaluate $\det A$, where

$$A = \begin{bmatrix} x^3 & x^2 & x & 1 \\ \alpha^3 & \alpha^2 & \alpha & 1 \\ \beta^3 & \beta^2 & \beta & 1 \\ \gamma^3 & \gamma^2 & \gamma & 1 \end{bmatrix}$$

Solution. Note that two rows of A become identical when x is replaced by α or β or γ in A. So

$$(x-\alpha), (x-\beta), (x-\gamma)$$

are factors of $\det A$.

Similarly, two rows become identical when β and γ are replaced by α in A. So

$$(\beta-\alpha), (\gamma-\alpha)$$

are factors of $\det A$. Also, if β is replaced by γ, then two rows of A become identical. Hence $(\beta-\gamma)$ is a factor of $\det A$. Thus

$$\det A = k(x-\alpha)(x-\beta)(x-\gamma)(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha) \qquad (1)$$

where k is to be determined. The coefficient of x^3 in $\det A$ is

$$\begin{vmatrix} \alpha^2 & \alpha & 1 \\ \beta^2 & \beta & 1 \\ \gamma^2 & \gamma & 1 \end{vmatrix} = \begin{vmatrix} \alpha^2 & \alpha & 1 \\ \beta^2-\alpha^2 & \beta-\alpha & 0 \\ \gamma^2-\alpha^2 & \gamma-\alpha & 0 \end{vmatrix} \quad \text{by } R_2 - R_1 \text{ and } R_3 - R_1$$

$$= (\beta-\alpha)(\gamma-\alpha)\begin{vmatrix} \beta+\alpha & 1 \\ \gamma+\alpha & 1 \end{vmatrix}, \quad \text{expending by 3rd column}$$

$$= (\beta-\alpha)(\gamma-\alpha)(\beta-\gamma)$$

The coefficient of x^3 on the right hand side of (1) is

$$k(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha) = -k(\beta-\alpha)(\beta-\gamma)(\gamma-\alpha)$$

So $\quad k = -1$

Hence $\quad \det A = -(x-\alpha)(x-\beta)(x-\gamma)(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)$

(5.26) Theorem. If A is an $n \times n$ skew symmetric matrix and n is odd, then $\det A = 0$

Proof. Suppose that A is an $n \times n$ skew symmetric matrix and n is odd

Now $\quad a_{ji} = -a_{ij}$

Since $\quad A^T = -A$,

$$\det A = \det A^T = \det(-A) = (-1)^n \det A = -\det A, \; n \text{ being odd}$$

Hence $\; 2\det A = 0$. So $\det A = 0$

Example 16. Let

$$B = \begin{bmatrix} 0 & b & c \\ -b & 0 & a \\ -c & -a & 0 \end{bmatrix}. \text{ Then } B^T = \begin{bmatrix} 0 & -b & -c \\ b & 0 & -a \\ c & a & 0 \end{bmatrix} = -B$$

So $\det B = \det B^T = \det(-B) = -\det B$

Thus $\det B = 0$.

Example 17. Find the values of λ so that

$\det(A - \lambda I) = 0$ when

(i) $A = \begin{bmatrix} 3 & 5 \\ 4 & 4 \end{bmatrix}$, (ii) $A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & 0 & -1 \\ 1 & -1 & 2 \end{bmatrix}$

Solution.

(i) For $A = \begin{bmatrix} 3 & 5 \\ 4 & 4 \end{bmatrix}$, $A - \lambda I = \begin{bmatrix} 3 & 5 \\ 4 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$

$$= \begin{bmatrix} 3-\lambda & 5 \\ 4 & 4-\lambda \end{bmatrix}$$

So

$$\begin{aligned} \det(A - \lambda I) &= \lambda^2 - 7\lambda + 12 - 20 \\ &= \lambda^2 - 7\lambda - 8 \\ &= (\lambda - 8)(\lambda + 1) \end{aligned}$$

Thus

$$\det(A - \lambda I) = 0 \quad \text{when } \lambda = 8 \text{ or } \lambda = -1.$$

(ii) For

$$A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & 0 & -1 \\ 1 & -1 & 2 \end{bmatrix},$$

$$A - \lambda I = \begin{bmatrix} -1 & 1 & 2 \\ 3 & 0 & -1 \\ 1 & -1 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$= \begin{bmatrix} -1-\lambda & 1 & 2 \\ 3 & -\lambda & -1 \\ 1 & -1 & 2-\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} -(1+\lambda) & 1 & 2 \\ 3 & -\lambda & -1 \\ 1 & -1 & 2-\lambda \end{vmatrix}$$

$$= -(1+\lambda)\begin{vmatrix} -\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} - \begin{vmatrix} 3 & -1 \\ 1 & 2-\lambda \end{vmatrix} + 2\begin{vmatrix} 3 & -\lambda \\ 1 & -1 \end{vmatrix}$$

$$= -(1+\lambda)(\lambda^2 - 2\lambda - 1) - (6 - 3\lambda + 1) + 2(-3 + \lambda)$$

$$\begin{aligned} &= -(\lambda^3 - \lambda^2 - 3\lambda - 1) - (7 - 3\lambda) + (-6 + 2\lambda) \\ &= -\lambda^3 + \lambda^2 + 8\lambda - 12 \end{aligned}$$

$$\begin{aligned} \det(A - \lambda I) = 0 &\Rightarrow \lambda^3 - \lambda^2 - 8\lambda + 12 = 0 \\ &\Rightarrow \lambda = 2, -3 \end{aligned}$$

Note. The values of λ obtained from $\det(A - \lambda I) = 0$ are called **eigenvalues** of the matrix A [See (7.13)]

(5.27) Theorem. For any $n \times n$ matrix $A = [a_{ij}]$, show that

$$\det A^T = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{i=1}^{n} a_{ij}A_{ij}$$

(The expansion given above is called the **expansion** of det A by the **cofactors** of the jth column of A.)

Proof. Since $\det A^T = \det A = \sum_{i=1}^{n} a_{ji}A_{ji}$

and the jth row of A becomes jth column of A^T, we have

$$\begin{aligned} \det A^T &= \sum_{i=1}^{n} a_{ji}A_{ji} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots a_{nj}A_{nj} \\ &= \sum_{i=1}^{n} a_{ij}A_{ij} \end{aligned}$$

EXERCISE 5.2

1. Solve for x each of the following equations

(i) $\begin{vmatrix} 1 & 2+x & 3 \\ 2 & 1 & 3+x \\ 3 & 2+x & 1 \end{vmatrix} = 0$

(ii) $\begin{vmatrix} 1 & 1 & 2 & 3 \\ 1 & 2-x^2 & 2 & 3 \\ 2 & 3 & 1 & 5 \\ 2 & 3 & 9 & 9-x^2 \end{vmatrix} = 0.$

(iii) $\begin{vmatrix} 1 & x & x^2 & x^3 \\ 1 & 2 & 2^2 & 2^3 \\ 1 & 3 & 3^2 & 3^3 \\ 1 & 4 & 4^2 & 4^3 \end{vmatrix} = 0$

(iv) $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1+x & 3 & 4 & 5 \\ 1 & 2 & x+1 & 4 & 5 \\ 1 & 2 & 3 & x+1 & 5 \\ 1 & 2 & 3 & 4 & x+1 \end{vmatrix} = 0$

(v) $\begin{vmatrix} x & a & a & a & a \\ a & x & a & a & a \\ a & a & x & a & a \\ a & a & a & x & a \\ a & a & a & a & x \end{vmatrix} = 0$

(vi) $\begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1+x & 1 & 1 \\ 1 & 1 & 1+x & 1 \\ 1 & 1 & 1 & 1+x \end{vmatrix} = 0.$

2. Evaluate each of the following $n \times n$ determinants:

(i) $\begin{vmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ b & b & b & \cdots & a \end{vmatrix}$

(ii) $\begin{vmatrix} 1-n & 1 & 1 & \cdots & 1 \\ 1 & 1-n & 1 & \cdots & 1 \\ 1 & 1 & 1-n & \cdots & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & 1 & \cdots & 1-n \end{vmatrix}$

(iii) $\begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 2 & 3 & 4 & \cdots & n & 1 \\ 3 & 4 & 5 & \cdots & 1 & 2 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ n & 1 & 2 & \cdots & n-2 & n-1 \end{vmatrix}$

(iv) $\begin{vmatrix} x+1 & x & \cdots & x \\ x & x+2 & \cdots & x \\ \vdots & \vdots & \cdots & \vdots \\ x & x & \cdots & x+n \end{vmatrix}$

3. If A and B are 3×3 matrices such that $\det(A^2B^2) = 108$ and $\det(A^3B^2) = 72$, find $\det(2A)$ and $\det(B^{-1})$.

4. Let A be an $n \times n$ matrix. Show that

(i) $\det A^m = (\det A)^m$ for any positive integer m.

(ii) if $\det A^m = 1$, then $\det A = \pm 1$.

(iii) if $\det A^m = 0$, then $\det A = 0$.

5. For any nonsingular matrix A, show that

(i) $\det(A^{-1}) = (\det A)^{-1}$

(ii) $\det(ABA^{-1}) = \det B$.

6. For what value of α is the matrix

$$A = \begin{bmatrix} -\alpha & \alpha-1 & \alpha+1 \\ 1 & 2 & 3 \\ 2-\alpha & \alpha+3 & \alpha+7 \end{bmatrix} \text{ singular?}$$

7. For

$$A = \begin{bmatrix} 1 & -1 & 3 \\ 4 & 1 & 6 \\ 2 & 0 & -2 \end{bmatrix}, \text{ verify that } A^{-1} = \frac{\operatorname{adj} A}{\det A}.$$

8. Let

$$P_5 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Show that $\det(P_5) = 1$.

Write P_n and evaluate $\det(P_n)$.

9. Find the eigenvalues of the given matrices:

(i) $\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$ (ii) $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

(iii) $\begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & -3 & 11 \\ 2 & -6 & 16 \\ 1 & -3 & 7 \end{bmatrix}$

⑤

Chapter 6

VECTOR SPACES

Vector spaces form a very important topic in both pure and applied mathematics Essentially it is a part of Linear Algebra In this chapter we discuss it and some of the concepts related to it Mainly we give a brief account of subspaces, their spanning sets, linear dependence and linear independence of subsets of a vector space, basis and dimension of a vector space, row and column spaces, solution space of a linear system of equations, linear transformations, rank and nullity of a linear transformation, vector space of linear transformations and matrix of a linear transformation

DEFINITION AND EXMAPLES

(6.1) Definition. Let F be a field and V a nonempty set on whose elements an operation of addition is defined Suppose that for every $a \in F$ and every $v \in V$, av is an element of V. Then V is called a **vector space** over F if the following axioms hold:

(i) V is an abelian group under addition

(ii) $a(v + w) = av + aw$

(iii) $(a + b)v = av + bv$

(iv) $a(bv) = (ab)v$

(v) $1v = v$, 1 being the multiplicative identity of F

for all $a, b \in F$; $v, w \in V$.

P^2, P^3 are [illegible] over the field.

The expression av, $a \in F$, $v \in V$ is called the scalar multiplication of v by a. The elements of F are called **scalars** and the elements of V are called **vectors** The vector space V over the field F is sometimes written as $V(F)$ If the field F, over which V is a vector space, is clearly understood from the context, then we simply write V instead of $V(F)$.

It may be noted that the scalar multiplication for a vector space V over the field F is a function $f : F \times V \to V$ defined by

$$f(a, v) = av, \quad \text{for all } a \in F, \ v \in V.$$

The additive identity of V – **the zero vector** – will be denoted by the bold faced $\mathbf{0}$ and the additive identity of F – **the zero scalar** – will be denoted by 0.

A vector space V is called a **real** or **complex** vector space according as F is the field of real or complex numbers respectively.

Example 1. The set $R^3 = \{(x, y, z) : x, y, z \in R\}$ is a vector space under addition and scalar multiplication defined by:

(i) $u + u' = (x, y, z) + (x', y', z') = (x + x', y + y', z + z'),$

for $u = (x, y, z), \ u' = (x', y', z') \in R^3.$

(ii) $au = (ax, ay, az), \quad a \in R.$

An element of R^3 is usually written as

$$x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \qquad (1)$$

where $\mathbf{i} = (1, 0, 0), \ \mathbf{j} = (0, 1, 0), \ \mathbf{k} = (0, 0, 1)$

are the unit vectors in the directions of the axes.

Physical quantities such as force, velocity, momentum etc., having both magnitude and direction are represented by (1). As such we call these quantities also as vectors.

Example 2. Let $(F, +, \cdot)$ be a field. By definition $(F, +)$ is an abelian group. For scalar multiplication, let $x \in F$. Then, for each $a \in F$, $ax \in F$ by the closure law in F and

$a(x + y) = ax + ay$	**(Distributive Law in F)**
$(a + b)x = ax + ay$	**(Distributive Law in F)**
$(ab)x = a(bx)$	**(Associative Law in F)**
$1 . x = x$	**(1 is Identity Element of F)**

for all $a, b, x, y \in F$.

Hence $(F, +, \cdot)$ is a vector space (over itself)

In particular the fields R, C and Q of real numbers, complex numbers and of rational numbers respectively are examples of vector spaces over themselves respectively.

Example 3. Let V be the set of all numbers of the form $a + b\sqrt{2}$, where $a, b \in Q$. Clearly, V is an additive abelian group. The identity of V is $0 + 0\sqrt{2}$ and the additive inverse of $a + b\sqrt{2}$ is $-a - b\sqrt{2}$. If $r \in Q$, then $r(a + b\sqrt{2}) = ra + rb\sqrt{2}$ is in V. Thus V is closed under scalar multiplication defined for members of Q. It is easy to verify that V satisfies the axioms (i) to (v) of Definition 6.1. Hence V is a vector space over Q.

Example 4. Let F be a field. Consider the set

$$F^n = \{(x_1, x_2, \cdots, x_n) : x_i \in F, \ 1 \le i \le n\}$$

Define addition and scalar multiplication in F^n as follows.

For $x = (x_1, x_2, \cdots, x_n), \quad y = (y_1, y_2, \cdots, y_n) \in F^n,$

we define

$$x + y = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n) \qquad (1)$$

and for $a \in F$, we put

$$ax = (ax_1, ax_2, \cdots, ax_n) \qquad (2)$$

Then $(F^n, +)$ is an abelian group under addition defined by (1) with $\mathbf{0} = (0, 0, \cdots, 0)$ as the additive identity and for each $x \in F^n$,

$$-x = (-x_1, -x_2, \cdots, -x_n)$$

is the additive inverse.

Also
$$\begin{aligned} a(x + y) &= a(x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n) \\ &= (a(x_1 + y_1), a(x_2 + y_2), \cdots, a(x_n + y_n)), \text{ by (2)} \\ &= (ax_1 + ay_1, ax_2 + ay_2, \cdots ax_n + ay_n), \\ &\qquad \text{by Distributive Law in } F \\ &= (ax_1, ax_2, \cdots, ax_n) + (ay_1, ay_2, \cdots, ay_n) \qquad \text{by (1)} \\ &= a(x_1, x_2, \cdots, x_n) + a(y_1, y_2, \cdots, y_n), \qquad \text{by (2)} \\ &= ax + ay \end{aligned}$$

for all $a \in F$ and $x, y \in F^n$. So (ii) is satisfied.

The conditions (iii) – (v) can be verified similarly.

As a special case when $F = R$, then we write R^n for F^n. So R^n is a vector space for each $n = 1, 2, \cdots$. We call R^n the **n-dimensional Euclidean space**.

We usually write vectors in R^n in the form of row vectors. If $x \in R^n$, we write $x = (x_1, x_2, \cdots, x_n)$ where $x_1, x_2, \cdots, x_n$ belong to R.

Example 5. Let M_{mn} denote the set of all $m \times n$ matrices with entries from the field of real numbers. Then M_{mn} is a vector space over R under the operations of matrix addition and scalar multiplication of matrices by real numbers. The $m \times n$ zero matrix is the zero vector of M_{mn}.

Example 6. Recall that a polynomial of degree n over a field F, in a variable x, is an expression of the form

$$p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n, \quad a_i \in F \tag{1}$$

Consider the set $P(x)$ of all polynomials of arbitrary degree. Elements of $P(x)$ are of the form (1). For $p(x), q(x) \in P(x)$,

$$q(x) = b_0 + b_1 x + b_2 x^2 + \dots + b_m x^m, \; b_i \in F,$$

we define addition and scalar multiplication in $P(x)$ by

$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a^m + b^m)x^m + a_{m+1}x^{m+1} + \dots + a_n x^n, \; (m < n) \tag{2}$$

and $\quad kp(x) = (ka_0) + (ka_1)x + (ka_2)x^2 + \dots + (ka_n)x^n \quad (3)$

for $\quad k \in F$

The zero polynomial, which is assigned no degree, is the additive identity of $P(x)$.

Then, under the addition and scalar multiplication defined by (2) and (3) respectively, $P(x)$ is a vector space over F.

In particular, the set $P_n(x)$ of all polynomials of degree $\le n$, n fixed, together with the zero polynomial, is also a vector space over F.

Example 7. Let X be a nonempty subset of R. Let V be the set of all functions $f: X \to R$. We define addition and scalar multiplication in V as follows:

For $\quad f, g \in V$ put

$$(f+g)(x) = f(x) + g(x) \tag{1}$$

and $\quad (af)(x) = a\, f(x) \quad (2)$

for all $x \in X$ and $a \in R$

(Note that for each $x \in X$ and $f \in V$, $f(x) \in R$)

We show that V is a vector space over R.

Let $f, g, h \in V$. Then

$((f+g)+h)(x) = (f+g)(x) + h(x)$, by (1)

$= (f(x) + g(x)) + h(x)$, by (1)

$= f(x) + (g(x) + h(x))$, by Associative Law in R

$= f(x) + (g+h)(x)$, by (1)

$= (f+(g+h))(x)$, by (1)

for all $x \in X$. Hence, by the definition of equality of two functions, we have

$(f+g)+h = f+(g+h)$, for all $f, g, h \in V$

The zero function $0: X \to R$ defined by

$$0(x) = 0 \quad \text{for all } x \in X \tag{3}$$

is the additive identity in V. Here, for any $f \in V$,

$(f+0)(x) = f(x) + 0(x)$, by (1)

$= f(x) + 0$, by (3)

$= f(x)$, since 0 is the additive identity of R

So $\quad f+0 = f$, for all $f \in V$

Similarly, $\; 0+f = f$

Next, for each $f \in V$, the function $(-f): X \to R$ defined by

$$(-f)(x) = -f(x), \quad \text{for all } x \in X \tag{4}$$

is the additive inverse of f in V. Here

$(f+(-f))(x) = f(x) + (-f)(x)$, by (1)

$= f(x) - f(x)$, by (4)

$= 0$

$= 0(x)$, by (3)

for all $x \in X$. Hence

$$f+(-f) = 0$$

Similarly, $\; (-f)+f = 0$

Moreover, for any $f, g \in V$,

$(f+g)(x) = f(x) + g(x)$, by (1)

$= g(x) + f(x)$ (Commutative Law in R)

$= (g+f)(x)$, by (1)

for all $x \in X$. Hence

$$f+g = g+f$$

So $(V, +)$ is an abelian group.

(ii) For the other axioms of a vector space for V, let $f, g \in V$, $a \in R$

Then, for all $x \in X$,

$$\begin{aligned}
(a(f+g))(x) &= a(f+g)(x), && \text{by (2)}\\
&= a(f(x)+g(x)), && \text{by (1)}\\
&= af(x) + ag(x) && \text{(Distributive Law in } R)\\
&= (af)(x) + (ag)(x), && \text{by (2)}\\
&= (af+ag)(x), && \text{by (1)}
\end{aligned}$$

and, therefore, $a(f+g) = af + ag$

(iii) For $a, b \in R$, $f \in V$ and all $x \in X$,

$$\begin{aligned}
((a+b)f)(x) &= (a+b)f(x), && \text{by (2)}\\
&= af(x) + bf(x) && \text{(Distributive Law in } R)\\
&= (af+bf)(x), && \text{by (1)}
\end{aligned}$$

Thus $(a+b)f = af + bf$

(iv) Again,

$$\begin{aligned}
(a(bf))(x) &= a(bf(x)), && \text{by (2)}\\
&= ab\,f(x) && \text{(Associative Law in } R)\\
&= (ab)f(x) && \text{(Associative Law in } R)\\
&= ((ab)f)x, && \text{by (2)}
\end{aligned}$$

for all $x \in X$, showing that $a(bf) = (ab)f$

(v) $(1f)(x) = 1f(x) = f(x)$

and so $1f = f$

Thus V is a vector space over R

(6.2) Theorem. Let V be a vector space over a field F. Then

(i) $a\mathbf{0} = \mathbf{0}$, for all $a \in F$

(ii) $0v = \mathbf{0}$, for all $v \in V$

(iii) $(-a)v = a(-v) = -av$, for all $a \in F$, $v \in V$

(iv) If $av = \mathbf{0}$ then either $a = 0$ or $v = \mathbf{0}$, for all $a \in F$, $v \in V$

(v) $a(u-v) = au - av$, for all $a \in F$, $u, v \in V$

Proof. The equation $a\mathbf{0} = \mathbf{0}$ should be read as "a (zero vector) $= \mathbf{0}$ vector" and likewise $0v = \mathbf{0}$ should be read as "(0 scalar) $v = \mathbf{0}$ vector".

(i) Since $\mathbf{0}$ is the additive identity of V, we have

$$\begin{aligned}
a\mathbf{0} &= a(\mathbf{0}+\mathbf{0})\\
&= a\mathbf{0} + a\mathbf{0}, && \text{by (ii) of (6.1)}
\end{aligned}$$

Since $a\mathbf{0} + \mathbf{0} = a\mathbf{0}$, we have, from $a\mathbf{0} + a\mathbf{0} = a\mathbf{0} + \mathbf{0}$ and by the cancellation law, $\mathbf{0} = a\mathbf{0}$

Thus, $a\mathbf{0} = \mathbf{0}$

(ii)

$$\begin{aligned}
0v &= (0+0)v\\
&= 0v + 0v, && \text{by (iii) of (6.1)}
\end{aligned}$$

Since $\mathbf{0}$ is the additive identity of V, it follows that

$$0v + \mathbf{0} = 0v + 0v$$

whence by the cancellation law, $\mathbf{0} = 0v$ or $0v = \mathbf{0}$

(iii) In this case

$$\begin{aligned}
\mathbf{0} &= 0v, && \text{by (ii) above}\\
&= (a+(-a))v, && \text{(0 is additive identity in } F)\\
&= av + (-a)v, && \text{by (iii) of (6.1)}
\end{aligned}$$

Adding $-av$ to both sides, and using the fact that $-av + av = \mathbf{0}$, we obtain

$$-av = (-a)v \tag{1}$$

Again,

$$\begin{aligned}
\mathbf{0} &= a\mathbf{0}\\
&= a(v+(-v))\\
&= av + a(-v), && \text{by (ii) of (6.1)}
\end{aligned}$$

As before, we get

$$-av = a(-v) \tag{2}$$

From (1) and (2) is follows that

$$(-av) = a(-v) = -av$$

(iv) Let $av = \mathbf{0}$ with $a \neq 0$. Then there exists an element $a^{-1} \in F$ such that $aa^{-1} = 1$. Now

$$\begin{aligned}
v &= 1v, && \text{by (v) of Definition 6.1}\\
&= (aa^{-1})v\\
&= a^{-1}(av), && \text{by (iv) of (6.1)}\\
&= a^{-1}\mathbf{0}\\
&= \mathbf{0}, && \text{by (i) above}
\end{aligned}$$

showing that v is the zero vector

Next, let $av = 0$ but $v \neq 0$. We have to show that $a = 0$.

Suppose $a \neq 0$. Then a^{-1} exists and

$$0 = a^{-1}(0) = a^{-1}(av) = (aa^{-1})v = 1v = v$$

contradicting that $v \neq 0$. Thus

$$a = 0$$

(v) Let $u, v \in V, a \in F$. Then

$$\begin{aligned} a(u - v) &= a(u + (-v)) \\ &= au + a(-v), && \text{by (ii) of (6.1)} \\ &= au + (-av), && \text{by (iii) above} \\ &= au - av, && \text{by (iii) above} \end{aligned}$$

SUBSPACES

(6.3) Definition. A nonempty subset W of a vector space V over a field F is said to be **a subspace** of V if W is itself a vector space over F under the same operations as defined in V.

Every vector space V has two subspaces, namely V itself and the set $\{0\}$ which consists of the 0 vector only.

These are called **trivial** subspaces of V. Any other subspace of V is called a **nontrivial** subspace.

Example 8. The space $P_n(x)$ of all polynomials of degree $\leq n$, together with the zero polynomial, is a subspace of the space $P(x)$ of all polynomials.

Example 9. Let V be the Euclidean space R^3. The set W consisting of those vectors whose third coordinate is zero, i.e., vectors of the form $(x, y, 0)$, $x, y \in R$ is a subspace of V.

(6.4) Theorem. Let V be a vector space over a field F and let W be a nonempty subset of V. Then W is a subspace of V if and only if

(i) $w_1, w_2 \in W \quad \Rightarrow w_1 + w_2 \in W$

(ii) $w \in W, a \in F \quad \Rightarrow aw \in W$.

In other words, W is a subspace of V, if and only if, it is closed under the operations of addition and scalar multiplication as defined in V.

Proof. Suppose that W is a nonempty subset of V satisfying (i) and (ii). Then we show that W is a vector space.

Let $w_1, w_2 \in W$. Then by **(ii)**, $(-1)w_2 \in W$, i.e., $-w_2 \in W$. Now $w_1, -w_2 \in W$ and so by **(i)**, $w_1 - w_2 \in W$. Therefore, W is a subgroup of the group V under addition. Since V is an abelian group, W is also an abelian group. Thus axiom (i) of (6.1) is satisfied for W. The axioms **(ii)** to **(v)** of a vectors space hold in W since they hold in the larger set V. Hence W is a vector space over F.

Conversely, if W is a subspace of V, then conditions **(i)** and **(ii)** hold by Definition 6.1.

The two necessary and sufficient conditions of Theorem 6.3 for W to be a subspace of V can be combined into one, which we state in the following corollary.

(6.5) Corollary. A nonempty subset W of V is a subspace of V if and only if $w_1, w_2 \in W$ and $a, b \in F$ imply $aw_1 + bw_2 \in W$.

Proof. Suppose that $w_1, w_2 \in W$ and $a, b \in F$ imply $aw_1 + bw_2 \in W$. Taking $a = b = 1$, we have $w_1 + w_2 \in W$. Taking $b = 0$, we get $aw_1 \in W$. Hence, by Theorem 6.4, W is a subspace of V.

Conversely, let W be a subspace of V. Then $aw_1, bw_2 \in W$ by (ii) of (6.4). Hence $aw_1 + bw_2 \in W$ by (i) of Theorem 6.4. Thus the condition is satisfied.

Example 10. Consider the homogeneous system $Ax = 0$ of n linear equations in n real unknown variables. The system has a nontrivial solution x^* if and only if the matrix A is singular (Theorem 4.12). So suppose that A is a singular $n \times n$ matrix. Then the system

$$Ax = 0$$

has an infinite number of solutions. Let S be set of all such solutions. The members of S are row vectors from R^n. Thus $S \subseteq R^n$.

If $x, y \in S$, then

$$Ax = 0, \quad Ay = 0$$

So $\quad A(x + y) = Ax + Ay = 0$

Hence $x + y \in S$.

Next, for any $c \in R$ and $x \in S$,

$$A(cx) = cAx = c \cdot 0 = 0$$

So $\quad cx \in S$.

Hence S is a subspace of R^n.

(6.6) Theorem. Let U and W be subspaces of a vector space V over a field F. Then $U \cap W$ is also a subspace of V.

Proof. Let $a, b \in F$ and $w_1, w_2 \in U \cap W$. Then $w_1, w_2 \in U$ and $w_1, w_2 \in W$. Since U and W are subspaces of V, by Corollary (6.5), $aw_1 + bw_2$ belongs to both U and W. Thus $aw_1 + bw_2 \in U \cap W$. Hence $U \cap W$ is a subspace of V.

Theorem 6.6 can be generalized as follows:

(6.7) Theorem. The intersection of any number of subspaces of a vector space V is a subspace of V.

Proof. Left as an exercise.

(6.8) Definition. (Sum of Two Subspaces). Let U and W be two subspaces of a vector space V. Define

$$U + W = \{u + w : u \in U \text{ and } w \in W\}.$$

Clearly $\quad U + W \subset V$.

We next show that $U + W$ is a subspace of V.

(6.9) Theorem. If U, W are subspaces of a vector space V, then $U + W$ is a subspace of V containing both U and W. Further, $U + W$ is the smallest subspace containing both U and W.

Proof. Let $v_1, v_2 \in U + W$. Then we can write

for subspace proofs

$$v_1 = u_1 + w_1,\ v_2 = u_2 + w_2, \text{ where } u_1, u_2 \in U \text{ and } w_1, w_2 \in W.$$

Hence $v_1 + v_2 = (u_1 + w_1) + (u_2 + w_2) = (u_1 + u_2) + (w_1 + w_2)$.

But since U, W are subspaces, so $u_1 + u_2 \in U$ and $w_1 + w_2 \in W$. Thus $v_1 + v_2 \in U + W$. Also $au_1 \in U$ and $aw_1 \in W$ for all $a \in F$. Therefore, $au_1 + aw_1 = a(u_1 + w_1) \in U + W$. Hence by Theorem 6.4, $U + W$ is a subspace of V.

Since $0 \in W$, $u + 0 \in U + W$ for every $u \in U$. Thus $U \subset U + W$. Similarly $W \subset U + W$. Thus U, W are both contained in $U + W$.

Let S be any subspace of V containing both U and W. Then, for every $u \in U$ and $w \in W$, we have $u \in S$ and $w \in S$ so that $u + w \in S$ and hence $U + W \subset S$. Thus $U + W$ is the smallest subspace containing both U and W.

(6.10) Definition. Let U, W be subspaces of a vector space V. If $U \cap W = \{0\}$, then $U + W$ is called the **direct sum** of U and W and is written as $U \oplus W$.

Thus a vector space V is called the direct sum of its subspaces U and W if

(i) $\quad V = U + W \quad$ and $\quad$ (ii) $U \cap W = \{0\}$.

vector x scalar = linear combination.

LINEAR COMBINATIONS AND SPANNING SETS

(6.11) Definition. Let V be a vector space over a field F and $v_1, v_2, \cdots, v_m \in V$. Any vector in V of the form

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m,$$

set of vector or scalar $\quad a_i \in F$ scalar

where $a_i \in F\ (1 \le i \le m)$, is called a **linear combination** (over F) of $v_1, v_2, \cdots, v_m$.

(6.12) Definition. Let S be a nonempty subset of a vector space V. The set of all linear combinations of elements of S is called the **linear span** of S and is denoted by $< S >$. (If S is infinite, take linear combinations of all finite subsets of S).

(6.13) Theorem. Let S be a nonempty set of vectors in a vector space V over a field F. Then $< S >$ is a subspace of V containing S and it is the smallest subspace of V containing S.

Proof. Let $u, v \in < S >$ and $a, b \in F$. Then there are vectors $u_1, u_2, \cdots, u_m; v_1, v_2, \cdots, v_n$ in S and scalars $a_1, a_2, \cdots, a_m; b_1, b_2, \cdots, b_n$ in F such that

$$u = a_1u_1 + a_2u_2 + \cdots + a_mu_m,\quad v = b_1v_1 + b_2v_2 + \cdots b_nv_n$$

So, for $a, b \in F$,

$$\begin{aligned} au + bv &= (aa_1)u_1 + (aa_2)u_2 + \cdots + (aa_m)u_m + (bb_1)v_1 \\ &\quad + (bb_2)u_2 + \cdots + (bb_n)v_n \end{aligned}$$

being a linear combination of elements of S, is in $< S >$. By Corollary 6.5, $< S >$ is a subspace of V.

Each vector v_i in S is a linear combination of $v_1, v_2, \cdots, v_k$ in S since

$$v_i = 0v_1 + 0v_2 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_k$$

Thus $< S >$ contains each of the vectors $v_1, v_2, \cdots, v_n$ belonging to S i.e.,

$S \subset < S >$.

If W is any other subspace of V containing S, then it contains all vectors of the form $\sum_{i=1}^{k} a_iv_i$, where $a_i \in F$, $v_i \in V$ and k is a natural number. Hence $< S > \subset W$. Thus $< S >$ is the smallest subspace of V containing S.

$< S >$ is said to be spanned (or generated) by S and S is called the spanning set for $< S >$.

(6.14) Corollary. If S, T are subsets of V then

$$S \subset T \text{ implies } < S > \subset < T >.$$

Proof. Left as an exercise.

(6.15) Definition. A vector space V is said to be **finite dimensional** if there is a finite subset S in V such that $V = < S >$

Example 11. In $V = R^3$, let S be the subset consisting of vectors

$$e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1), \text{ i.e., } S = \{e_1, e_2, e_3\}$$

Then $< S >$ contains all vectors of the form $a_1e_1 + a_2e_2 + a_3e_3$, where $a_1, a_2, a_3 \in R$

$$\begin{aligned} \text{Now} \quad a_1e_1 + a_2e_2 + a_3e_3 &= a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) \\ &= (a_1, a_2, a_3) \end{aligned}$$

So $< S > \subseteq R^3$.

But (a_1, a_2, a_3) is an arbitrary vector in R^3. So $R^3 \subseteq < S >$. Thus $< S > = R^3$.

Since S consists of a finite number of elements of R^3, by Definition 6.15, R^3 is finite dimensional (over R).

Example 12. Write the vector $v = (1, -2, 5) \in R^3$ as a linear combination of the vector $v_1 = (1, 1, 1)$, $v_2 = (1, 2, 3)$ and $v_3 = (2, -1, 1)$

Solution. We have to express v as

$$v = a_1v_1 + a_2v_2 + a_3v_3, \quad \text{for } a_1, a_2, a_3 \in R$$

$$\begin{aligned} \text{i.e.} \quad (1, -2, 5) &= a_1(1, 1, 1) + a_2(1, 2, 3) + a_3(2, -1, 1) \\ &= (a_1 + a_2 + 2a_3, a_1 + 2a_2 - a_3, a_1 + 3a_2 + a_3) \end{aligned}$$

$$\begin{aligned} \text{Therefore,} \quad a_1 + a_2 + 2a_3 &= 1 \\ a_1 + 2a_2 - a_3 &= -2 \\ a_1 + 3a_2 + a_3 &= 5 \end{aligned}$$

Solving the above system of equations by the methods of Chapter 4, we obtain

$$a_1 = -6, \quad a_2 = 3, \quad a_3 = 2$$

Hence $v = (1, -2, 5) = -6v_1 + 3v_2 + 2v_3$ is the desired linear combination.

Example 13. Find equations defining the subspace W of R^3 spanned by the vector $(2, 3, 4)$

Solution. Every vector $v = (x, y, z)$ of W is a linear combination of $(2, 3, 4)$. Therefore,

$$(x, y, z) = t(2, 3, 4), \quad t \in R$$

$$\text{or} \quad x = 2t, \quad y = 3t, \quad z = 4t$$

which are parametric equations of a straight line through $(0, 0, 0)$ with direction vector $(2, 3, 4)$. This straight line is the subspace W spanned by $(2, 3, 4)$

Example 14. Find an equation defining the subspace W of R^3 spanned by

$$v_1 = (1, -3, 2), \quad v_2 = (-2, 1, 2), \quad v_3 = (-3, -1, 6)$$

by expressing an arbitrary elements $(x, y, z) \in R^3$ as a linear combination of v_1, v_2, v_3

Solution. Every vector $v = (x, y, z)$ of $W \subset R^3$ is a linear combination of v_1, v_2, v_3. Hence there are real numbers a, b, c such that

$$\begin{aligned} (x, y, z) &= av_1 + bv_2 + cv_3 \\ &= a(1, -3, 2) + b(-2, 1, 2) + c(-3, -1, 6) \\ &= (a - 2b - 3c, -3a + b - c, 2a + 2b + 6c) \end{aligned}$$

i.e., the system of equations in a, b, c

$$\begin{aligned} x &= a - 2b - 3c \quad &(1) \\ y &= -3a + b - c \quad &(2) \\ z &= 2a + 2b + 6c \quad &(3) \end{aligned}$$

is consistent.

i.e., the equivalent system

$$\begin{aligned} a - 2b - 3c &= x \\ b + 2c &= x + y + z, \quad \text{(adding the above three equations)} \\ 6(b + 2c) &= z - 2x, \quad \text{(subtracting 2 times the equation (1) from (3))} \end{aligned}$$

is consistent. But then the last two equations give

$$\begin{aligned} 6(x + y + z) &= z - 2x \\ \text{i.e.} \quad 8x + 6y + 5z &= 0 \end{aligned}$$

which is an equation of the plane through $(0, 0, 0)$. Thus

$$W = \{(x, y, z): 8x + 6y + 5z = 0\}.$$

Example 15. Let $V = P(x)$ be the vector space of all polynomials over a field (Example 6). $S = \{1, x, x^2, x^3, \cdots\}$, then $< S > = V$ because any polynomial is a linear combination of 1 and powers of x. Since S is not a finite subset of V, we infer that V is not finite dimensional. It is easy to see that a finite number of elements in S cannot generate V.

A vector space which is not finite dimensional is called **infinite dimensional**. Thus $P(x)$ is an infinite dimensional vector space.

Example 16. Show that the set

$$S = \{(x, y, z) \in R^3 : 2x + 3y - 4z = 2\}$$

is not a subspace of R^3.

Solution. The point $(1, 0, 0) \in S$. But $3(1, 0, 0) = (3, 0, 0)$ does not belong to S. Thus (ii) of (6.4) is not satisfied and so S is not a subspace of R^3.

EXERCISE 6.1

1. Let V be the set of all infinite sequences in a field F with addition and scalar multiplication defined as below:

For
$$u = \{a_n\} = a_1, a_2, a_3, \cdots, a_n, \cdots \in V$$
$$v = \{b_n\} = b_1, b_2, b_3, \cdots, b_n, \cdots \in V,$$
$$u + v = \{a_n\} + \{b_n\} = a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n, \cdots$$
and
$$ku = k\{a_n\} = ka_1, ka_2, ka_3, \cdots, ka_n, \cdots$$

where a_n, b_n and k are all in F, $n = 1, 2, 3, \cdots$

Show that V is a vector space over F.

2. Let V be the set of all ordered pairs of real numbers. Check whether V is a vector space over R with respect to the indicated operations. If not, state the axioms which fail to hold.

(i) $(a, b) + (c, d) = (a + c, b + d)$ and $k(a, b) = (ka, b)$

(ii) $(a, b) + (c, d) = (a, b)$ and $k(a, b) = (ka, kb)$

(iii) $(a, b) + (c, d) = (a + c, b + d)$ and $k(a, b) = (k^2a, k^2b)$

(iv) $(a, b) + (c, d) = (a + c, b + d)$ and $k(a, b) = (ka, 0)$

3. Check whether each of the following is a real vector space.

(i) The set $C[a, b]$ of all continuous real-valued functions defined on $[a, b]$ with the usual operations on functions as defined in Example 7.

(ii) The set of all functions $f \in C[a, b]$ such that $f(a) = f(b)$

(iii) The set of all solutions of the differential equation

$$\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 6y = 0$$

(iv) The set of all 2×2 real matrices of the form $\begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix}$

4. Check whether each of the following subsets is a subspace of the indicted vector space:

(i) Q the set of rational numbers in R

(ii) All 2×2 nonsingular real matrices in M_{22}

(iii) The set $B[a, b]$ of all bounded real functions defined on $[a, b]$ in the space of all real functions defined on $[a, b]$

5. Show that the union of two subspaces of a vector space need not be a subspace. Let X and Y be subspaces of a vector space V. Prove that $X \cup Y$ is a subspace of V if and only if either $X \subset Y$ or $Y \subset X$.

6. Which of the following are subspaces of R^3?

(i) $W = \{(x, y, z) : x + y + z = 0\}$

(ii) $W = \{(x, y, z) : x \geq 0\}$

(iii) $W = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$

(iv) $W = \{(x, y, z) : x, y, z \text{ are rationals}\}$

(v) $W = \{(x, 0, z) : x, y, z \in R\}$

(vi) $W = \{(x, y, z) : x, y, z \in R, y^2 = x + z\}$

(vii) $W = \{(x, y, z) : x, y, z \in R, 2x - 3y - 4z = 0\}$

7. Let V be the vector space of all real-valued functions defined on R. State which of the following are subspaces of V.

(i) The set of all even functions

(ii) The set of all differentiable functions

(iii) The set $W = \{f : f(x) = kf(-x), k \in R \text{ fixed}\}$

(iv) The set $W = \left\{f \in V : \int_0^1 f(x)\,dx = 0\right\}$

8. Let V be the vector space of all real polynomials of degree $\le n$ together with the zero polynomial. Determine whether or not W is a subspace of V, where W consists of the zero polynomial and all polynomials

(i) with integral coefficients and of degree $\le n$

(ii) of degree ≤ 3

(iii) with only even powers of x and of degree $\le n$.

9. Express the vector $(2, -5, 3)$ in R^3 as a linear combination of the vectors $(1, -3, 2), (2, -4, -1)$ and $(1, -5, 7)$

10. For what value of k will the vector $(1, -2, k)$ in R^3 be a linear combination of the vectors $(3, 0, -2)$ and $(2, -1, -5)$?

11. Let U and W be the subspaces of R^3 defined by

$$U = \{(x, y, z) : x = y = z\} \quad \text{and}$$
$$W = \{(0, y, z) : y, z \in R\}$$

Show that $R^3 = U \oplus W$

12. Show that each of the following sets of vectors generates R^3

(i) $\{(1, 2, 3), (0, 1, 2), (0, 0, 1)\}$

(ii) $\{(1, 1, 1), (0, 1, 1), (0, 1, -1)\}$

13. Determine whether the set $S = \{(1, 1, 2), (1, 0, 1), (2, 1, 3)\}$ spans R^3

14. Show that the yz-plane

$$W = \{(0, y, z) : y, z \in R\} \quad \text{is spanned by}$$

(i) $(0, 1, 1)$ and $(0, 2, -1)$

(ii) $(0, 1, 2), (0, 2, 3)$ and $(0, 3, 1)$

15. Find an equation (or equations) of the subspace W of R^3 spanned by each of the following sets of vectors

(i) $\{(1, -3, 5), (-2, 6, -10)\}$

(ii) $\{(1, -3, 2), (-2, 0, 3)\}$

(iii) $\{(1, -2, 1), (-2, 0, 3), (3, -2, -2)\}$

16. Show that the complex numbers $2 + 3i$ and $1 - 2i$ generate the vector space C over R

17. Let S and T be subsets of a vector space V. Show that

(i) $<S> \cup <T> \subset <S \cup T>$

(ii) $<S \cap T> \subset <S> \cap <T>$

Give an example to show that set equality need not hold in either case

LINEAR DEPENDENCE AND BASIS

(6.16) Definition. Let V be a vector space over a field F. The vectors $v_1, v_2, \cdots, v_m \in V$ are said to be **linearly dependent** over F if

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0$$

and not all a_i are zero, $a_i \in F$, $i = 1, 2, \cdots, m$

(6.17) Definition. If the vectors $v_1, v_2, \cdots, v_m \in V$ are not linearly dependent then they are said to be **linearly independent** over F. That is $v_1, v_2, \cdots, v_m$ are linearly independent over F if

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0 \quad \text{implies each } a_i = 0, i = 1, 2, \ldots m$$

In this case we say that $\{v_1, v_2, \cdots, v_m\}$ is a linearly independent set.

An infinite set S of vectors is linearly independent if each of its finite subsets is linearly independent. Otherwise it is linearly dependent.

The empty set $\emptyset$ is defined to be linearly independent.

It may be noted that if one of the vectors $v_1, v_2, \cdots, v_m$ is zero, say $v_i = 0$, then the vectors are linearly dependent, because

$$0v_1 + 0v_2 + \cdots + 1v_i + 0v_{i+1} + \cdots + 0v_m = 0$$

and the coefficient of v_i is not zero

On the other hand, any nonzero vector v, by itself, is linearly independent, for

$$av = 0 \quad \text{and} \quad v \ne 0 \text{ imply } a = 0 \text{ by (6.2) (iv)}$$

Example 17. Show that the vectors $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ belonging to R^3 are linearly independent over R

Solution. $a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) = (0, 0, 0)$ implies

$$(a, b, c) = (0, 0, 0)$$

i.e., $a = 0, b = 0, c = 0$

Thus the given vectors are linearly independent over R

Example 18. Show that the vectors $(3, 0, -3), (-1, 1, 2), (4, 2, -2)$ and $(2, 1, 1)$ are linearly dependent over R

Solution. Let

$$a(3, 0, -3) + b(-1, 1, 2) + c(4, 2, -2) + d(2, 1, 1) = (0, 0, 0)$$

where $a, b, c, d \in R$

i.e., $(3a - b + 4c + 2d, b + 2c + d, -3a + 2b - 2c + d) = (0, 0, 0)$

Therefore,
$$3a - b + 4c + 2d = 0 \quad (1)$$
$$b + 2c + d = 0 \quad (2)$$
$$-3a + 2b - 2c + d = 0 \quad (3)$$

Subtracting equation (2) from (3) we have
$$-3a + b - 4c = 0 \quad (4)$$

By adding (1) and (4), we get $d = 0$. So (2) becomes
$$b + 2c = 0. \qquad \text{That is } b = -2c. \quad (5)$$

Equations (4) and (5) with $d = 0$ give $-3a = 6c$ so that
$$a = b = -2c,\ d = 0$$

So $-2(3, 0, -3) - 2(-1, 1, 2) + 1(4, 2, -2) + 0(2, 1, 1) = (0, 0, 0)$

Therefore the given vectors are linearly dependent.

Example 19. Let $V = P_3(x)$ be the vector space of polynomials of degree ≤ 3 together with the zero polynomial over R. Determine whether the vectors
$$x^3 - 3x^2 + 5x + 1,\ x^3 - x^2 + 8x + 2 \quad \text{and} \quad 2x^3 - 4x^2 + 9x + 5$$
belonging to $P_3(x)$ are linearly independent.

Solution. Suppose that for $a, b, c \in R$,
$$a(x^3 - 3x^2 + 5x + 1) + b(x^3 - x^2 + 8x + 2) + c(2x^3 - 4x^2 + 9x + 5) = 0$$

i.e., $$(a + b + 2c)\,x^3 + (-3a - b - 4c)x^2 + (5a + 9b + 9c)\,x + (a + 2b + 5c) = 0 \quad (1)$$

Since, $1, x, x^2, x^3$ are linearly independent, (1) implies
$$\left.\begin{aligned} a + b + 2c &= 0 \\ -3a - b - 4c &= 0 \\ 5a + 8b + 9c &= 0 \\ a + 2b + 5c &= 0 \end{aligned}\right\} \quad (2)$$

The system (2) has a nontrivial solution if rank of the matrix
$$A = \begin{bmatrix} 1 & 1 & 2 \\ -3 & -1 & -4 \\ 5 & 8 & 9 \\ 1 & 2 & 5 \end{bmatrix}$$
is less than 3 (Theorem 4.12). It is easy to see that Rank $A = 3$. So the system (2) cannot have a nontrivial solution and therefore, it has the unique trivial solution $a = 0 = b = c$.

Hence the given vectors are linearly independent.

Example 20. What conditions must a, b, c and d satisfy so that the matrices
$$\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ -2 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
in M_2 are linearly dependent ?

Solution. Since the matrices are to be linearly dependent, there must exist scalars c_1, c_2, c_3, not all zero, such that
$$c_1\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 2 & 3 \\ -2 & 1 \end{bmatrix} + c_3\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (1)$$

This implies
$$\begin{bmatrix} c_1 + 2c_2 + ac_3 & 2c_1 + 3c_2 + bc_3 \\ -c_1 - 2c_2 + cc_3 & c_2 + dc_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

i.e., the system of homogeneous equations
$$\left.\begin{aligned} c_1 + 2c_2 + ac_3 &= 0 \\ 2c_1 + 3c_2 + bc_3 &= 0 \\ -c_1 - 2c_2 + cc_3 &= 0 \\ c_2 + dc_3 &= 0 \end{aligned}\right\} \quad (2)$$
must have a nontrivial solution.

By Theorem 4.12, the system (2) will have a nontrivial solution if rank of the matrix
$$A = \begin{bmatrix} 1 & 2 & a \\ 2 & 3 & b \\ -1 & -2 & c \\ 0 & 1 & d \end{bmatrix}$$
is less than 3 i.e., Rank $A = 2$ at the most.

N

$$A \underset{R}{\sim} \begin{bmatrix} 1 & 2 & a \\ 0 & -1 & b-2a \\ 0 & 0 & c+a \\ 0 & 1 & d \end{bmatrix} \quad \text{by} \quad \begin{matrix} R_2 - 2R_1 \\ \text{and} \quad R_3 - R_1 \end{matrix}$$

$$\underset{R}{\sim} \begin{bmatrix} 1 & 2 & a \\ 0 & -1 & b-2a \\ 0 & 0 & c+a \\ 0 & 0 & b+d-2a \end{bmatrix} \quad \text{by} \quad R_4 + R_2$$

The last matrix has rank 2 at the most only if

$$c + a = 0 \quad \text{and} \quad b + d - 2a = 0$$

Thus the required conditions on a, b, c, d are

$$a + c = 0 \quad \text{and} \quad b + d - 2a = 0$$

Note that, for $a = b = d = 1$, $c = -1$ and $c_1 = 1$, $c_2 = -1$, $c_3 = 1$, the equation (…) satisfied.

Example 21. Let V be the vector space of all functions defined on R to R. C… whether the vectors $2, 4\sin^2 x, \cos^2 x$ are linearly independent in V.

Solution. A linear combination of the given vectors is

$$2a + 2b\sin^2 x + c\cos^2 x, \qquad \text{where } a, b, c \in R$$

Now let $\quad 2a + 4b\sin^2 x + c\cos^2 x = \theta$

This equation is true for all $x \in R$

Set $x = 0, \frac{\pi}{4}, \frac{\pi}{2}$ in (1) to obtain the equations

$$\left.\begin{aligned} 2a \qquad + c &= 0 \\ 2a + 2b + \frac{c}{2} &= 0 \\ 2a + 4b \qquad &= 0 \end{aligned}\right\}$$

From these equations, we have

$$a = -\frac{c}{2}, \qquad b = \frac{c}{4} \qquad \text{and} \qquad c = c.$$

Thus the system (2) has a nontrivial solution $\left(-\frac{c}{2}, \frac{c}{4}, c\right)$ and so the given vectors are linearly dependent.

(6.18) Theorem. Let V be a vector space over a field F and

$$S = \{v_1, v_2, \ldots, v_m\} \text{ be a set of vectors in } V. \text{ Then}$$

(i) If S is linearly independent, then any subset of S is also linearly independent.

(ii) If S is linearly dependent, then the set $\{v, v_1, v_2, \ldots, v_m\}$ is linearly dependent for all $v \in V$, i.e., every superset of S is also linearly dependent.

Proof. (i) Let $\{v_1, v_2, \ldots, v_i\}$ $(1 \le i < m)$ be a subset of S and suppose that

$$a_1v_1 + a_2v_2 + \ldots + a_iv_i = \theta, \qquad a_1, a_2, \ldots, a_i \in F$$

Then $$a_1v_1 + a_2v_2 + \ldots + a_iv_i + 0\,v_{i+1} + \ldots + 0\,v_m = \theta \tag{1}$$

Since $S = \{v_1, v_2, \ldots, v_i, v_{i+1}, \ldots, v_m\}$ is linearly independent, from (1) it follows that

$$a_1 = a_2 = a_3 = \ldots = a_i = 0$$

Thus if $a_1v_1 + a_2v_2 + \ldots + a_iv_i = \theta$, then all a_i's are zero and hence the subset $\{v_1, v_2, \ldots, v_i\}$ is linearly independent.

(ii) Suppose that $a_1v_1 + a_2v_2 + \ldots + a_mv_m = \theta$ with $a_i \neq 0$, where i is some integer such that $1 \le i \le m$. Then

$$a_1v_1 + a_2v_2 + \ldots + a_iv_i + \ldots + a_mv_m + 0\,v = \theta$$

and since $a_i \neq 0$, it follows that the set $\{v, v_1, v_2, \ldots, v_m\}$ is linearly dependent.

(6.19) Theorem. (i) A set $S = \{v_1, v_2, \ldots, v_n\}$ of n vectors $(n \ge 2)$ in a vector space V is linearly dependent if and only if at least one of the vectors in S is a linear combination of the remaining vectors of the set. $a_1v_1 = -k(v_1 + a_2v_2 + \ldots a_{n-1}v_{n-1})$ $v_i = -\frac{1}{k}(\ldots)$

(ii) A set $S = \{v_1, v_2, \ldots, v_n\}$ of vectors, $n \ge 2$, in a vector space V is linearly dependent if and only if one of the vectors in S is a linear combination of the vectors preceding it. $\underbrace{a_1v_1 + \ldots + a_kv_k}_{\text{non-zero}} + 0v_{k+1} + \ldots 0v_t = 0$

Proof. (i) Suppose the set

$$S = \{v_1, v_2, \ldots, v_n\}$$

is linearly dependent. Then there are scalars $a_1, a_2, \ldots, a_n$, at least one of them, say a_i, is nonzero, such that

$$a_1v_1 + a_2v_2 + \ldots + a_iv_i + \ldots + a_nv_n = \theta$$

$0 = v_j = a_1v_1 + a_2v_2 + \dots + (-1)v_j + a_{j+1}v_{j+1} + \dots + a_nv_n$

$a_1v_1 + a_2v_2 + \dots$

Dividing through by a_i and transposing, we have

$$v_i = -\frac{1}{a_i}(a_1v_1 + a_2v_2 + \dots + a_{i-1}v_{i-1} + a_{i+1}v_{i+1} + \dots + a_nv_n)$$

expressing v_i as a linear combination of the remaining vectors of the set.

Conversely, let some vector v_j of the given set be a linear combination of the remaining vectors, i.e.,

$$v_j = a_1v_1 + a_2v_2 + \dots + a_{j-1}v_{j-1} + a_{j+1}v_{j+1} + \dots + a_nv_n$$

Then $a_1v_1 + a_2v_2 + \dots + a_{j-1}v_{j-1} - v_j + a_{j+1}v_{j+1} + \dots + a_nv_n = \theta$ and there is at least one coefficient, namely -1 of v_j, which is nonzero and therefore, by definition, the set

$$\{v_1, v_2, \dots, v_{j-1}, v_j, \dots, v_n\}$$

is linearly dependent.

(ii) Suppose that the set $S = \{v_1, v_2, \dots, v_n\}$ is linearly dependent. Then there are scalars $a_1, a_2, \dots, a_n$, not all zero, such that

$$a_1v_1 + a_2v_2 + \dots + a_nv_n = \theta. \quad (1)$$

Let a_k be the last nonzero scalar in (1). Then the terms $a_{k+1}v_{k+1}, a_{k+2}v_{k+2}, \dots, a_nv_n$ are all zero. So (1) can be written as

$$a_1v_1 + a_2v_2 + \dots + a_kv_k = \theta, \quad a_k \neq 0,$$

so that

$$v_k = -\frac{1}{a_k}(a_1v_1 + a_2v_2 + \dots + a_{k-1}v_{k-1})$$

Hence v_k is a linear combination of the vectors preceding it.

Conversely, suppose that in $S = \{v_1, v_2, \dots, v_n\}$ some one of the vectors, v_k say, is a linear combination of the vectors preceding it. That is

$$v_k = b_1v_1 + b_2v_2 + \dots + b_{k-1}v_{k-1}$$

Then $\quad b_1v_1 + b_2v_2 + \dots + b_{k-1}v_{k-1} + (-1)v_k = \theta$,

so that

$$b_1v_1 + b_2v_2 + \dots + b_{k-1}v_{k-1} + (-1)v_k + 0\,v_{k+1} + \dots + 0\,v_n = \theta \quad (2)$$

with not all coefficients in (2) zero. Hence S is linearly dependent.

(6.20) Definition. A set of linearly independent vectors spanning a vector space V is called a **basis** for V.

Unless $V = \{\theta\}$, the basis vectors are nonzero, since we can ignore any zero

Example 22. The vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ span R^3 and since these vectors are linearly independent, it follows that the set $\{e_1, e_2, e_3\}$ forms a basis of R^3. [$\{e_1, e_2, e_3\}$ is called the **standard basis for** R^3]

Example 23. Show that the 2×2 matrices

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

form a basis for the vector space V of all 2×2 symmetric matrices.

Solution. Recall that a 2×2 symmetric matrix is of the form

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

$\{E_1, E_2, E_3\}$ is a basis of the vector space V of all 2×2 symmetric matrices if and only if every symmetric matrix A is a linear combination of E_1, E_2, E_3. But

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and E_1, E_2, E_3 are linearly independent. (Verify!)

So $\{E_1, E_2, E_3\}$ is a basis of V.

Can this result be generalized to:

(i) the space of 3×3 symmetric matrices,

(ii) the space of 4×4 symmetric matrices,

(iii) to the space of $n \times n$ symmetric matrices?

(6.21) Theorem. Any finite dimensional vector space contains a basis.

Proof. Let $v_1, v_2, \dots, v_r$ be a finite spanning set for the vectors space V assumed not to be $\{\theta\}$. If $v_1, v_2, \dots, v_r$ are linearly independent, they form a basis for V and there is nothing to prove.

Suppose that $v_1, v_2, \dots, v_r$ are not linearly independent. Then these vectors are linearly dependent. Therefore, by Theorem **6.19**, one of the vectors say v_r is a linear combination of the remaining ones. We cast out this vector from the set and, re-numbering if necessary, we obtain a set of $r - 1$ vectors $v_1, v_2, \dots, v_{r-1}$. Clearly any linear combination of $v_1, v_2, \dots, v_r$ is also a linear combination of $\{v_1, v_2, \dots, v_{r-1}\}$. Thus the set $\{v_1, v_2, \dots, v_{r-1}\}$ is also a spanning set for V. Continuing in this way we arrive at a linearly independent spanning set $\{v_1, v_2, \dots, v_n\}$ $(1 \le n < r)$ and this forms a basis for V. Thus every finite dimensional vector space contains a basis.

(6.22) Definition. The dimension of a vector space V over a field F, written dim V, is the number of elements in a basis for V.

It follows from Theorem 6.21 that the dimension of a vector space V cannot exceed the number of elements in a spanning set for V.

(6.23) Theorem. All bases of a finite dimensional vector space contain the same number of elements.

Proof. Suppose that the vector space V has two bases A and B with m and n elements respectively. Since A spans V and B is linearly independent, it follows from the previous ~~remark that~~ the number of elements in B cannot exceed the number of elements in A. Thus $n \geq m$. By interchanging the roles of A and B, we have $m \leq n$. Hence $n = m$.

Example 24. Find a basis for the plane $y - z = 0$ considering it as a subspace of R^3.

Solution. The general solution of the equation $y - z = 0$ is $y = z$. In vector form the general solution is

$$(x, y, z) = (x, y, y) = x\,(1, 0, 0) + y\,(0, 1, 1)$$

Thus the plane $y - z = 0$ is spanned by the vectors $(1, 0, 0)$ and $(0, 1, 1)$. The dimension of the given subspace is two, since the vectors $(1, 0, 0)$ and $(0, 1, 1)$ form a basis for the given plane. [The vectors $(1, 0, 0)$, $(0, 1, 1)$ are linearly independent.]

(6.24) Theorem. Let V be a vector space such that dim $V = n < \infty$. A set of vectors

$$\{v_1, \ v_2, \ \ldots, \ v_n\} \subset V$$

is a basis for V if and only if each vector in V is uniquely expressible as a linear combination of $v_1, \ v_2, \ \ldots, \ v_n$.

Proof. Let the set $\{v_1, \ v_2, \ \ldots, \ v_n\}$ be a basis for V. Then every vector v in V can be expressed in at least one way as a linear combination of the basis vectors i.e.,

$$v = a_1v_1 + a_2v_2 + \ldots + a_nv_n \tag{1}$$

where a_i, $1 \leq i \leq n$, are scalars.

Suppose v is also expressible as

$$v = b_1v_1 + b_2v_2 + \ldots + b_nv_n \tag{2}$$

$b_1, b_2, \ldots, b_n$ being scalars.

Subtracting equation (2) from (1), we get

$$0 = (a_1 - b_1)v_1 + (a_2 - b_2)\,v_2 + \ldots + (a_n - b_n)\,v_n$$

Since $v_1, v_2, \ldots, v_n$ are linearly independent, we have

$$a_1 - b_1 = a_2 - b_2 = \ldots = a_n - b_n = 0$$

i.e., $\quad a_i = b_i, \quad 1 \leq i \leq n$

Thus every vector v in V can be expressed in a unique way as a linear combination of the basis vectors.

Conversely, let every vector of V be uniquely expressible as a linear combination of $v_1, v_2, \ldots, v_n$. Then these vectors span V. We prove that they are linearly independent. Suppose that, for scalars $a_1, a_2, \ldots, a_n$

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$$

Since

$$0v_1 + 0v_2 + \cdots + 0v_n = 0,$$

the uniqueness of representation (1) implies

$$0 = a_1 = a_2 = \cdots = a_n$$

Thus $v_1, v_2, \ldots, v_n$ are linearly independent. Since $v_1, v_2, \cdots, v_n$ span V and are linear independent, they form a basis for V.

(6.25) Theorem. Let $S = \{v_1, v_2, \cdots, v_n\}$ be a basis for an n-dimensional vector space V over a field F. Then every set with more than n vectors is linearly dependent.

Proof. Let $B = \{u_1, u_2, \cdots, u_r\}$ be any set of r vectors in V where $r > n$. We shall show that the set B is linearly dependent.

In order to show that the set B is linearly dependent we must find scalars $c_1, c_2, \ldots$, not all zero, such that

$$c_1u_1 + c_2u_2 + \cdots + c_ru_r = 0 \tag{1}$$

Since the set $S = \{v_1, v_2, \cdots, v_n\}$ is linearly independent, by (6.24) each u_i can be expressed in a unique way as a linear combination of $v_1, v_2, \cdots, v_n$. Thus

$$\left.\begin{array}{rcl} u_1 &=& a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ u_2 &=& a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots & & \vdots \qquad\quad \vdots \qquad \cdots \quad \vdots \\ u_r &=& a_{r1}v_1 + a_{r2}v_2 + \cdots + a_{rn}v_n \end{array}\right\} \tag{2}$$

where a_{ij} are scalars $(1 \leq i \leq r, 1 \leq j \leq n)$.

Substituting the values of u_i, $i = 1, 2, \cdots, r$, from (2) into (1), we have

$$c_1\,(a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n) + c_2\,(a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n) + \cdots + c_r\,(a_{r1}v_1 + a_{r2}v_2 + \cdots + a_{rn}v_n) = 0$$

i.e.,

$$(c_1a_{11} + c_2a_{21} + \cdots + c_ra_{r1})\,v_1 + (c_1a_{12} + c_2a_{22} + \cdots + c_ra_{r2})\,v_2 + \cdots + (c_1a_{1n} + c_2a_{2n} + \cdots + c_ra_{rn})\,v_n = 0 \tag{3}$$

Since $v_1, v_2, \ldots, v_n$ are linearly independent, from (3) we have

$$a_{11}c_1 + a_{21}c_2 + \cdots + a_{r1}c_r = 0$$
$$a_{12}c_1 + a_{22}c_2 + \cdots + a_{r2}c_r = 0$$
$$\vdots \qquad \vdots \qquad \cdots \qquad \vdots \qquad \vdots$$
$$a_{1n}c_1 + a_{2n}c_2 + \cdots + a_{rn}c_r = 0$$

$m < n$
$n < r \Rightarrow r > n$.

which is a homogeneous system of n equations in r unknowns $c_1, c_2, \ldots, c_r$. Since $n <$ … this system has a nontrivial solution. [See (4.11)]. Thus, at least one of the c_i's is not zero and so the set B is linearly dependent.

(6.26) Theorem. Let $v_1, v_2, \ldots, v_n$ be linearly independent in a vector space V over a field F. If v is any nonzero vector in V, then the set $\{v_1, v_2, \ldots, v_n, v\}$ is linearly independent if and only if v is not in the linear span $< v_1, v_2, \ldots, v_n >$.

Proof. Suppose that v is not in the linear span $< v_1, v_2, \ldots, v_n >$.

Consider any equation of the form

$$a_1v_1 + a_2v_2 + \ldots + a_nv_n + a\,v = 0, \tag{1}$$

where $a_1, a_2, \ldots, a_n, a$ are scalars. If $a \neq 0$, then

$$v = -\frac{1}{a}(a_1v_1 + a_2v_2 + \ldots + a_nv_n)$$

showing that $v \in < v_1, v_2, \ldots, v_n >$, contrary to the supposition unless $a = 0$. But since $v_1, v_2, \ldots, v_n$ are linearly independent, equation (1), with $a = 0$, implies $a_i = 0$, $1 \le i \le n$. Thus $v_1, v_2, \ldots, v_n, v$ are linearly independent.

Conversely, let $v_1, v_2, \ldots, v_n, v$ be linearly independent but

$$v \in < v_1, v_2, \ldots, v_n >$$

Then, for some scalars $a_1, a_2, \ldots, a_n$, v can be expressed as a linear combination of $v_1, v_2, \ldots, v_n$

i.e., $$v = a_1v_1 + a_2v_2 + \ldots + a_nv_n$$

So, $$a_1v_1 + a_2v_2 + \ldots + a_n v_n + (-1)\,v = 0.$$

That is, $v, v_1, v_2, \ldots, v_n$ are linearly dependent, contrary to our assumption.

Hence $$v \notin < v_1, v_2, \ldots, v_n >.$$

(6.27) Theorem. Let V be an n dimensional vector space over a field F. Then every $S = \{v_1, v_2, \ldots, v_n\}$ of n linearly independent vectors in V is a basis for V.

Proof. Each of the vectors $v_1, v_2, \ldots, v_n$ is a linear combination of these n vectors, since every vector is equal to itself. Let $v \in V$ be any vector. The set of vectors

$$\{ v_1, v_2, \ldots, v_n, v \}$$

is linearly dependent, by Theorem 6.25. Therefore, $a_1v_1 + a_2v_2 + \ldots + a_nv_n + av$ … implies at least one of the scalars $a_1, a_2, \ldots, a_n, a$ belonging to F is not zero. But $a \neq$ … since $v_1, v_2, \ldots, v_n$ are linearly independent. Therefore, v can be expressed as a linear combination of the vectors $v_1, v_2, \ldots, v_n$. Thus $V = < S >$ and so S is a basis for V.

(6.28) Theorem. (i) Any linearly independent set of vectors in a finite dimensional vector spave V can be extended to a basis for V.

(ii) If W is a subspace of a finite dimensional vector space V, then $\dim W \le \dim V$. Moreover, if $\dim W = \dim V$, then $W = V$.

Proof. (i) Let $S = \{v_1, v_2, \ldots, v_r\}$ be a linearly independent set of vectors in an n-dimensional vector space V and $r < n$. Since $\dim V = n$, the set S cannot span V. There is a vector, say $v_{r+1} \in V$, such that $v_{r+1} \notin < S >$. By Theorem 6.26, the set $\{v_{r+1}\} \cup S$ is linearly independent. The process can be repeated $n - r$ times to get a larger set $\{v_1, v_2, \ldots, v_{r+1}, \ldots, v_n\}$ which is linearly independent. By Theorem 6.27, this set is a basis for V.

(ii) Let V be of dimension n. Then, by Theorem 6.25, any set of $n + 1$ or more vectors is linearly dependent. Furthermore, since a basis of W consists of linearly independent vectors, it cannot contain more than n elements. Hence $\dim W \le n = \dim V$.

If $\dim W = \dim V$, then every basis of W is also a basis for V. Thus $W = V$.

The following theorem gives an equivalent definition of direct sum (See **6.10**).

(6.29) Theorem. A vector space V is the direct sum of its subspaces U and W if and only if each $v \in V$ can be written uniquely as

(i)' $$v = u + w, \quad u \in U, \quad w \in W.$$

Proof. Recall that a vector space V is the direct sum of its subspaces U and W if

(i) $V = U + W$ and **(ii)** $U \cap W = \{ 0 \}$.

Suppose that V is the direct sum of its subspaces U and W. Then conditions (i) and (ii) are satisfied.

Now let $v \in V$. Then

$$v = u + w \tag{1}$$

for some $u \in U$, $w \in W$. To see that the expression for v in (1) is unique, let

$$v = u_1 + w_1, \quad u_1 \in U, \quad w_1 \in W$$

be another expression for v. Then

$$u - u_1 = w_1 - w \in U \cap W = \{ 0 \}.$$

Hence $u - u_1 = 0 = w - w_1$. That is $u = u_1$, $w = w_1$. So the expression for v in (1) is unique and (i)' is satisfied.

Conversely, suppose that (i)' is satisfied. Then (i) is a part of (i)'.

To see that (ii) is satisfied, let $v \in U \cap W$. Then v can be written as

$$\begin{aligned} v &= u + 0 , & u \in U \\ &= 0 + w , & w \in W \end{aligned}$$

By (i)', the expression for v is unique. Therefore,

$$u = 0 = w, \text{ so that } v = 0$$

Hence $\quad U \cap W = \{0\}$.

(6.30) Theorem. If U and W are finite dimensional subspaces of a vector space V over a field F, then

(i) $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$

(ii) If $\quad U \cap W = \{0\}$ then $V = U \oplus W$ and

$\dim V = \dim U + \dim W$

Proof. (i) Suppose $U \cap W \neq \{0\}$ and $U \cap W$ has a basis $\{v_1, v_2, \ldots, v_r\}$. Since $U \cap W \subset U$, this set of vectors forms a part of a basis of U. Let $\{v_1, v_2, \ldots, v_r, u_1, u_2, \ldots, u_s\}$ be a basis of U. Similarly, let $\{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_t\}$ be a basis of W. Thus $U \cap W$, U, W have dimensions r, $r + s$, $r + t$ respectively. Clearly $\{v_1, v_2, \ldots, v_r, u_1, u_2, \ldots, u_s, w_1, w_2, \ldots, w_t\}$ is a spanning set for $U + W$, so that $\dim(U + W) \leq r + s + t$. We need to show that equality must hold. This will follow if we show that this set is linearly independent and hence forms a basis for $U + W$. Suppose that

$$a_1v_1 + a_2v_2 + \ldots + a_rv_r + b_1u_1 + b_2u_2 + \ldots + b_su_s + c_1w_1 + c_2w_2 + \ldots + c_tw_t = 0 \qquad (1)$$

where a's, b's and c's are all in the field F. Then

$$a_1v_1 + a_2v_2 + \ldots + a_rv_r + b_1u_1 + b_2u_2 + \ldots + b_su_s = -(c_1w_1 + c_2w_2 + \ldots + c_tw_t) \qquad (2)$$

Now the vector on the right of (2) is in W (since the w_i's are the basis of W) and that on the left is in U. It follows that each is in $U \cap W$ and therefore, is of the form $d_1v_1 + d_2v_2 + \ldots + d_rv_r$. Thus

$$d_1v_1 + d_2v_2 + \ldots + d_rv_r + (c_1w_1 + c_2w_2 + \ldots + c_tw_t) = 0.$$

But $\{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_t\}$ is a basis for W so that this set is linearly independent. Hence

$$d_1 = \ldots = d_r = c_1 = \ldots = c_t = 0$$

Equation (1) then becomes

$$a_1v_1 + \ldots + a_rv_r + b_1u_1 + \ldots + b_su_s = 0.$$

Again $a_1 = \ldots = a_r = b_1 = \ldots = b_s = 0$, since $\{v_1, v_2, \ldots, v_r, u_1, u_2, \ldots, u_s\}$ is a basis for U and therefore, is a linearly independent set. Consequently, equation (1) implies that the a's, b's and c's are zero. Thus

$$\{v_1, v_2, \ldots, v_r, u_1, u_2, \ldots, u_s, w_1, w_2, \ldots, w_t\}$$

is a linearly independent set and hence is a basis for $U + W$. So

$$\begin{aligned} \dim(U + W) &= r + s + t \\ &= (r + s) + (r + t) - r \\ &= \dim U + \dim W - \dim(U \cap W) \end{aligned} \qquad (3)$$

(ii) Next suppose that $U \cap W = \{0\}$. Let $\{u_1, u_2, \ldots, u_s\}$ be a basis for U and $\{w_1, w_2, \ldots, w_t\}$ be a basis for W. Then $\{u_1, u_2, \ldots, u_s, w_1, w_2, \ldots, w_t\}$ is a spanning set for $U + W = U \oplus W$.

Now if

$$a_1u_1 + \ldots + a_su_s + b_1w_1 + \ldots + b_tw_t = 0$$

then $\quad a_1u_1 + a_2u_2 + \ldots + a_su_s = -(b_1w_1 + b_2w_2 + \ldots + b_tw_t) = 0$,

since each vector is in $U \cap W$. This in turn implies that

$$a_1 = \ldots = a_s = b_1 = \ldots = b_t = 0$$

Hence $\{u_1, u_2, \ldots, u_s, w_1, w_2, \ldots, w_t\}$ is a basis for $U \oplus W$ and

$$\dim(U \oplus W) = s + t = \dim U + \dim W.$$

This agrees with the general result (3) with

$$\dim(U \cap W) = 0.$$

Here we define the dimension of the vector space $\{0\}$ to be zero.

EXERCISE 6.2

1. Determine whether the following vectors in R^4 are linearly independent or linearly dependent:

   (i) $(1, 3, -1, 4), (3, 8, -5, 7), (2, 9, 4, 23)$

   (ii) $(1, -2, 4, 1), (2, 1, 0, -3), (1, -6, 1, 4)$

2. Let $V = P_3(x)$ be the vector space of all polynomials of degree ≤ 3 over R together with the zero polynomial. Determine whether $u, v, w \in V$ are linearly dependent or linearly independent:

   (i) $u = x^3 - 4x^2 + 2x + 3$, $v = x^3 + 2x^2 + 4x - 1$, $w = 2x^3 - x^2 + 3x + 3$

   (ii) $u = x^3 - 3x^2 - 2x + 3$, $v = x^3 - 4x^2 - 3x + 4$, $w = 2x^3 - 7x^2 - 7x + 9$.

3. Show that the vectors $(1-i, i)$ and $(2, -1+i)$ in C^2 are linearly dependent over C but linear independent over R

4. Show that the vectors $(3+\sqrt{2}, 1+\sqrt{2})$ and $(7, 1+2\sqrt{2})$ in R^2 are linearly dependent over R but linearly independent over Q

5. Suppose u, v and w are linearly independent vectors. Prove that

   (i) $u+v-2w,\ u-v-w,\ u+w$ are linearly independent

   (ii) $u+v-3w,\ u+3v-w,\ u+w$ are linearly independent

6. Determine k so that the vectors $(1, -1, k-1)$, $(2, k, -4)$, $(0, 2+k, -8)$ in R^3 are linearly dependent.

7. Using the technique of Theorem 6.21 of casting out vectors which are linear combination of others, find a linearly independent subset of the given set spanning the same subspace

   (i) $\{(1,-3,1), (2,1,-4), (-2,6,-2), (-1,10,-7)\}$ in R^3

   (ii) $\{1, \sin^2 x, \cos 2x, \cos^2 x\}$ in the space of all functions from R to R.

   (iii) $\{1, 3x-4, 4x+3, x^2+2, x-x^2\}$ in the space $P_2(x)$ of polynomials.

8. Verify that the polynomials $2-x^2$, x^3-x, $2-3x^2$ and $3-x^3$ form a basis for $P_3(x)$. Express each of (i) $1+x$ and (ii) $x+x^2$ as a linear combination of these basis vectors

9. Determine whether or not the given set of vectors is a basis for R^2

   (i) $\{(1,1), (3,1)\}$ (ii) $\{(2,1), (1,-1)\}$

10. Determine whether or not the given set of vectors is a basis for R^3

   (i) $\{(1,2,-1), (0,3,1), (1,-5,3)\}$

   (ii) $\{(2,4,-3), (0,1,1), (0,1,-1)\}$

11. Let V be the real vector space of all functions defined on R into R. Determine whether the given vectors are linearly independent or linearly dependent in V

   (i) $x, \cos x$

   (ii) $\sin^2 x, \cos^2 x, \cos 2x$

   (iii) $\sin x, \cos x, \sinh x, \cosh x$

   (iv) $\sin x, \sin x + \cos x, \sin x - \cos x$

   (v) e^{ax}, e^{bx}, e^{cx}, a, b, c being distinct real numbers

12. Determine a basis for each of the following subspaces of R^3

   (i) The plane $x-2y+5z=0$

   (ii) The line $\dfrac{x}{-2}=\dfrac{y}{1}=\dfrac{z}{6}$

   (iii) All vectors of the form (a, b, c), where $3a-2b+c=0$

13. Find the dimension of the subspace

   $\{(x_1, x_2, x_3, x_4) : x_2 = x_3\}$ of R^4. Also determine a basis

14. A subspace U of R^4 is spanned by the vectors $(1,0,2,3)$ and $(0,1,-1,2)$ and a subspace W is spanned by $(1,2,3,4)$, $(-1,-1,5,0)$ and $(0,0,0,1)$. Find the dimensions of U and W.

15. Suppose U and W are distinct four dimensional subspaces of a vector space V of dimension six. Find the possible dimension of $U \cap W$.

16. Find a basis and dimension of the subspace W of R^4 spanned by

   (i) $(1,4,-1,3)$, $(2,1,-3,-1)$ and $(0,2,1,-5)$

   (ii) $(1,-4,-2,1)$, $(1,-3,-1,2)$ and $(3,-8,-2,7)$

17. Let U and W be 2-dimensional subspaces of R^3. Show that

   $U \cap W \neq \{0\}$ J.S

18. If two or more rows (columns) of an $n \times n$ matrix A are linearly dependent, then show that $\det A = 0$.

   J.S

ROW AND COLUMN SPACES

(6.31) Definition. Let $A = [a_{ij}]$ be an $m \times n$ matrix over R, $m \le n$. As defined in (3.1), rows of A

$$P_1 = [a_{11}\ a_{12}\ \dots\ a_{1n}],\ P_2 = [a_{21}\ a_{22}\ \dots\ a_{2n}], \dots,$$
$$P_m = [a_{m1}\ a_{m2}\ \dots\ a_{mn}]$$

are called **row vectors** of A. The subspace of R^n spanned by $\{P_1, P_2, \dots, P_m\}$ is called the **row space** of the matrix A.

Similarly, the subspace spanned by the column vectors

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, \begin{bmatrix} \\ \\ \\ a_{mn} \end{bmatrix}$$

of A is called the **column space** of A. The dimension of the row space is called the **row rank** of the matrix A. Thus the row rank is the maximum number of linearly independent row vectors of A. Similarly, the **column rank** of A is the maximum number of linearly independent column vectors of A.

(6.32) Theorem. Row equivalent matrices have the same row space.

Proof. Let $B \overset{R}{\sim} A$. Then, since rows of B are obtained from the rows of A by elementary row operations each row of B is clearly a row of A or a linear combination of rows of A. Hence the row space of B is contained in the row space of A. On the other hand, we can apply **inverse elementary row operations**[1] to B and obtain A. Hence the row space of A is contained in the row space of B.

Thus row space of A = row space of B.

(6.33) Corollary. Row equivalent matrices have the same row rank.

(6.34) Corollary. If P is nonsingular, then the matrices A and PA have the same row rank.

As $PA \overset{R}{\sim} A$, the row rank of PA is equal to the row rank of A.

(6.35) Corollary. The row rank of a nonsingular matrix A of order $n \times n$ is n.

Proof. Since A is nonsingular, $A \overset{R}{\sim} I_n$. The n rows of I_n are linearly independent because

$$a_1 [1\ 0\ 0\ \dots\ 0] + a_2 [0\ 1\ 0\ 0\ \dots\ 0] + \dots + a_n [0\ 0\ \dots\ 0\ 1] = [0\ 0\ \dots\ 0]$$

$$\Rightarrow \quad a_1 = 0, \quad a_2 = 0, \dots, a_n = 0$$

Hence row rank of I_n is n and so row rank of A is n (By Corollary 6.33).

(6.36) Theorem. The row rank of a matrix equals its column rank.

Proof. Let A be any nonzero matrix. Then there exist nonsingular matrices P and Q such that

$$PAQ = D = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Maximum number of linearly independent rows (columns) of D is r. Hence row rank of $D = r =$ column rank of D.

Now row rank of A = row rank of $PA = r_1$ (say). Thus row rank of PA = row rank of $DQ^{-1} = r_1$. But DQ^{-1} has at most r nonzero rows since if

$$Q^{-1} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}, \quad DQ^{-1} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} I_r Q_1 \\ 0 \end{bmatrix} = \begin{bmatrix} Q_1 \\ 0 \end{bmatrix}$$

1. Inverse elementary row operation of R_{ij} is R_{ji}, of kR_i is $k^{-1} R_i$ and of $R_i + kR_j$ is $R_i - kR_j$

where Q_1 has r rows. Since Q^{-1} is nonsingular, its rows are linearly independent and thus the rows of Q_1 are linearly independent (since a subset of a linearly independent set is linearly independent set). Hence row rank of DQ^{-1} is r. Thus $r_1 = r$.

Consider $Q^T A^T P^T = D^T$

Now column rank of A = row rank of A^T = row rank of $Q^T A^T = r_2$ (say). Proceeding as above, we can prove that row rank of $Q^T A^T$ = row rank of $D^T (P^T)^{-1} = r$. This means $r_2 = r$. Thus row rank of A = column rank of A.

(6.37) Definition. The **rank** of a matrix A is defined to be the common value of the row rank and the column rank of A.

Corollary 6.33 can now be restated in the following form.

(6.38) Theorem. The ranks of equivalent matrices are equal.

(6.39) Theorem. The nonzero rows of a matrix in echelon form are linearly independent.

Proof. Suppose the nonzero rows $P_n, P_{n-1}, \dots, P_1$ of a matrix in echelon form are linearly dependent. Then one of the rows, say P_m is a linear combination of the preceding ones, i.e.,

$$P_m = a_{m+1} P_{m+1} + a_{m+2} P_{m+2} + \dots + a_n P_n$$

Suppose that the kth component of P_m is its first nonzero element. Then since the matrix is in echelon form, the kth components of $P_{m+1}, P_{m+2}, \dots, P_n$ are all zero and so the kth component of P_m is

$$a_{m+1}(0) + \dots + a_n(0) = 0$$

But this is contrary to the assumption that the kth component of P_m is nonzero. Hence the nonzero rows are linearly independent.

For computation of rank of a matrix, the following result is useful.

(6.40) Corollary. The rank of a matrix is equal to the number of nonzero rows in its echelon form. This agrees with our earlier Definition **3.46** of the rank of a matrix.

RANK-ALTERNATIVE METHOD

The rank of a matrix A was found by reducing it to an echelon form. To reduce A to an echelon form many row operations are needed and in the process fractions creep in which make the computations awkward and cumbersome.

The Alternative Method is very elegant involving no fractions and it also yields an echelon matrix that is row equivalent to A and whose nonzero rows constitute a basis for the row space of A.

(6.41) Definition. Let $A = [a_{ij}]$ be an $m \times n$ matrix. We define determinants d_{ij} of order 2 as under:

For $2 \le i \le m$ and $2 \le j \le n$,

$$d_{ij} = \det\begin{bmatrix} a_{11} & a_{1j} \\ a_{i1} & a_{ij} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{1j} \\ a_{i1} & a_{ij} \end{vmatrix} = a_{11}\,a_{ij} - a_{1j}\,a_{i1}$$

(6.42) Theorem. Let $A = [a_{ij}]$ be an $m \times n$ matrix with $a_{11} \ne 0$ and d_{ij} be as in (6.41). Then

$$\text{Rank } A = 1 + \text{Rank}\begin{bmatrix} d_{22} & d_{23} & \cdots & d_{2n} \\ d_{32} & d_{33} & \cdots & d_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ d_{m2} & d_{m3} & \cdots & d_{mn} \end{bmatrix}$$

Proof. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{bmatrix}$$ be an $m \times n$ matrix and $a_{11} \ne 0$.

Multiply R_i, $(2 \le i \le m)$ by a_{11} so that

$$A \underset{\sim}{R} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{11}a_{21} & a_{11}a_{22} & \cdots & a_{11}a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{11}a_{m1} & a_{11}a_{m2} & \cdots & a_{11}a_{mn} \end{bmatrix}$$

$$\underset{\sim}{R} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & \cdots & a_{11}a_{2n} - a_{1n}a_{21} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & a_{11}a_{m2} - a_{12}\,a_{m1} & \cdots & a_{11}a_{mn} - a_{1n}a_{m1} \end{bmatrix} \quad \text{by } \begin{matrix} R_i - a_{i1}R_1 \\ 2 \le i \le m \end{matrix}$$

$$\underset{\sim}{R} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & d_{22} & \cdots & d_{2n} \\ 0 & d_{32} & \cdots & d_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & d_{m2} & \cdots & d_{mn} \end{bmatrix}$$

Since $a_{11} \ne 0$, the first row in the last matrix is not a linear combination of the other rows and so the rank of A is one more than the rank of

$$\begin{bmatrix} d_{22} & d_{23} & \cdots & d_{2n} \\ d_{32} & d_{33} & \cdots & d_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ d_{m2} & d_{m3} & \cdots & d_{mn} \end{bmatrix}$$

as desired.

The theorem will be used recursively until we obtain a matrix having one row (column).

An echelon matrix row equivalent to A will be obtained by retaining R_1 of A and replacing the rows R_i, $2 \le i \le m$, by the first rows of the successive matrices obtained by recursive use of the theorem provided that the first $i - 1$ entries in R_i are zero $(2 \le i \le m)$.

The method is illustrated by examples.

Example 25. Find the rank of the matrix

$$A = \begin{bmatrix} -3 & 5 & 1 & 2 \\ 7 & 2 & 0 & -4 \\ -8 & 3 & 1 & 6 \end{bmatrix}.$$

Also write an echelon matrix row equivalent to A.

Solution.

$$\text{Rank } A = 1 + \text{Rank}\begin{bmatrix} \begin{vmatrix} -3 & 5 \\ 7 & 2 \end{vmatrix} & \begin{vmatrix} -3 & 1 \\ 7 & 0 \end{vmatrix} & \begin{vmatrix} -3 & 2 \\ 7 & -4 \end{vmatrix} \\ \begin{vmatrix} -3 & 5 \\ -8 & 3 \end{vmatrix} & \begin{vmatrix} -3 & 1 \\ -8 & 1 \end{vmatrix} & \begin{vmatrix} -3 & 2 \\ -8 & 6 \end{vmatrix} \end{bmatrix}$$

$$= 1 + \text{Rank}\begin{bmatrix} -41 & -7 & -2 \\ 31 & 5 & -2 \end{bmatrix}$$

$$= 2 + \text{Rank}\left[\begin{vmatrix} -41 & -7 \\ 31 & 5 \end{vmatrix} \quad \begin{vmatrix} -41 & -2 \\ 31 & -2 \end{vmatrix}\right]$$

$$= 2 + \text{Rank}\,[12 \quad 144] = 3$$

An echelon matrix row equivalent to A is

$$\begin{bmatrix} -3 & 5 & 1 & 2 \\ 0 & -41 & -7 & -2 \\ 0 & 0 & 12 & 144 \end{bmatrix}$$

Example 26. Find the rank of the matrix

$$A = \begin{bmatrix} 1 & -2 & 1 & 0 & 5 \\ -1 & 0 & 1 & -2 & 2 \\ 1 & -6 & 3 & -2 & 12 \\ 2 & -3 & 0 & 2 & 3 \end{bmatrix}$$

and write an echelon matrix row equivalent to A

Solution.

$$\text{Rank } A = 1 + \text{Rank}\begin{bmatrix} \begin{vmatrix} 1 & -2 \\ -1 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ -1 & -2 \end{vmatrix} & \begin{vmatrix} 1 & 5 \\ -1 & 2 \end{vmatrix} \\ \begin{vmatrix} 1 & -2 \\ 1 & -6 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 1 & -2 \end{vmatrix} & \begin{vmatrix} 1 & 5 \\ 1 & 12 \end{vmatrix} \\ \begin{vmatrix} 1 & -2 \\ 2 & -3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 2 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 2 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 5 \\ 2 & 3 \end{vmatrix} \end{bmatrix}$$

$$= 1 + \text{Rank}\begin{bmatrix} -2 & 2 & -2 & 7 \\ -4 & 2 & -2 & 7 \\ 1 & -2 & 2 & -7 \end{bmatrix}$$

$$= 2 + \text{Rank}\begin{bmatrix} \begin{vmatrix} -2 & 2 \\ -4 & 2 \end{vmatrix} & \begin{vmatrix} -2 & -2 \\ -4 & 2 \end{vmatrix} & \begin{vmatrix} -2 & 7 \\ -4 & 7 \end{vmatrix} \\ \begin{vmatrix} -2 & 2 \\ 1 & -2 \end{vmatrix} & \begin{vmatrix} -2 & -2 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} -2 & 7 \\ 1 & -7 \end{vmatrix} \end{bmatrix}$$

$$= 2 + \text{Rank}\begin{bmatrix} 4 & -4 & 14 \\ 2 & -2 & 7 \end{bmatrix}$$

$$= 3 + \text{Rank}\left[\begin{vmatrix} 4 & -4 \\ 2 & -2 \end{vmatrix} \quad \begin{vmatrix} 4 & 14 \\ 2 & 7 \end{vmatrix}\right]$$

$$= 3 + \text{Rank}\,[0 \quad 0] = 3$$

An echelon matrix row equivalent to A is

$$\begin{bmatrix} 1 & -2 & 1 & 0 & 5 \\ 0 & -2 & 2 & -2 & 7 \\ 0 & 0 & 4 & -4 & 14 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Example 27. Find the rank of the matrix

$$A = \begin{bmatrix} 2 & 1 & 3 & 5 \\ 15 & 8 & 1 & 12 \\ 11 & 5 & 8 & 6 \\ 12 & 8 & 7 & 10 \end{bmatrix}$$

Solution. $\text{Rank } A = 1 + \text{Rank}\begin{bmatrix} 1 & -43 & -51 \\ -1 & -17 & -43 \\ 4 & -22 & -40 \end{bmatrix}$

$$= 2 + \text{Rank}\begin{bmatrix} -60 & -94 \\ 150 & 164 \end{bmatrix} = 3 + \text{Rank}\,[4260] = 4$$

An echelon matrix row equivalent to A is $\begin{bmatrix} 2 & 1 & 3 & 5 \\ 0 & 1 & -43 & -51 \\ 0 & 0 & -60 & -94 \\ 0 & 0 & 0 & 4260 \end{bmatrix}$

PROOFS OF THEOREMS IN CHAPTER 4

(6.43) Proof of Corollary 4.9. If A is nonsingular, then by Theorem 4.8

$$x = A^{-1}b = A^{-1}\mathbf{0} = \mathbf{0}$$

is the unique solution.

Conversely, if the system $Ax = \mathbf{0}$ has a unique solution $x = \mathbf{0}$, then the columns a_j of A satisfy the equation

$$\sum_{j=1}^{n} x_j \mathbf{a}_j = 0$$

only when $x_j = 0$, $j = 1, 2, \ldots, n$ i.e., the columns of A are linearly independent and so A is nonsingular.

(6.44) Proof of Theorem 4.10. (Consistency Criterion). The matrix of the coefficients and the augmented matrix are

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad A_b = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

We shall represent the columns of the matrices by bold faced letters. Thus $\mathbf{a}_j$ will be the jth column of the matrix A. Now the ranks of A and A_b are the maximum number of linearly independent columns in A and A_b, respectively. Since every column of A is also a column of A_b, the rank of A cannot exceed the rank of A_b. There are, therefore, only two possibilities

(i) Rank A < Rank A_b

(ii) Rank A = Rank A_b

We note that the rank of A_b cannot be greater than 1 + Rank A. If the first possibility is true, then any set of linearly independent columns of A_b must include the column $\mathbf{b}$ so that $\mathbf{b}$ is not a linear combination of the columns of A. For otherwise Rank A = Rank A_b. Thus it follows that there do not exist any x_j's satisfying

$$\sum_{j=1}^{n} x_j \mathbf{a}_j = \mathbf{b},$$

where $\mathbf{a}_j$ represents the jth column of A, that is, there exist no x_j's satisfying $Ax = b$. Hence there is no solution and the given system is inconsistent.

Next consider the second possibility that Rank A = Rank $A_b = k$. In this case there is a set of k linearly independent columns of A. Without any loss of generality, we can suppose that the first k columns of A are linearly independent. Then every other column of A_b is expressible as a linear combination of these k columns. Thus there exist numbers $x_1, x_2, \ldots, x_k$ such that

$$\sum_{j=1}^{n} x_j \mathbf{a}_j = \mathbf{b}$$

Adding zero times the remaining $n - k$ columns of A, we get $Ax = \mathbf{b}$, where $x = (x_1, x_2, \ldots, x_k, 0, \ldots, 0)$. Hence in this case there is at least one solution.

Conversely, if a solution $y = (y_1, y_2, \ldots, y_n)$ of the system of equations $Ax = \mathbf{b}$ exists, then b must be a linear combination of columns of A. Thus there exist scalars y_j, not all zero, such that

$$\sum_{j=1}^{n} y_j \mathbf{a}_j = \mathbf{b} \quad \text{or} \quad Ay = \mathbf{b}$$

Taking transpose of both the sides, we have $y^T A^T = b^T$

Thus b^T is a linear combination of the rows of A^T. Hence

$$\text{Rank } A_b = \text{Rank } A_b^T = \text{Rank } A^T = \text{Rank } A.$$

The theorem is proved.

(6.45) Proof of Theorem 4.11. The proof of the theorem will be given by induction on m.

If $m = 1$, there is only one equation

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \tag{1}$$

with $n \geq 2$. If $a_{11} = 0$, a nontrivial solution of (1) is

$$x_1 = 1, \; x_2 = x_3 = \cdots = x_n = 0.$$

If $a_{11} \neq 0$, we take $x_2 = \cdots = x_n = 1$ so that (1) becomes

$$a_{11}x_1 + a_{12} + \cdots + a_{1n} = 0$$

giving $x_1 = -\dfrac{1}{a_{11}}(a_{12} + \cdots + a_{1n}), \; x_2 = \cdots = x_n = 1.$

which is a nontrivial solution of (1). Hence the theorem is true for $m = 1$ and $n \geq 2$.

Now assume that the theorem is true for some positive integer $m = k$. We shall show that it is true for $m = k + 1$ and $k + 1 < n$.

Consider the set of $k+1$ equations

$$\left.\begin{array}{ccccccccc} a_{11} x_1 & + & a_{12} x_2 & + & \cdots & + & a_{1n} x_n & = & 0 \\ a_{21} x_1 & + & a_{22} x_2 & + & \cdots & + & a_{2n} x_n & = & 0 \\ \vdots & & \vdots & & \cdots & & \vdots & & \vdots \\ a_{k1} x_1 & + & a_{k2} x_2 & + & \cdots & + & a_{kn} x_n & = & 0 \\ a_{k+1\,1} x_1 & + & a_{k+1\,2} x_2 & + & \cdots & + & a_{k+1\,n} x_n & = & 0 \end{array}\right\}$$

If $a_{11} = a_{21} = \ldots = a_{k1} = a_{k+1\,1} = 0$, then

$x_1 = 1,\ x_2 = 0 = x_3 = \ldots = x_n$

is a nontrivial solution of the system (2). Hence we may assume that at least one of the coefficients is nonzero. Without any loss of generality, let $a_{11} \neq 0$, for otherwise we consider the equation with the nonzero coefficient of x_1 and relabel it as the first equation. Then

$$x_1 = -\frac{1}{a_{11}}(a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n).$$

Substituting this value of x_1 in the second through the $(k+1)$st equations of system (2), we obtain the equivalent system

$$\begin{array}{ccccccccccc} a_{11}x_1 & + & a_{12} x_2 & + & a_{13} x_3 & + & \cdots & + & a_{1n} x_n & = & 0 \\ & & b_{22} x_2 & + & b_{23} x_3 & + & \cdots & + & b_{2n} x_n & = & 0 \\ & & \vdots & & \vdots & & \cdots & & \vdots & & \vdots \\ & & b_{k2} x_2 & + & b_{k3} x_3 & + & \cdots & + & b_{kn} x_n & = & 0 \\ & & b_{k+1\,2} x_2 & + & b_{k+1\,3} x_3 & + & \cdots & + & b_{k+1\,n} x_n & = & 0 \end{array}$$

where $b_{ij} = a_{ij} - \dfrac{a_{i1}\, a_{1j}}{a_{11}}$, $i = 2, 3, \ldots, k+1$, $j = 2, 3, \ldots, n$.

The last k equations of the system (3) are in $(n-1)$ unknowns $x_2, x_3, \ldots, x_n$ and $k < n-1$, since $k+1 < n$. Therefore, by the induction hypothesis, these equations have a nontrivial solution $x_2, x_3, \ldots, x_n$. Once $x_2, x_3, \ldots, x_n$ are known, we obtain

$$x_1 = -\frac{1}{a_{11}}(a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n)$$

Thus the theorem is true for $m = k+1 < n$ and hence for all $m < n$, by induction.

(6.46) Proof of Theorem 4.12. Let the rank of A be k such that $k < n$. There are k linearly independent columns, and every set of $k+1$ columns of A is linearly dependent. If we transfer these columns to be the first k columns, then every remaining column is linearly dependent on these so that there exist scalars $x_1, x_2, \ldots, x_k$, not all zero, such that

$$\sum_{j=1}^{n} x_j a_j = a_{k+1} \qquad i = 1, 2, \ldots, n-k$$

Thus $(x_1, x_2, \ldots, x_k, 0, 0, \ldots, 0)$ is a nontrivial solution of $Ax = 0$. In fact the last $n-k$ elements of the solution vector can be given arbitrary values.

Conversely, if there is a nontrivial solution $y \neq 0$ of $Ax = 0$ and Rank $A = n$, then A^{-1} exists and so, from $Ay = 0$, we obtain

$$A^{-1}Ay = A^{-1}0 = 0$$

i.e. $y = 0$ which is a contradiction. Hence Rank A is less than n.

The second part is now obvious.

(6.47) Definition. (**The Solution Space.**) Consider Example 7 of Chapter 4. The solution of the system

$$\begin{aligned} x_1 - x_2 + 2x_3 + x_4 &= 0 \\ 3x_1 + 2x_2 + x_4 &= 0 \\ 4x_1 + x_2 + 2x_3 + 2x_4 &= 0 \end{aligned}$$

is given by

$$x_1 = -\frac{4}{5}x_3 - \frac{3}{5}x_4$$

$$x_2 = \frac{6}{5}x_3 + \frac{2}{5}x_4$$

Let $x_3 = a$ and $x_4 = b$. Then

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\frac{4}{5}a - \frac{3}{5}b \\ \frac{6}{5}a + \frac{2}{5}b \\ a \\ b \end{bmatrix} = a\begin{bmatrix} -\frac{4}{5} \\ \frac{6}{5} \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} -\frac{3}{5} \\ \frac{2}{5} \\ 0 \\ 1 \end{bmatrix}$$

Thus every solution of the given system of equations is a linear combination of

$$\begin{bmatrix} -\frac{4}{5} \\ \frac{6}{5} \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -\frac{3}{5} \\ \frac{2}{5} \\ 0 \\ 1 \end{bmatrix}$$

It is easy to check that the set of all solutions forms a vector space, called the **solution space** of the given system.

The vectors $\begin{bmatrix} -\frac{4}{5} \\ \frac{6}{5} \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -\frac{3}{5} \\ \frac{2}{5} \\ 0 \\ 1 \end{bmatrix}$ form a basis of this space

Note that the two vectors are linearly independent

[Also see Example 10]

LINEAR TRANSFORMATIONS

(6.48) Definition. Let U and V be two vector spaces over the same field F and let $T: U \to V$ be a function. Then T is called a **linear transformation** if the following conditions are satisfied:

(i) $T(u_1 + u_2) = T(u_1) + T(u_2)$, for all $u_1, u_2 \in U$

(ii) $T(au) = aT(u)$, for all $u \in U$ and $a \in F$

In other words $T: U \to V$ is linear if it preserves the two operations of a vector space, namely vector addition and scalar multiplication.

The definition includes the possibility that U and V are both the same space.

Substituting $a = 0$ in (ii), we obtain $T(\theta) = \theta$ so that a linear transformation maps the θ vector of U into the θ vector of V.

It is easy the see that $T: U \to V$ is a linear transformation if and only if

(i)' $T(a_1u_1 + a_2u_2) = a_1 T(u_1) + a_2 T(u_2)$ for all $a_1, a_2 \in F$, $u_1, u_2 \in U$

For, if (i) and (ii) are satisfied, then

$$\begin{aligned} T(a_1u_1 + a_2u_2) &= T(a_1u_1) + T(a_2u_2), && \text{by (i)} \\ &= a_1 T(u_1) + a_2 T(u_2), && \text{by (ii)} \end{aligned}$$

and form (i)', we have (i) and (ii) by putting $a_1 = a_2 = 1$ and $a_2 = 0$ in (i)' respectively.

So T is a linear transformation.

Clearly, this extends to a linear combination of any number of vectors. Thus if $u_1, u_2, \ldots, u_n$ are vectors in U and $a_1, a_2, \ldots, a_n$ are in F, then

$$T(a_1u_1 + a_2u_2 + \ldots + a_nu_n) = a_1 T(u_1) + a_2 T(u_2) + \ldots + a_n T(u_n)$$

Thus a linear transformation T preserves linear combinations.

Example 28. Let $T: U \to V$ be a linear transformation where U and V are vector spaces over the same field F. Show that for all u, v belonging to U

$$T(-u) = -T(u) \quad \text{and} \quad T(u - v) = T(u) - T(v)$$

Solution. Since $T(au) = a\,T(u)$ for all $a \in F$, we put $a = -1$ to obtain

$$T(-u) = T(-1u) = -1\,T(u) = -T(u)$$

Also

$$\begin{aligned} T(u - v) &= T(u + (-v)) \\ &= T(u) + T(-v), && \text{by (6.48(i))} \\ &= T(u) - T(v) \end{aligned}$$

Example 29. Let U and V be two vector spaces over the same field F. The zero transformation $\theta: U \to V$ is defined as

$$\theta(u) = \theta \quad \text{for all } u \in U$$

Thus the zero transformation maps every vector in U onto the zero vector of V.

It is easy to see that the zero transformation is linear.

Example 30. Let $T: R^2 \to R^2$ be given by

$$T(x_1, x_2) = (-x_2, x_1)$$

i.e. T is a 90° counterclockwise rotation of the points of the plane. Then T is linear.

Solution. Let $x = (x_1, x_2)$, $y = (y_1, y_2) \in R^2$ and $a, b \in F$. Then

$$\begin{aligned} T(ax + by) &= T(a(x_1, x_2) + b(y_1, y_2)) \\ &= T(ax_1 + by_1, ax_2 + by_2) \\ &= (-(ax_2 + by_2), (ax_1 + by_1)) \\ &= a(-x_2, x_1) + b(-y_2, y_1) \\ &= aT(x_1, x_2) + b\,T(y_1, y_2) = aT(x) + bT(y). \end{aligned}$$

Hence T in linear.

Example 31. Let $T: R^2 \to R^2$ be given by

$$T(x_1, x_2) = (x_1 + 1, x_2 + 2).$$

Then T is not linear.

Solution. If T were linear then it should map $(0, 0) \in R^2$ onto $(0, 0) \in R^2$. But

$$T(0, 0) = (1, 2) \neq (0, 0).$$

Thus T is not linear.

Example 32. Let $V = P(x)$ be the vector space of all polynomials in x over R (Example 6). Then the transformations

(i) $D : V \to V$ defined by $D(v) = \dfrac{dv}{dx}$ for all $v \in V$

(ii) $I : V \to V$ defined by $I(v) = \displaystyle\int_0^x v\,dx$

are linear.

Solution. (i) Let $u, v \in V$ and $a \in R$. Then

$$D(u+v) = \frac{d}{dx}(u+v)$$

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

$$D(au) = \frac{d(au)}{dx} = a\frac{du}{dx} = a\,D(u)$$

Therefore D is linear.

(ii)

$$I(u+v) = \int_0^x (u+v)\,dx$$

$$= \int_0^x u\,dx + \int_0^x v\,dx = I(u) + I(v)$$

$$I(au) = \int_0^x au\,dx = a\int_0^x u\,dx = aI(u)$$

Therefore I is linear.

Example 33. Let U and V be two vector spaces over the same field F. Let $S : U \to V$ and $T : U \to V$ be two linear transformations. Define

$S + T : U \to V$ and $aS : U \to V$ by

$$(S+T)(u) = S(u) + T(u) \qquad (1)$$

$$(aS)(u) = a\,S(u) \qquad (2)$$

for all $u \in U$ and $a \in F$.

Show that $S + T$ and aS are linear transformations.

Solution. Let $u_1, u_2 \in U$ and a_1, a_2 be scalars. Then

$$(S+T)(a_1u_1 + a_2u_2) = S(a_1u_1 + a_2u_2) + T(a_1u_1 + a_2u_2), \qquad \text{by (1)}$$

$$= a_1S(u_1) + a_2S(u_2) + a_1T(u_1) + a_2T(u_2), \qquad \text{since } S \text{ and } T \text{ are linear}$$

$$= a_1(S(u_1) + T(u_1)) + a_2(S(u_2) + T(u_2)), \qquad \text{(Commutative property in } V)$$

$$= a_1(S+T)(u_1) + a_2(S+T)(u_2), \qquad \text{by (1)}$$

showing that $S + T$ is linear.

Again,

$$(aS)(a_1u_1 + a_2u_2) = aS(a_1u_1 + a_2u_2)$$

$$= aa_1S(u_1) + aa_2S(u_2) = a_1(aS)(u_1) + a_2(aS)(u_2)$$

Thus aS is linear.

It can be shown that if $L(U, V)$ denotes the set of all linear transformations from U into V, then $L(U, V)$ is a vector space (See Problem 14 EXERCISE 6.3) with $\mathbf{0} : U \to V$ given by $\mathbf{0}(u) = \mathbf{0}$ for all $u \in U$ as the additive identity and for each $S \in L(U, V)$ the function $(-S) : U \to V$ defined by $(-S)(u) = -S(u)$ as the additive inverse of S.

(6.49) Definition. Let $T : U \to V$ be a linear transformation. The subset of U consisting of all vectors that T maps onto the $\mathbf{0}$ vector of V is called the **null space (or kernel)** of T and is denoted by $N(T)$. In symbols

$$N(T) = \{u \in U : T(u) = \mathbf{0}\}$$

As with functions in general, the subset of V consisting of all functional values of T is called the **range** of T and is denoted by $R(T)$, i.e.,

$$R(T) = \{v \in V : \text{there exists } u \in U \text{ with } T(u) = v\}$$

(6.50) Theorem. Let $T : U \to V$ be a linear transformation, then

(i) $R(T)$ is a subspace of V

(ii) $N(T)$ is a subspace of U

Proof. (i) Let $v_1, v_2 \in R(T)$ and a, b be scalars. Since v_1, v_2 belong to the range of T, there exist $u_1, u_2 \in U$ such that

$$T(u_1) = v_1, \qquad T(u_2) = v_2$$

Now $a_1u_1 + bu_2 \in U$, since U is a vector space, and

$$av_1 + bv_2 = a\,T(u_1) + b\,T(u_2) = T(au_1 + bu_2)$$

since T is linear.

Thus $av_1 + bv_2 \in R(T)$.

So by (6.5), $R(T)$ is a subspace of V.

(ii) As in **(i)** let $u_1, u_2 \in N(T)$ and a, b be scalars. Since

$$T(u_1) = 0 \quad \text{and} \quad T(u_2) = 0,$$

we have

$$\begin{aligned} T(au_1 + bu_2) &= aT(u_1) + bT(u_2) \\ &= a\,0 + b\,0 = 0 \end{aligned}$$

Hence $au_1 + bu_2 \in N(T)$ and so $N(T)$ is a subspace of U.

Example 34. Let $v_1 = (1, 1)$ and $v_2 = (1, 0)$ be a basis of R^2. Find a formula for the linear transformation $T : R^2 \to R^3$ for which

$$T(v_1) = (1, 2, 1) \quad \text{and} \quad T(v_2) = (-1, 0, 2)$$

Solution. Let $x = (x_1, x_2)$ be any vector in R^2. Then x is uniquely expressible as a linear combination of the basis vectors, i.e., for $a_1, a_2 \in R$,

$$\begin{aligned} x = (x_1, x_2) &= a_1v_1 + a_2v_2 \\ &= a_1(1, 1) + a_2(1, 0) \\ &= (a_1 + a_2, a_1) \end{aligned}$$

Therefore,

$$x_1 = a_1 + a_2 \quad \text{or} \quad a_2 = x_1 - a_1$$
$$x_2 = a_1$$

and so

$$x = a_1v_1 + a_2v_2 = x_2v_1 + (x_1 - x_2)v_2$$

$$\begin{aligned} T(x) = T(x_1, x_2) &= T(x_2v_1 + (x_1 - x_2)v_2) \\ &= x_2\,T(v_1) + (x_1 - x_2)\,T(v_2) \\ &= x_2\,(1, 2, 1) + (x_1 - x_2)\,(-1, 0, 2) \\ &= (2x_2 - x_1, 2x_2, 2x_1 - x_2) \end{aligned}$$

which is the required formula for the linear transformation T.

(6.51) Definition. A linear transformation $T : U \to V$ is called **one-to-one** (or **injective**) if and only if, for all $u_1, u_2 \in U$, if $u_1 \neq u_2$ then $T(u_1) \neq T(u_2)$.

Equivalently, $T(u_1) = T(u_2)$ implies $u_1 = u_2$.

T is called an **onto** transformation if $R(T) = V$.

A linear transformation that is both one-to-one and onto is called an **isomorphism** (or a **nonsingular transformation**).

(6.52) Theorem. A linear transformation $T : U \to V$ is one-to-one if and only if $N(T) = \{0\}$.

Proof. Suppose that T is one-to-one. Let $u \in N(T)$. Then

$$T(u) = 0 = T(0)$$

So

$$u = 0 \qquad \text{Thus } N(T) = \{0\}$$

Conversely, let $N(T) = \{0\}$. To see that T is then one-to-one, let

$$T(u_1) = T(u_2) \qquad \text{for } u_1, u_2 \in U$$

Then

$$T(u_1 - u_2) = 0, \qquad \text{so that } u_1 - u_2 \in N(T) = \{0\}. \text{ So}$$
$$u_1 - u_2 = 0 \qquad \text{That is } u_1 = u_2. \text{ Hence } T \text{ is one-to-one}$$

(6.53) Corrollary. A one-to-one linear transformation preserves basis and dimension.

Proof. Let $T : U \to V$ be a one-to-one linear transformation where U and V are vector spaces. Let $B = \{u_1, u_2, \ldots, u_n\}$ be a basis of U. We show that $B' = \{T(u_1), T(u_2), \ldots, T(u_n)\}$ is a basis of $T(U)$. For this we have to prove that

(i) B' is linearly independent and (ii) B' spans $T(U)$

For (i), suppose that for scalars $a_1, a_2, \ldots, a_n$

$$a_1\,T(u_1) + a_2\,T(u_2) + \ldots + a_n\,T(u_n) = 0 \tag{1}$$

Then, as T is a linear transformation, (1) can be written as

$$T(a_1u_1 + a_2u_2 + \ldots + a_nu_n) = 0 = T(0) \tag{2}$$

Since T is one-to-one, (2) implies

$$a_1u_1 + a_2u_2 + \ldots + a_nu_n = 0 \tag{3}$$

As $u_1, u_2, \ldots, u_n$ are linearly independence, (3) implies

$$a_1 = a_2 = \ldots = a_n = 0$$

Hence B' is linearly independent.

Next, let $v \in T(U)$. Then there is a $u \in U$ such that $v = T(u)$.

But $u = b_1u_1 + b_2u_2 + \ldots + b_nu_n$, for some scalars $b_1, b_2, \ldots, b_n$

Hence

$$\begin{aligned} v = T(u) &= T(a_1u_1 + a_2u_2 + \ldots + a_nu_n) \\ &= a_1\,T(u_1) + a_2\,T(u_2) + \ldots + a_n\,T(u_n) \end{aligned}$$

Thus B' spans $T(U)$ and so is a basis of $T(U)$.

Clearly,

$$\dim U = n = \dim T(U)$$

Dimensional theorem:-

(6.54) Theorem. Let $T : U \to V$ be a linear transformation from an n dimensional vector space U to a vector space V over the same field F. Then

$$\dim N(T) + \dim R(T) = n$$

Proof. Assume that $\dim N(T) = r$ and let $\{u_1, u_2, \ldots, u_r\}$ be a basis for $N(T)$. Since the set $\{u_1, u_2, \ldots, u_r\}$ is linearly independent, by Theorem **6.28**, it can be extended to form a basis for U. Therefore, there are $n - r$ vectors $u_{r+1}, u_{r+2}, \ldots, u_n$ such that $\{u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ is a basis for U. We shall show that the $n - r$ vectors $T(u_{r+1}), T(U_{r+2}), \ldots, T(u_n)$ form a basis for $R(T)$.

Let v be any vector in $R(T)$. Then $v = T(u)$ for some $u \in U$. Since

$\{u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ is a basis for U, the vector u can be uniquely expressed as a linear combination of the basis vectors, i.e.,

$$u = a_1u_1 + a_2u_2 + \ldots + a_ru_r + a_{r+1}u_{r+1} + \ldots + a_nu_n$$

where $a_1, a_2, \ldots, a_r, a_{r+1}, \ldots, a_n$ belong to F. So

$$\begin{aligned} v &= T(u) = T(a_1u_1 + a_2u_2 + \ldots + a_ru_r + a_{r+1}u_{r+1} + \ldots + a_nu_n) \\ &= a_1T(u_1) + a_2T(u_2) + \ldots + a_rT(u_r) + a_{r+1}T(u_{r+1}) + \ldots + a_nT(u_n) \end{aligned}$$

since T is linear

$$= 0 + 0 + \ldots + 0 + 0 + a_{r+1}T(u_{r+1}) + \ldots + a_nT(u_n)$$

because $u_1, u_2, \cdots, u_r$ are in $N(T)$ and so

$$T(u_1) = T(u_2) = \cdots = T(u_r) = 0$$

Thus any $v \in R(T)$ is a linear combination of the vectors $T(u_{r+1}), T(u_{r+2}), \cdots, T(u_n)$. Hence these vectors span $R(T)$.

The proof will be complete if we show that the vectors $T(u_{r+1}), T(u_{r+2}), \cdots, T(u_n)$ are linearly independent in $R(T)$.

Let $\quad b_{r+1}T(u_{r+1}) + b_{r+2}T(u_{r+2}) + \cdots + b_nT(u_n) = 0. \quad (1)$

where $b_{r+1}, b_{r+2}, \cdots, b_n \in F$

We need to show that $b_{r+1} = b_{r+2} = \cdots = b_n = 0$

Since T is linear, (1) can be written as

$$T(b_{r+1}u_{r+1} + b_{r+2}u_{r+2} + \cdots + b_nu_n) = 0$$

This means that the vector $b_{r+1}u_{r+1} + b_{r+2}u_{r+2} + \cdots + b_nu_n$ is $N(T)$. Hence it can be written as a linear combination of the basis vectors $u_1, u_2, \cdots, u_r$. Therefore,

$b_{r+1}u_{r+1} + b_{r+2}u_{r+2} + \cdots + b_nu_n = b_1u_1 + b_2u_2 + \cdots + b_ru_r$ where $b_1, b_2, \cdots, b_r \in F$

i.e. $\quad b_1u_1 + b_2u_2 + \cdots + b_ru_r - b_{r+1}u_{r+1} - b_{r+2}u_{r+2} - \cdots - b_nu_n = 0 \quad (2)$

Since $\{u_1, u_2, \cdots, u_r, u_{r+1}, \cdots, u_n\}$ is a linearly independent set, from (2) we must have

$$b_1 = b_2 = \cdots = b_r = b_{r+1} = \cdots = b_n = 0$$

Hence from (1) we see that $T(u_{r+1}), T(u_{r+2}), \cdots, T(u_n)$ are linearly independent and so they form a basis for $R(T)$. Thus

$$\begin{aligned} \dim R(T) = n - r &\Rightarrow n = r + \dim R(T) \\ &\Rightarrow \dim N(T) + \dim R(T) = n \end{aligned}$$

$\dim R(T)$ and $\dim N(T)$ are respectively called **rank** and **nullity** of T. This theorem is usually referred to as the **Dimensional Theorem**.

Example 35. Let $T : R^3 \to R^3$ be the linear transformation defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2, x_1 + x_3, x_2 + x_3)$$

Find a basis and dimension of **(i)** $R(T)$ **(ii)** $N(T)$.

Solution. (i) R^3 is generated by (1, 0, 0), (0, 1, 0), (0, 0, 1). Therefore, $R(T)$ is generated by

$$T(1, 0, 0), \quad T(0, 1, 0), \quad T(0, 0, 1)$$

Now

$$\begin{aligned} T(1, 0, 0) &= (1, 1, 0) \\ T(0, 1, 0) &= (-1, 0, 1) \\ T(0, 0, 1) &= (0, 1, 1) \end{aligned}$$

We check whether the spanning set $\{(1, 1, 0), (-1, 0, 1), (0, 1, 1)\}$ of $R(T)$ is linearly independent. For this we consider the following matrix formed by these vectors as its columns and reduce it to the echelon form. The matrix is

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \underset{\sim}{R} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{by } R_2 - R_1$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{by } R_3 - R_2$$

Thus the vectors in the spanning set are linearly dependent. But the vectors $(1, 1, 0)$ and $(-1, 0, 1)$ are linearly independent. Hence they form a basis for $R(T)$ and $\dim R(T) = 2$

(ii) A vector $x = (x_1, x_2, x_3) \in R^3$ is in $N(T)$ if $T(x_1, x_2, x_3) = (0, 0, 0)$

i.e., $(x_1 - x_2, x_1 + x_3, x_2 + x_3) = (0, 0, 0)$

$$x_1 - x_2 = 0 \quad (1)$$

or

$$x_1 + x_3 = 0 \quad (2)$$

$$x_2 + x_3 = 0 \quad (3)$$

From (1), $x_1 = x_2$ and from (3), $x_3 = -x_2 = -x_1$

So $$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_2 \\ -x_2 \end{bmatrix} = x_2\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Thus the vector $(1, 1, -1)$ generates $N(T)$. $\{(1, 1, -1)\}$ is a basis for $N(T)$, since $(1, 1, -1)$ is nonzero and, therefore, linearly independent. Hence $\dim N(T) = 1$

EXERCISE 6.3

1. Check which of the following define linear transformations from R^3 to R^2?

(i) $T(x_1, x_2, x_3) = (x_1 - x_2, x_1 - x_3)$

(ii) $T(x_1, x_2, x_3) = (|x_1|, x_2 - x_3)$

(iii) $T(x_1, x_2, x_3) = (x_1 + 1, x_2 + x_3)$

(iv) $T(x_1, x_2, x_3) = (0, x_3)$

(v) $T(x_1, x_2, x_3) = \left(x_1 + \dfrac{x_2}{x_3}, x_3\right)$

(vi) $T(x_1, x_2, x_3) = (3x_1 - 2x_2 + x_3, x_3 - 3x_2 - 2x_1)$

2. Show that each of the following defines linear transformation from R^3 to R^3?

(i) $T(x_1, x_2, x_3) = (x_1 - x_2, x_2 - x_3, x_1)$

(ii) $T(x_1, x_2, x_3) = (x_1 + x_2, -x_1 - x_2, x_3)$

(iii) $T(x_1, x_2, x_3) = (x_2, -x_1, -x_3)$

(iv) $T(x_1, x_2, x_3) = (x_1 - 3x_2 - 2x_3, x_2 - 4x_3, x_3)$

(v) $T(x_1, x_2, x_3) = (x_1 + x_3, x_1 - x_3, x_2)$

3. Show that each of the following transformations is not linear

(i) $T: R^2 \to R$ defined by $T(x_1, x_2) = x_1 x_2$

(ii) $T: R^2 \to R^3$ defined by $T(x_1, x_2) = (x_1 + 1, 2x_2, x_1 + x_2)$

(iii) $T: R^3 \to R^2$ defined by [illegible]

(iv) $T: R^2 \to R^2$ defined by [illegible]

(v) $T: R^3 \to R^3$ defined by $T(x_1, x_2, x_3) = (x_1, x_2, x_3) + (1, 1, 1)$

4. Determine which of the following transformations are linear

(a) $T: M_{22} \to R$ defined by

(i) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + d$

(ii) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$

(b) $T: P_2(x) \to P_2(x)$ defined by

(i) $T(a + bx + cx^2) = a + (b + c)x + (2a - 3b)x^2$

(ii) $T(a + bx + cx^2) = (a + 1) + bx + cx^2$

5. If A is an $m \times n$ matrix, show that $T(x) = Ax$ is a linear transformation from R^n into R^m, where $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in R^n$

6. Determine which of the following linear transformations are one-to-one

(i) $T: R^2 \to R^3$ defined by $T(x_1, x_2) = (x_1 + x_2, x_1 - x_2, x_1 + 2x_2)$

(ii) $T: R^3 \to R^2$ defined by $T(x_1, x_2, x_3) = (x_1 - x_2, x_3)$

(iii) $T: R^2 \to R^3$ defined by $T(x_1, x_2) = (x_1, x_1 + x_2, x_1 - x_2)$

7. Let C be the vector space of complex numbers over R and $T: C \to C$ be defined by $T(z) = \bar{z}$, where $\bar{z}$ denotes the complex conjugate of z. Show that T is linear

8. Let V be the vector space $P_n(x)$ of polynomials $p(x)$ with real coefficients and of degree not exceeding n together with the zero polynomial. Let $T: V \to V$ be defined by

$$T(p(x)) = p(x + 1)$$ Show that T is linear

9. Let $v_1 = (1, 1, 1)$, $v_2 = (1, 1, 0)$ and $v_3 = (1, 0, 0)$ be a basis for R^3. Find a linear transformation $T : R^3 \to R^2$ such that

$$T(v_1) = (1, 0), \quad T(v_2) = (2, -1) \quad \text{and} \quad T(v_3) = (4, 3).$$

10. Let $T : R^2 \to R$ be the linear transformation for which

$$T(1, 1) = 3 \quad \text{and} \quad T(0, 1) = -2.$$

Find $T(x_1, x_2)$.

11. Let $D : P_2(x) \to P_2(x)$ be the differentiation operator

$$D(p(x)) = p'(x) \text{ for all } p(x) \in P_2(x). \text{ Find } N(D).$$

12. Define $T : R^3 \to R^3$ by $T(x_1, x_2, x_3) = (-x_3, x_1, x_1 + x_3)$. Find $N(T)$. Is T one-to-one ?

13. Suppose U, V and W are vector spaces over the same field F. Let $T : U \to V$ and $S : V \to W$ be linear transformations. The transformation $S \circ T : U \to W$ is defined by $(S \circ T)(u) = S(T(u))$ for all $u \in U$. Show that $S \circ T$ is a linear transformation.

14. Let U and V be two vector spaces over the same field F. Denote the set of all linear transformations from U into V by $L(U, V)$. Show that $L(U, V)$ is a vector space over F with vector space operations as defined in **Example 33.**

15. Find a basis and dimension of each of $R(T)$ and $N(T)$, where

(i) $T : R^3 \to R^3$ is defined by

$$T(x_1, x_2, x_3) = (x_1 + 2x_2 - x_3, x_2 + x_3, x_1 + x_2 - 2x_3)$$

(ii) $T : R^3 \to R^4$ is defined by

$$T(x_1, x_2, x_3) = (2x_1 + x_3, 4x_1 + x_2, x_1 + x_3, x_1 - 4x_2)$$

16. Show that linear transformations preserve linear dependence.

17. Find the rank of each matrix in **Problem 8** of **EXERCISE 3.2** by the method of **(6.42)**.

MATRIX OF A LINEAR TRANSFORMATION

(6.55) Definition. Let V and W be finite dimensional vector spaces over the same field F with $\dim V = n$ and $\dim W = m$. Let $B = \{v_1, v_2, \ldots, v_n\}$ and $E = \{w_1, w_2, \ldots, w_m\}$ be any bases for V and W respectively. Any vector v in V can be expressed in a unique way as a linear combination of $v_1, v_2, \ldots, v_n$. Thus

$$v = x_1v_1 + x_2v_2 + \ldots + x_n v_n, \quad x_i \in F, \; i = 1, 2, \ldots, n.$$

We call $(x_1, x_2, \ldots, x_n)$ the **coordinate vector** of v relative to the basis B.

Let $T : V \to W$ be a linear transformation. The image $T(v_1), T(v_2), \ldots, T(v_n)$ are elements of W and each can be expressed uniquely as a linear combination of the basis vectors $w_1, w_2, \ldots, w_m$. Therefore,

$$\begin{aligned}
T(v_1) &= a_{11}w_1 + a_{21}w_2 + \ldots + a_{m1}\,w_m \\
T(v_2) &= a_{11}w_1 + a_{21}w_2 + \ldots + a_{m1}\,w_m \\
&\vdots \\
T(v_j) &= a_{1j}w_1 + a_{2j}w_2 + \ldots + a_{mj}\,w_m \\
&\vdots \\
T(v_n) &= a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{mn}\,w_m
\end{aligned}$$

where $a_{ij} \in F$.

The $m \times n$ matrix whose jth column is the coordinate vector of $T(v_j)$ is called the matrix of T with respect to the bases B and E. Thus the matrix A of $T : V \to W$ relative to the bases B and E is

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

Conversely, if $A = [\, a_{ij} \,]$ is an $m \times n$ matrix with entries $a_{ij} \in F$, then A represents a linear transformation

$$T : F^n \longrightarrow F^m$$

defined by the equation

$$T(x) = Ax$$

where $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a column vector of F^n.

Example 36. Find the matrix of the linear transformation $T : R^3 \to R^4$ defined by

$$T(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3, x_1 - x_3, x_1)$$

with respect to the standard bases for R^3 and R^4.

Solution. Standard bases B and E for R^3 and R^4 are

$$B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} = \{e_1, e_2, e_3\}$$

$$E = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}.$$

$$\begin{aligned}
T(e_1) &= (1, 0, 1, 1) \\
&= 1(1, 0, 0, 0) + 0(0, 1, 0, 0) + 1(0, 0, 1, 0) + (0, 0, 0, 1) \\
T(e_2) &= (1, 1, 0, 0) \\
&= 1(1, 0, 0, 0) + 1(0, 1, 0, 0) + 0(0, 0, 1, 0) + 0(0, 0, 0, 1) \\
T(e_3) &= (0, 1, -1, 0) \\
&= 0(1, 0, 0, 0) + 1(0, 1, 0, 0) - 1(0, 0, 1, 0) + 0(0, 0, 0, 1)
\end{aligned}$$

The matrix A of T is

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}$$

Example 37. Let $T: R^3 \to R^4$ be the linear transformation defined by

$$T(x_1, x_2, x_3) = (x_2 + x_3, x_1 + x_2, x_1, x_1 - x_2)$$

Find the matrix of T with respect to the base $B = \{v_1, v_2, v_3\}$ for R^3 and the base $E = \{w_1, w_2, w_3, w_4\}$ for R^4 where

$$\begin{aligned}
v_1 &= (1, 1, 0), \quad v_2 = (1, 0, -1), \quad v_3 = (0, 1, 0) \\
w_1 &= (1, -1, 0, 0), \quad w_2 = (1, 0, 1, 0) \\
w_3 &= (0, 1, 0, 0), \quad w_4 = (0, 0, 1, 1)
\end{aligned}$$

Solution. From the definition of T, we have

$$\begin{aligned}
T(v_1) &= T(1, 1, 0) = (1, 2, 1, 0) \\
&= a_1 w_1 + a_2 w_2 + a_3 w_3 + a_4 w_4 \\
&= a_1(1, -1, 0, 0) + a_2(1, 0, 1, 0) + a_3(0, 1, 0, 0) + a_4(0, 0, 1, 1) \\
&= (a_1 + a_2, -a_1 + a_3, a_2 + a_4, a_4)
\end{aligned}$$

Comparing the components, we have

$$\begin{aligned}
a_1 + a_2 &= 1, & -a_1 + a_3 &= 2, \\
a_2 + a_4 &= 1, & a_4 &= 0
\end{aligned}$$

which give

$$a_1 = 0, \quad a_2 = 1, \quad a_3 = 2, \quad a_4 = 0$$

Hence,

$$T(v_1) = 0w_1 + w_2 + 2w_3 + 0w_4$$

Similarly,

$$\begin{aligned}
T(v_2) &= T(1, 0, -1) = (-1, 1, 1, 1) = -w_1 + 0w_2 + 0w_3 + w_4 \\
T(v_3) &= (0, 1, 0) = (1, 1, 0, -1) = 0w_1 + w_2 + w_3 - w_4
\end{aligned}$$

The matrix A of T with respect to the given bases is

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

Note that the columns of A are formed from the coefficients of w_1, w_2, w_4 in $T(v_1), T(v_2), T(v_3)$

Example 38. Let V be the vector space $P_3(x)$ of polynomials $p(x)$ in x with real coefficients of degree not exceeding 3 together with the zero polynomial. Let $T: V \to V$ be the linear transformation defined by

$$T(p(x)) = \frac{d}{dx}(p(x)), \quad \text{for all } p(x) \in V$$

Find the matrices of T with respect to the bases

(i) $\{1, x, x^2, x^3\}$ **(ii)** $\{1, 1+x, 1+x^2, 1+x^3\}$

Solution. (i) We have

$$\begin{aligned}
T(1) &= \frac{d}{dx}(1) = 0 = 0 \times 1 + 0 \times x + 0 \times x^2 + 0 \times x^3 \\
T(x) &= \frac{d}{dx}(x) = 1 = 1 \times 1 + 0 \times x + 0 \times x^2 + 0 \times x^3 \\
T(x^2) &= \frac{d}{dx}(x^2) = 2x = 0 \times 1 + 2x + 0 \times x^2 + 0 \times x^3 \\
T(x^3) &= \frac{d}{dx}(x^3) = 3x^2 = 0 \times 1 + 0 \times x + 3x^2 + 0 \times x^3
\end{aligned}$$

The matrix of T with respect to the given basis is

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (1)$$

(ii) Here as before, we obtain

$$\begin{aligned}
T(1) &= \frac{d}{dx}(1) = 0 \\
&= 0(1) + 0(1 + x) + 0(1 + x^2) + 0(1 + x^3)
\end{aligned}$$

$$T(1+x) = \frac{d}{dx}(1+x) = 1$$
$$= 1(1) + 0(1+x) + 0(1+x^2) + 0(1+x^3)$$
$$T(1+x^2) = \frac{d}{dx}(1+x^2) = 2x$$
$$= 2(x+1-1) = -2+2(1+x)$$
$$= -2(1) + 2(1+x) + 0(1+x^2) + 0(1+x^3)$$
$$T(1+x^3) = \frac{d}{dx}(1+x^3) = 3x^2$$
$$= 3(1+x^2-1) = -3+3(1+x^2)$$
$$= -3(1) + 0(1+x) + 3(1+x^2) + 0(1+x^3)$$

Hence the matrix of T with respect to the basis (ii) is

$$\begin{bmatrix} 0 & 1 & -2 & -3 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (2)$$

The reader should note that the matrix (2) is equivalent to the matrix (1).

Example 39. The matrix of a linear transformation $T : R^3 \to R^3$ is

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$$

Determine T in terms of coordinates of a vector $x \in R^3$.

Solution. Let $x = (x_1, x_2, x_3)$. T is given by

$$T(x) = Ax = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2 + 2x_3 \\ 0x_1 + x_2 + x_3 \\ -x_1 + x_2 + x_3 \end{bmatrix}$$

i.e.,
$$T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2 + 2x_3 \\ 0x_1 + x_2 + x_3 \\ -x_1 + x_2 + x_3 \end{bmatrix}$$

or
$$T(x_1, x_2, x_3) = (3x_1 + x_2 + 2x_3,\ 0x_1 + x_2 + x_3,\ -x_1 + x_2 + x_3).$$

EXERCISE 6.4

1. Find the matrix of each of the following linear transformations from R^3 to R^3 with respect to the standard basis for R^3

(i) $T(x_1, x_2, x_3) = (x_1, x_2, 0)$

(ii) $T(x_1, x_2, x_3) = (x_1 + x_2, -x_1 - x_2, x_3)$

(iii) $T(x_1, x_2, x_3) = (x_2, -x_1, -x_1)$

(iv) $T(x_1, x_2, x_3) = (x_1, x_2 + x_3, x_1 + x_2 + x_3)$

2. Find the matrix of each of the following linear transformations with respect to the standard bases of the given spaces

(i) $T : R \to R^2$ defined by $T(x) = (3x, 5x)$

(ii) $T : R^3 \to R^2$ defined by $T(x_1, x_2, x_3) = (3x_1 - 4x_2 + 9x_3, 5x_1 + 3x_2 - 2x_3)$

(iii) $T : R^2 \to R^4$ defined by $T(x_1, x_2) = (3x_1 + 4x_2, 5x_1 - 2x_2, x_1 + 7x_2, 4x_1)$

(iv) $T : R^4 \to R$ defined by $T(x_1, x_2, x_3, x_4) = 2x_1 + 3x_2 - 7x_3 + x_4$

3. Each of the following is the matrix of a linear transformation $T : R^n \to R^m$. Determine m, n and express T is terms of coordinates:

(i) $\begin{bmatrix} 6 & -1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 1 & 2 \\ 2 & 5 & 6 \\ -2 & 3 & -1 \end{bmatrix}$

(iii) $\begin{bmatrix} 3 & 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 & 1 \end{bmatrix}$

4. The matrix of a linear transformation $T: R^3 \to R^3$ is

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

Find T in terms of coordinates and its matrix with respect to the basis

$$v_1 = (0, 1, 2),\ v_2 = (1, 1, 1),\ v_3 = (1, 0, -2)$$

5. A linear transformation $T: R^2 \to R^3$ maps the vector $(1, 1)$ into $(0, 1, 2)$ and the vector $(-1, 1)$ into $(2, 1, 0)$. What matrix does T represent with respect to the standard bases for R^2 and R^3?

6. Let $T: M_{23} \to M_{32}$ be defined by $T(A) = A^T$, $A \in M_{23}$. Find the matrix of T relative to the standard bases for M_{23} and M_{32}.

(6)

Chapter 7

INNER PRODUCT SPACES

Inner Product Spaces form an important topic of Functional Analysis. These are simply vector spaces over the field F of real or complex numbers and with an inner product defined on them.

In this chapter we discuss some important properties of inner products spaces. We shall define a norm induced by an inner product. The concepts of orthogonal and orthonormal systems have been included in this chapter. We shall briefly describe the Gram-Schmidt orthonormalization process, eigenvalues and eigenvectors of matrices, similar matrices and orthogonal diagonalization of symmetric matrices.

DEFINITIONS AND EXAMPLES

(7.1) Definition. (Inner Product). Let V be a vector space over the field F of real or complex numbers. A mapping $< , >: V \times V \to F$ is said to be an **inner product on** V if the following conditions are satisfied.

(i) $<v, v> \geq 0$ and $<v, v> = 0$ if and only if $v = 0$, for all $v \in V$.

(ii) $<v_1, v_2> = \overline{<v_2, v_1>}$ for all $v_1, v_2 \in V$.

Here $\overline{<v_2, v_1>}$ denotes the complex conjugate of $<v_2, v_1>$.

(iii) $<av_1 + bv_2, v_3> = a<v_1, v_3> + b<v_2, v_3>$ for all $v_1, v_2, v_3 \in V$, $a, b \in F$.

If F is taken as the field of real numbers only then

$$\overline{<v_2, v_1>} = <v_2, v_1>,$$

so that condition (ii) becomes

(ii)' $<v_1, v_2> = <v_2, v_1>$ for all $v_1, v_2 \in V$

The pair $(V, <\ ,\ >)$, where V is a vector space over the field F of real or complex numbers and $<\ ,\ >$ an inner product on V, is called an inner product space.

Here-in-after we shall consider inner product spaces over R only.

Using condition (ii)' we have, for all $a, b \in R$ and $u, v, w \in V$,

(iii)' $<u, av + bw> = <av + bw, u>$, by (ii)'

$= a<v, u> + b<w, u>$, by (iii)

$= a<u, v> + b<u, w>$, by (ii)'

Example 1. Let $u, v \in R^n$, where

$$u = (u_1, u_2, \ldots, u_n),\quad v = (v_1, v_2, \ldots, v_n).$$

Then the dot product

$$<u, v> = u_1v_1 + u_2v_2 + \cdots + u_nv_n \qquad (1)$$

is an inner product on R^n. It is easy to verify conditions (i) – (iii) of Definition 7.1. The inner product is called the **Educlidean inner product** on R^n and will be frequently used in the following work.

R^n with this inner product is called the n-dimensional **Euclidean space**.

R^n will represent n-dimensional Eculidean space if no inner product is specified.

Example 2. Let V be the vector space of all $n \times 1$ matrices over R. $C_1, C_2 \in V$ are written as:

$$C_1 = [x_1\ x_2\ \ldots\ x_n]^T \quad\text{and}\quad C_2 = [y_1\ y_2\ \ldots\ y_n]^T \quad\text{etc.}$$

Such matrices are called **column vectors**.

Then, $C_1^T = [x_1\ x_2\ \ldots\ x_n]$. Define

$$<C_1, C_2> = \det(C_1^TC_2).$$

Now, $C_1^T C_2 = [x_1\ x_2\ \ldots\ x_n]\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \left[x_1 y_1 + x_2 y_2 + \cdots + x_n y_n\right]$.

Therefore,

$$<C_1, C_2> = \det(C_1^T C_2) = \det\left[x_1 y_1 + x_2 y_2 + \cdots + x_n y_n\right]$$
$$= x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

It is an inner product on the vector space V of all $n \times 1$ column vectors over the field of real numbers.

Example 3. Let $u, v \in R^2$, $u = (x_1, x_2)$, $v = (y_1, y_2)$

Then $<u, v> = x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2$ is an inner product on R^2

For all $u = (x_1, x_2) \in R^2$

(i)
$$<u, u> = x_1^2 - x_1x_2 - x_2x_1 + 3x_2^2$$
$$= x_1^2 - 2x_1x_2 + 3x_2^2$$
$$= (x_1 - x_2)^2 + 2x_2^2 \geq 0.$$

Also, $<u, u> = 0 \Leftrightarrow (x_1 - x_2)^2 = 0$ and $x_2^2 = 0$

$\Leftrightarrow x_1 = x_2 = 0 \Leftrightarrow u = 0$

(ii) Here, $<u, v> = x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2$
$$= y_1 x_1 - y_2 x_1 - y_1 x_2 + 3y_2 x_2$$
$$= y_1 x_1 - y_1 x_2 - y_2 x_1 + 3y_2 x_2$$
$$= <v, u>, \text{ for all } u, v \in R^2.$$

(iii) Let $w = (z_1, z_2)$. Then

$$<au + bv, w> = <(ax_1 + by_1, ax_2 + by_2), (z_1, z_2)>$$
$$= (ax_1 + by_1) z_1 - (ax_1 + by_1) z_2 - (ax_2 + by_2) z_1 + 3(ax_2 + by_2) z_2$$
$$= ax_1z_1 + by_1z_1 - ax_1z_2 - by_1z_2 - ax_2z_1 - by_2z_1 + 3ax_2z_2 + 3by_2z_2$$
$$= a(x_1 z_1 + x_1 z_2 - x_2 z_1 + 3x_2 z_2) + b(y_1 z_1 - y_1 z_2 - y_2 z_1 + 3y_2 z_2)$$
$$= a<u, w> + b<v, w>, \text{ for all } u, v, w \in R \text{ and } a, b \in F.$$

Thus all conditions of an inner product are satisfied.

So $(R^2, <\ ,\ >)$ is an inner product space.

Remarks. One may note that we defined an inner product on R^n in Example 1 by (1). For R^2 the Euclidean inner product is given by:

$$<u, v> = x_1 y_1 + x_2 y_2 \text{ where } u = (x_1, y_1),\ v = (x_2, y_2) \in R^2.$$

In Example 3 we defined an inner product on R^2 differently. Thus more than one inner products can be defined on a vector space over F.

Example 4. Let V be the vector space of all real-valued continuous functions on the interval $a \le t \le b$. Then for $f, g \in V$,

$$<f, g> = \int_a^b f(t)\ g(t)\,dt$$

is an inner product on V.

We leave the verifications as an exercise.

(7.2) Definition. (Norm or Length of a Vector). Let V be an inner product space and $v \in V$. Then the real number $\sqrt{<v, v>}$ is called the **norm** of v and is denoted by the symbol $\|v\|$.

If $\|v\| = 1$, i.e., if $<v, v> = 1$, then v is called a **unit vector** or is said to be a **normalized vector**.

Any nonzero vector $u \in V$ can be normalized by multiplying it by $\frac{1}{\|u\|}$. Thus,

$u = \frac{1}{\|u\|}\ u = \frac{u}{\|u\|}$ is a unit vector and we say that the vector u has been normalized, (as $\|u\| = \left\|\frac{1}{\|u\|}u\right\| = \frac{1}{\|u\|}\|u\| = 1$)

Example 5. Find the norm of $v = (3, 4) \in R^2$ with respect to the Euclidean inner product and the inner product defined in Example 3.

Solution. In the first case, $<v, v> = \|v\|^2 = 3 \times 3 + 4 \times 4 = 9 + 16 = 25$

Therefore, $\|v\| = \sqrt{25} = 5$.

In the second case,

$<v, v> = \|v\|^2 = 9 - 24 + 48 = 33$ so that

$\|v\| = \sqrt{33}$

Example 6. Normalize $v = (1, 2, 1) \in R^3$

Solution. Here $\|v\| = \sqrt{1+4+1} = \sqrt{6}$.

Thus the normalized vector $= \left(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$

(7.3) Theorem. (The Cauchy-Schwarz Inequality). Let u, v be elements of an inner product space V over R. Then

$$|<u, v>| \le \|u\|\ \|v\| \qquad (1)$$

Proof. If $v = 0$, then $<u, v> = <u, 0> = <u, 0w> = 0<u, w> = 0$ for all $w \in V$. So both sides of (1) are zero and the equality holds in (1).

Now let $v \ne 0$. Then, for all $t \in R$,

$0 \le \|u - tv\|^2 = <u - tv, u - tv>$

$= <u, u> - t<u, v> - t<v, u> + t^2<v, v>$

$= <u, u> - 2t<u, v> + t^2<v, v>$ as $<u, v> = <v, u>$ (2)

Let $t = \frac{<u, v>}{\|v\|^2}$. Then (2) becomes

$0 \le \|u\|^2 - \frac{2<u, v><u, v>}{\|v\|^2} + \frac{|<u, v>|^2}{\|v\|^4}\|v\|^2$, since $t^2 = |t|^2$

$$= \|u\|^2 - \frac{|<u, v>|^2}{\|v\|^2} \qquad (3)$$

Multiplying both sides of (3) by $\|v\|^2 > 0$, we have

$$0 \le \|u\|^2\ \|v\|^2 - |<u, v>|^2$$

That is,

$$|<u, v>|^2 \le \|u\|^2\ \|v\|^2 \qquad (4)$$

Taking square root of both the sides in (4), we get

$$|<u, v>| \le \|u\|\,\|v\|$$

as required.

(7.4) Theorem. The norm in an inner product space V satisfies the following axioms: For all $u, v \in V$ and $k \in R$

(i) $\|v\| \ge 0$ and $\|v\| = 0$ if and only if $v = 0$

(ii) $\|kv\| = |k|\,\|v\|$

(iii) $\|u + v\| \le \|u\| + \|v\|$ **(The Triangle Inequality)**

Proof. (i) Since $\|v\| = \sqrt{<v, v>}$ and $<v, v> \ge 0$,

$\|v\| \ge 0$ for all $v \in V$. $\|v\| \ge 0$ for all $v \in V$

Now $\|v\| = 0 \Leftrightarrow v = 0$. Thus (i) is proved.

(ii) Here $\| kv \|^2 = <kv, kv> = k^2 <v, v> = |k|^2 \| v \|^2$.

Taking square root of both sides, we get

$\| kv \| = |k| \| v \|$ for all $k \in R$ and $v \in V$, which proves (ii)

(iii) In this case

$$\begin{aligned}
\| u + v \|^2 &= <u + v, u + v> \\
&= <u, u> + <u, v> + <v, u> + <v, v> \\
&= \| u \|^2 + 2 <u, v> + \| v \|^2, \text{ since } <u, v> = <v, u> \\
&\le \| u \|^2 + 2 |<u, v>| + \| v \|^2 \\
&\le \| u \|^2 + 2 \| u \| \| v \| + \| v \|^2, \text{ (by the Cauchy-Schwarz Inequality)} \\
&= (\| u \| + \| v \|)^2.
\end{aligned}$$

Taking square root of both sides, we have

$\| u + v \| \le \| u \| + \| v \|$ for all $u, v \in V$,

proving (iii).

ORTHOGONALITY

(7.5) Definition. (Angle Between Two Vectors). Let V be an inner product space and $u, v \in V$. By the Cauchy-Schwarz inequality, we have

$$|<u, v>| \le \| u \| \| v \|, \text{ so that}$$

$$\frac{|<u, v>|}{\| u \| \| v \|} \le 1.$$

The real number θ defined by

$$\cos \theta = \frac{<u, v>}{\| u \| \| v \|}, \quad 0 \le \theta \le \pi$$

is called the angle between u *and* v.

The vectors u and v are said to be **orthogonal** if and only if $\theta = 90^\circ$, in which case $\frac{<u, v>}{\| u \| \| v \|} = 0$. So u and v are orthogonal if and only if $<u, v> = 0$.

If u is orthogonal to v then we write $u \perp v$ and read 'u is orthogonal to v'.

The relation of being orthogonal between vectors is clearly symmetric, for $<u, v> = 0 \Leftrightarrow <v, u> = 0$. Thus if u is orthogonal to v, then v is orthogonal to u.

The vector 0 is orthogonal to every $v \in V$, for $<0, v> = <0v, v> = 0 <v, v> = 0$.

Conversely, if u is orthogonal to every $v \in V$, then $<u, v> = 0$ for all $v \in V$ and in particular $<u, u> = 0$. That is $\| u \|^2 = 0$ so that $\| u \| = 0$. Thus $u = 0$ by (i) of (7.4).

Example 7. Let $x = (1, -1, 2)$ and $y = (-1, 1, 1) \in R^3$. Then

$$<x, y> = -1 - 1 + 2 = 0.$$

So $x \perp y$.

Similarly, $x = (1, -1, 1, -1)$, $y = (-1, 2, 2, -1) \in R^4$ are orthogonal.

Example 8. If u is orthogonal to v, then every scalar multiple of u is also orthogonal to v.

Solution. Here if $<u, v> = 0$ then $<ku, v> = k <u, v> = k0 = 0$, where $k \in R$. Therefore, ku is orthogonal to v.

Example 9. Find a unit vector orthogonal to both $(1, 1, 2)$ and $(0, 1, 3)$ in R^3.

Solution. Let $(x, y, z) \in R^3$ be a vector orthogonal to the given vectors.

Since $<(x, y, z), (1, 1, 2)> = 0$, we have $x + y + 2z = 0$ (1)

Again, $<(x, y, z), (0, 1, 3)> = 0 \Rightarrow y + 3z = 0$ (2)

From (2), $y = -3z$. Substituting into (1), we get

$$x - 3z + 2z = 0 \quad \text{i.e.} \quad x = z.$$

$$\text{Thus} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} z \\ -3z \\ z \end{bmatrix} = z \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$$

So $(1, -3, 1)$ is a vector satisfying

Now $\|(1, -3, 1)\| = \sqrt{1 + 9 + 1}$

The required unit vector is $\left(\frac{1}{\sqrt{11}}, \frac{-3}{\sqrt{11}}, \frac{1}{\sqrt{11}}\right)$.

Alternative Method:

The vector given by

$$\begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 1 & 2 \\ 0 & 1 & 3 \end{vmatrix} = e_1 - 3e_2 + e_3$$

is perpendicular to both the given vectors with $\frac{1}{\sqrt{11}}e_1 - \frac{3}{\sqrt{11}}e_2 + \frac{1}{\sqrt{11}}e_3$ as the unit vector.

Here $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$

Can we extend this method to higher dimensions?

(7.6) Definition. (Orthogonal Complement). Let W be a subset of an inner product space V over R. The **orthogonal complement** of W, denoted by $W^\perp$ and read as 'W perp', consists of those vectors in V which are orthogonal to every $w \in W$. Thus

$$W^\perp = \{v \in V : \langle v, w \rangle = 0 \quad \text{for all } w \in W\}.$$

$W^\perp$ is a subspace of V.

For, suppose $u, v \in W^\perp$. Then for any $a, b \in R$ and any $w \in W$, we have

$$\langle (au + bv), w \rangle = a\langle u, w \rangle + b\langle v, w \rangle = a\,0 + b\,0 = 0$$

Therefore $au + bv \in W^\perp$. Thus W is a subspace of V.

(7.7) Definition. (Orthogonal and Orthonormal Systems). A set $S = \{u_i : i \in I\}$ of vectors in an inner product space V over R is said to be an **orthogonal system** if its distinct vectors are orthogonal, i.e., if $\langle u_i, u_j \rangle = 0$ for all $u_i, u_j \in S$, $i \neq j$.

S is said to be an **orthonormal system** if $\langle u_i, u_j \rangle = \delta_{ij}$, where

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

Example 10. The set

$$\{e_1 = (1, 0, 0),\ e_2 = (0, 1, 0),\ e_3 = (0, 0, 1)\}$$

is an orthonormal system in R^3.

Similarly, the system $\{u_1 = (1, -1, 1, -1),\ u_2 = (3, 1, -1, 1),\ u_3 = (0, 2, 1, -1),\ u_4 = (0, 0, 1, 1)\}$ is an orthonormal system in R^4.

Example 11. Let V be the vector space of real-valued continuous functions on the interval $-\pi \le t \le \pi$ with inner product defined by

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)\, g(t)\, dt, \qquad \text{for all } f, g \in V. \text{ Then}$$

$$\{1, \cos t, \cos 2t, \ldots, \sin t, \sin 2t, \ldots\}$$

is an orthogonal system in V.

Here one has to prove that

$$\int_{-\pi}^{\pi} \cos mt \sin nt\, dt = 0 \quad \text{for all } m, n$$

$$\int_{-\pi}^{\pi} \sin mt \sin nt\, dt = \begin{cases} 0 & \text{if } m \neq n \\ \pi & \text{if } m = n \geq 1 \end{cases}$$

$$\int_{-\pi}^{\pi} \cos mt \cos nt\, dt = \begin{cases} 0 & \text{if } m \neq n \\ \pi & \text{if } m = n \geq 1 \\ 2\pi & \text{if } m = n = 0 \end{cases}$$

and is left as an exercise.

The following theorem gives a general method of finding a nonzero vector orthogonal to a given vector. **This method is due to** Arif Rafiq.

(7.8) Theorem. Let $v = x_1u_1 + x_2u_2 + \ldots + x_nu_n$ be a vector in an inner product space V with any basis $\{u_1, u_2, \ldots, u_n\}$. Then the vector

$$v^* = \begin{vmatrix} u_1 & u_2 \\ x_1 & x_2 \end{vmatrix} + \begin{vmatrix} u_2 & u_3 \\ x_2 & x_3 \end{vmatrix} + \cdots + \begin{vmatrix} u_{n-2} & u_{n-1} \\ x_{n-2} & x_{n-1} \end{vmatrix} + \begin{vmatrix} u_{n-1} & u_n \\ x_{n-1} & x_n \end{vmatrix} + \begin{vmatrix} u_n & u_1 \\ x_n & x_1 \end{vmatrix}$$

$$= (x_2 - x_n)\, u_1 + (x_3 - x_1)\, u_2 + \cdots + (x_n - x_{n-2})\, u_{n-1} + (x_1 - x_{n-1})\, u_n$$

is orthogonal to v provided $v^* \neq \mathbf{0}$. If $v^* = \mathbf{0}$ then the vector

$$v^* = (x_2 - x_1, 0, \ldots, 0)$$

is orthogonal to v.

Proof. Here

$$\langle v, v^* \rangle = x_1(x_2 - x_n) + x_2(x_3 - x_1) + \ldots + x_{n-1}(x_n - x_{n-2}) + x_n(x_1 - x_{n-1}) = 0.$$

So v^* is orthogonal to v. The second case is obvious.

(7.9) Theorem. Every orthonormal system $\{u_1, u_2, \ldots, u_n\}$ is linearly independent. Moreover, for all $v \in V$, the vector

$$w = v - \sum_{k=1}^{n} < v, u_k > u_k$$

is orthogonal to each u_i, $1 \le i \le n$.

Proof. Suppose that, for scalars $a_1, a_2, \ldots, a_n$,

$$a_1u_1 + a_2u_2 + \ldots + a_nu_n = 0.$$

Taking inner product of both sides with u_i and using the fact that $< u_k, u_i > = 0$ for all $k \ne i$, we have

$$\begin{aligned}
0 &= < 0, u_i > = < a_1u_1 + a_2u_2 + \ldots + a_n u_n, u_i > \\
&= a_1 < u_1, u_i > + \ldots + a_{i-1} < u_{i-1}, u_i > + a_i < u_i, u_i > \\
&\quad + a_{i+1} < u_{i+1}, u_i > + \ldots + a_n < u_n, u_i > \\
&= a_i \,.\, 1 = a_i \qquad \text{for all } i = 1, 2, \ldots, n.
\end{aligned}$$

Hence $\{u_1, u_2, \ldots, u_n\}$ is linearly independent.

Next to see that w is orthogonal to each u_i, $1 \le i \le n$, we ahve

$$\begin{aligned}
<w, u_i> &= < v - \sum_{k=1}^{n} < v, u_k > u_k \,, u_i > \\
&= < v, u_i > - < \sum_{k=1}^{n} < v, u_k > u_k, u_i > \\
&= < v, u_i > - \sum_{k=1}^{n} < v, u_k > < u_k, u_i > \\
&= < v, u_i > - < v, u_i >, \text{ because } <u_k, u_i> = \begin{cases} 0 & \text{for } k \ne i \\ 1 & \text{for } k = i \end{cases} \\
&= 0
\end{aligned}$$

Hence w is orthogonal to u_i, $1 \le i \le n$.

(7.10) The Gram[1]-Schmidt[2] Orthonormalization Process. Given a basis $\{v_1, \ldots, v_n\}$ of an inner product space V over R, an orthonormal basis $\{u_1, u_2, \ldots, u_n\}$ of V can be constructed as follows:

STEP 1. Let $u_1 = \dfrac{v_1}{\| v_1 \|}$

STEP 2. Let $w_2 = v_2 - < v_2 , u_1 > u_1$ and $u_2 = \dfrac{w_2}{\| w_2 \|}$

STEP 3. Let $w_3 = v_3 - < v_3, u_1 > u_1 - < v_3, u_2 > u_2$ and $u_3 = \dfrac{w_3}{\| w_3 \|}$

and so on

$$w_n = v_n - < v_n, u_1 > - < v_n, u_2 > u_2 - \ldots - < v_n, u_{n-1} > u_{n-1}$$

and

$$u_n = \frac{w_n}{\| w_n \|} \qquad w_n = v_n - \sum_{k=1}^{n-1} \langle v_n, e_k \rangle e_k.$$

The set $\{u_1, u_2, \ldots, u_n\}$ is orthonormal (hence linearly independent) and span $\{u_1, u_2, \ldots, u_n\}$ = span $\{v, v_2, \ldots, v_n\}$.

Example 12. Show that $\{(1, 1, 1), (0, 1, 1), (0, 0, 1)\}$ is a basis of R^3. Using Gram-Schmidt orthonormalization process, transform this basis into an orthonormal basis.

Solution. Since the matrix

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

is a nonsingular echelon matrix, the given three vectors of R^3 are linearly independent and so form a basis of R^3.

First Method (Gram-Schmidt Process)

Let $v_1 = (1, 1, 1)$, $v_2 = (0, 1, 1)$, $v_3 = (0, 0, 1)$

$$\text{Now } u_1 = \frac{v_1}{\| v_1 \|} = \frac{1}{\sqrt{3}}(1, 1, 1) = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

$$\begin{aligned}
w_2 &= v_2 - < v_2, u_1 > u_1 \\
&= (0, 1, 1) - \frac{2}{\sqrt{3}}\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \\
&= (0, 1, 1) - \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right) = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)
\end{aligned}$$

1. German mathematician (1876-1959).
2. Danish actuary (1850-1916).

$$u_2 = \frac{w_2}{\|w_2\|} = \frac{1}{\sqrt{\frac{4}{9}+\frac{1}{9}+\frac{1}{9}}}\left(-\frac{2}{3},\frac{1}{3},\frac{1}{3}\right)$$

$$= \frac{3}{\sqrt{6}}\left(-\frac{2}{3},\frac{1}{3},\frac{1}{3}\right) = \left(-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}}\right)$$

$$w_3 = v_3 - <v_3, u_1> u_1 - <v_3, u_2> u_2$$

$$= (0,0,1) - \frac{1}{\sqrt{3}}\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right) - \left(\frac{1}{\sqrt{6}}\right)\left(-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}}\right)$$

$$= (0,0,1) - \left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) - \left(-\frac{2}{6},\frac{1}{6},\frac{1}{6}\right) = \left(0,-\frac{1}{2},\frac{1}{2}\right)$$

$$u_3 = \frac{w_3}{\|w_3\|} = \frac{1}{\sqrt{\frac{1}{4}+\frac{1}{4}}}\left(0,-\frac{1}{2},\frac{1}{2}\right)$$

$$= \frac{2}{\sqrt{2}}\left(0,-\frac{1}{2},\frac{1}{2}\right) = \left(0,-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$$

Thus the orthonormal basis is

$$\left\{\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right),\left(-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}}\right),\left(0,-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)\right\}$$

Second Method:

The process described here is a special case of a general theorem of Arif Rafiq which gives another method to find an orthonormal system just from one given vector in R^n, $(n \geq 3)$, and in fact in any n-dimensional inner product space.

Take $v_1 = (0, 1, 1)$. Then, by Theorem 7.8,

$$v_1^{\bullet} = \begin{vmatrix} e_1 & e_2 \\ 0 & 1 \end{vmatrix} + \begin{vmatrix} e_2 & e_3 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} e_3 & e_1 \\ 1 & 0 \end{vmatrix}$$

$$= 0e_1 + e_2 - e_3 = (0, 1, -1)$$

is orthogonal to v_1 where $\{e_1, e_2, e_3\}$ is the standard basis of R^3.

Now take $v_2 = (0, 1, -1)$.

The vector

$$v_3 = \begin{vmatrix} e_1 & e_2 & e_3 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{vmatrix} = -2e_1 = (-2, 0, 0)$$

is then orthogonal to both v_1 and v_2. Thus the vectors

$$u_1 = \frac{v_1}{\|v_1\|} = \left(0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right); \quad u_2 = \frac{v_2}{\|v_2\|} = \left(0,\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)$$

and $u_3 = \frac{v_3}{\|v_3\|} = (-1, 0, 0)$

form an orthonormal basis for R^3.

Remarks:

1. The orthonormal sets obtained by the two methods are different. This shows that the orthonormal system are not uniquely determined.
2. In the second method we used only one vector from the gives set of vectors to obtain an orthonormal system.
3. By the second method, using any of the other two vectors yield the following sytems of orthonormal vectors.
   (i) For $v_1 = (0, 0, 1)$, the orthonormal system is
   $$(0,0,1), \left(0,\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)$$
   (ii) For $v_2 = (1, 1, -1)$ and in general, for such a vector in the higher dimensions, we choose the vector $v_2 = (1, -1, 0)$ as a vector perpendicular to v_1 and then find the third vector by the given method. In this case the third vector is $(1, 1, -2)$. So the orthonormal system is
   $$\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right),\left(\frac{1}{\sqrt{2}},\frac{-1}{\sqrt{2}},0\right),\left(\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{-2}{\sqrt{6}}\right)$$
4. With slight modifications the method extends to higher dimensions.

EXERCISE 7.1

1. Let $u = (u_1, u_2)$ and $v = (v_1, v_2)$ belong to R^2.
   (i) Verify that $< u, v > = u_1v_1 - 2u_1v_2 - 2u_2v_1 + 5u_2v_2$ is an inner product on R^2.
   (ii) For what value of k
   $$< u, v > = u_1 v_1 - 3u_1 v_2 - 3u_2 v_1 + ku_2 v_2$$
   is an inner product on R^2?

2. Find the norm of $(2, 3) \in R^2$ w.r.t.
   (i) The Euclidean inner product on R^2.
   (ii) The inner product of Problem 1 (i) above.

3. Find the norm of $\left(\frac{1}{2}, -\frac{1}{4}, \frac{1}{3}, \frac{1}{6}\right) \in R^4$ w.r.t. the Euclidean inner product on R^4.

4. Let V denote the vector space of 2×2 matrices over R. If $A, B \in V$ and $\text{Tr}(A)$ (called the **trace** of A) denotes the sum of the main diagonal entries of A, show that
   $$< A, B > = \text{Tr}(B^T A)$$
   is an inner product on V. Also find the norm of $A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$.

5. Let V be the vector space $P(x)$ of all polynomials over R. Show that
   $< f, g > = \int_0^1 f(t)\, g(t)\, dt$ defines an inner product on V.

6. Let $u_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, u_3 = \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}$.
   Using Example 2, find
   (i) the inner product of each pair of above column vectors
   (ii) the norm of each vector
   (iii) a vector orthogonal to u_1 and u_2
   (iv) a vector orthogonal to u_1 and u_3.

7. Show that (1, 1) (0, 1) is a basis of R^2. Using the Gram-Schmidt process, find an orthonormal basis of R^2.

8. Show that (1, 0, 1), (0, 1, 1), (0, 0, 1) is a basis of R^3. Using the Gram-Schmidt process, find an orthonormal basis of R^3 by taking
   (i) $u_1 = (0, 0, 1)$
   (ii) $u_1 = (1, 0, 1)$.

9. Show that $\{(1, -1, 0), (2, -1, -2), (1, -1, -2)\}$ is a basis of R^3. Find an orthonormal basis of R^3 using the Gram-Schmidt process.

10. (i) Let $\{e_1, e_2, \ldots, e_n\}$ be an orthonormal basis of an inner product space V over R. Show that for any $v \in V$,
    $$v = < v, e_1 > e_1 + < v, e_2 > e_2 + \ldots + < v, e_n > e_n$$
    (ii) If $T : V \to V$ is a linear transformation, show that $< T(e_j), e_i >$ is the ijth entry of the matrix representing T in the given basis $\{e_1, e_2, \ldots, e_n\}$.

11. Let W be a subspace of an inner product space V. Show that there is an orthonormal basis of W which is part of an orthonormal basis of V.

ORTHOGONAL MATRICES

(7.11) Definition. A square matrix A over R for which $A^T = A^{-1}$, or equivalently $AA^T = A^TA = I$, is called an **orthogonal matrix**. Clearly, an orthogonal matrix is nonsingular.

(7.12) Theorem. The following conditions for a square matrix A are equivalent:
(i) A is orthogonal.
(ii) The rows of A form an orthonormal set.
(iii) The columns of A form an orthonormal set.

Proof. (i) $\Leftrightarrow$ (ii). Let A [a_{ij}] be an $n \times n$ matrix. Suppose A is orthogonal, R_i and C_i, $1 \le i \le n$, denote the rows and columns of A respectively. Then $A^T = [\, b_{ij} \,]$ where $b_{ij} = a_{ji}$.

$$\begin{aligned} \text{Now} \quad AA^T &= [\, c_{ij} \,], \\ \text{where} \quad c_{ij} &= \sum_{k=1}^{n} a_{ik}\, b_{kj} = \sum_{k=1}^{n} a_{ik}\, a_{jk} \\ &= a_{i1}\, a_{j1} + a_{i2}\, a_{j2} + \cdots + a_{in}\, a_{jn} \\ &= < [a_{i1} \;\; a_{i2} \; \ldots \; a_{in}], [a_{j1} \;\; a_{j2} \; \ldots \; a_{jn}] > \\ &= < R_i, R_j >. \end{aligned}$$

Since A is orthogonal, $AA^T = I$. Thus

$$< R_i, R_j > = c_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

So the rows of A form an orthonormal set. Thus (i) $\Rightarrow$ (ii)

Conversely, if rows of A form an orthonormal set, then

$$< R_i, R_j > = c_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

So $\quad AA^T = [c_{ij}] = I \qquad$ Thus (ii) $\Rightarrow$ (i)

(i) $\Leftrightarrow$ (iii) Let $A^TA = [d_{ij}]$, then

$$d_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj} = \sum_{k=1}^{n} a_{ki}a_{kj} = \sum_{k=1}^{n} a_{kj}a_{ki}$$

$$= a_{1j}a_{1i} + a_{2j}a_{2i} + \ldots + a_{nj}a_{ni}$$

$$= < [a_{1j}\ a_{2j}\ \ldots\ a_{nj}]^T, [a_{1i}\ a_{2i}\ \ldots\ a_{ni}]^T >$$

$$= < C_j, C_i >$$

Since $\quad A^TA = < C_j, C_i > = d_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$

Thus columns of A form an orthonormal set. So (i) $\Rightarrow$ (iii)

Conversely, if columns of A form an orthonormal set, then

$A^TA = [d_{ij}] = [< C_j, C_i >] = I$. Therefore, (iii) $\Rightarrow$ (i)

Thus (i) $\Leftrightarrow$ (ii), (i) $\Leftrightarrow$ (iii), (ii) $\Leftrightarrow$ (iii) and the proof is complete.

Note: As (ii) $\Leftrightarrow$ (iii), it follows that A is orthogonal if and only if A^T is orthogonal.

Example 13. Show that the rows (columns) of the matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

form an orthonormal set.

Solution. Since the rows (columns) of A form an orthonormal set if and only if A is orthogonal, we have only to show that $AA^T = I$

Now $\quad AA^T = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$= \begin{bmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \cos\theta\sin\theta & 0 \\ \sin\theta\cos\theta - \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

Similarly, $\quad A^TA = I$.

Thus A is orthogonal and so rows (columns) of A form an orthonormal set.

Example 14. Find an orthogonal matrix A whose first row is $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$

Solution. Let $v_1 = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$ and $w_2 = (x, y, z)$ be orthogonal to v_1

Then $\quad < v_1, w_2 > = 0$

$$\Rightarrow \quad \frac{1}{3}x + \frac{2}{3}y + \frac{2}{3}z = 0 \text{ or } x + 2y + 2z = 0 \qquad (1)$$

If we take $\quad x = 0,\ 2y + 2z = 0$ or $z = -y$, then

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ y \\ -y \end{bmatrix} = y\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

So we may take $w_2 = (0, 1, -1)$.

Normalize w_2 to obtain $v_2 = \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ as second row of A.

Take $\quad w_3 = (x, y, z)$ so that w_3 is orthogonal to both v_1 and v_2

Now $\quad < v_1, w_3 > = 0 \Rightarrow \quad \frac{1}{3}x + \frac{2}{3}y + \frac{2}{3}z = 0 \qquad (2)$

$\qquad < v_2, w_3 > = 0 \Rightarrow \quad \frac{1}{\sqrt{2}}y - \frac{1}{\sqrt{2}}z = 0 \qquad (3)$

$$(2) \Rightarrow x + 2y + 2z = 0 \qquad (2')$$
$$(3) \Rightarrow y - z = 0 \qquad (3')$$
$$(3') \Rightarrow y = z$$
$$(2') \Rightarrow x + 4z = 0 \Rightarrow x = -4z$$

Thus $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -4z \\ z \\ z \end{bmatrix} = -z\begin{bmatrix} 4 \\ -1 \\ -1 \end{bmatrix}$

Thus we can take $w_3 = (4, -1, -1)$. Normalizing w_3, we get

$v_3 = \left(\frac{4}{\sqrt{18}}, \frac{-1}{\sqrt{18}}, \frac{-1}{\sqrt{18}}\right)$ as third row of A.

Hence $A = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{4}{3\sqrt{2}} & \frac{-1}{3\sqrt{2}} & \frac{-1}{3\sqrt{2}} \end{bmatrix}$

Note that v_3 could also be obtained as:

$$v_3 = \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{vmatrix} = \frac{-4}{3\sqrt{2}}e_1 + \frac{1}{3\sqrt{2}}e_2 + \frac{1}{3\sqrt{2}}e_3$$
$$= \left(\frac{-4}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}\right)$$

instead of solving the linear equations.

Note: The above matrix A is not unique.

Alternative Method

Let $R_1 = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$. A vector (row) orthogonal to R_1 is

$$R_1^* = \begin{vmatrix} e_1 & e_2 \\ \frac{1}{3} & \frac{2}{3} \end{vmatrix} + \begin{vmatrix} e_2 & e_3 \\ \frac{2}{3} & \frac{2}{3} \end{vmatrix} + \begin{vmatrix} e_3 & e_1 \\ \frac{2}{3} & \frac{1}{3} \end{vmatrix} = \frac{1}{3}e_2 - \frac{1}{3}e_3 = \left(0, \frac{1}{3}, \frac{-1}{3}\right)$$

and $$R'_3 = \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & \frac{1}{3} & -\frac{1}{3} \end{vmatrix} = -\frac{4}{9}e_1 + \frac{1}{9}e_2 + \frac{1}{9}e_3 = \left(\frac{-4}{9}, \frac{1}{9}, \frac{1}{9}\right)$$

Now $R_1 = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$, $R_2 = \frac{R_1^*}{\|R_1^*\|} = \left(0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$

and $R_3 = \frac{R'_3}{\|R'_3\|} = \left(\frac{-4}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}\right)$

give the required matrix as

$$A = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{-4}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} \end{bmatrix}$$

EXERCISE 7.2

1. If A is an orthogonal matrix, show that $\det A = 1$ or -1.
2. Find an orthogonal matrix whose first row is
$$\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$$
3. Find an orthogonal matrix whose first row is a multiple of $(1, 1, 1)$.
4. Find an orthogonal matrix whose first row is
$$\left(0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right).$$
5. Show that the products and inverses of orthogonal matrices are orthogonal. Hence show that orthogonal matrices form a group under multiplication.

EIGENVALUES AND EIGENVECTORS

(7.13) Definition. If A is an $n \times n$ matrix over R, then a scalar $\lambda \in R$ is called an **eigenvalue** of A if there exists a nonzero column vector $v \in R^n$ such that $Av = \lambda v$. In this case v is called an **eigenvector** of A corresponding to the eigenvalue λ.

Now $Av = \lambda v \Rightarrow Av = \lambda I v$, where I is identity matrix of the size of A.

$$\Rightarrow (A - \lambda I)v = 0.$$

Since $v \neq 0$, by Theorem **4.12**, $A - \lambda I$ is singular so that

$$\det(A - \lambda I) = |A - \lambda I| = 0. \qquad (1)$$

Remarks:

1. For the $n \times n$ matrix A, the equation
$$p(\lambda) = |A - \lambda I| = 0 \qquad (2)$$
is a polynomial equation of degree n in λ and so it has n roots
$$\lambda_1, \lambda_2, \cdots, \lambda_n$$
(some of these roots may have multiplicity > 1).

2. The polynomial $p(\lambda)$ can be written as
$$p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda) \qquad (3)$$
$$= (-1)^n[\lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + (-1)^n \lambda_1\lambda_2\cdots\lambda_n] \qquad (4)$$

3. For $\lambda = 0$, the equations (1) and (2) become:
$$p(0) = |A| = \det A = \lambda_1 \lambda_2 \lambda_3 \cdots \lambda_n$$

4. The coefficient of $(-1)^{n-1}\lambda^{n-1}$ is $\lambda_1 + \lambda_2 + \cdots + \lambda_n$ which is the sum $\sum_{i=1}^{n} a_{ii}$ of the main diagonal elements of the matrix A.

5. The equation
$$p(\lambda) = |A - \lambda I| = 0$$
is called the **characteristic equation of A.**
The roots of $p(\lambda) = 0$ are called the **eigenvalues** (or **proper values, characteristic values)** of A.

6. The vector v satisfying the equation $Av = \lambda v$ are called the **eigenvectors** or **characteristics vectors** of A, corresponding to the eigenvalue λ.

Example 15. Find the eigenvalues and eigenvectors of the matrix
$$A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$$

Solution. The characteristic equation of A is
$$p(\lambda) = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 \\ -2 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) + 2 = 0$$
or
$$\lambda^2 - 5\lambda + 6 = 0 \qquad (1)$$

The roots of (1) are
$$\lambda_1 = 2, \quad \lambda_2 = 3$$
which are eigenvalues of A.

The eigenvector of A corresponding to the eigenvalue $\lambda_1 = 2$ is given by:
$$Av_1 = 2v_1$$

That is
$$\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = 2\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \text{ where } v_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

Thus
$$x_1 + y_1 = 2x_1 \qquad (2)$$
$$-2x_1 + 4y_1 = 2y_1 \qquad (3)$$

From (2), $x_1 = y_1$, and (3) gives $y_1 = x_1$.

Take $x_1 = y_1 = 1$. Then $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector.

All other eigenvectors are scalar multiples of $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

The eigenspace E_2 of A corresponding to the eigvenvalue $\lambda_1 = 2$ is generated by $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. It is a subspace of R^2.

The eigenvector v_2 of A, corresponding to the eigenvalue $\lambda_2 = 3$, is given by:
$$Av_2 = 3v_2$$

That is

$$\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = 3\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

So

$$\left.\begin{aligned} x_2 + y_2 &= 3x_2, \\ -2x_2 + 4y_2 &= 3y_2 \end{aligned}\right] \Rightarrow y_2 = 2x_2 .$$

For $x_2 = 1, y_2 = 2$. $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is eigenvector for $\lambda_2 = 3$.

Other eigenvectors corresponding to $\lambda_2 = 3$ are all scalar multiples of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

The eigenspace E_3 corresponding to $\lambda_2 = 3$ is generated by $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Example 16. Find the eigenvalues and corresponding eigenvectors of the matrix

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$$

Solution. Here

$$A - \lambda I = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2-\lambda & 2 & 1 \\ 1 & 3-\lambda & 1 \\ 1 & 2 & 2-\lambda \end{bmatrix}$$

So,

$$\det(A - \lambda I) = -\lambda^3 + 7\lambda^2 - 11\lambda + 5.$$

Now $|A - \lambda I| = 0 \Rightarrow \lambda^3 - 7\lambda^2 + 11\lambda - 5 = 0.$

The roots of this equation are

$$\lambda_1 = 5, \quad \lambda_2 = 1, \quad \lambda_3 = 1.$$

When $\lambda = \lambda_1 = 5$, $(A - \lambda I)\,v = 0$, where $v = [x_1 \; x_2 \; x_3]^T$

$$\Rightarrow \begin{bmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Now $\begin{bmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{bmatrix} \underset{\sim}{R} \begin{bmatrix} 1 & -2 & 1 \\ -3 & 2 & 1 \\ 1 & 2 & -3 \end{bmatrix}$ by R_{12}

$\underset{\sim}{R} \begin{bmatrix} 1 & -2 & 1 \\ 0 & -4 & 4 \\ 0 & 4 & -4 \end{bmatrix}$ by $R_2 + 3R_1$ and $R_3 - R_1$

$\underset{\sim}{R} \begin{bmatrix} 1 & -2 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$ by $\frac{1}{4}R_2$ and $\frac{1}{4}R_3$

$\underset{\sim}{R} \begin{bmatrix} 1 & -2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ by $R_3 + R_2$

So $\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

which implies

$$x_1 - 2x_2 + x_3 = 0$$

and $$x_2 + x_3 = 0.$$

Let $x_3 = a$. Then $x_2 = -a$, $x_1 = -3a$.

Thus $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3a \\ -a \\ a \end{bmatrix} = -a\begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$

Hence an eigenvector corresponding to $\lambda = 5$ is $[3 \;\; 1 \;\; -1]^T$.

Scalar multiples of $[3 \;\; 1 \;\; -1]^T$ are all eigenvectors corresponding to $\lambda = 5$ and they form a subspace of R^3 generated by $[3 \;\; 1 \;\; -1]^T$ which is the eigenspace of A corresponding to $\lambda = 5$.

Now $\lambda_2 = 1$ and $\lambda_3 = 1$. In this case,

$$(A - I)\,v = 0 \Rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow \quad x_1 + 2x_2 + x_3 = 0.$

Let $x_2 = a$ and $x_3 = b$. Then $x_1 = -2a - b$

and
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2a - b \\ a \\ b \end{bmatrix} = a\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

In this case we get two linearly independent vectors $[-2 \;\; 1 \;\; 0]^T$ and $[-1 \;\; 0 \;\; 1]^T$ which are eigenvectors corresponding to $\lambda = 1$.

Any linear combination of $[-2 \;\; 1 \;\; 0]^T$ and $[-1 \;\; 0 \;\; 1]^T$ is also an eigenvector corresponding to $\lambda = 1$. These linear combinations form a subspace of R^3 with a basis $\{[-2 \;\; 1 \;\; 0]^T, [-1 \;\; 0 \;\; 1]^T\}$. This subspace is the **eigenspace** of A corresponding to $\lambda = 1$.

(7.14) Theorem. Nonzero eigenvectors of a matrix A corresponding to distinct eigenvalues are linearly independent.

Proof. Let $v_1, v_2, \ldots, v_n$ be nonzero eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ respectively.

We prove the theorem by induction on n.

If $n = 1$, then v_1 is linearly independent as $v_1 \neq 0$.

Suppose $n > 1$ and the result is true for all $k \leq n - 1$.

Let $\quad a_1v_1 + a_2v_2 + \ldots + a_nv_n = 0, \qquad (1)$

where $a_i \in R$.

Then $A(a_1v_1 + a_2v_2 + \ldots + a_nv_n) = A0 = 0$

$\Rightarrow \quad a_1Av_1 + a_2Av_2 + \ldots + a_nAv_n = 0$

$\Rightarrow \quad a_1\lambda_1v_1 + a_2\lambda_2v_2 + \ldots + a_n\lambda_nv_n = 0 \qquad (2)$

Multiplying (1) by λ_n, we get

$a_1\lambda_nv_1 + a_2\lambda_nv_2 + \ldots + a_n\lambda_nv_n = 0 \qquad (3)$

Subtracting (3) from (2), we have

$a_1(\lambda_1 - \lambda_n)\,v_1 + a_2(\lambda_2 - \lambda_n)\,v_2 + \cdots a_{n-1}(\lambda_{n-1} - \lambda_n)\,v_{n-1} = 0$

By induction hypothesis, $v_1, v_2, \ldots, v_{n-1}$ are linearly independent.

Therefore, $a_1(\lambda_1 - \lambda_n) = a_2(\lambda_2 - \lambda_n) = \ldots = a_{n-1}(\lambda_{n-1} - \lambda_n) = 0$

Since $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all distinct so $\lambda_i - \lambda_n \neq 0$, for $i = 1, 2, \ldots, n-1$.

Hence $\quad a_1 = a_2 = \ldots = a_{n-1} = 0$

So (1) becomes $a_nv_n = 0$. But $v_n \neq 0$. Therefore, $a_n = 0$.

Thus $v_1, v_2, \ldots, v_n$ are linearly independent.

(7.15) Theorem. If λ is an eigenvalue of an orthogonal matrix, then $|\lambda| = 1$.

Proof. Let A be an $n \times n$ orthogonal matrix and λ be an eigenvalue of A. Then there exists a nonzero column vector $v \in R^n$ such that

$Av = \lambda v \qquad (1)$

Now $\quad v^TA^T = \lambda v^T \qquad (2)$

From (1) and (2), we have

$v^TA^TAv = \lambda^2 v^Tv$

or $\quad v^TIv = \lambda^2 v^Tv, \quad$ since A is orthogonal, $A^TA = I$.

Thus $\quad (1 - \lambda^2)\,v^Tv = 0$

Since $v \neq 0$, $v^Tv \neq 0$, we have $1 - \lambda^2 = 0$ or $\lambda^2 = 1$

i.e., $\quad |\lambda|^2 = 1 \quad$ or $\quad |\lambda| = 1$.

(7.16) Theorem. Any two eigenvectors corresponding to two distinct eigenvalues of an orthogonal matrix are orthogonal.

Proof. Let A be an $n \times n$ orthogonal matrix. Suppose

$Av_1 = \lambda_1v_1$ and $Av_2 = \lambda_2v_2$ where $\lambda_1 \neq \lambda_2$.

Then $\quad v_2^TA^T = \lambda_2v_2^T$ and so

$v_2^TA^TAv_1 = \lambda_1\lambda_2v_2^Tv_1$

or $\quad v_2^TIv_1 = \lambda_1\lambda_2v_2^Tv_1, \quad$ since $A^TA = I$

Thus $\quad (1 - \lambda_1\lambda_2)\,v_2^Tv_1 = 0.$

By Theorem 7.15, $|\lambda_2| = 1$, so $\lambda_2^2 = 1$.

Thus $(\lambda_2^2 - \lambda_1\lambda_2)\, v_2^T v_1 = \boldsymbol{0}$

or $\lambda_2(\lambda_2 - \lambda_1)\, v_2^T v_1 = \boldsymbol{0}$.

Now $\lambda_2(\lambda_2 - \lambda_1) \neq 0$ because $\lambda_1 \neq \lambda_2$, $\lambda_2 \neq 0$, (since $|\lambda_2| = 1$).

Therefore, $v_2^T v_1 = \boldsymbol{0}$. In other words, $< v_2, v_1 > = 0$ and v_1, v_2 are orthogonal vectors.

EXERCISE 7.3

1. For each of the following matrices, find the characteristic polynomial, all eigenvalues and a basis of each eigenspace:

(i) $\begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$

(iii) $\begin{bmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{bmatrix}$ **(iv)** $\begin{bmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{bmatrix}$

(v) $\begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$ **(vi)** $\begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}$

2. Show that eigenvalues of a diagonal matrix are its diagonal elements and the eigenvectors are the standard basis vectors.

3. Show that A and A^T have the same eigenvalues. (A is a square matrix).

4. Show that an eigenvector of a square matrix cannot correspond to two distinct eigenvalues.

5. If λ is an eigenvalue of a nonsingular matrix A, then show that λ^{-1} is an eigenvalue of A^{-1}.

6. If A and B are square matrices, show that AB and BA have the same eigenvalues.

7. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of a square matrix of order n, then $k\lambda_1, k\lambda_2, \ldots, k\lambda_n$ where k is a scalar, are eigenvalues of kA.

8. Suppose v is an eigenvector of $n \times n$ matrices A and B. Show that v is also an eigenvector of $aA + bB$, where a and b are any scalars.

9. Let λ be an eigenvalue of a square matrix A. Let E_λ denote the set of all eigenvectors of A corresponding to eigenvalue λ. Show that E_λ is a subspace of V (E_λ is the **eigenspace** of A corresponding to λ)

10. For each of the following linear transformations $T: R^2 \to R^2$, find all eigenvalues and a basis for each eigenspace.

(i) $T(x, y) = (3x + 3y, x + 5y)$

(ii) $T(x, y) = (y, x)$

(iii) $T(x, y) = (y, -x)$

11. For each of the following operators $T: R^3 \to R^3$, find all eigenvalues and a basis for each eigenspace. Also reduce the matrix of T to a diagonal matrix.

(i) $T(x, y, z) = (x + y + z, 2y + z, 2y + 3z)$

(ii) $T(x, y, z) = (x + y, y + z, -2y - z)$

SIMILAR MATRICES

(7.17) Definition. If A and B are two $n \times n$ matrices over R, then A is said to be **similar to** B if there exists a nonsingular matrix P such that

$$B = P^{-1} AP.$$

Similarity of matrices is an equivalence relation on the set of all $n \times n$ matrices.

To see this, we first note that for the nonsingular identity matrix I, $A = I^{-1} AI$. So A is similar to A.

If A is similar to B then $B = P^{-1} AP$. But then, for some nonsingular matrix P^{-1}, we have $(P^{-1})^{-1} B P^{-1} = A$. Therefore, B is similar to A.

Again, let A be similar to B and B be similar to C. Then for nonsingular matrices P_1, P_2, we have

$$B = P_1^{-1} A P_1 \quad \text{and} \quad C = P_2^{-1} B P_2$$

So $$C = P_2^{-1} P_1^{-1} A P_1 P_2 = (P_1P_2)^{-1} A (P_1P_2)$$

That is, A is similar to C.

Thus, being similar is an equivalence relation on the set of $n \times n$ matrices.

(7.18) Theorem. Similar matrices have the same eigenvalues.

Proof. Let the matrix A be similar to B. Then there exists a nonsingular matrix P such that $B = P^{-1} AP$. So

$$|B-\lambda I| = |P^{-1}AP - \lambda I| = |P^{-1}AP - \lambda P^{-1}P|$$
$$= |P^{-1}(A-\lambda I)P|$$
$$= |P^{-1}||P||A-\lambda I|, \quad \text{since } |AB| = |A||B|$$
$$= |A-\lambda I|.$$

Thus A and B have the same characteristic polynomials and so they have the same eigenvalues. Hence the result.

(7.19) Theorem. If u is an eigenvector of $B = P^{-1}AP$ corresponding to the eigenvalue λ of B, then $v = Pu$ is an eigenvector of A corresponding to the same eigenvalue λ of A.

Proof. By the given hypothesis $Bu = \lambda u$ and $PB = AP$.

Therefore, $\quad Av = APu \quad$ (since $v = Pu$)
$$= PBu = P(\lambda u) = \lambda(Pu) = \lambda v.$$

Thus $v = Pu$ is an eigenvector of A corresponding to the eigenvalue λ.

(7.20) Theorem. Any square matrix A of order n similar to a diagonal matrix B with diagonal $(b_1, b_2, \ldots, b_n)$, where all b_i are distinct and nonzero, has n linearly independent eigenvectors.

Proof. Suppose $P^{-1}AP = \text{diag}(b_1, b_2, \ldots, b_n) = B$. The eigenvalues of B and A are the same. Since

$$B = \begin{bmatrix} b_1 & 0 & \ldots & 0 \\ 0 & b_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & b_n \end{bmatrix};$$

$$|B-\lambda I| = \det\begin{bmatrix} b_1-\lambda & 0 & \ldots & 0 \\ 0 & b_2-\lambda & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & b_n-\lambda \end{bmatrix}$$
$$= (b_1-\lambda)(b_2-\lambda)\ldots(b_n-\lambda) = 0.$$

So $\quad \lambda = b_1, b_2, \ldots, b_n$ are the n eigenvalues of B.

When $\quad \lambda = b_1$ and v is the eigenvector corresponding to $\lambda = b_1$, we have

$$(B-\lambda I)v = 0 \Rightarrow \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & b_2-b_1 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & b_n-b_1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\text{or} \quad \begin{bmatrix} 0 \\ (b_2-b_1)x_2 \\ (b_3-b_1)x_3 \\ \vdots \\ (b_n-b_1)x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

So $\quad (b_k - b_1)x_k = 0 \quad$ for $\quad k = 2, 3, \ldots, n$

Since $b_k \neq b_1$ for any $k = 2, 3, \ldots, n$, we have $x_k = 0$

Thus $[x_1, 0, 0, \cdots, 0]^T = x_1[1, 0, 0, \cdots, 0]^T$ is an eigenvector corresponding to $\lambda = b_1$.

or $\quad e_1 = [1\ 0\ 0\ \ldots\ 0]^T$ is an eigenvector corresponding to the eigenvalue b_1

Similarly, $e_2 = [0\ 1\ 0\ \ldots\ 0]^T$ is an eigenvector corresponding to $\lambda = b_2$, and so on, $e_n = [0\ 0\ \ldots\ 1]^T$ is an eigenvector corresponding to $\lambda = b_n$.

By Theorem 7.19, $e^T_j = Pe_j$ are eigenvectors of A for $j = 1, 2, \ldots, n$. Since P is nonsingular and $e_1, e_2, \ldots, e_n$ are linearly independent, so $e^T_1, e^T_2, \ldots, e^T_n$ are linearly independent eigenvectors of A.

(7.21) Theorem. An $n \times n$ matrix A has n linearly independent eigenvectors, if and only if, it is similar to a diagonal matrix.

Proof. Suppose that A has n linearly independent eigenvectors. We shall show that A is similar to a diagonal matrix.

Let the n linearly independent eigenvectors of A be the vectors $v_1, v_2, \ldots, v_n$ which correspond to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively. Each $v_i \in R^n$ and

$$Av_i = \lambda_i v_i, \quad (i = 1, 2, \ldots, n).$$

Let $\quad P = [v_1\ v_2\ \ldots\ v_n]$ (i.e., $v_1, v_2, \ldots, v_n$ are columns of P)

Then $\quad AP = [Av_1\ Av_2\ \ldots\ Av_n]$
$$= [\lambda_1 v_1\ \lambda_2 v_2\ \ldots\ \lambda_n v_n]$$
$$= [v_1\ v_2\ \ldots\ v_n]\begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$
$$= P\,\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Thus $\quad P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$

Conversely, suppose that A is similar to a diagonal matrix. Then there is a nonsingular $n \times n$ matrix P such that

$$P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) = D.$$

Since A and D are similar matrices, their eigenvalues are the same. Now the eigenvalues of D are given by

$$|D - \lambda I| = 0 = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\ldots(\lambda_n - \lambda).$$

Thus $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of D and so also of A. Therefore, A has n eigenvectors corresponding to these eigenvalues. By Theorem 7.14 these eigenvectors are linearly independent.

SYMMETRIC MATRICES

Recall that a square matrix A is symmetric if $A^T = A$.

We now state some theorems about symmetric matrices. The proofs of these theorems are beyond the scope of this book.

(7.22) Theorem. The eigenvalues of a symmetric matrix are all real.

Proof. Let λ be a nonzero eigenvalue of a symmetric matrix A so that

$$Av = \lambda v \tag{1}$$

for some eigenvector.

Suppose that λ is not real. Since λ is a root of the characteristic polynomial, and complex roots occur in conjugate pairs, $\bar{\lambda}$ is also an eigenvalue of A with an eigenvector $\bar{v}$. So

$$A\bar{v} = \bar{\lambda}\bar{v} \quad \text{i.e.,} \quad \bar{v}^T A = \bar{\lambda}\bar{v}^T, \tag{2}$$

because $A^T = A$.

Using (1) and (2), we ahve

$$\bar{v}^T Av = \bar{v}^T(Av) = \bar{v}^T(\lambda v) = \lambda(\bar{v}^T v). \tag{3}$$

and $$\bar{v}^T Av = (\bar{v}^T A)v = (\bar{\lambda}\bar{v}^T)v = \bar{\lambda}(\bar{v}^T v). \tag{4}$$

From (3) and (4), we get

$$\lambda(\bar{v}^T v) = \bar{\lambda}(\bar{v}^T v)$$

or $$(\lambda - \bar{\lambda})(\bar{v}^T v) = 0. \tag{5}$$

Now if $v = [x_1\ x_2\ \cdots\ x_n]^T$, then $\bar{v} = [\bar{x}_1\ \bar{x}_2\ \cdots\ \bar{x}_n]^T$

so that
$$\begin{aligned} \bar{v}^T v &= [\bar{x}_1 x_1 + \bar{x}_2 x_2 + \cdots + \bar{x}_n x_n] \\ &= [|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2] \\ &\neq 0, \quad \text{since } v \neq 0 \end{aligned}$$

Hence (5) gives

$$\lambda - \bar{\lambda} = 0 \qquad \text{That is } \lambda = \bar{\lambda}.$$

Thus λ is real.

(7.23) Theorem. If an $n \times n$ matrix A has only real eigenvalues then there is an orthogonal matrix P such that $P^T AP$ is an upper triangular matrix.

(7.24) Definition. A matrix A is said to be **orthogonally similar** to a matrix B if there is an orthogonal matrix P such that $P^T AP = B$.

(7.25) Theorem. An $n \times n$ real matrix A is orthogonally similar to a diagonal matrix if and only if A is symmetric.

(7.26) Theorem. Eigenvectors of a symmetric matrix A corresponding to distinct eigenvalues are orthogonal.

Proof. Let v_1, v_2 be eigenvectors of a symmetric matrix A corresponding to two distinct eigenvalues λ_1, λ_2. Then

$$Av_1 = \lambda_1 v_1, \quad Av_2 = \lambda_2 v_2$$

So $$v_2^T Av_1 = v_2^T(Av_1) = v_2^T \lambda_1 v_1 = \lambda_1 v_2^T v_1$$

and $$v_2^T Av_1 = (v_2^T A)v_1 = (\lambda_2 v_2^T)v_1 = \lambda_2 v_2^T v_1.$$

Hence $$\lambda_1 v_2^T v_1 = \lambda_2 v_2^T v_1$$

That is, $$(\lambda_1 - \lambda_2)v_2^T v_1 = 0.$$

Since $\lambda_1 \neq \lambda_2$, $v_2^T v_1 = 0$, i.e., $\langle v_2, v_1 \rangle = 0$. Thus v_1, v_2 are orthogonal.

Note: If A is an $n \times n$ symmetric matrix and has n distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, then the eigenvectors $v_1, v_2, \ldots, v_n$ of A corresponding to these eigenvalues are mutually orthogonal. Let $v_i^* = \frac{v_i}{\|v_i\|}$. Then $\{v_1^*, v_2^*, \ldots, v_n^*\}$ is an orthonormal set of eigenvectors of A. the matrix

$$P = [v_1^*\ v_2^*\ \ldots\ v_n^*]$$

whose columns are the vectors v_i^*, $i = 1, 2, \ldots, n$, is an orthogonal matrix such that

$$P^T AP = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

If a symmetric matrix A has a repeated eigenvalue then the eigenvectors corresponding to this eigenvalue may not be orthogonal to each other. We then have to use the Gram-Schmidt method or some other method to get an orthogonal set of eigenvectors to form the matrix P. This is illustrated in Example 18.

DIAGONALIZATION OF MATRICES

(7.27) Definition. An $n \times n$ matrix A is said to be **diagonalizable** if it is similar to a diagonal matrix. That is, A is diagonalizable if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. The matrix P is said to **diagonalize** A.

If P is an orthogonal matrix i.e. $P^T = P^{-1}$, and such that $P^{-1}AP = P^TAP$ is a diagonal matrix then A is called **orthogonally diagonalizable** and P is said to **orthogonally diagonalize** A.

Theorem 7.21 gives a necessary and sufficient condition for an $n \times n$ matrix to be diagonalizable.

In this section, we give an algorithm to diagonalize a matrix. However for symmetric matrices, we have

(7.28) Theorem. A square matrix A is orthogonally diagonalizable and has real eigenvalues if and only if, A is symmetric.

Proof of this theorem is beyond the scope of this text.

(7.29) An Algorithim for Diagonalization. Given an $n \times n$ matrix A, the following procedure will diagonalize A:

(i) Find the distinct n eigenvalues of A.

(ii) Find n linearly independent eigenvectors $v_1, v_2, \ldots, v_n$ corresponding to the eigenvalues of A.

(iii) Form the matrix P with v_i as columns, $i = 1, 2, \ldots, n$.

(iv) If P is orthogonal, then $P^TAP = P^{-1}AP$ is the required diagonal matrix. The diagonal entries being the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ corresponding to the eigenvectors $v_1, v_2, \ldots, v_n$ respectively.

The method will be illustrated by examples.

Example 16. Find a real orthogonal matrix P for which $P^{-1}AP$ is a diagonal matrix where

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

Solution. Eigenvalues of A are given by

$$|A - \lambda I| = 0.$$

Therefore, $\begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = 0$ or $(1-\lambda)^2 - 4 = 0$

Thus $1 - 2\lambda + \lambda^2 - 4 = 0$

$\Rightarrow \lambda^2 - 2\lambda - 3 = 0$

$\Rightarrow (\lambda - 3)(\lambda + 1) = 0$

$\Rightarrow \lambda = 3, \quad \lambda = -1.$

When $\lambda = 3$, $A - 3I = \begin{bmatrix} 1-3 & 2 \\ 2 & 1-3 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$

$(A - 3I)\,v = 0 \Rightarrow \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\Rightarrow -2x_1 + 2x_2 = 0$

and $2x_1 - 2x_2 = 0.$

Both of these equations give $x_1 - x_2 = 0$. So

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_1 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Thus $[1 \quad 1]^T$ is an eigenvector. We normalize it to get $\left[\frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}}\right]^T$.

When $\lambda = -1$ $(A + I)\,v = 0 \Rightarrow \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

i.e., $2x_1 + 2x_2 = 0$ or $x_1 + x_2 = 0.$

Therefore, $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_1 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$

Thus $[1 \quad -1]^T$ is another eigenvector. We normalize it to get $\left[\frac{1}{\sqrt{2}} \quad \frac{-1}{\sqrt{2}}\right]^T$.

Let $P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$. Then $P^{-1} = P^T = P$ and

$$P^{-1}AP=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\end{bmatrix}$$

$$=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\frac{3}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ \frac{3}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}=\begin{bmatrix}3 & 0\\ 0 & -1\end{bmatrix}$$

Note: In the above example we have two linearly independent eigenvectors.

Example 17. Find an orthogonal matrix P such that $P^T AP$ is diagonal matrix and whose diagonal elements are eigenvalues of A, where

$$A=\begin{bmatrix}7 & -2 & 1\\ -2 & 15 & -2\\ 1 & -2 & 7\end{bmatrix}$$

Solution. Here

$$0=|A-\lambda I|=\begin{vmatrix}7-\lambda & -2 & 1\\ -2 & 15-\lambda & -2\\ 1 & -2 & 7-\lambda\end{vmatrix}$$

or $\quad \lambda^3-29\lambda^2+250\lambda-672=0 \qquad (1)$

To find the roots of (1), we use the fact that each integral root of (1) divides the constant term 672. Now 672 is divisible by ±1, ±2, ±3, ±4, ±6, ±7, ±12, ±14, ±16, etc. It is easily checked by substitution into (1) that ±1, ±2, ±3, ±4 are not roots of (1).

Now, using synthetic division method and choosing 6 as a possible root of (1), we see that

	1	−29	250	−672
6		6	−208	672
	1	−23	42	0

Thus $\lambda-6$ is a factor of L.H.S. of (1) and so 6 is a root of (1). The remaining roots are the solutions of the equation

$2\lambda^2-3\lambda+42=0$, which are 7 and 16.

Now when $\lambda=6$,

$$(A-6I)\,v_1=\begin{bmatrix}1 & -2 & 1\\ -2 & 9 & -2\\ 1 & -2 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

So, using the indicated elementary row operations, we have

$$\begin{bmatrix}1 & -2 & 1\\ -2 & 9 & -2\\ 1 & -2 & 1\end{bmatrix}\underset{\sim}{R}\begin{bmatrix}1 & -2 & 1\\ 0 & 5 & 0\\ 0 & 0 & 0\end{bmatrix}\quad \text{by } R_2+2R_1 \text{ and } R_3-R_1$$

$$\underset{\sim}{R}\begin{bmatrix}1 & -2 & 1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}\quad \text{by } \tfrac{1}{5}R_2.$$

So $\quad (A-6I)\,v_1=\begin{bmatrix}0\\0\\0\end{bmatrix}\Rightarrow\begin{bmatrix}1 & -2 & 1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix},$

which gives

$x_1-2x_2+x_3=0$ and $x_2=0$. So $x_1=-x_3=a$, say and $x_2=0$.

Hence $v_1=a\begin{bmatrix}1\\0\\-1\end{bmatrix}$ is an eigenvector corresponding to $\lambda=6$.

Next, for the eigenvalue $\lambda=7$, we have

$$(A-7I)\,v_2=\begin{bmatrix}0 & -2 & 1\\ -2 & 8 & -2\\ 1 & -2 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

and $\begin{bmatrix}0 & -2 & 1\\ -2 & 8 & -2\\ 1 & -2 & 0\end{bmatrix}\underset{\sim}{R}\begin{bmatrix}1 & -2 & 0\\ -2 & 8 & -2\\ 0 & -2 & 1\end{bmatrix}$ by R_{13}

$$\underset{\sim}{R}\begin{bmatrix}1 & -2 & 0\\ 0 & 4 & -2\\ 0 & -2 & 1\end{bmatrix}\quad \text{by } R_2+2R_1$$

$$\underset{\sim}{R}\begin{bmatrix} 1 & -2 & 0 \\ 0 & 4 & -2 \\ 0 & 0 & 0 \end{bmatrix} \text{ by } R_3 + \frac{1}{2}R_2$$

$$\underset{\sim}{R}\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \text{ by } \frac{1}{4}R_2.$$

Therefore,

$$(A - 7I)\,v_2 = 0 \Rightarrow \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $x_1 - 2x_2 = 0,\ x_2 - \frac{1}{2}x_3 = 0$. That is, $x_1 = 2x_2 = x_3 = b$, say.

Thus $v_2 = b\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ is an eigenvector corresponding to $\lambda = 7$.

Lastly, for $\lambda = 16$, we have

$$(A - 16I)\,v_3 = \begin{bmatrix} -9 & -2 & 1 \\ -2 & -1 & -2 \\ 1 & -2 & -9 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

and $\begin{bmatrix} -9 & -2 & 1 \\ -2 & -1 & -2 \\ 1 & -2 & -9 \end{bmatrix} \underset{\sim}{R} \begin{bmatrix} 1 & -2 & -9 \\ -2 & -1 & -2 \\ -9 & -2 & 1 \end{bmatrix}$ by R_{13}

$$\underset{\sim}{R}\begin{bmatrix} 1 & -2 & -9 \\ 0 & -5 & -20 \\ 0 & -20 & -80 \end{bmatrix} \text{ by } R_2 + 2R_1 \text{ and } R_3 + 9R_1$$

$$\underset{\sim}{R}\begin{bmatrix} 1 & -2 & -9 \\ 0 & -5 & -20 \\ 0 & 0 & 0 \end{bmatrix} \text{ by } R_3 - 4R_2$$

$$\underset{\sim}{R}\begin{bmatrix} 1 & -2 & -9 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix} \text{ by } -\frac{1}{5}R_2.$$

Therefore,

$$(A - 16I)\,v_3 = 0 \Rightarrow \begin{bmatrix} 1 & -2 & -9 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $x_1 - 2x_2 - 9x_3 = 0,\ x_2 + 4x_3 = 0$

That is, $x_1 = 2x_2 + 9x_3,\ x_2 = -4x_3$. Let $x_3 = c$,

Then $x_1 = x_3 = c,\ \ x_2 = -4c$ and

$$v_3 = c\begin{bmatrix} 1 \\ -4 \\ 1 \end{bmatrix} \text{ is an eigenvector corresponding to } \lambda = 16.$$

The vectors v_1, v_2, v_3 are orthogonal. We normalize these vectors. Thus

$$\frac{v_1}{\|v_1\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{bmatrix}^T, \qquad \frac{v_2}{\|v_2\|} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}^T$$

and $\dfrac{v_3}{\|v_3\|} = \begin{bmatrix} \frac{1}{\sqrt{18}} & \frac{-4}{\sqrt{18}} & \frac{1}{\sqrt{18}} \end{bmatrix}^T$

The orthogonal matrix is

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{2}{3} & \frac{1}{\sqrt{18}} \\ 0 & \frac{1}{3} & \frac{-4}{\sqrt{18}} \\ \frac{-1}{\sqrt{2}} & \frac{2}{3} & \frac{1}{\sqrt{18}} \end{bmatrix}$$

$$P^TAP = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{\sqrt{18}} & \frac{-4}{\sqrt{18}} & \frac{1}{\sqrt{18}} \end{bmatrix} \begin{bmatrix} 7 & -2 & 1 \\ -2 & 15 & -2 \\ 1 & -2 & 7 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{2}{3} & \frac{1}{\sqrt{18}} \\ 0 & \frac{1}{3} & \frac{-4}{\sqrt{18}} \\ \frac{-1}{\sqrt{2}} & \frac{2}{3} & \frac{1}{\sqrt{18}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{6}{\sqrt{2}} & 0 & \frac{-6}{\sqrt{2}} \\ \frac{14}{3} & \frac{7}{3} & \frac{14}{3} \\ \frac{16}{\sqrt{18}} & \frac{-64}{\sqrt{18}} & \frac{16}{\sqrt{18}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{2}{3} & \frac{1}{\sqrt{18}} \\ 0 & \frac{1}{3} & \frac{-4}{\sqrt{18}} \\ \frac{-1}{\sqrt{2}} & \frac{2}{3} & \frac{1}{\sqrt{18}} \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 16 \end{bmatrix}$$

Example 18. (The Case of Repeated Roots.)

For the matrix

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}$$

find a matrix P such that P^TAP is a diagonal matrix.

Solution. First we find the eigenvectors of A. The characteristic polynomial of A is

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 & -2 \\ 2 & 1-\lambda & -2 \\ -2 & -2 & 1-\lambda \end{vmatrix} = -\lambda^3 + 3\lambda^2 + 9\lambda + 5$$

So the eigenvalues of A are roots of the equation

$$\lambda^3 - 3\lambda^2 - 9\lambda - 5 = 0. \qquad (1)$$

By inspection, -1 is a root of the equation (1). By synthetic division, we have

	1	−3	−9	−5
−1		−1	4	5
	1	−4	−5	0

so that the remaining roots are the solutions of $\lambda^2 - 4\lambda - 5 = 0$. These are -1 and 5. Thus -1 is a repeated root of (1).

Now the eigenvector corresponding to 5 is the solution vector of

$$(A - \lambda I)\,v = 0, \qquad \lambda = 5.$$

Here $A - 5I = \begin{bmatrix} -4 & 2 & -2 \\ 2 & -4 & -2 \\ -2 & -2 & -4 \end{bmatrix} \underset{\sim}{R} \begin{bmatrix} 2 & -4 & -2 \\ -4 & 2 & -2 \\ -2 & -2 & -4 \end{bmatrix}$ by R_{12}

$$\underset{\sim}{R} \begin{bmatrix} 2 & -4 & -2 \\ 0 & -6 & -6 \\ 0 & -6 & -6 \end{bmatrix} \text{ by } R_2 + 2R_1 \text{ and } R_3 + R_1$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \text{ by } \frac{1}{2}R_1, -\frac{1}{6}R_2 \text{ and } -\frac{1}{6}R_3$$

$$\underset{\sim}{R} \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \text{ by } R_3 - R_2.$$

Hence $(A - 5I)\,v = 0 \Rightarrow \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

So v is given by:

$$x_1 - 2x_2 - x_3 = 0 \text{ and } x_2 + x_3 = 0$$

Take $x_3 = -a$ Then $x_2 = a$ and $x_1 = 2a - a = a$ That is,

$$v'_1 = \begin{bmatrix} a \\ a \\ -a \end{bmatrix} = a\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

For the eigenvectors corresponding to the eigenvalue -1, we reduce the matrix

$$A - \lambda I = A + I = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 2 & -2 \\ -2 & -2 & 2 \end{bmatrix}$$

to the echelon form as follows

$$\begin{bmatrix} 2 & 2 & -2 \\ 2 & 2 & -2 \\ -2 & -2 & 2 \end{bmatrix} \underset{\sim}{R} \begin{bmatrix} 2 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{by } R_2 - R_1 \text{ and } R_3 + R_1$$

So the eigenvector corresponding to $\lambda = -1$ is given by

$$\begin{bmatrix} 2 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

i.e., $2x_1 + 2x_2 - 2x_3 = 0$ or $x_1 + x_2 - x_3 = 0$.

Let $x_1 = a$, $x_2 = b$. Then $x_3 = a + b$. So

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a \\ b \\ a+b \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Although v'_1 is orthogonal to both $v'_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $v'_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, v'_2 and v'_3 are not orthogonal to each other. So we determine a vector orthogonal to v'_2, as follows. (See also an alternative method at the end of the solution).

A vector in the subspace $< v'_2, v'_3 >$ generated by v'_2 and v'_3 is of the form $av'_2 + v'_3 = v''_3$, say, $a \in F$.

We determine a by the condition that v'_2 an v''_3 are orthogonal. So

$$\begin{aligned} < v'_2, v''_3 > &= < v'_2, av'_2 + v'_3 > \\ &= a < v'_2, v'_2 > + < v'_2, v'_3 > \\ &= a \det\left[v_2'^T v'_2 \right] + \det\left[v_2'^T v'_3 \right] \\ &= a \det\left\{ [1 \;\; 0 \;\; 1]\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\} + \det\left\{ [1 \;\; 0 \;\; 1]\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} \\ &= 2a + 1 \end{aligned}$$

Since v'_2 and v''_3 are orthogonal, we have $2a + 1 = 0$ which gives $a = -\frac{1}{2}$

Thus $v''_3 = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$

Take $v_3 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = 2v''_3$

Then $v_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = v'_2$ and $v_3 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$ are the required orthogonal eigenvectors.

The orthogonal matrix P is formed by normalizing the orthogonal eigenvectors v_1, v_2, v_3. Thus

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

$$P^T AP = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{5}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{-5}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$ is the required diagonal matrix.

Also if B is the matrix formed by v_1, v_2, v_3; then

$$B^{-1}AB = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 2 \\ -1 & 1 & 1 \end{bmatrix}$$

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

is the required diagonal matrix.

Note: For the matrix B formed by the orthogonal eigenvectors v_1, v_2, v_3,

i.e., $$B = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 2 \\ -1 & 1 & 1 \end{bmatrix}.$$

It is easy to see that $B^T \neq B^{-1}$

So, in general, the matrix formed by an orthognonal set of eigenvectors may not be an orthogonal matrix.

Also, it is not necessary to orthonormalize the matrix formed by the orthogonal eigenvectors. Just form the matrix P whose columns are the orthogonal eigenvectors then find $P^{-1}AP$ to diagonalise A. The answer may be different because of det P being ignored. But this will save tedious calculations.

Moreover, changing the order of the eigenvectors also changes the required diagonal matrix. **The Examiner has to take care while marking such questions.**

EXERCISE 7.4

1. For each of the following symmetric matrices, find the eigenvalues, the eigenvectors and an orthogonal matrix P for which $P^T AP$ is diagonal. Find also the diagonal matrix.

(i) $A = \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}$ (ii) $A = \begin{bmatrix} 5 & 4 \\ 4 & -1 \end{bmatrix}$

(iii) $A = \begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$

2. Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. Find a (real) orthogonal matrix P for which $P^T AP$ is daigonal.

3. For each of the following symmetric matrices, find the eigenvalues, eigenvectors and an orthogonal matrix P for which $P^T AP$ is diagonal and whose diagonal elements are eigenvalues of A:

(i) $\begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}$ (ii) $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$

(iii) $\begin{bmatrix} 2 & -4 & -2 \\ -4 & 2 & -2 \\ 2 & -2 & -1 \end{bmatrix}$ (iv) $\begin{bmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix}$

4. Find the eigenvalues and eigenvectors of each of the following matrices. Find P such that $P^T AP$ is a diagonal matrix.

(i) $A = \begin{bmatrix} a & a \\ a & a \end{bmatrix}$ (ii) $A = \begin{bmatrix} a & a & a \\ a & a & a \\ a & a & a \end{bmatrix}$

5. Show that the eigenvalues of the matrix

$$A = \begin{bmatrix} a & a & a & a \\ a & a & a & a \\ a & a & a & a \\ a & a & a & a \end{bmatrix} \text{ are}$$

$\lambda = 4a, 0, 0, 0$ and the corresponding eigenvectors are $[\,1\;1\;1\;1\,]^T, [\,1\;-1\;0\;0\,]^T$, $[\,1\;0\;-1\;0\,]^T, [\,1\;0\;0\;-1\,]^T$. Form the matrix P with columns as these vectors and show that

$$P^{-1} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{-3}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{-3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{-3}{4} \end{bmatrix}$$

Also check that

$$P^{-1}AP = \begin{bmatrix} 4a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Chapter 8

1, 2, 3, ..., n, ...

INFINITE SERIES

Infinite series are of great importance in both pure and applied mathematics. They play a significant role in physics and engineering. In fact many functions can be represented by infinite series.

The theory of infinite series is developed through the use of a special kind of function called sequence.

SEQUENCES

(8.1) Definition. Let X be a nonempty set. A function $f : N \to X$, whose domain is the set of natural numbers, is called an **infinite sequence** in X. If the domain of f is the finite set of numbers $\{1, 2, 3, \cdots, n\}$ then it is called a **finite sequence**. The range of a sequence may be a subset of real or complex numbers. A sequence is called a **real sequence** if its range is a subset of real numbers. In what follows, only real sequences will be studied.

Let $a : N \to R$ be a sequence. The values (images) of the function a at the points 1, 2, 3, $\cdots$ will be denoted by $a_1, a_2, a_3, \cdots$ instead of $a(1), a(2), a(3), \cdots$. Although a sequences is a function but it is customary that the range

$$a_1, a_2, a_3, \cdots, a_n, \cdots$$

of a sequence $a : N \to R$ is often called a **sequence**. $a_1, a_2, a_3, \cdots, a_n, \cdots$ are called **terms** of the sequence. A convenient device to specify a sequences is to state a formula for its nth term. The sequence

$$1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots$$

may be specified by writing its nth term $a_n = \frac{1}{n}$ and the sequence is symbolically written as $\left\{\frac{1}{n}\right\}$. The sequence $a_1, a_2, a_3, \cdots a_n \cdots$ is written in bracket notation as $\{a_n\}$

A sequence $\{a_n\}$ is said to **converge** if its nth term approaches a definite number as n increases without bound. Formally, a sequence $\{a_n\}$ is said to have the limit L, if given any $\varepsilon > 0$, there exists a positive integer N such that

$$|a_n - L| < \varepsilon \qquad \text{for all} \quad n \geq N.$$

This is equivalent to saying that all terms of the sequence, when $n \geq N$, belong to the open interval

$$]L - \varepsilon, L + \varepsilon[$$

i.e., $a_N, a_{N+1}, a_{N+2}, \cdots$ lie in the interval $]L - \varepsilon, L + \varepsilon[$

If a sequence $\{a_n\}$ has limit L, we say that the sequence $\{a_n\}$ **converges** to L (or the sequence is **convergent** and has the limit L) and write symbolically as $a_n \to L$ when $n \to \infty$.

As in the case of a function, if a sequence converges then its limit is unique.

The sequence $\{a_n\} = \left\{\frac{1}{n}\right\}$ converges to 0, for given $\varepsilon = 0.01$, we note that

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} < 0.01 \qquad \text{if } n \geq 101$$

Thus all the terms of the sequence after the 100th term lie in $]-0.01, 0.01[$ and the requirements of the definition of limit are satisfied.

(8.2) Divergent Sequence. A sequence that does not have a definite number as its limit is said to diverge or a divergent sequence.

A sequence $\{a_n\}$ is said **diverge** to ∞ as $n \to \infty$ if corresponding to any positive real number G, however large, there is a positive integer N such that

$$a_n > G \qquad \text{for all} \quad n > N$$

In this case we write $a_n \to \infty$ as $n \to \infty$.

Similarly, $\{a_n\}$ is said to diverge to $-\infty$ if, for any negative number G (however large in magnitude), there is a positive integer N such that

$$a_n > G \qquad \text{for all} \quad n > N.$$

In symbols, we write $a_n \to -\infty$ as $n \to \infty$.

(8.3) Properties of Sequences.

1. Two sequences $\{a_n\}$ and $\{b_n\}$ are said to be equal if and only if $a_n = b_n$ for each positive integer n

2. If $\{a_n\}$ and $\{b_n\}$ are convergent sequences such that $a_n \to a$ and $b_n \to b$ and c is a constant, then

(i) $\lim\limits_{n \to \infty} c = c$ (ii) $\lim\limits_{n \to \infty} c\, a_n = ca$

(iii) $\lim\limits_{n \to \infty} (a_n \pm b_n) = \lim\limits_{n \to \infty} a_n \pm \lim\limits_{n \to \infty} b_n = a \pm b$

(iv) $\lim\limits_{n \to \infty} (a_n\, b_n) = \lim\limits_{n \to \infty} a_n \lim\limits_{n \to \infty} b_n = ab$

(v) $\lim\limits_{n \to \infty} \left(\frac{a_n}{b_n}\right) = \dfrac{\lim\limits_{n \to \infty} a_n}{\lim\limits_{n \to \infty} b_n} = \frac{a}{b}$ (provided that $b \neq 0$)

3. If $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ are sequences such that $a_n \leq b_n \leq c_n$ for all n and if $\lim\limits_{n \to \infty} a_n = L = \lim\limits_{n \to \infty} c_n$, then $\lim\limits_{n \to \infty} b_n = L$ **(Sandwich Theorem)**

4. Let $\{a_n\}$ be a sequence and f a continuous function defined on $[1, \infty[$ such that $f(n) = a_n$. Then

$$\lim_{n \to \infty} a_n = \lim_{x \to \infty} f(x)$$

(8.4) Definition. A subset X of real (or complex) numbers is said to be **bounded**, if for some positive $M \in R$, $|x| \leq M$ for all $x \in X$. A sequence $\{a_n\}$ is said to be **bounded** if and only if there is a positive number M such that

$$|a_n| \leq M \qquad \text{for all } n.$$

i.e., the set $A = \{a_1, a_2, \cdots, a_n, \cdots\}$, whose elements are the terms of the sequence, is a bounded set.

(8.5) Theorem. A convergent sequence is bounded.

Proof. Let a sequence $\{a_n\}$ be such that $a_n \to a$. By definition, for given $\varepsilon > 0$, there is a positive integer N such that.

$$|a_n - a| < \varepsilon \qquad \text{for all} \quad n \geq N.$$

Since $\quad |a_n| - |a| \leq |a_n - a|$,

we have $\quad |a_n| - |a| < \varepsilon \qquad \text{for all} \quad n \geq N.$

Let $M = \max \{ |a_1|, |a_2|, \cdots |a_{N-1}|, \varepsilon + |a| \}$. Then

$|a_n| \leq M$ for all n and so $\{a_n\}$ is a bounded sequence.

Note. The converse of this theorem is false. A bounded sequence need not be a convergent sequence.

Consider the sequence

$$\{a_n\} = \{1 + (-1)^n\}.$$

Clearly, $|a_n| \leq 2$ for all n and so $\{a_n\}$ is bounded.

Here $\lim\limits_{n \to \infty} a_n = 0$ if n is odd

and $\lim\limits_{n \to \infty} a_n = 2$ if n is even.

Thus the limit is not unique and so $\lim\limits_{n \to \infty} a_n$ does not exist. Hence $\{a_n\}$ is not convergent. By definition, it is a divergent sequence.

(8.6) A Divergence Test. An unbounded sequence is divergent.

As already seen above, a bounded sequence may also be divergent.

(8.7) Definition (Monotonic Sequences). A sequence $\{a_n\}$ is said to be

(i) **nondecreasing if** $a_n \leq a_{n+1}$ for all n

(ii) **nonincreasing if** $a_n \geq a_{n+1}$ for all n

(iii) **strictly increasing if** $a_n < a_{n+1}$ for all n

(iv) **Strictly decreasing if** $a_n > a_{n+1}$ for all n

All such sequences are called **monotone (or monotonic)**.

To check the monotonicity of a sequence $\{a_n\}$, we also make use of the following.

(8.8) Theorem. If a differentiable function f defined on $[1, \infty[$ is such that $f(n) = a_n$, then $\{a_n\}$ is increasing or decreasing according as f is increasing or decreasing on $[1, \infty[$.

Consider the sequence $\{a_n\}$ whose nth terms is $\frac{\ln n}{n}$. The function f defined on $[1, \infty[$ by

$$f(x) = \frac{\ln x}{x}, \quad x \geq 1$$

is differentiable and $\quad f(n) = \frac{\ln n}{n}, \quad n \geq 1.$

Now $f'(x) = \frac{1 - \ln x}{x^2} < 0$ because, for all $x \geq 3$, $\ln x > 1$. Thus, by a theorem of calculus, f is a decreasing function on $[3, \infty[$. Hence $\frac{\ln n}{n}$ is a decreasing function on $[3, \infty[$. So $\left\{\frac{\ln n}{n}\right\}$ is a decreasing sequence except for the first two terms.

It may be noted that $a_1 = 0$ and $a_2 = \frac{\ln 2}{2} = 0.35$, so that the sequence is increasing for the first two terms. But it is not a monotonic sequence.

We have already seen in (8.5) that a convergent sequence is bounded, but the converse need not be true. However, for monotonic sequences, we have

(8.9) Theorem. A bounded monotonic sequence is convergent.

Proof. Let $\{a_n\}$ be a bounded nondecreasing sequence i.e.,

$$a_1 \leq a_2 \leq a_3 \leq \cdots \leq a_n \leq \cdots$$

Since the sequence is bounded, the set

$$A = \{ a_1, a_2, \cdots, a_n, \cdots \},$$

whose elements are terms of the sequence, is a bounded set. a_1 is a lower bound of A. Since A is also bounded above, it has the supremum. Let M be the supremum of A. $a_n \leq M$ for all n and any number smaller than M is not an upper bound of A. That is, for $\varepsilon > 0$, $M - \varepsilon$ is not an upper bound of A. So for some natural number N, $M - \varepsilon < a_N$. Since $\{a_n\}$ is a nondecreasing sequence, we have

$$M - \varepsilon < a_N \leq a_{N+1} \leq a_{N+2} \leq \ldots \leq M < M + \varepsilon$$

or $\quad M - \varepsilon < a_n < M + \varepsilon \quad$ for all $n \geq N$.

or $\quad |a_n - M| < \varepsilon \quad$ for all $n \geq N$.

Hence, by definition, $a_n \to M$ as $n \to \infty$.

Similarly, a bounded nonicreasing sequence $\{b_n\}$ converges and its limit is the infimum of the set

$$B = \{b_1, b_2, \cdots, b_n, \cdots\}.$$

Example 1. Find the limit of the sequence $\left\{\dfrac{\ln n}{n}\right\}$ as $n \to \infty$.

Solution. We have

$$\frac{\ln n}{n} = \frac{1}{n}\int_1^n \frac{dt}{t} \le \frac{1}{n}\int_1^n \frac{dt}{\sqrt{t}} = \frac{2}{n}\left[\sqrt{t}\right]_1^n$$

$$= \frac{2}{n}(\sqrt{n}-1) = 2\left(\frac{1}{\sqrt{n}}-\frac{1}{n}\right) \to 0 \text{ as } n \to \infty.$$

The limit of this sequence could also be evaluated by L' Hospital's Rule. Let $x \ge 1$ be real. The function f defined by $f(x) = \dfrac{\ln x}{x}$ is such that $f(n) = \dfrac{\ln n}{n}$. Hence by IV of (8.3), we have

$$\lim_{n \to \infty} \frac{\ln n}{n} = \lim_{x \to \infty} \frac{\ln x}{x} \qquad \left(\frac{\infty}{\infty}\right)$$

$$= \lim_{x \to \infty} \frac{1/x}{1} = 0.$$

EXERCISE 8.1

The nth term of a sequence is given. Determine whether the sequence converges or diverges. If it converges find its limit:

1. $\dfrac{2}{\sqrt{n^2+3}}$ 2. $\dfrac{(n-3)!}{(n-1)!}$ 3. $1+\dfrac{(-1)^n}{n}$

4. $\dfrac{\sqrt{n+1}}{n}$ 5. $n^{\frac{1}{n}}$ $n^{\frac{1}{n}}$ 6. $\dfrac{2^n}{(2n)!}$

7. $\dfrac{3n^4+1}{4n^2-1}$ 8. $\dfrac{\ln(n+1)}{\sqrt{n}}$ 9. $\dfrac{e^n}{n}$

10. $\ln n - \ln(n+1)$ 11. $\dfrac{\sin^2 n}{n}$ 12. $\dfrac{(2n)!}{(n!)^2}$

13. $\left(\dfrac{2-n^2}{3+n^2}\right)^n$ 14. $\dfrac{(\ln n)^2}{n}$ 15. $\sqrt{n}\left(\sqrt{n+1}-\sqrt{n}\right)$

16. $\dfrac{5^n+(-1)^n}{5^{n+1}+(-1)^{n+1}}$ 17. $(c^n+d^n)^{\frac{1}{n}},\ d>c>0$ 18. $\dfrac{5^n}{(n+1)^2}$

19. $1+\dfrac{1}{2!}+\dfrac{1}{3!}+\cdots+\dfrac{1}{n!}$ 20. $a_1=\sqrt{2},\ a_{n+1}=\sqrt{2+a_n},\ n\ge 1$

INFINITE SERIES

(8.10) Definition. Let $\{a_n\}$ be a sequence. An expression of the form

$$a_1+a_2+a_3+\cdots+a_n+\cdots \qquad (1)$$

containing infinitely many terms of the sequence $\{a_n\}$ is called an infinite series.

The series (1) is symbolically written as

$$\sum_{n=1}^{\infty} a_n \quad \text{or} \quad \sum_{1}^{\infty} a_n$$

or simply $\sum a_n$, when there is no ambiguity about the number of terms. a_n is called the nth term of the series (1).

Let S_n denote the sum of first n terms of the series (1). Then we write

$$S_1 = a_1$$
$$S_2 = a_1 + a_2$$
$$S_3 = a_1 + a_2 + a_3$$
$$\vdots \quad \vdots \quad \vdots \quad \vdots$$
$$S_n = a_1 + a_2 + a_3 + \cdots + a_n$$

The sequence $\{S_n\}$ is called the sequence of partial sums of the series (1) and the number S_n is called the **nth partial sum** of the series $\sum_{1}^{\infty} a_n$.

(8.11) Definition. Let $\{S_n\}$ be the sequence of partial sums of the series $\sum_{1}^{\infty} a_n$. If the sequence $\{S_n\}$ converges to the limit S, then $\sum_{1}^{\infty} a_n$ is said to **converge** and S is called the **sum** of the series. In this case we write

$$\sum_{1}^{\infty} a_n = S$$

If the sequence $\{S_n\}$ diverges, then the series $\sum_{1}^{\infty} a_n$ is said to be **divergent**.

Example 2. Find the sum of the series

$$\sum_{1}^{\infty} a_n = \frac{3}{10} + \frac{3}{10^2} + \frac{3}{10^3} + \cdots$$

Solution. Here, $S_n = \frac{3}{10} + \frac{3}{10^2} + \frac{3}{10^3} + \cdots + \frac{3}{10^n}$

$$= \frac{\frac{3}{10}\left(1 - \frac{1}{10^n}\right)}{1 - \frac{1}{10}} = \frac{1}{3}\left(1 - \frac{1}{10^n}\right)$$

Now $\lim_{n \to \infty} S_n = \lim_{n \to \infty} \frac{1}{3}\left(1 - \frac{1}{10^n}\right) = \frac{1}{3}$, since $\frac{1}{10^n} \to 0$ as $n \to \infty$.

Thus

$$\frac{3}{10} + \frac{3}{10^2} + \frac{3}{10^3} + \cdots = \frac{1}{3}$$

Example 3. The reader is familiar with the geometric series

$$\sum_{1}^{\infty} a_n = a + ar + ar^2 + \cdots + ar^{n-1} + \cdots \qquad (1)$$

We investigate the behaviour of this series for different values of r.

Here $S_n = a + ar + ar^2 + \cdots + ar^{n-1} = \frac{a - ar^n}{1 - r}, r \neq 1.$

Now, $\lim_{n \to \infty} S_n = \lim_{n \to \infty} \frac{a - ar^n}{1 - r}$

$$= \frac{a - a \lim_{n \to \infty} r^n}{1 - r}$$

But $\lim_{n \to \infty} r^n = 0$ if $|r| < 1$ and ∞ if $|r| > 1$.

Thus the series (1)

(i) converges to $S = \frac{a}{1-r}$, if $|r| < 1$.

(ii) diverges, if $|r| > 1$.

(iii) if $|r| = 1$, i.e., $r = \pm 1$, then for $r = 1$ the series is

$$a + a + a + \cdots$$

so that the nth partial sum is $S_n = na$ and $\lim_{n \to \infty} na = \pm\infty$ according as a is positive or negative.

If $r = -1$, the series is

$$a - a + a - a + \cdots$$

so the sequence of partial sums is $a, 0, a, 0, \cdots$ which has limit a or 0 according as the partial sums contain an odd or an even number of terms. Thus $\{S_n\}$ does not converge to unique limit and so (1) diverges for $r = -1$. Thus the series (1) diverges for $|r| = 1$.

Example 4. Determine whether the following series converges or diverges.

$$\sum_{n=1}^{\infty} \frac{1}{(n+2)(n+3)} = \frac{1}{3 \cdot 4} + \frac{1}{4 \cdot 5} + \frac{1}{5 \cdot 6} + \cdots$$

If it converges, find its sum.

Solution. The nth partial sum of the series is

$$S_n = \sum_{k=1}^{n} \frac{1}{(k+2)(k+3)}$$

$$= \frac{1}{3 \cdot 4} + \frac{1}{4 \cdot 5} + \cdots + \frac{1}{(n+2)(n+3)}$$

Now $\frac{1}{(n+2)(n+3)} = \frac{1}{n+2} - \frac{1}{n+3}$

Therefore,

$$S_n = \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{4} - \frac{1}{5}\right) + \cdots + \left(\frac{1}{n+2} - \frac{1}{n+30}\right)$$

$$= \frac{1}{3} + \left(-\frac{1}{4} + \frac{1}{4}\right) + \left(-\frac{1}{5} + \frac{1}{5}\right) + \cdots + \left(-\frac{1}{n+2} + \frac{1}{n+2}\right) - \frac{1}{n+3}$$

This is a **telescopic sum** which means that each term cancels part of the next term so that the sum reduces to only two terms. Thus

$$S_n = \frac{1}{3} - \frac{1}{n+3} = \frac{n}{3(n+3)}$$

$$\lim_{n \to \infty} = \lim_{n \to \infty} \frac{n}{(3n+3)} = \frac{1}{3}.$$

Hence the series converges and its sum is $\frac{1}{3}$.

Example 5. The Euler's Series. Investigate the behaviour of the series

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2} + \cdots$$

Solution. We have

$$S_n = \frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2}$$

The sequence $\{S_n\}$ is monotonically increasing since

$$S_1 = 1, \qquad S_2 = 1 + \frac{1}{4}, \qquad S_3 = 1 + \frac{1}{4} + \frac{1}{9}, \ldots$$

We check whether $\{S_n\}$ is bounded.

$$\begin{aligned} S_n &= 1 + \frac{1}{2\cdot 2} + \frac{1}{3\cdot 3} + \cdots + \frac{1}{n\,.\,n} \\ &< 1 + \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots + \frac{1}{(n-1)\,.\,n} \\ &= 1 + \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 2 - \frac{1}{n} \end{aligned}$$

Hence $|S_n| < 2$ for all n and $\{S_n\}$ converges by Theorem 8.9. Since the sequence of partial sums of the series is convergent, $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent by (8.11).

(8.12) Theorem. (Cauchy Criterion). If a series $\sum_{1}^{\infty} a_n$ is convergent, then for every $\varepsilon > 0$ there exists a positive integer N such that

$$|S_m - S_n| < \varepsilon \qquad \text{for all } m \geq N,\ n \geq N.$$

Proof. Let $\sum_{1}^{\infty} a_n$ be convergent and its sum be S. Then

$$\begin{aligned} |S_m - S_n| &= |S_m - S + S - S_n| \\ &\leq |S_m - S| + |S_n - S| \end{aligned} \tag{1}$$

Since $\lim_{m \to \infty} S_m = S = \lim_{n \to \infty} S_n$, for every $\varepsilon > 0$ there exist positive integers n_1 and n_2 such that

$$|S_m - S| < \frac{\varepsilon}{2} \quad \text{and} \quad |S_n - S| < \frac{\varepsilon}{2} \tag{2}$$

for all $m \geq n_1$ and $n \geq n_2$.

Let $N = \max\{n_1, n_2\}$. Then from (1) and the inequalities (2), we have

$$|S_m - S_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \text{for all } m, n \geq N.$$

as required.

Example 6. Test for convergence the harmonic series

$$\sum_{1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots$$

(The name 'harmonic series' is given due to the reason that $\frac{1}{n}$ is the harmonic mean of $\frac{1}{n-1}$ and $\frac{1}{n+1}$).

Solution. We have

$$S_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

$$S_{2n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}.$$

If $n > 1$, then

$$|S_{2n} - S_n| = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}$$
$$> \frac{1}{2n} + \frac{1}{2n} + \cdots + \frac{1}{2n}$$
$$= n\left(\frac{1}{2n}\right) = \frac{1}{2}. \qquad (1)$$

If the series were convergent, then by Theorem **8.12**, with $\varepsilon = \frac{1}{2}$, $m = 2n$ and $n = n$, we have

$$|S_{2n} - S_n| < \varepsilon \qquad \text{for } n \geq N.$$

where N is sufficiently large. But from (1) it is clear that this inequality is never true and so the series is divergent.

(8.13) Theorem. If the series $\sum_1^\infty a_n$ converges then $\lim_{n\to\infty} a_n = 0$.

Proof. We have

$$S_n = a_1 + a_2 + \cdots + a_{n-1} + a_n$$
$$S_{n-1} = a_1 + a_2 + \cdots + a_{n-1}$$
so that $$S_n - S_{n-1} = a_n \qquad (1)$$

Since the series converges, $\lim_{n\to\infty} S_n$ exists. Let $\lim_{n\to\infty} S_n = S$. Now $n - 1 \to \infty$ as $n \to \infty$, so we also have $\lim_{n\to\infty} S_{n-1} = S$. Hence from (1), we obtain

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1}) = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = S - S = 0.$$

(8.14) Remark. This theorem states only a necessary condition for convergence, which means that for a convergent series its nth term must converge to zero. But the nth term approaching zero does not imply convergence of the series as can be seen from the harmonic series whose nth term $\frac{1}{n} \to 0$ as $n \to \infty$ but the series is divergent.

(8.15) Corollary. (A Divergence Test). If $\lim_{n\to\infty} a_n \neq 0$, then the series $\sum_1^\infty a_n$ diverges.

Example 7. Examine the series $\sum_1^\infty \frac{5n+2}{3n-1}$ for convergence.

Solution. If the series converges we must have

$$\lim_{n\to\infty} a_n = 0$$

But

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{5n+2}{3n-1} = \frac{5}{3} \neq 0$$

Hence the given series diverges by (8.15).

(8.16) Theorem.

(i) If $\sum_1^\infty a_n$ and $\sum_1^\infty b_n$ are convergent series with sums S and T, then the series $\sum_1^\infty (a_n + b_n)$ and $\sum_1^\infty (a_n - b_n)$ are convergent and sums of these series are

$$\sum_1^\infty (a_n + b_n) = \sum_1^\infty a_n + \sum_1^\infty b_n = S + T$$

and $$\sum_1^\infty (a_n - b_n) = \sum_1^\infty a_n - \sum_1^\infty b_n = S - T.$$

(ii) If $\sum_1^\infty a_n$ converges and $\sum_1^\infty b_n$ diverges then $\sum_1^\infty (a_n + b_n)$ diverges.

(iii) If c is a nonzero real number, then the series $\sum_1^\infty a_n$ and $\sum_1^\infty ca_n$ either both converge or both diverge. In case of convergence.

$$\sum_1^\infty ca_n = c\sum_1^\infty a_n = c\,S, \text{ where } \sum_1^\infty a_n \text{ converges to } S.$$

(iv) The **insertion** (or **deletion**) of a finite number of terms into (or from) an infinite series does not alter its convergence or divergence. However, if the series is convergent then its sum is changed by such insertion (or deletion).

Proof. (i) Let S_n and T_n denote the nth partial sums of $\sum_1^\infty a_n$ and $\sum_1^\infty b_n$ respectively.

Then

$$\sum_{k=1}^{n} (a_k + b_k) = (a_1 + b_1) + (a_2 + b_2) + \cdots + (a_n + b_n)$$
$$= (a_1 + a_2 + \cdots + a_n) + (b_1 + b_2 + \cdots + b_n)$$
$$= S_n + T_n$$

Therefore,

$$\sum_1^\infty (a_n + b_n) = \lim_{n \to \infty} (S_n + T_n)$$
$$= \lim_{n \to \infty} S_n + \lim_{n \to \infty} T_n = S + T$$

The proof of the remaining part is similar

(ii) Suppose that the new series $\sum_1^\infty (a_n + b_n)$ converges Consider the series

$$\sum_1^\infty [(a_n + b_n) - a_n]$$

Since $\sum_1^\infty a_n$ is convergent by hypothesis and $\sum_1^\infty (a_n + b_n)$ is convergent by our supposition, using (i) above the series $\sum_1^\infty [(a_n + b_n) - a_n] = \sum_1^\infty b_n$ is convergent, which contradicts the hypothesis Hence our supposition is wrong and therefore, the series $\sum_1^\infty (a_n + b_n)$ diverges

The proofs of the other parts are similar and are left as an exercise

(8.17) Definition. A series $\sum_1^\infty a_n$ is said to be a **positive term series** (or a series of positive terms) if $a_n > 0$ for all positive integers n

In order to examine the convergence or divergence of an infinite series we test the nth partial sum S_n of the series But for most of the series, it is not possible to find a simple formula for the nth partial sum In the following, we shall formulate tests to check the behaviour of a series by considering its nth term However, these tests will not provide formulas to find the sum of a convergent series

If $\sum_1^\infty a_n$ is a positive term series, then its sequence $\{S_n\}$ of partial sums is monotonically increasing and the series converges if $\{S_n\}$ is bounded **(Theorem 8.9)**

(8.18) Theorem. Let $\sum_1^\infty a_n$ be a convergent positive term series The order of the terms can be rearranged in any manner and the resulting series remains convergent having the same sum as the given series

Proof. Let $\{S_n\}$ be the sequence of partial sums of the given convergent series and let $\lim_{n \to \infty} S_n = S$. Suppose $\sum_1^\infty b_n$ is the series formed by rearranging the terms of $\sum_1^\infty a_n$ and $\{T_n\}$ is the sequence of partial sums of $\sum_1^\infty b_n$. Clearly, $T_n \le S$ for all n, since T_n is the sum of n terms of the series $\sum_1^\infty a_n$. Thus $\{T_n\}$ is a bounded sequence. It is also monotonically increasing as the series is of positive terms Hence $\{T_n\}$ converges by (8.9) and, therefore, the series $\sum_1^\infty b_n$ also converges. Since each term in $\sum_{n=1}^\infty b_n$ occurs in $\sum_{n=1}^\infty a_n$ and vice-versa, both the series have the same sum S

THE BASIC COMPARISON TEST

(8.19) Theorem. Let $\sum_1^\infty a_n$ and $\sum_1^\infty b_n$ be series of positive terms with $a_n \le b_n$ for each $n = 1, 2, 3, \ldots$ Then

(i) If $\sum_1^\infty b_n$ converges, so does $\sum_1^\infty a_n$ (ii) If $\sum_1^\infty a_n$ diverges, so does $\sum_1^\infty b_n$

Proof.

(i) Let S_n and T_n denote respectively the nth partial sums of the series $\sum_1^\infty a_n$ and $\sum_1^\infty b_n$

Then $S_n \le T_n$ for all $n = 1, 2, 3, \ldots$ Since the series $\sum_1^\infty b_n$ converges, the sequence

$\{T_n\}$ of its partial sums converges and is, therefore, bounded. Hence $\{S_n\}$ is also bounded. But $\{S_n\}$ is an increasing sequence since $a_n > 0$ for all n. Thus the sequence $\{S_n\}$ is convergent and so $\sum_1^\infty a_n$ is also convergent.

(ii) Assume that $\sum_1^\infty a_n$ diverges but $\sum_1^\infty b_n$ converges. Since $a_n \le b_n$ for each $n = 1, 2, 3, \cdots$, by (i) $\sum_1^\infty a_n$ must converge, contrary to the hypothesis. Thus our supposition is wrong and $\sum_1^\infty b_n$ diverges.

Example 8. Determine whether the series $\sum_1^\infty \frac{1}{1+n^2}$ converges or diverges.

Solution. We have $\frac{1}{1+n^2} < \frac{1}{n^2}$.

But $\sum_1^\infty \frac{1}{n^2}$ converges. Hence, by the Basic Comparison Test, the given series converges.

Example 9. Show that $\sum_1^\infty \frac{n+5}{n^2+4}$ diverges.

Solution. We note that

$$\frac{n+5}{n^2+4} \ge \frac{n+5}{n^2+4n} \quad \text{for all } n \ge 1$$

$$= \left(\frac{n+5}{n+4}\right)\frac{1}{n} > \frac{1}{n}, \text{ since } \frac{n+5}{n+4} \text{ is greater than 1.}$$

But $\sum_1^\infty \frac{1}{n}$ diverges and so the given series also diverges by **(8.19) (ii)**.

THE LIMIT COMPARISON TEST

(8.20) Theorem. Let $\sum_1^\infty a_n$ and $\sum_1^\infty b_n$ be series of positive terms.

(i) If $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = L \ne 0$, then either both the series converge or both diverge.

(ii) If $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = 0$ and $\sum_1^\infty b_n$ converges, then $\sum_1^\infty a_n$ also converges.

(iii) If $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \infty$ and $\sum_1^\infty b_n$ diverges, then $\sum_1^\infty a_n$ also diverges.

Proof. Since all the terms of $\sum_1^\infty a_n$ and $\sum_1^\infty b_n$ are positive and $L \ne 0$, we must have $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = L > 0$. By definition, given $\varepsilon > 0$ there exists a positive integer N such that

$$\left|\frac{a_n}{b_n} - L\right| < \varepsilon \qquad \text{for all } n \ge N$$

or $$L - \varepsilon < \frac{a_n}{b_n} < L + \varepsilon \qquad \text{for all } n \ge N$$

or $$\frac{1}{2}L < \frac{a_n}{b_n} < \frac{3}{2}L \qquad \text{when } n \ge N, \text{ if we take } \varepsilon = \frac{1}{2}L.$$

Thus, $$\frac{1}{2}L\,b_n < a_n < \frac{3}{2}L\,b_n, \qquad \text{for all } n \ge N. \qquad (1)$$

If $\sum_1^\infty a_n$ converges, then $\sum_N^\infty a_n$ converges. From the left half of inequalities in (1), it follows that $\sum_N^\infty \frac{1}{2}L\,b_n$ converges. Hence $\sum_1^\infty b_n$ converges. The right half of inequalities (1) shows that if $\sum_1^\infty b_n$ converges, then so does $\sum_N^\infty \frac{3}{2}b_n$ and by the Basic Comparison Test, $\sum_N^\infty a_n$ converges. Thus $\sum_1^\infty a_n$ converges if and only if $\sum_1^\infty b_n$ converges. Reasoning as before, $\sum_1^\infty a_n$ diverges if and only if $\sum_1^\infty b_n$ diverges.

(ii) If $\lim\limits_{n \to \infty} \dfrac{a_n}{b_n} = 0$, then, given $\varepsilon > 0$, there exists a positive integer N such that

$$\left| \frac{a_n}{b_n} \right| < \varepsilon \qquad \text{for all } n \geq N$$

or
$$\frac{a_n}{b_n} < \varepsilon \qquad \text{for all } n \geq N$$

or
$$a_n < \varepsilon\, b_n \qquad \text{when } n \geq N$$

Hence if $\sum\limits_{1}^{\infty} b_n$ converges, so does $\sum\limits_{1}^{\infty} a_n$ by **(8.19) (i)**

(iii) If $\lim\limits_{n \to \infty} \dfrac{a_n}{b_n} = \infty$, then by definition, given $G > 0$, however large, there exists a positive integer N such that

$$\left| \frac{a_n}{b_n} \right| > G \qquad \text{for all } n \geq N$$

or
$$\frac{a_n}{b_n} > G \qquad \text{for all } n \geq N$$

or
$$a_n > G\, b_n \qquad \text{for all } n \geq N$$

Hence we conclude that if $\sum\limits_{1}^{\infty} b_n$ diverges, so does $\sum\limits_{1}^{\infty} a_n$ by **(8.19) (ii)**

Example 10. Use the Limit Comparison Test to determine whether each of the following series converges or diverges:

(i) $\displaystyle\sum_{1}^{\infty} \frac{n+1}{2n^2+1}$ (ii) $\displaystyle\sum_{1}^{\infty} \frac{n-4}{n^3+n+3}$

Solution. (i) The general term of the series is a quotient containing a highest exponent 1 of n in the numerator and a highest exponent 2 of n^2 in the denominator. This suggests a comparison with the series whose general term is

$$b_n = \frac{n}{n^2} = \frac{1}{n}$$

Therefore,

$$\lim_{n \to \infty} \left[\frac{\dfrac{n+1}{2n^2+1}}{\dfrac{1}{n}} \right] = \lim_{n \to \infty} \left(\frac{n^2+n}{2n^2+1} \right) = \frac{1}{2} > 0$$

But $\sum\limits_{1}^{\infty} \dfrac{1}{n}$ diverges. Hence the given series also diverges by **(8.20) (i)**.

We compare the given series with the series whose general term is $b_n = \dfrac{1}{n^2}$. We have

$$\lim_{n \to \infty} \left[\frac{\dfrac{n-4}{n^3+n+3}}{\dfrac{1}{n^2}} \right] = \lim_{n \to \infty} \left(\frac{n^3-4n^2}{n^3+n+5} \right) = 1 > 0$$

Since both the series $\sum\limits_{1}^{\infty} \dfrac{n-4}{n^3+n+3}$ and $\sum\limits_{1}^{\infty} \dfrac{1}{n^2}$ behave alike. But $\sum\limits_{1}^{\infty} \dfrac{1}{n^2}$ converges and so $\sum\limits_{1}^{\infty} \dfrac{n-4}{n^3+n+3}$ also converges by **(8.20) (i)**

THE INTEGRAL TEST

(8.21) Theorem. (Cauchy's Integral Test)[1]. Let $\sum\limits_{1}^{\infty} a_n$ be a positive term series. If f is a continuous and nonincreasing function on $[1, \infty[$ such that $f(n) = a_n$ for all positive integers n, then

(i) $\sum\limits_{1}^{\infty} a_n$ converges if $\displaystyle\int_1^{\infty} f(x)\,dx$ converges.

(ii) $\sum\limits_{1}^{\infty} a_n$ diverges if $\displaystyle\int_1^{\infty} f(x)\,dx$ diverges.

[1] Due to A.L. Cauchy (1789 – 1867). A French mathematician.

Proof. (i)

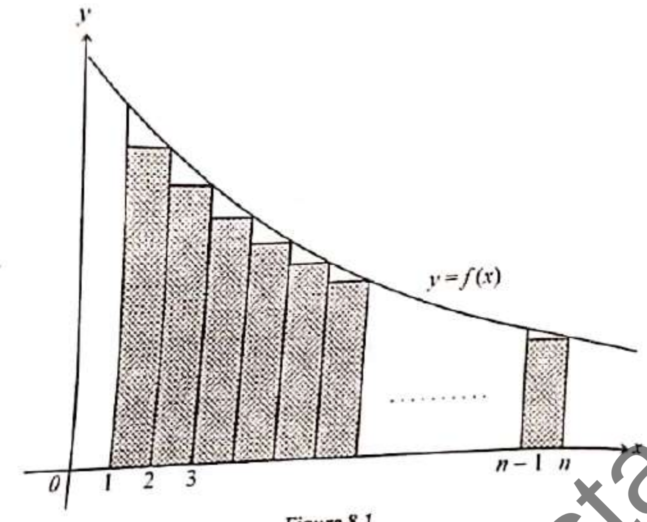


Figure 8.1

In Figure 8.1, each rectangle has height $f(i) = a_i$, $i = 2, 3, \cdots, n$ and width 1 so that each one has an area of measure a_i, $i = 2, 3, \cdots, n$. The total area of the $(n-1)$ rectangles is less than or equal to the area under the curve $y = f(x)$ from $x = 1$ to $x = n$. That is

$$0 \le a_2 + a_3 + \cdots + a_n \le \int_1^n f(x)\,dx$$

or $\quad a_1 + a_2 + \cdots + a_n \le a_1 + \int_1^n f(x)\,dx \qquad (1)$

If $\int_1^\infty f(x)\,dx$ converges, then $\lim\limits_{n\to\infty} \int_1^n f(x)\,dx$ exists and equals some number L.

Since $f(x) \ge 0$, we must have

$$\int_1^n f(x)\,dx < L, \qquad \text{for all } n. \qquad (2)$$

If S_n denotes the nth partial sum of the given series, (1) and (2) yield

$$S_n < a_1 + L \qquad \text{for all } n.$$

Thus $\{S_n\}$ is bounded. By (8.9) it is convergent and hence the series $\sum_1^\infty a_n$ is convergent.

(ii)

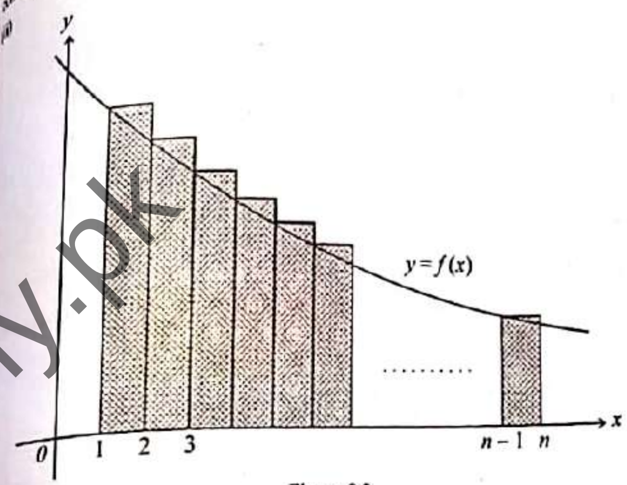


Figure 8.2

In Figure 8.2, the $(n-1)$ rectangles have total area greater than or equal to the area under the graph of $y = f(x)$ from $x = 1$ to $x = n$. Thus

$$a_1 + a_2 + \cdots + a_{n-1} > \int_1^n f(x)\,dx$$

i.e., $\quad S_n \ge \int_1^n f(x)\,dx$, since $a_n > 0$.

If $\int_1^\infty f(x)\,dx$ diverges, then $\lim\limits_{n\to\infty} \int_1^n f(x)\,dx \to \infty$ and so $\{S_n\}$ is divergent. Consequently the given series is divergent.

Example 11. Consider the harmonic series $\sum_1^\infty \frac{1}{n}$. We use the Integral Test to show its divergence. Here $f(x) = \frac{1}{x}$ which satisfies the hypothesis of **Theorem 8.21** and

$$\int_1^{\infty} \frac{dx}{x} = \lim_{t \to \infty} \int_1^t f(x)\, dx$$

$$= \lim_{t \to \infty} \int_1^t \frac{dx}{x} = \lim_{t \to \infty} \Big| \ln x \Big|_1^t = \lim_{t \to \infty} \ln t = \infty$$

Since the improper integral diverges, the given series also diverges.

Example 12. Test for convergence the p-series (or hyper harmonic series)

$$\sum_{1}^{\infty} \frac{1}{n^p} = \frac{1}{1^p} + \frac{1}{2^p} + \frac{1}{3^p} + \cdots$$

Solution. If $p \le 0$, the series diverges since in this case $\lim\limits_{n \to \infty} \frac{1}{n^p} \ne 0$ and the **Divergence Test (8.15)** applies.

For $p > 0$, the function $f(x) = x^{-p}$ is positive and decreasing on $[1, \infty[$ and we can apply the **Integral Test**. The integral

$$\int_1^{\infty} \frac{1}{x^p}\, dx = \lim_{t \to \infty} \int_1^t x^{-p}\, dx$$

$$= \lim_{t \to \infty} \left| \frac{x^{-p+1}}{1-p} \right|_1^t, (p \ne 1) = \lim_{t \to \infty} \left[\frac{t^{1-p}}{1-p} - \frac{1}{1-p} \right]$$

This limit exists only if the exponent of t is negative, that is, if $-p + 1 < 0$ or $p > 1$. Thus, the improper integral converges if $p > 1$ and diverges for $p < 1$. If $p = 1$ the p-series becomes the harmonic series which diverges. Hence the p-series converges for $p > 1$ and diverges for $p \le 1$.

Example 13. Determine whether the series $\sum\limits_{1}^{\infty} \frac{\arctan n}{1+n^2}$ converges or diverges.

Solution. If we replace n by x in the formula for a_n, we get the function

$$f(x) = \frac{\arctan x}{1+x^2}$$

For $x \ge 1$, the function has positive values and is differentiable. Moreover

$$f'(x) = \frac{1}{(1+x^2)^2}(1 - 2x \arctan x) < 0 \text{ for } x \ge 1.$$

Thus f is a decreasing function for $x \ge 1$. The integral

$$\int_1^{\infty} \frac{\arctan x}{1+x^2}\, dx = \lim_{t \to \infty} \int_1^t \frac{\arctan x}{1+x^2}\, dx$$

$$= \lim_{t \to \infty} \left| \frac{(\arctan t)^2}{2} \right|_1^t$$

$$= \lim_{t \to \infty} \left[\frac{(\arctan t)^2}{2} - \frac{(\arctan 1)^2}{2} \right] = \frac{\pi^2}{8} - \frac{\pi^2}{32} = \frac{3\pi^2}{32}$$

Hence the series $\sum\limits_{1}^{\infty} \frac{\arctan n}{1+n^2}$ converges by the Integral Test.

EXERCISE 8.2

Determine whether the given series converges or diverges. If it converges find its sum **(Problems 1-5):**

1. $\sum\limits_{1}^{\infty} \cos \pi n$
2. $\sum\limits_{0}^{\infty} \frac{1}{(2+x)^n}$, $|x| < 1$
3. $\sum\limits_{0}^{\infty} \frac{2^{n+2}}{3^n}$
4. $\sum\limits_{1}^{\infty} \left(\frac{1}{2^n} - \frac{1}{2^{n+1}} \right)$
5. $\sum\limits_{1}^{\infty} \frac{1}{9n^2 + 3n - 2}$

Each of the following is the nth partial sums of an infinite series. Determine the series and check whether it converges **(Problems 6-8):**

6. $S_n = \frac{3n}{4n+1}$
7. $S_n = \frac{n^2}{n+1}$
8. $S_n = \frac{1}{2^n}$

9. Prove that if a positive term series $\sum\limits_{1}^{\infty} a_n$ converges then the series $\sum\limits_{1}^{\infty} \sqrt{a_n\, a_{n-1}}$ converges.

10. Give an example in which $\sum\limits_{1}^{\infty} a_n$ and $\sum\limits_{1}^{\infty} b_n$ both diverge but $\sum\limits_{1}^{\infty} (a_n + b_n)$ converges.

Using the Comparison Tests, investigate convergence or divergence of the series in Problems 11-20:

11. $\sum_1^\infty \frac{1}{n^{1/2}+n^{3/2}}$ 12. $\sum_1^\infty \frac{\sqrt{n}}{1+n}$ 13. $\sum_1^\infty \frac{2}{\sqrt{n+1}}$

14. $\sum_1^\infty \frac{1}{n^n}$ 15. $\sum_0^\infty \frac{1}{n!}$ 16. $\sum_1^\infty \frac{e^{2n}+e^{-2n}}{e^{n}+e^{-n}}$

17. $\sum_1^\infty \frac{\ln n}{n}$ 18. $\sum_1^\infty \left(1+\frac{1}{n}\right)^n$ 19. $\sum_1^\infty \sin\left(\frac{\pi}{n}\right)$

20. $\sum_1^\infty \frac{1}{e^{n^2}}$

In Problems 21-40, test each series for convergence or divergence.

Ratio

21. $\sum_1^\infty n^2 e^{-n^3}$ 22. $\sum_1^\infty \frac{1}{(2n-1)^{1/3}}$ 23. $\sum_1^\infty \frac{2^n}{3^n+1}$

24. $\sum_1^\infty \frac{\ln n}{1+\ln n}$ 25. $\sum_1^\infty \frac{\ln(n+1)}{n+2}$ 26. $\sum_1^\infty \frac{\ln n}{n^3}$

27. $\sum_1^\infty \frac{\sqrt{n}}{n^2+\cos(2n-6)}$ 28. $\sum_1^\infty \frac{1}{(n+1)\,[\ln(n+1)]^2}$ 29. $\sum_1^\infty n^2\sin^2\left(\frac{1}{n}\right)$

30. $\sum_2^\infty \frac{1}{n(\ln n)^n}$ 31. $\sum_1^\infty \frac{2n-1}{n(n+1)(n+2)}$ 32. $\sum_1^\infty \frac{n^{p-1}}{n^p+1}, p>2$

33. $\sum_1^\infty \frac{e^{\arctan n}}{1+n^2}$ 34. $\sum_2^\infty \frac{1}{n\sqrt{n^2-1}}$ 35. $\sum_1^\infty \frac{n^2}{e^n}$

36. $\sum_1^\infty \frac{2\cdot4\cdot6\cdots(2n)}{1\cdot3\cdot5\cdots(2n-1)}$ 37. $\sum_2^\infty \frac{1}{n-\sqrt{n}}$ 38. $\sum_0^\infty \frac{5^n+n}{6^n+n}$

39. $\sum_1^\infty \frac{\ln(n+1)}{n^2}$ 12 40. $\sum_{10}^\infty \frac{1}{n(\ln n)(\ln\ln n)}$

THE RATIO TEST

(8.22) **Theorem. (D' Alembert's[1] Test).** Let $\sum_1^\infty a_n$ be a series of positive terms and suppose that $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L$, where L is a nonnegative real number or ∞.

(i) If $L<1$, the series $\sum_1^\infty a_n$ converges

(ii) If $L>1$ or ∞, the series $\sum_1^\infty a_n$ diverges

(iii) If $L=1$ the test fails to determine convergence or divergence of the series.

Proof. (i) Suppose $L<1$. Choose r with $L<r<1$. Then $r-L>0$.

Let $\varepsilon = r-L$ be given. Since $L=\lim_{n\to\infty}\frac{a_{n+1}}{a_n}$, it follows that there exists a positive integer k such that.

$$\left|\frac{a_{n+1}}{a_n}-L\right|<\varepsilon \qquad \text{for all } n\ge k$$

$$\text{or} \quad L-\varepsilon < \frac{a_{n+1}}{a_n} < L+\varepsilon \qquad \text{, for all } n\ge k.$$

$$\text{or} \quad \frac{a_{n+1}}{a_n} < r \qquad \text{for all } n\ge k, \text{ since } \varepsilon = r-L.$$

$$\begin{aligned} \text{Thus,}\quad a_{n+1} &< r\,a_n, \qquad \text{for all } n\ge k \\ a_{k+1} &< r\,a_k \\ a_{k+2} &< r\,a_{k+1} < r^2 a_k \\ a_{k+3} &< r\,a_{k+2} < r^3 a_k \\ \vdots \quad & \quad \vdots \qquad \vdots \end{aligned}$$

On adding these inequalities, we

$$a_{k+1}+a_{k+2}+a_{k+3}+\cdots < \qquad a_k+\cdots \tag{1}$$

1. Pronounced "Da-lem-bair". A French mathematician (1717 – 1783)

But the geometric series on the right of (1), with ratio $r < 1$, converges. Therefore, by the Basic Comparison Test, the series $a_{k+1} + a_{k+2} + a_{k+3} + \cdots$ converges.

Thus $\sum_{1}^{\infty} a_n$ converges by Theorem 8.16 (iv).

(ii) Assume $L > 1$. Take $\varepsilon = L - 1 > 0$. There exists a positive integer k such that

$$\left| \frac{a_{n+1}}{a_n} - L \right| < \varepsilon \qquad \text{for all } n \geq k$$

or $$L - \varepsilon < \frac{a_{n+1}}{a_n} < L + \varepsilon \qquad \text{for all } n \geq k.$$

or $$\frac{a_{n+1}}{a_n} > L - \varepsilon = 1 \qquad \text{for all } n \geq k$$

i.e. $$a_{n+1} > a_n \qquad \text{for all } n \geq k$$

So $\{a_n\}$ is a monotonically increasing sequence. Thus $\lim_{n \to \infty} a_n \neq 0$. Therefore by (8.15) the series $\sum_{1}^{\infty} a_n$ diverges.

If $\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = \infty$, then however large $G > 0$ may be, we have

$$\frac{a_{n+1}}{a_n} > G, \qquad \text{for all } n \geq k.$$

or $$a_{n+1} > G a_n, \qquad \text{for all } n \geq k$$

Reasoning as above, we conclude that $\sum_{1}^{\infty} a_n$ diverges.

(iii) It is easy to see that for each of the two series $\sum_{1}^{\infty} \frac{1}{n}$ and $\sum_{1}^{\infty} \frac{1}{n^2}$, we have $L = 1$.

But $\sum_{1}^{\infty} \frac{1}{n}$ diverges and $\sum_{1}^{\infty} \frac{1}{n^2}$ converges. Thus, the Ratio Test fails to give any information regarding convergence or divergence of the series when $L = 1$.

Example 14. Let $\sum_{1}^{\infty} a_n = \sum_{1}^{\infty} \frac{n^4}{2^n}$. Then

$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = \lim_{n \to \infty} \frac{(n+1)^4}{2^{n+1}} \cdot \frac{2^n}{n^4} = \lim_{n \to \infty} \left(\frac{n+1}{n}\right)^4 \frac{1}{2} = \frac{1}{2} < 1$$

Hence $\sum_{1}^{\infty} a_n$ converges.

Example 15. Investigate the behaviour of the series $\sum_{1}^{\infty} a_n = \sum_{1}^{\infty} \frac{(n!)^2}{(n^2)!}$

Solution. We have

$$\frac{a_{n+1}}{a_n} = \frac{[(n+1)!]^2}{[(n+1)^2]!} \cdot \frac{(n^2)!}{(n!)^2} = \frac{(n+1)^2 (n!)^2 (n^2)!}{(n+1)^2 [(n+1)^2 - 1]! (n!)^2}$$

$$= \frac{(n^2)!}{[(n+1)^2 - 1][(n+1)^2 - 2] \ldots [(n+1)^2 - 2n] (n^2)!}$$

$$\to 0 < 1, \text{ as } n \to \infty$$

Hence the series converges by the Ratio Test.

Example 16. Examine the series $\sum_{1}^{\infty} \frac{n!}{n^2}$ for convergence or divergence.

Solution. $$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = \lim_{n \to \infty} \frac{(n+1)!}{(n+1)^2} \times \frac{n^2}{n!} = \lim_{n \to \infty} \frac{(n+1) n^2}{(n+1)^2}$$

$$= \lim_{n \to \infty} \frac{n^2}{n+1} = \infty.$$

Therefore, by the Ratio Test, the series $\sum_{1}^{\infty} \frac{n!}{n^2}$ diverges.

Example 17. Determine whether the series $\sum_{1}^{\infty} a_n = \sum_{1}^{\infty} \frac{n^n}{n!}$ converges or diverges.

Solution. We have

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)^{n+1}}{(n+1)!} \times \frac{n!}{n^n} = \frac{(n+1)^{n+1}}{(n+1)\, n!} \times \frac{n!}{n^n}$$

$$= \frac{(n+1)^n}{n^n} = \left(1 + \frac{1}{n}\right)^n \to e > 1 \text{ as } n \to \infty.$$

Thus, by the Ratio Test, the series diverges.

CAUCHY'S ROOT TEST

(8.23) Theorem. Let $\sum_{1}^{\infty} a_n$ be a series of positive terms and suppose that $\lim_{n\to\infty} (a_n)^{1/n} = L$, where L is a nonnegative real number or ∞

(i) If $L < 1$, the series $\sum_{1}^{\infty} a_n$ converges

(ii) If $L > 1$ or ∞, the series $\sum_{1}^{\infty} a_n$ diverges.

(iii) If $L = 1$ the test fails i.e., no conclusion can be made about the convergence or divergence of the series.

Proof.

(i) We proceed as in the proof of the Ratio Test **(8.22)**

Suppose $L < 1$. We choose r with $L < r < 1$. Since $\lim_{n\to\infty} (a_n)^{1/n} = L$, given $\varepsilon > 0$, there exists a positive integer N such that

$$\left|(a_n)^{1/n} - L\right| < \varepsilon \quad \text{for all } n \ge N.$$

or $$L - \varepsilon < (a_n)^{1/n} < L + \varepsilon \quad \text{for all } n \ge N.$$

or $$(a_n)^{1/n} < r \quad \text{for all } n \ge N, \ (\text{take } \varepsilon = r - L)$$

i.e., $$a_n < r^n \quad \text{for all } n \ge N,$$

Thus $$\sum_{N}^{\infty} a_n < \sum_{N}^{\infty} r^n$$

But $\sum_{N}^{\infty} r^n$ converges since $r < 1$.

Therefore, the series $\sum_{N}^{\infty} a^n$ converges by the Basic Comparison Test. The convergence of the series $\sum_{1}^{\infty} a^n$ follows from Theorem **8.16 (iv)**.

(ii) Let $L > 1$. Take r such that $L > r > 1$. We have

$$L - \varepsilon < (a_n)^{1/n} \quad \text{for all } n \ge N$$

or $$r < (a_n)^{1/n} \quad \text{for all } n \ge N, \ (\text{take } \varepsilon = L - r)$$

i.e. $$a_n > r^n \quad \text{for all } n \ge N$$

Thus $$\sum_{N}^{\infty} a_n > \sum_{N}^{\infty} r^n.$$

But $\sum_{N}^{\infty} r^n$ diverges since $r > 1$.

Hence $\sum_{N}^{\infty} a_n$ diverges. The series $\sum_{N}^{\infty} a_n$ then diverges by Theorem **8.16 (iv)**.

The proof is similar if $\lim_{n\to\infty} (a_n)^{1/n} = \infty$.

(iii) We have

$$\lim_{n\to\infty} \frac{1}{n^{1/n}} = 1, \left\{ \text{since } \lim_{n\to\infty} n^{1/n} = 1 \right\}$$ and the series $\sum_{1}^{\infty} \frac{1}{n}$ diverges.

Again, $\lim_{n\to\infty} \left(\frac{1}{n^2}\right)^{1/n} = 1$ and the series $\sum_{1}^{\infty} \frac{1}{n^2}$ converges

Thus no conclusion can be drawn if $L = 1$.

Note: This test is useful for series involving terms having exponents.

Example 18. Determine whether the series $\sum_{1}^{\infty} \left(\frac{n}{2n+1}\right)^n$ converges or diverges.

Solution. We have

$$\lim_{n\to\infty} \left[\left(\frac{n}{2n+1}\right)^n\right]^{1/n} = \lim_{n\to\infty} \frac{n}{2n+1} = \frac{1}{2} < 1.$$

Hence by Cauchy's Root Test, the given series converges.

Example 19. Test the series $\sum_{1}^{\infty} n^n$ for convergence or divergence.

Solution. Since the series involves powers of n, we use Cauchy's Root Test.

$$\lim_{n \to \infty} (n^n)^{1/n} = \lim_{n \to \infty} n = \infty.$$

Hence the series diverges by (8.23(ii)).

Example 20. Determine whether the series $\sum_{1}^{\infty} \left(\frac{2^n}{n^3}\right)$ converges or diverges.

Solution. We Have

$$\lim_{n \to \infty} \left(\frac{2^n}{n^3}\right)^{1/n} = \lim_{n \to \infty} \frac{2}{n^{3/n}} = \frac{2}{\lim\limits_{n \to \infty} n^{3/n}} = \frac{2}{1} = 2 > 1,$$

$$\left(\text{since } \lim_{n \to \infty} n^{1/n} = 1\right)$$

The series diverges by Cauchy's Roots Test.

EXERCISE 8.3

Apply the Ratio Test to determine whether the given series converges or diverges **(Problems 1-10):**

1. $\sum_{0}^{\infty} \frac{2^n}{(2n)!}$ 2. $\sum_{1}^{\infty} \frac{n!}{n^n}$ 3. $\sum_{1}^{\infty} \frac{7^n}{n(5^{n+1})}$

4. $\sum_{1}^{\infty} \frac{n}{n^2+1}$ 5. $\sum_{1}^{\infty} \frac{(2n)!}{4^n}$ 6. $\sum_{1}^{\infty} \frac{2^n}{n(n+2)}$

7. $\sum_{1}^{\infty} \frac{(n+1)(n+2)}{n!}$ 8. $\sum_{1}^{\infty} n^3 e^{-n^4}$ 9. $\sum_{1}^{\infty} \frac{(n!)^2}{(3n)!}$

10. $\sum_{1}^{\infty} \frac{(n+2)!}{4!\, n!\, 2^n}$

Apply Cauchy's Root Test to determine whether the given series converges or diverges **(Problems 11-16):**

11. $\sum_{1}^{\infty} \left(\frac{3n+2}{2n-1}\right)^n$ 12. $\sum_{1}^{\infty} \frac{1}{n^n}$ 13. $\sum_{1}^{\infty} \left(\frac{n}{10}\right)^n$

14. $\sum_{1}^{\infty} \left(\frac{n}{1+n^3}\right)^n$ 15. $\sum_{1}^{\infty} \frac{3^n}{n^3}$ 16. $\sum_{1}^{\infty} n\left(\frac{\pi}{n}\right)^n$

In **Problems 17-36,** apply any appropriate test to determine the convergence or divergence of the series:

17. $\sum_{1}^{\infty} \frac{e^n}{(\ln n)^n}$ 18. $\sum_{1}^{\infty} \frac{1^n + 2^n}{3^n}$ 19. $\sum_{1}^{\infty} \left(1 + \frac{1}{\sqrt{n}}\right)^{-n^{3/2}}$

20. $\sum_{1}^{\infty} \frac{\ln n}{e^n}$ 21. $\sum_{1}^{\infty} \frac{(n!)^2\, 2^n}{(2n+2)!}$ 22. $\sum_{1}^{\infty} \frac{\arctan n}{n^2}$

23. $\sum_{1}^{\infty} \left(\frac{5n}{2n+1}\right)^{3n}$ 24. $\sum_{1}^{\infty} \left(\frac{n!}{n^n}\right)^n$ 25. $\sum_{1}^{\infty} \frac{5\sqrt{n}+1}{\sqrt{n^3 - 2n^2 + 3}}$

26. $\sum_{0}^{\infty} \frac{2^n + n}{(n+1)!}$ 27. $\sum_{1}^{\infty} \frac{\sqrt{n}}{2^n}$ 28. $\sum_{1}^{\infty} \frac{n^n \cdot 2^n}{(n+2)!}$

29. $\sum_{1}^{\infty} \frac{(2n+1)!}{n^2\,(n+1)!}$ 30. $\sum_{1}^{\infty} \frac{(2n+1)(3^n+1)}{4^n+1}$ 31. $\sum_{1}^{\infty} \frac{2^{2n-1}}{(2n-1)!}$

32. $\sum_{1}^{\infty} \frac{n!}{e^{n^2}}$ 33. $\sum_{2}^{\infty} \frac{1}{\sqrt{n}\,(\ln n)^3}$

34. $1 + \frac{1\cdot 2}{1\cdot 3} + \frac{1\cdot 2\cdot 3}{1\cdot 3\cdot 5} + \frac{1\cdot 2\cdot 3\cdot 4}{1\cdot 3\cdot 5\cdot 7} + \cdots$ 35. $\frac{2}{5} + \frac{2\cdot 4}{5\cdot 8} + \frac{2\cdot 4\cdot 6}{5\cdot 8\cdot 11} + \frac{2\cdot 4\cdot 6\cdot 8}{5\cdot 8\cdot 11\cdot 14} + \cdots$

36. $\sum_{1}^{\infty} \frac{(a+1)(2a+1)\cdots(na+1)}{(b+1)(2b+1)\cdots(nb+1)}$, $a > 0$, $b > 0$.

37. If $x > 0$, show that the series $\sum_{1}^{\infty} \frac{(n!)^2}{(2n)!} x^n$ converges for $x < 4$.

38. If $x > 0$, prove that the series $\sum_{1}^{\infty} \frac{1\cdot 3\cdot 5 \cdots (2n-1)}{1\cdot 4\cdot 7 \cdots (3n-2)} x^n$ converges for $x < \frac{3}{2}$.

39. Show that $\sum_{1}^{\infty} \frac{x^n}{n!}$ converges for all positive values of x.

40. Find those positive values of x for which the series $1 + \sum_{1}^{\infty} \frac{x^{2n}}{2n}$ converges.

ALTERNATING SERIES

So far we have studied series with positive terms (of course the series with all negative terms are handled exactly in the same way). Now we shall discuss infinite series having both positive and negative terms. Such series are called **mixed series**. If only a finite number of terms have positive and negative signs, we can delete those terms and examine the resulting series by the methods of the previous sections.

An important type of mixed series is the one whose terms are alternately positive and negative. Such a series is called an **alternating series**.

(8.24) Definition. If $a_n > 0$ for all n, then the series

$$\sum_{1}^{\infty} (-1)^{n-1} a_n = a_1 - a_2 + a_3 - \cdots + (-1)^{n-1} a_n + \cdots$$

and the series

$$\sum_{1}^{\infty} (-1)^n a_n = -a_1 + a_2 - a_3 + \cdots + (-1)^n a_n + \cdots$$

are called alternating series.

The following theorem gives a test for the convergence of an alternating series.

(8.25) Theorem (Alternating Series Test)[1]. Let $a_n > 0$, $n = 1, 2, 3, \cdots$. The alternating series

$$\sum_{1}^{\infty} (-1)^{n-1} a^n = a_1 - a_2 + a_3 - \cdots + (-1)^{n-1} a_n + \cdots$$

converges if the following two conditions hold:

(i) $\{ a_n \}$ is a nonincreasing sequence, i.e., $a_1 \geq a_2 \geq a_3 \geq \cdots$

(ii) $\lim_{n \to \infty} a_n = 0$.

1. Also called Leibnitz Test.

Proof. Let S_n denote the nth partial sum of the series, i.e.

$$S_n = \sum_{k=1}^{n} (-1)^{k-1} a_k. \text{ Then}$$

$$S_{2n} = (a_1 - a_2) + (a_3 - a_4) + \cdots + (a_{2n-1} - a_{2n})$$

From condition (i), it follows that the expression in each of the parentheses is nonnegative. Hence

$$S_2 \leq S_4 \leq S_6 \leq \cdots$$

i.e. $\{ S_{2n} \}$ is a nondecreasing sequence.

Again,

$$S_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n} \leq a_1, \text{ since}$$

the expression in each of the parentheses is nonnegative.

Thus, $\{ S_{2n} \}$ is a monotonically increasing sequence which is bounded above. Therefore $\{ S_{2n} \}$ is a convergent sequence.

$$\lim_{n \to \infty} S_{2n} = S$$

$$S_{2n+1} = S_{2n} + a_{2n+1}$$

Hence

$$\lim_{n \to \infty} S_{2n+1} = \lim_{n \to \infty} S_{2n} + \lim_{n \to \infty} a_{2n+1} = S + 0 = S,$$

because by (ii), $\lim_{n \to \infty} a_{2n+1} = 0$.

Therefore, the sequence $\{ S_{2n} \}$ of even partial sums and the sequence $\{ S_{2n+1} \}$ of odd partial sums have the same limit S.

We shall show that $\lim_{n \to \infty} S_n = S$. Since $\lim_{n \to \infty} S_{2n} = S$, for any $\varepsilon > 0$ there is a positive integer n_1 such that

$$| S_{2n} - S | < \varepsilon, \qquad \text{for } 2n \geq n_1.$$

And, because $\lim_{n \to \infty} S_{2n+1} = S$, for $\varepsilon > 0$, there exists a positive integer n_2 such that

$$| S_{2n+1} - S | < \varepsilon, \qquad \text{for } 2n + 1 \geq n_2.$$

Let $N = \max \{ n_1, n_2 \}$. It follows that if n is any positive integer, even or odd, then

$$| S_n - S | < \varepsilon, \qquad \text{for } n \geq N.$$

Thus $S_n \to S$ as $n \to \infty$. The convergence of the series follows from **Definition 8.11**.

Example 21. The series

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\dots+(-1)^{n-1}\frac{1}{n}+\dots \quad \text{converges, since}$$

$$a_n = \frac{1}{n} > \frac{1}{n+1} = a_{n+1} \quad \text{and} \quad \lim_{n\to\infty} a_n = \lim_{n\to\infty}\frac{1}{n} = 0$$

and the two conditions of (8.25) are satisfied.

Example 22. Determine whether the alternating series

$$\sum_{1}^{\infty}(-1)^{n-1}\frac{n+1}{3n+1} \quad \text{converges or diverges.}$$

Solution. Here $\quad a_n = \dfrac{n+1}{3n+1}$

So, $$\lim_{n\to\infty} a_n = \lim_{n\to\infty}\frac{n+1}{3n+1} = \frac{1}{3} \neq 0$$

The second condition of Theorem 8.25 is not satisfied. Hence the given series does not converge.

ABSOLUTE AND CONDITIONAL CONVERGENCE

(8.26) Definition. A series $\sum_{1}^{\infty} a_n$ is said to converge absolutely if the series $\sum_{1}^{\infty} |a_n|$ converges.

(8.27) Definition. A series $\sum_{1}^{\infty} a_n$ is said to converge conditionally if $\sum_{1}^{\infty} a_n$ converges and $\sum_{1}^{\infty} |a_n|$ diverges.

Example 23. The series

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\dots \tag{1}$$

converges conditionally since the series $1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\dots$ of absolute values of the terms of (1) is divergent.

The series

$$\frac{1}{1^2}-\frac{1}{2^2}+\frac{1}{3^2}-\frac{1}{4^2}+\dots$$

is absolutely convergent since the series

$$\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\dots$$

converges.

(8.28) Theorem. If the series $\sum_{1}^{\infty}|a_n|$ converges then so does the series $\sum_{1}^{\infty} a_n$, Q 11

that is, if a series converges absolutely then it converges.

Proof. Let

$$\sum_{1}^{\infty}|a_n| = |a_1|+|a_2|+|a_n|+\dots \quad \text{be convergent.}$$

Consider the series $\sum_{1}^{\infty}(a_n+|a_n|)$.

We have

$$a_n+|a_n| = \begin{cases} 0 & \text{if } a_n \text{ is negative.} \\ 2|a_n| & \text{if } a_n \text{ is positive.} \end{cases}$$

Hence

$$0 \le a_n+|a_n| \le 2|a_n| \tag{1}$$

Since $\sum_{1}^{\infty}|a_n|$ converges by the given hypothesis, the series $\sum_{1}^{\infty}2|a_n| = 2\sum_{1}^{\infty}|a_n|$ converges. Therefore, by the Basic Comparison Test, we infer from (1) that the series $\sum_{1}^{\infty}(a_n+|a_n|)$ converges. But

$$\sum_{1}^{\infty} a_n = \sum_{1}^{\infty}[(a_n+|a_n|)-|a_n|].$$

From the convergence of the two series $\sum_{1}^{\infty}(a_n+|a_n|)$ and $\sum_{1}^{\infty}|a_n|$, it follows that $\sum_{1}^{\infty} a_n$ converges by Theorem 8.15 (i).

Note: The converse of this theorem is false as already seen in Example 23.

(8.29) Theorem. (The Ratio Test for Absolute Convergence). Let $\sum_{1}^{\infty} a_n$ be an infinite series and $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = L$, where L is a nonnegative real number or ∞.

(i) If $L < 1$, the series $\sum_{1}^{\infty} a_n$ converges absolutely.

(ii) If $L > 1$ or ∞, the series $\sum_{1}^{\infty} a_n$ diverges.

(iii) If $L = 1$, the test fails and the series may be absolutely convergent, conditionally convergent or divergent.

(8.30) Theorem. (The Root Test for Absolute Convergence). Let $\sum_{1}^{\infty} a_n$ be an infinite series and $\lim_{n\to\infty} (|a_n|)^{1/n} = L$, where L is a nonnegative real number or ∞

(i) If $L < 1$, the series $\sum_{1}^{\infty} a_n$ converges absolutely.

(ii) If $L > 1$ or ∞, the series $\sum_{1}^{\infty} a_n$ is divergent.

(iii) If $L = 1$, the test fails and the series may be absolutely convergent, conditionally convergent or divergent.

Example 24. Determine whether the series

$$\sum_{1}^{\infty} (-1)^n \left(\frac{n^3}{e^n}\right) \quad \text{converges absolutely or diverges}$$

Solution. By the Ratio Test for absolute convergence, we have

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{(n+1)^3 \cdot e^n}{e^{n+1} \cdot n^3}$$

$$= \lim_{n\to\infty} \frac{(n+1)^3}{n^3} \cdot \frac{1}{e} = \lim_{n\to\infty} \left(1+\frac{1}{n}\right)^3 \cdot \frac{1}{e} = 1 \cdot \frac{1}{e} < 1.$$

Hence the series converges absolutely.

Example 25. Test for convergence the series $\sum_{1}^{\infty} \frac{\sin n a}{n^2}$, where a is any real number.

Solution. Consider the series $\sum_{1}^{\infty} \left|\frac{\sin n a}{n^2}\right|$. Since $|\sin x| \le 1$, we have

$$\left|\frac{\sin n a}{n^2}\right| = \frac{|\sin n a|}{n^2} \le \frac{1}{n^2}$$

But $\sum_{1}^{\infty} \frac{1}{n^2}$ converges. By the Basic Comparison test, $\sum_{1}^{\infty} \left|\frac{\sin n a}{n^2}\right|$ converges. The convergence of $\sum_{1}^{\infty} \frac{\sin n a}{n^2}$ follows from **(8.28)**.

Example 26. Determine whether the series

$$\sum_{1}^{\infty} (-1)^{n-1} \frac{n+2}{n(n+3)} \quad \text{converges absolutely, converges conditionally or diverges}$$

Solution. The given series is an alternating series with

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{n+2}{n(n+3)} = 0.$$

Moreover,

$$\frac{|a_{n+1}|}{|a_n|} = \frac{(n+3)}{(n+1)(n+4)} \times \frac{n(n+3)}{n+2} = \frac{n^3+6n^2+9n}{n^3+7n^2+14n+8}$$

$$= \frac{n^3+6n+9}{n^3+7n^2+9n+(n^2+5n+8)} < 1.$$

Therefore, $|a_{n+1}| < |a_n|$.

Hence the series (1) converges by **(8.25)**

Next compare the series of absolute values with $\sum_{1}^{\infty} \frac{1}{n}$. We have

$$\lim_{n\to\infty} \frac{n+2}{n(n+3)} \times n = 1.$$

But $\sum_{1}^{\infty} \frac{1}{n}$ diverges. Hence the series $\sum_{1}^{\infty} \frac{n+2}{n(n+3)}$ also diverges.

So the series does not converge absolutely. Hence (1) is a conditionally convergent series.

161

27. Test the series $\sum_{1}^{\infty} (-1)^n \left(\frac{n+2}{5n+3}\right)^n$ for absolute convergence, conditional convergence or divergence.

Solution. Here, by the Root Test,

$$\lim_{n \to \infty} (|a_n|)^{1/n} = \lim_{n \to \infty} \left[\left(\frac{n+2}{5n+3}\right)^n\right]^{1/n} = \lim_{n \to \infty} \frac{n+2}{5n+3} = \frac{1}{5} < 1$$

Hence the series converges absolutely by (8.30).

EXERCISE 8.4

Use the Alternating Series Test to determine whether the given series converges (**Problems 1–6**):

1. $\sum_{2}^{\infty} (-1)^n \frac{1}{n \ln n}$ 2. $\sum_{1}^{\infty} \frac{\cos n\pi}{\sqrt{n\pi}}$ 3. $\sum_{1}^{\infty} (-1)^{n-1} \frac{n^2}{n^2+1}$

4. $\sum_{1}^{\infty} (-1)^{n-1} \frac{1}{e^n}$ 5. $\sum_{1}^{\infty} (-1)^{n-1} \frac{n+4}{n^2+n}$ 6. $\sum_{1}^{\infty} (-1)^n \frac{\ln n}{n}$

Test the given series for (i) absolute convergence (ii) conditional convergence (iii) divergence (**Problems 7-29**):

7. $\sum_{1}^{\infty} \frac{(-1)^{n-1} n!}{(2n)!}$ 8. $\sum_{1}^{\infty} \frac{\sin \sqrt{n}}{\sqrt{n^3+1}}$ 9. $\sum_{1}^{\infty} \frac{(-1)^n (n+2)}{n(n+1)}$

10. $\sum_{1}^{\infty} \frac{(-1)^n n!}{1 \cdot 3 \cdot 5 \cdots (2n-1)}$ 11. $\sum_{1}^{\infty} (-1)^{n-1} \left(\frac{n+2}{3n-1}\right)^n$ 12. $\sum_{1}^{\infty} (-1)^n \tan\left(\frac{1}{n}\right)$

13. $\sum_{1}^{\infty} (-1)^{n-1} \frac{n^2}{(n+2)!}$ 14. $\sum_{1}^{\infty} (-1)^{n-1} (\sqrt{n+1} - \sqrt{n})$ 15. $\sum_{2}^{\infty} (-1)^n \frac{1}{\ln(\ln n)}$

16. $\sum_{1}^{\infty} (-1)^{n+1} \arctan n$ 17. $\sum_{2}^{\infty} (-1)^{n-1} \frac{1}{n (\ln n)^2}$ 18. $\sum_{1}^{\infty} \frac{(-1)^{n-1} \sqrt{n}}{n+1}$

19. $\sum_{1}^{\infty} \frac{(-1)^{n-1} n^2}{(2n+1)(n+5)}$ 20. $\sum_{1}^{\infty} \frac{(-1)^{n-1} 2 \sinh n}{e^{2n}}$ 21. $\sum_{1}^{\infty} \frac{\cos(n\pi)}{n+3}$

22. $\sum_{1}^{\infty} \frac{n!}{(-2)^n}$ 23. $\sum_{0}^{\infty} \frac{(-2)^n}{3^n+1}$

24. $\sum_{0}^{\infty} (-1)^n \left[\frac{\pi}{2} - \arctan n\right]$ 25. $\sum_{1}^{\infty} (-1)^{n+1} \frac{3n^2-2n+5}{n^3+n^2+n+1}$

26. $\sum_{1}^{\infty} (-1)^n \frac{n^n}{n!}$ 27. $\sum_{1}^{\infty} (-1)^{n+1} \frac{e^n}{n}$

28. $\sum_{1}^{\infty} \frac{(-1)^n n^n}{a^{nb+c}}$, $a > 1$; b, c real 29. $\sum_{1}^{\infty} (-1)^{n+1} \frac{n^n}{3^{n^2}}$

Determine the values of x for which the given series (i) converges absolutely (ii) converges conditionally (iii) diverges (**Problems 30-35**):

30. $\sum_{1}^{\infty} \frac{nx^n}{3^n}$ 31. $\sum_{1}^{\infty} \frac{x^n}{\sqrt{n}}$ 32. $\sum_{0}^{\infty} \frac{4^n}{x^n}$

33. $\sum_{0}^{\infty} \frac{(-1)^n x^n}{3^n (n+1)!}$ 34. $\sum_{1}^{\infty} (-1)^{n-1} \frac{x^n}{n(n+1)}$

-2.728
0.5

35. $\sum_{0}^{\infty} \frac{(-1)^n x^{2n}}{2^{2n} (n!)^2}$ (**Bessel[1] function** of the first kind of order zero).

POWER SERIES

(8.31) Definition. An infinite series of the form

$$c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n + \cdots = \sum_{0}^{\infty} c_n x^n = 0, \qquad (1)$$

is called a **power series** in x. The coefficients c_n are real numbers and x is a real variable.

A series of the form

$$c_0 + c_1(x-a) + c_2(x-a)^2 + \cdots + c_n(x-a)^n + \cdots = c_n (x-a)^n \qquad (2)$$

is a power series in $x - a$, where a is a real number. However, the series (2) can always be reduced to a series of the form (1) by setting $x - a = y$. We shall study power series of the form (1).

1. Friedrich Wilhelm Bessel (1784 – 1846), a German astronomer.

CONVERGENCE OF POWER SERIES

If a numerical value of x is substituted into (1), then it is a seri s of constant terms. The behaviour (convergence or divergence) of such series has been discussed in the previous sections.

To check that for which values o... the series (1) converges, we use the Ratio Test for absolute convergence (Theorem 8.29) or the Root Test (Theorem 8.30).

Example 28. Find the value of x for which the power series $\sum_{n=0}^{\infty} (n)^{2n} x^{2n}$ converges.

Solution. Here, we let $a_n = (n)^{2n} x^{2n}$ and $a_{n+1} = [(n+1)]^{2(n+1)} x^{2(n+1)}$

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)^2 (n+1)^{2n} x^{2n} . x^2}{n^{2n} x^{2n}}$$

$$= (n+1)^2 \left(\frac{n+1}{n}\right)^{2n} x^2 = \left[(n+1)\left(1+\frac{1}{n}\right)^n x\right]^2$$

When $x = 0$, $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 0.$

When $x \neq 0$, $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \infty.$

By the Ratio Test for absolute convergence, the series converges for $x = 0$ and diverges for all nonzero values of x.

It is more convenient to use the Root Test here. We have

$$\lim_{n\to\infty} |a_n|^{1/n} = \lim_{n\to\infty} [(n)^{2n} x^{2n}]^{1/n}$$

$$= \lim_{n\to\infty} n^2 |x|^2$$

$$= \begin{cases} 0 & \text{if } x = 0 \\ \infty & \text{if } x \neq 0 \end{cases}$$

By the Root Test (8.30), the series converges for $x = 0$ and diverges for all nonzero values of x.

Example 29. Find the values of x for which the power series $\sum_{n=0}^{\infty} \frac{x^n}{(2n)!}$ converges.

Solution. We have

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{x^{n+1}}{(2n+2)!}\,\frac{(2n)!}{x^n}\right| = \frac{|x|}{(2n+2)(2n+1)} \to 0 \quad \text{as } n \to \infty.$$

Thus the series converges absolutely for all values of x.

Example 30. Determine the values of x for which the power series $\sum_{n=2}^{\infty} \frac{x^n}{\ln n}$ converges absolutely, converges conditionally and diverges.

Solution. By the Ratio Test for absolute convergence, we have

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{x^{n+1}}{\ln(n+1)}\,\frac{\ln n}{x^n}\right| = \frac{\ln n}{\ln(n+1)}\,|x|$$

Now, $$\lim_{n\to\infty} \frac{\ln n}{\ln(n+1)} = \lim_{n\to\infty} \frac{\frac{1}{n}}{\frac{1}{n+1}} = \lim_{n\to\infty} \frac{n+1}{n} = 1.$$

Thus, $$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = |x|.$$

Hence the series converges absolutely if $|x| < 1$ and diverges if $|x| > 1$.

When $|x| = 1$, the test fails.

For $x = -1$, the series becomes $\sum_{n=2}^{\infty} \frac{(-1)^n}{\ln n}$.

Here $\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{1}{\ln n} = 0.$

By the Alternating Series Test (8.25), the series converges.

For $x = 1$, the series becomes $\sum_{n=2}^{\infty} a_n = \sum_{n=2}^{\infty} \frac{1}{\ln n}$

Compare it with $\sum_{n=2}^{\infty} \frac{1}{n} = \sum_{n=2}^{\infty} b_n$.

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{n}{\ln n} = \lim_{n\to\infty} \frac{1}{\frac{1}{n}} = \infty.$$

Since $\sum\limits_{n=2}^{\infty} b_n$ diverges, $\sum\limits_{n=2}^{\infty} a_n$, also diverges by the Basic Comparison Test (8.19).

Hence the series converges conditionally for $x = -1$ and diverges for $x = 1$.

(8.32) Theorem. If the power series $\sum\limits_{n=0}^{\infty} c_n x^n$ converges for $x = x_1$, then it converges absolutely for all x such that $|x| < |x_1|$.

Proof. If $\sum\limits_{n=0}^{\infty} c_n x_1^n$ converges, then $\lim\limits_{n \to \infty} c_n x_1^n = 0$. Choose $\varepsilon = 1$. There exists a positive integer N such that.

$$|c_n x_1^n| < 1 \qquad \text{for all } n \geq N.$$

Let x be any real such that $|x| < |x_1|$ i.e. $\left|\dfrac{x}{x_1}\right| < 1$.

Now

$$|c_n x^n| = \left|\frac{c_n x^n x_1^n}{x_1^n}\right| = |c_n x_1^n| \left|\frac{x}{x_1}\right|^n < \left|\frac{x}{x_1}\right|^n, \text{ for all } n \geq N.$$

The series $\sum\limits_{n=0}^{\infty} \left|\dfrac{x}{x_1}\right|^n$ is a geometric series with common ratio $\left|\dfrac{x}{x_1}\right| < 1$ and so it is a convergent series. Since $\sum\limits_{n=0}^{\infty} |c_n x^n| < \sum\limits_{n=0}^{\infty} \left|\dfrac{x}{x_1}\right|^n$, by the Basic Comparison Test, $\sum\limits_{n=0}^{\infty} |c_n x^n|$ converges for $|x| < |x_1|$.

(8.33) Theorem. If the power series $\sum\limits_{n=0}^{\infty} c_n x^n$ diverges for $x = x_2$, then it diverges for all x such that $|x| > |x_2|$.

Proof. Suppose to the contrary that the series converges for some x such that $|x| > |x_2|$. Then by Theorem 8.32, the series must converge when $x = x_2$. This contradicts the hypothesis that the series diverges for $x = x_2$. Hence the given series is divergent for all values of x for which $|x| > |x_2|$.

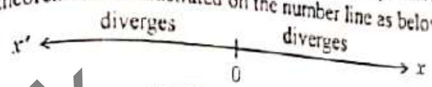
(8.34) Theorem. Let $\sum\limits_{n=0}^{\infty} c_n x^n$ be a given power series. Then exactly one of the following conditions holds:

(i) The series converges only for $x = 0$.

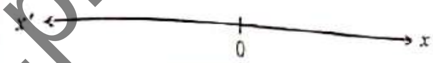
(ii) The series converges absolutely for all values of x.

(iii) The series converges absolutely for all values of x for which $|x| < R$ and diverges for all values of x for which $|x| > R$, where R is a positive number.

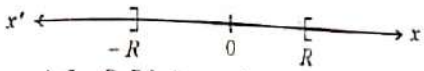
The theorem can be illustrated on the number line as below:

(i)


convergent at $x = 0$

(ii)


convergent for all x

(iii)


converges in $]-R, R[$, diverges for $x < -R$ and $x > R$. The series may converge or diverge at the end points $-R$ and R.

Proof. (i) If $x = 0$, the series becomes $c_0 + 0 + 0 + \cdots = c_0$. Thus every power series of the form $\sum\limits_{n=0}^{\infty} c_n x^n$ is convergent when $x = 0$. If this is the only point for which the series converges, then condition (i) is true.

(ii) Let the series be convergent for $x = x_1$, where $x_1 \neq 0$. By Theorem 8.32, the series is absolutely convergent for all x such that $|x| < |x_1|$. If the series is not divergent for any value of x, then we conclude that it converges absolutely for all values of x.

(iii) If the given series is convergent for $x = x_1$, where $x_1 \neq 0$ and is divergent for $x = x_2$, where $|x_2| > |x_1|$, then by Theorem 8.33, the series is divergent for all x for which $|x| > |x_2|$.

Hence $|x_2|$ is an upper bound for the set of values of $|x|$ for which the series is absolutely convergent. By the completeness axiom of real numbers, the set of upper bounds has the least upper bound which is the number R of condition (iii).

Thus we have proved that exactly one of the three conditions hold.

If the power series is of the form $\sum\limits_{n=0}^{\infty} c_n (x-a)^n$, then conditions **(i)** and **(iii)** of Theorem **8.34** become.

(i) The series converges only for $x = a$.

(iii) The series is absolutely convergent for all values of x for which $|x - a| < R$ i.e. for $x \in \left]a - R, a + R\right[$ converges absolutely and diverges if $x < a - R$ or $x > a + R$, R being a positive number.

(8.35) Definition. The set of all values of x for which a power series converges is called the **interval of convergence** of the power series. The number R of Theorem 8.34 **(iii)** is called the **radius of convergence** of the series.

Every power series has an interval of convergence and radius of convergence. If the power series $\sum\limits_{n=0}^{\infty} c_n x^n$ converges only for $x = 0$ then its interval of convergence reduces to the point 0 and its radius of convergence is zero.

If the power series converges for all values of x, then its interval of convergence is $\left]-\infty, \infty\right[$ and its radius of convergence is ∞.

If the power series converges for $|x| < R$ and diverges for $|x| > R$, then its radius of convergence is R and its interval for convergence is one of the intervals $\left]-R, R\right[$, $\left[-R, R\right]$, $\left]-R, R\right]$ or $\left[-R, R\right[$.

For the general power series $\sum\limits_{n=0}^{\infty} c_n (x - a)^x$, the interval of convergence is the point a if it converges only for $x = a$ and its radius of convergence is 0. If it converges for all values of x, then its interval of convergence is $\left]-\infty, \infty\right[$ and its radius of convergence is ∞. If the series converges for $|x - a| < R$ and diverges for $|x - a| > R$, then its radius of convergence is R and its interval of convergence is one of the following

$$\left]a - R, a + R\right[, \left[a - R, a + R\right], \left]a - R, a + R\right] \text{ or } \left[a - R, a + R\right[$$

Example 31. Find the interval of convergence of the power series $\sum\limits_{n=0}^{\infty} \frac{n!}{n^n} x^n$.

Solution.

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{(n+1)!\, x^{n+1}}{(n+1)^{n+1}} \cdot \frac{n^n}{n!\, x^n}\right|$$

$$= \left|\frac{(n+1)\, n!\, x^{n+1}}{(n+1)\,(n+1)^n} \cdot \frac{n^n}{n!\, x^n}\right|$$

$$= \left|\frac{x}{(n+1)^n} \cdot n^n\right| = \frac{1}{\left(\frac{n+1}{n}\right)^n}|x| = \frac{1}{\left(1 + \frac{1}{n}\right)^n}|x|$$

$$\lim_{n \to \infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n \to \infty}\frac{1}{\left(1 + \frac{1}{n}\right)^n}|x| = \frac{|x|}{\lim\limits_{n \to \infty}\left(1 + \frac{1}{n}\right)^n} = \frac{|x|}{e}.$$

The series converges absolutely for $\frac{|x|}{e} < 1$ i.e. for $|x| < e$

If $\frac{|x|}{e} = 1$, the series becomes $\sum\limits_{n=1}^{\infty} \frac{n!}{n^n} e^n$ or $\sum\limits_{n=1}^{\infty} (-1)^n \frac{n!}{n^n} e^n$ according as $x = e$ or $x = -e$.

Now $\frac{n!}{n^n} e^n = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot \frac{1}{n^n} e^n = \sqrt{2\pi n} \to \infty$ as $n \to \infty$.

[If n is large then by Stirling's formula $n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$]

Thus both of the above series diverge

Therefore, the interval of convergence of the given series is $\left]-e, e\right[$

Example 32. Find the radius of convergence and interval of convergence of the power series $\sum\limits_{n=0}^{\infty} \frac{(-1)^{n+1}(x+1)^{2n}}{(n+1)^2 5^n}$

Solution.

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{(-1)^{n+2}(x+1)^{2n+2}}{(n+2)^2 5^{n+1}} \cdot \frac{(n+1)^2 5^n}{(-1)^{n+1}(x+1)^{2n}}\right|$$

$$= \frac{(x+1)^2}{5}\,\frac{(n+1)^2}{(n+2)^2} \to \frac{(x+1)^2}{5} \quad \text{as } n \to \infty$$

The series converges absolutely for $(x+1)^2 < 5$ and diverges for $(x+1)^2 > 5$ i.e. converges for $|x+1| < \sqrt{5}$ and diverges for $|x+1| > \sqrt{5}$

The radius of convergence is $R = \sqrt{5}$.

Now $|x+1| < \sqrt{5}$ is equivalent to

$$-\sqrt{5} - 1 < x < \sqrt{5} - 1 \left[.\right.$$

or $\quad x \in \left]-\sqrt{5} - 1, \sqrt{5} - 1\right[$.

If $\quad x = -\sqrt{5} - 1$, the series becomes

$$\sum_{n=0}^{\infty} \frac{(-1)^{n+1}(\sqrt{5})^{2n}}{(n+1)^2\, 5n} = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(n+1)^2}$$

which is a convergent series.

Similarly, the series converges at $x = \sqrt{5} - 1$.

Thus interval of convergence is $\left[-\sqrt{5} - 1, \sqrt{5} - 1\right]$.

DIFFERENTIATION AND INTEGRATION OF POWER SERIES

(8.36) Each power series $\sum\limits_{n=0}^{\infty} c_n x^n$ defines a function f, where

$$f(x) = \sum_{n=0}^{\infty} c_n x^n$$

for each x in the interval of convergence of the power series $\sum\limits_{n=0}^{\infty} c_n x^n$ is called a **power series representation** of $f(x)$. The function f is continuous and differentiable. Taylor's and Maclaurin's formulas are used to obtain a power series representation of a function having derivatives of any order on some interval. It is known that the geometric series $1 + x + x^2 + \cdots$ converges for $-1 < x < 1$ and its sum is $\frac{1}{1-x}$. Thus a power series representation of $\frac{1}{1-x}$ is the power series $1 + x + x^2 + \cdots$ whose radius of convergence is 1.

Having obtained the function f represented by power series, a natural question arises whether this function can be differentiated and integrated. The answer is in affimative and new power series can be obtained by term by term differentiation and integration of a given power series within its interval of convergence. The main properties of such a function are stated in the next theorem.

(8.37) Theorem. Let the power series $\sum\limits_{n=0}^{\infty} c_n x^n$ have a nonzero radius of convergence R and $f(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n + \ldots$

Then for every number x within the interval of convergence of the given series, the following properties hold:

(i) f is continuous, i.e.

$$\lim_{x \to a} \left(\sum_{n=0}^{\infty} c_n x^n \right) = \sum_{n=0}^{\infty} \lim_{x \to a} c_n x^n = \sum_{n=0}^{\infty} c_n a^n, \text{ where } |a| < R$$

(ii) f is differentiable and

$$f'(x) = \frac{d}{dx}\left(\sum_{n=0}^{\infty} c_n x^n \right) = \sum_{n=0}^{\infty} \frac{d}{dx}(c_n x^n) = \sum_{n=1}^{\infty} n c_n x^{n-1}$$

i.e., the power series can be differentiated term by term.

(iii) f is integrable and

$$\int_0^x \left(\sum_{n=0}^{\infty} c_n t^n \right) dt = \sum_{n=0}^{\infty} \left(\int_0^x c_n t^n \, dt \right) = \sum_{n=0}^{\infty} \frac{c_n x^{n+1}}{n+1}$$

i.e. the power series can be integrated term by term.

The proof of this theorem is complicated and is omitted.

It is easy to check that the new series obtained in (ii) and (iii) after differentiation and integration have the same radius of convergence R as the given power series $\sum\limits_{n=0}^{\infty} c_n x^n$.

Example 33. Show that

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

Solution. $\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{x^{n+1}}{(n+1)!} \times \frac{n!}{x^n}\right| = \left|\frac{x}{n+1}\right| \to 0$ as $n \to \infty$.

Thus $\sum\limits_{n=0}^{\infty} \frac{x^n}{n!}$ converges absolutely for all values of x.

Let $$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (1)$$

By **8.37 (ii)**, we have

$$f'(x) = \sum_{n=1}^{\infty} \frac{n x^{n-1}}{n!}$$

$$= \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (2)$$

From (1) and (2), we get

$$f'(x) = f(x)$$

i.e. $$\frac{f'(x)}{f(x)} = 1$$

or $$\ln f(x) = x + \ln c$$

or $$\frac{f(x)}{c} = e^x \quad \text{or} \quad f(x) = c e^x. \quad (3)$$

From (1), $f(0) = 1$ (We take $x^0 = 1$ even when $x = 0$)

Thus from (3), $f(0) = 1 = c$.

Hence $f(x) = e^x$ as desired.

Example 34. Use Theorem 8.37 and the geometric series $1 + x + x^2 + \cdots$ to find a power series representation of $\ln \frac{1}{1-x}$. Hence show that $\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$

Solution. We have

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots, \qquad (-1 < x < 1) \qquad (1)$$

Integrating both sides of (1), we get

$$\int_0^x \frac{1\,dt}{1-t} = x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots, \qquad \text{by 8.37 (iii)}$$

Now $\int_0^x \frac{dt}{1-t} = \left[-\ln(1-t)\right]_0^x = -\ln(1-x) = \ln\frac{1}{1-x}$

Thus, $\ln \frac{1}{1-x} = x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^n}{n} + \cdots$

Putting $x = -1$ into the above equation, we obtain

$$-\ln 2 = -1 + \frac{1}{2} - \frac{1}{3} + \cdots$$

or $\quad \ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \cdots \quad$ as required.

Example 35. Derive the series

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + \cdots, \quad -1 \le x \le 1$$

This is known as **Gregory's**[1] **Series**.

Solution. In (1) of Example 34, insert $-x^2$ for x. The result is

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots \qquad -1 < x < 1$$

Integrating, we have

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots \qquad (1)$$

The radius of convergence of the above series is $R = 1$ and we test the points $x = 1, x = -1$. For $x = 1$, the series is

$$\arctan 1 = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}$$

which converges by the Alternating Series Test.

1. Named after the Scottish mathematician James Gregory (1638 – 1675)

For $x = -1$, the series becomes

$$\arctan(-1) = -1 + \frac{1}{3} - \frac{1}{5} + \frac{1}{7} - \cdots$$

which is also convergent.

Thus the interval of convergence of (1) is $[-1, 1]$

Example 36. Compute $\int_0^1 e^{-x^2}\,dx$

Solution. The integrand does not have an elementary function as its antiderivative. But it has a power series expansion. The integral may be computed by integrating the series term by term. We have

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

Replace x by $-x^2$ in the above equation, then

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} - \cdots$$

Using Theorem 8.37 (iii), we get

$$\int_0^1 e^{-x^2}\,dx = \left[x - \frac{x^3}{3} + \frac{x^5}{5\cdot 2!} - \frac{x^7}{7\cdot 3!} + \cdots\right]_0^1$$

$$= 1 - \frac{1}{3} + \frac{1}{5\cdot 2!} - \frac{1}{7\cdot 3!} + \cdots$$

ADDITION AND MULTIPLICATION OF POWER SERIES

(8.38) Let $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ be two power series with a common interval of convergence. Then

(i) $\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n + x_n) x^n$

(ii) $\left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty} c_n x^n$

where $c_n = a_0 b_n + a_1 b_{n-1} + \ldots + a_{n-1} b_1 + a_n b_0$

i.e., power series may be added and multiplied together much the same way as polynomials.

A power series may also be divided by another power series in a manner similar to the process of division of polynomials.

The interval of convergence of each of the new power series obtained by algebraic operations is the common interval of convergence of the two given power series.

Example 37. Using the power series expansions of $\ln(1+x)$ and $\frac{1}{1+x^2}$, find the power series of $\frac{\ln(1+x)}{1+x^2}$.

Solution. We have

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots \quad (1)$$

and $$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots \quad (2)$$

Integrating (1), we get

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \quad (3)$$

The power series (2) and (3) converge for $|x| < 1$. Multiply (2) and (3) as below.

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}$$

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots$$

$$\begin{aligned}\frac{\ln(1+x)}{1+x^2} &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots \\ &\quad - x^3 + \frac{x^4}{2} - \frac{x^5}{3} + \cdots \\ &\quad + x^5 - \cdots \\ &= x - \frac{x^2}{2} - \frac{2}{3}x^3 + \frac{x^4}{4} + \frac{13}{15}x^5 - \cdots\end{aligned}$$

This series converges for $|x| < 1$.

EXERCISE 8.5

In each of the following, find the radius of convergence and interval of convergence (Problems 1-24).

1. $\sum_{n=0}^{\infty} \frac{x^n}{(2n)!}$
2. $\sum_{n=0}^{\infty} \frac{2^n x^n}{\ln(n+2)}$
3. $\sum_{n=2}^{\infty} \frac{(x-5)^n \ln n}{n+1}$
4. $\sum_{n=1}^{\infty} \frac{\sin n\pi x}{n^2}$
5. $\sum_{n=1}^{\infty} n^2 (x-2)^n$
6. $\sum_{n=1}^{\infty} \frac{n!\, x^n}{(2n)!}$
7. $\sum_{n=1}^{\infty} \frac{n\, 2^n (x-1)^n}{n+1}$
8. $\sum_{n=0}^{\infty} \frac{(x-2)^n}{2^n \sqrt{n+1}}$
9. $\sum_{n=1}^{\infty} x^{2n}$
10. $\sum_{n=2}^{\infty} \frac{x^n}{(\ln n)^n}$
11. $\sum_{n=1}^{\infty} n^n (x+1)^n$
12. $\sum_{n=1}^{\infty} \frac{2^n (x-3)^n}{n^2}$
13. $\sum_{n=1}^{\infty} \frac{1+(-1)^n}{n} x^n$
14. $\sum_{n=0}^{\infty} \left(\frac{x^2+3}{12}\right)^n$
15. $\sum_{n=2}^{\infty} \frac{(-1)^n x^n}{n(\ln n)^2}$
16. $\sum_{n=0}^{\infty} \frac{n x^n}{(n+1)(n+2)2^n}$
17. $\sum_{n=0}^{\infty} (-1)^n 2^n \sin^n x, \left(-\frac{\pi}{2} \le x \le \frac{\pi}{2}\right)$
18. $\sum_{n=2}^{\infty} \frac{(x-c)^n \ln n}{c^n}$
19. $\sum_{n=0}^{\infty} \frac{(ax+b)^n}{c^n}, a > 0, c > 0$
20. $\sum_{n=1}^{\infty} \frac{n(x-1)^n}{2^n(3n-1)}$
21. $\sum_{n=1}^{\infty} \frac{n^n}{n!} x^n$
22. $\sum_{n=1}^{\infty} (-1)^n \frac{1\cdot 3\cdot 5 \cdots (2n-1)}{2\cdot 4\cdot 6 \cdots 2n} x^{2n+1}$
23. $\sum_{n=1}^{\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) x^n$
24. $\sum_{n=1}^{\infty} \frac{n!\, x^n}{1\cdot 4\cdot 7 \cdots (3n-2)}$
25. Obtain a power series representation of $\frac{x}{(1+x^2)^2}$ if $|x| < 1$.

26. Find the sum of the series $2 + 6x + 2x^2 + 20x^3 + \cdots$, $|x| < 1$

27. Using power series representation of $\frac{e^x - 1}{x}$, show that $\sum_{n=1}^{\infty} \frac{n}{(n+1)!} = 1$

28. Find a series of powers of x that converges to $\tan x$, $-\frac{\pi}{2} < x < \frac{\pi}{2}$

29. Use power series to find the value of $\int_0^{\frac{1}{2}} \frac{dx}{1+x^4}$ to three places of decimal

30. Estimate $\int_0^1 x^2 e^{-x^2}\,dx$ to three places of decimal

Chapter 9

FIRST-ORDER DIFFERENTIAL EQUATIONS

Mathematical models for real world phenomena often take the form of equations ...ing various quantities and their rates of change (derivatives). For example, the ...tion of a particle involves the distances covered in time t and velocity v and/or ...celeration a. Now the rate of change $\frac{ds}{dt}$ of s with respect to t is the velocity v and rate of change $\frac{dv}{dt}$ of velocity with respect to t is the acceleration a. A particle moving in a ...aight line has an equation of motion as $s = f(t)$, where t is in seconds and s is in meters. ...s velocity satisfies the equation

$$v = \frac{ds}{dt} = 4t^2 + 5t - 3.$$

This leads us to the definition of a differential equation (D.E.).

D.E. AND THEIR CLASSIFICATIONS

(9.1) Definition. An equation involving one dependent variable and its derivatives with respect to one or more independent variables, is called differential equation. For example

(i) $\frac{dy}{dx} + y\cos x = \sin x$

(ii) $\dfrac{d^2y}{dx^2} + xy\left(\dfrac{dy}{dx}\right)^2 = 0$

(iii) $\left[1 + \left(\dfrac{dy}{dx}\right)^2\right]^{3/2} = \dfrac{d^2y}{dx^2}$

(iv) $x\dfrac{\partial z}{\partial x} + y\dfrac{\partial z}{\partial y} = nx$

(v) $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2} = 0$

are differential equations

(9.2) Definition. A differential equation, in which ordinary derivatives of the dependent variable with respect to a single independent variable occur, is called an **ordinary differential equation (O.D.E.)** Equations **(i)**, **(ii)** and **(iii)** above are examples of ordinary differential equation.

(9.3) Definition. A differential equation involving partial derivatives of the dependent variable with respect to more than one independent variable is called a **partial differential equation**. Equations **(iv)** and **(v)** given above are partial differential equations.

(9.4) Definition. The **order** of a differential equation is the order of the highest derivative that occurs in the equation.

(9.5) Definition. The **degree** of a differential equation is the greatest exponent of the highest order derivative that appears in the equation. (The dependent variable and its derivatives should be expressed in a form free from radicals and fractions.)

The differential equations given in **(9.1)** have the following orders and degrees

(i) order 1, degree 1

(ii) order 2, degree 1

(iii) order 2, degree 2, exponent of $\dfrac{d^2y}{dx^2}$ is 2 after removing the radical by squaring both sides of the equation.

(iv) order, 1, degree 1

(v) order 2, degree 1.

We shall study ordinary differential equations only.

Recall that a function $T : U \to V$, where U, V, are vector spaces over the same field F, is called **linear** if, for $\alpha, \beta \in F, x, y \in U$,

$$T(\alpha x + \beta y) = \alpha\, T(x) + \beta\, T(y).$$

(9.6) Definition. (Linear Differential Equations). An ordinary differential equation

$$F\left(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \dots, \frac{d^ny}{dx^n}\right) = 0$$

is said to be linear if F is a linear function of the variables

$$x, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \dots, \frac{d^ny}{dx^n}.$$

A similar definition applies to partial differential equations.

Thus the general linear ordinary differential equation of order n is

$$a_0(x)\frac{d^ny}{dx^n} + a_1(x)\frac{d^{n-1}y}{dx^{n-1}} + \dots + a_{n-1}(x)\frac{dy}{dx} + a_n(x)\, y = F(x)$$

where $a_0(x)$ is not identically zero.

The equation $\dfrac{d^3y}{dx^3} + 2e^x\dfrac{d^2y}{dx^2} + y\dfrac{dy}{dx} = x^3$

$\dfrac{d}{dx} \cdot y\, g$

$\dfrac{d}{dx} \cdot \dfrac{dy}{dx}$

is not linear because of the term $y\dfrac{dy}{dx}$.

It should be carefully noted that in a linear ordinary differential equation

$\dfrac{d^2y}{dx^2}$

(i) the dependent variable y and its derivatives are all of degree one.

(ii) no products of y or any of its derivatives appear

(iii) no transcedental function of y and / or its derivatives occur.

A differential equation which is not linear is called a **nonlinear differential equation**.

Differential equations occur in the mathematical formulation of many problems in science and engineering. Some such problems are.

(i) Determining the motion of projectile, rocket, satellite or planet.

(ii) Finding the charge or current in an electric circuit.

(iii) Study of chemical reactions.

(iv) Determination of curves with given geometrical properties.

(9.7) Definition. A **solution** (or **integral**) of a differential equation is a relation between the variables, not containing derivatives, such that this relation and the derivatives obtained from it satisfy the given differential equation identically. For example,

The equation

$$\frac{dy}{dx} = -\lambda y \quad \text{has a solution}$$

$$y = c\, e^{-\lambda x}, \quad \text{where } c \text{ is an arbitrary constant.}$$

The equation

$$\frac{d^2y}{dx^2} + y = 0 \qquad \text{has solutions}$$

$$y = A\cos x,\ \ y = B\sin x \ \text{ and } \ y = A\cos x + B\sin x,$$

where A and B are arbitrary constants.

A solution of a differential equation which contains as many arbitrary constants as the order of the equation is called **general solution (or integral)** of the differential equation. A solution obtained from the general solution by giving particular values to the constants is called a **particular solution** (or integral). The graph of a particular integral is called an **integral curve** of the differential equation.

FORMATION OF DIFFERENTIAL EQUATIONS

(9.8) Given a relation

$$f(x, y, c_1, c_2, \ldots, c_n) = 0 \qquad (1)$$

between variables x, y and containing n constants $c_1, c_2, \ldots, c_n$, it is always possible to form a differential equation of order n such that the given relation (1) is the general solution of the equation. This is done by differentiating (1) n times thereby obtaining n equations and then eliminating the n constants from the original relation and n derived equations. The method is illustrated by means of examples.

Example 1. The equation

$$y = x + a \qquad (1)$$

represents a family of parallel straight lines for different values of a. Elimination of one constant 'a' requires two equations. The second equation is obtained by differentiating (1)

Thus $\frac{dy}{dx} = 1$ is the differential equation of the relation (1), with a eliminated.

Example 2. Form the differential equation by eliminating the two constants A and B from the relation

$$y = A\sin x + B\cos x \qquad (1)$$

Solution. It is clear that three equations are required to eliminate two unknowns A and B. We obtain two other needed equations by successive differentiation of (1). Thus from (1), we have

$$\frac{dy}{dx} = A\cos x - B\sin x \qquad (2)$$

$$\frac{d^2y}{dx^2} = -A\sin x - B\cos x = -y, \text{ using (1)} \qquad (3)$$

So $$\frac{d^2y}{dx^2} + y = 0 \qquad (4)$$

is the required differential equation and (1) is its general solution.

Example 3. Find the differential equation of all parabolas whose axes are parallel to the y-axis.

Solution. General equation of a parabola whose axis is parallel to the y-axis is

$$y = ax^2 + bx + c \qquad (1)$$

In order to obtain its differential equation, we have to eliminate a, b, c from (1). For that we need three more equations. Differentiating (1) successively, we have

$$\frac{dy}{dx} = 2ax + b,$$

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

$$\frac{d^3y}{dx^3} = 0$$

The last equation does not contain any of the constants a, b and c. Thus

$$\frac{d^3y}{dx^3} = 0$$

is the differential equation of all parabolas whose axes are parallel to the y-axis.

INITIAL AND BOUNDARY CONDITIONS

We have observed that general solution of a differential equation contains the same number of arbitrary constants as is the order of the equation. Sometimes we need to find the solutions of differential equations subject to supplementary conditions. Two types of conditions will be often encountered.

(9.9) Definition. (Initial Conditions). It is often required to find the solution of a differential equation subject to certain conditions. If the conditions relate to one value of the independent variable such as $y = y_0$ at $x = x_0$ (written as $y(x_0) = y_0$) and $\frac{dy}{dx} = y'(x_0)$ at $x = x_0$, where x_0 belongs to some interval $]\,\alpha, \beta\,[$ then they are called **initial conditions** (or **one-point boundary conditions**) and x_0 is called the **initial point**. An **initial value problem** consists of a differential equation (of any order) together with a collection of initial conditions that must be satisfied by the solution of the differential equation and derivatives at the initial point.

(9.10) Definition. (Boundary Conditions). The problem of finding the solution of a differential equation such that all the associated constraints relate to two different values of the independent variable is called a **two-point boundary value problem** (or simply a **boundary value problem**). The associated supplementary **boundary conditions** are called **two-point boundary conditions**.

Example 4. Solve $\frac{dy}{dx} = 2x$ (1)

such that $y(1) = 4$.

Solution. This is an initial value problem. We note that $y = x^2 + c$, c being arbitrary constant, is the general solution of (1). Since $y(1) = 4$, we have

$$y(1) = 1^2 + c.$$

Therefore, $4 = 1 + c$ or $c = 3$.

Thus $y = x^2 + 3$ is the solution of the initial value problem (1).

Note that the general solution represents a family of parabolas for different values of c. The solution $y = x^2 + 3$ is a particular member of the family that passes through (1, 4).

Example 5. Solve: $\frac{d^2y}{dx^2} + y = 0$ (1)

subject to the conditions

$$y(\pi/4) = \sqrt{2},\ y'(\pi/4) = \frac{1}{\sqrt{2}}.$$

Solution. Since both the conditions relate to one value of x, namely $x = \pi/4$, this is an initial value problem. We have already noted in Example 2 that

$$y = A\sin x + B\cos x \quad (2)$$

is the general solution of (1). Differentiating (2) w.r.t. x, we get

$$y' = A\cos x - B\sin x. \quad (3)$$

By the given conditions, we have respectively from (2) and (3)

$$y(\pi/4) = \sqrt{2} = A\sin\pi/4 + B\cos\pi/4 = \frac{A}{\sqrt{2}} + \frac{B}{\sqrt{2}}$$

$$y'(\pi/4) = \frac{1}{\sqrt{2}} = A\cos\pi/4 - B\sin\pi/4 = \frac{A}{\sqrt{2}} - \frac{B}{\sqrt{2}}.$$

Hence
$$A + B = 2$$
$$A - B = 1.$$

Solving for A and B, we find that

$$A = \frac{3}{2},\quad B = \frac{1}{2}.$$

With these values of A and B, the particular solution of (1) is

$$y = \frac{3}{2}\sin x + \frac{1}{2}\cos x.$$

Example 6. Verify that $y = c_1\cos x$ and $y = c_2\sin x$ are solutions of $\frac{d^2y}{dx^2} + y = 0$. Find particular solution of the equation satisfying the boundary conditions

$$y(0) = 1,\ y(\pi/2) = 2$$

Solution. We have

$$y = c_1\cos x$$

Differentiating twice, we get

$$\frac{dy}{dx} = -c_1\sin x$$

$$\frac{d^2y}{dx^2} = -c_1\cos x = -y$$

or
$$\frac{d^2y}{dx^2} + y = 0.$$

Thus $y = c_1\cos x$ is a solution of

$$\frac{d^2y}{dx^2} + y = 0. \quad (1)$$

Similarly, it can be checked that $y = c_2\sin x$ is also a solution of (1).

General solution of (1) is

$$y = c_1\cos x + c_2\sin x \quad (2)$$

Applying the boundary conditions, we obtain from (2)

$$y(0) = 1 = c_1 + 0$$

and
$$y\left(\frac{\pi}{2}\right) = 2 = 0 + c_2.$$

Thus $c_1 = 1$ and $c_2 = 2$. Hence the particular solution of (1) satisfying the given conditions is

$$y = \cos x + 2\sin x.$$

Example 7. Solve the boundary value problem

$$\frac{d^2y}{dx^2} + y = 0,\quad y(0) = 1,\quad y(\pi) = 5.$$

Solution. Applying the boundary conditions to the general solution

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

of the given equation, we have

$$y(0) = 1 = B$$
$$y(\pi) = 5 = -B.$$

Scanned with CamScanner

Thus we obtain two values of B and we are unable to determine any definite value of A. Hence the boundary value problem has no solution.

It follows that a boundary value problem need not always have a solution.

EXERCISE 9.1

1. Classify each of the following equations as ordinary or partial differential equations, state the order and degree of each equation and determine whether the equation is linear or nonlinear.

   (i) $\frac{d^3y}{dx^3} + 4\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 3y = \cos x$

   (ii) $x^2\,dy + y^2\,dx = 0$

   (iii) $\frac{\partial^3 u}{\partial x^3} + \frac{\partial^2 u}{\partial y^2} + \left(\frac{\partial u}{\partial z}\right)^2 + ux^3 + uy^2 + uz = 0$

   (iv) $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + u\frac{\partial u}{\partial x} + u\frac{\partial u}{\partial y} + u = 0$

   (v) $\left(\frac{dy}{dx}\right)^2 = \left(\frac{d^2y}{dx^2} + y\right)^{\frac{3}{2}}$.

2. Form the differential equation of which the given function is a solution:

   (i) $y = x + 3e^{-x}$

   (ii) $y = (x^3 + c)\,e^{-3x}$, c being an arbitrary constant

   (iii) $ax + \ln|y| = y + b$

   (iv) $y = ae^x + b\ln|x| + cx + d$

   (v) $x^2 + y^2 + 2gx + 2fy + c = 0$, f, g and c being arbitrary constants.

   (vi) $u = f(x, y, z) = (x^2 + y^2 + z^2)^{-1/2}$; $(x, y, z) \neq (0, 0, 0)$

   (vii) $u = f(x - ay) + g(x + ay)$, where f and g are twice differentiable functions.

3. Find the differential equation of all

   (i) circles of radius a

   (ii) circles that pass through the origin

   (iii) ellipses in standard form

   (iv) parabolas, each of which has a latus rectum $4a$ and whose axes are parallel to the x-axis.

   (v) hyperbolas in standard form

   (vi) coincs whose axes coincide with the axes of coordinates.

Solve the following initial value problems

(i) $\frac{dy}{dx} = -\frac{x}{y}$, $y(3) = 4$,

given that the differential equation has $x^2 + y^2 = c^2$ as the general solution

(ii) $\frac{dy}{dx} + y = 2xe^{-x}$, $y(-1) = e + 3$,

given that the differential equation has $y = (x^2 + c)e^{-x}$ as the general solution

(iii) $\frac{d^2y}{dx^2} - \frac{dy}{dx} - 12y = 0$, $y(0) = -2$, $y'(0) = 6$

where $y = Ae^{4x} + Be^{-3x}$ is the general solution of the given differential equation

(iv) $x\frac{dy}{dx} + 2y = 4x^2$, $y(1) = 2$,

given that $y = x^2 + \frac{c}{x^2}$ is the general solution of the differential equation

(v) $x^3\frac{d^3y}{dx^3} - 3x^2\frac{d^2y}{dx^2} + 6x\frac{dy}{dx} - 6y = 0$, $y(2) = 0$, $y'(2) = 2$, $y''(2) = 6$,

given that $y = c_1x + c_2x^2 + c_3x^3$ is the general solution of the given differential equation.

Solve the boundary value problems:

(i) $\frac{d^2y}{dx^2} + y = 0$, $y(0) = 1$, $y'\left(\frac{\pi}{2}\right) = +1$

given that $y = c_1\sin x + c_2\cos x$ is the general solution of the given equation.

(ii) $\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = 0$, $y(0) = 0$, $y(1) = 1$,

where $y = c_1e^x + c_2e^{2x}$ is the general solution of the given equation.

In the previous section we classified the differential equations and also saw how a differential equation can be formed by eliminating constants from a given functional equation. The problem of finding general solution (integral) of a given differential equation will now be considered. The solution of any differential equation may or may not exist. Even if the integral of a given equation exists, it may not be easy to find. We shall only discuss methods of solutions of special types of differential equations. The study of existence of solutions is beyond the scope of this book.

SEPARABLE EQUATIONS

(9.11) Definition. A differential equation of the type

$$F(x)\,G(y)\,dx + f(x)\,g(y)\,dy = 0 \qquad (1)$$

is called an equation with separable variables or simply a separable equation. Equation (1) may be written as

$$\frac{F(x)}{f(x)}\,dx + \frac{g(y)}{G(y)}\,dy = 0$$

which can be easily integrated.

Example 8. Solve $\dfrac{dy}{dx} = \dfrac{x^2}{y}$ (1)

Solution. Equation (1) can be written as

$$y\,dy = x^2\,dx$$

Integrating both the sides, we get

$$\frac{y^2}{2} = \frac{x^3}{3} + c_1$$

or $3y^2 = 2x^2 + c$

which is the required solution of (1).

Example 9. Solve: $\dfrac{dy}{dx} = \dfrac{1}{x\tan y}$

Solution. The given equation is separable and can be written as

$$\frac{dx}{x} = \tan y\,dy.$$

Integrating, we have

$$\ln|x| = -\ln|\cos y| + c$$

or $\ln|x| + \ln|\cos y| = c$

i.e., $\ln|x\cos y| = c$

or $|x\cos y| = e^c = a$ (say)

or $x\cos y = a$ because $a > 0$.

Example 10. Solve: $x\sin y\,dx + (x^2+1)\cos y\,dy = 0.$ (1)

Solution. Dividing (1) by $(x^2+1)\sin y$, we have

$$\frac{x}{x^2+1}\,dx + \cot y\,dy = 0$$

which is a separable equation. Therefore,

$$\int\frac{x\,dx}{x^2+1} + \int\cot y\,dy = c_1 = \ln|c_2|$$

i.e., $\frac{1}{2}\ln(x^2+1) + \ln|\sin y| = \ln|c_2|$

or $\ln(x^2+1) + 2\ln|\sin y| = 2\ln|c_2|$

Now $2\ln|\sin y| = \ln|\sin y|^2 = \ln(\sin y)^2 = \ln\sin^2 y$

And $2\ln|c_2| = \ln|c_2|^2 = \ln c_2^2 = \ln c$, where $c_2^2 = c > 0$

Hence $\ln(x^2+1) + \ln\sin^2 y = \ln c.$

From this it follows that

$(x^2+1)\sin^2 y = c$ is the required solution

Example 11. Solve the initial value problem

$$\frac{dy}{dx} = \frac{2x}{y + x^2 y}, \qquad y(0) = -2$$

Solution. We have

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

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

Integrating, we obtain

$$\frac{y^2}{2} = \ln(1+x^2) + \ln c, \qquad c \text{ being a constant.}$$

$$= \ln(c(1+x^2))$$

or $$y^2 = \ln(c^2(1+x^2)^2) \qquad (1)$$

Now setting $y(0) = -2$ into (1), we have

$$4 = \ln c^2 \quad \text{or} \quad c^2 = e^4.$$

So (1) becomes

$$y^2 = \ln(e^4(1+x^2)^2)$$

which is the required solution.

EXERCISE 9.2

Solve (Problems 1–15).

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

2. $\dfrac{dy}{dx} + y^2 \sin x = 0$

3. $\dfrac{dy}{dx} = 1 + x + y^2 + xy^2$

4. $(xy + 2x + y + 2)\,dx + (x^2 + 2x)\,dy = 0$

5. $\dfrac{dy}{dx} = 2x^2 + y - x^2 y + xy - 2x - 2$

6. $\csc y\,dx + \sec x\,dy = 0$

7. $y(1+x)\,dx + x(1+y)\,dy = 0$

8. $y\sqrt{1+x^2}\,dx + x\sqrt{1+y^2}\,dy = 0$

9. $\dfrac{dy}{dx} + \sqrt{\dfrac{1-y^2}{1-x^2}} = 0$

10. $(e^x + 1)\,y\,dy = (y+1)\,e^x\,dx$

11. $\dfrac{dy}{dx} = \dfrac{y^3 + 2y}{x^2 + 3x}$

12. $(\sin x + \cos x)\,dy + (\cos x - \sin x)\,dx = 0$

13. $e^x\left(1 + \dfrac{dy}{dx}\right) = xe^{-y}$

14. $xe^{x^2+y}\,dx = y\,dy$

15. $(2x\cos y)\,dx + x^2(\sec y - \sin y)\,dy = 0$.

Solve the initial value problems:

16. $2(y-1)\,dy = (3x^2 + 4x + 2)\,dx$, $\quad y(0) = -1$

17. $(3x+8)(y^2+4)\,dx - 4y(x^2+5x+6)\,dy = 0$, $\quad y(1) = 2$

18. $(1+2y^2)\,dy = y\cos x\,dx$, $\quad y(0) = 1$

19. $8\cos^2 y\,dx + \csc^2 x\,dy = 0$, $\quad y\left(\dfrac{\pi}{12}\right) = \dfrac{\pi}{4}$

20. $\dfrac{dy}{dx} = \dfrac{x(x^2+1)}{4y^3}$, $\quad y(0) = -\dfrac{1}{\sqrt{2}}$.

HOMOGENEOUS EQUATIONS

(9.12) Definition. A function $f(x, y)$ is called homogeneous of degree n if

$$f(tx, ty) = t^n f(x, y),$$

where t is a nonzero real number. Thus $\sqrt{xy}$, $\dfrac{x^{10}+y^{10}}{x^2+y^2}$ and $\sin\left(\dfrac{x}{y}\right)$ are homogeneous function of degree 1, 8 and 0 respectively. (Check!)

A first order differential equation

$$\frac{dy}{dx} = f(x, y) \qquad (1)$$

is said to be **homogeneous** if f is a homogeneous function of any degree. If (1) is written in the form

$$M(x, y)\,dx + N(x, y)\,dy = 0$$

then it is called homogeneous if $M(x, y)$ and $N(x, y)$ are homogeneous functions of the same degree.

The equation

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

is homogeneous, but

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

is not homogeneous.

(9.13) Theorem. A homogeneous equation

$$\frac{dy}{dx} = g\left(\frac{y}{x}\right) \qquad (1)$$

can be transformed into a separable equation (in the variables v and x) by the substitution $y = vx$.

Proof. Put $y = vx$ into (1). Then $\dfrac{dy}{dx} = v + x\dfrac{dv}{dx}$ and (1) becomes

$$v + x\frac{dv}{dx} = g(v)$$

or

$$v - g(v) + x\frac{dv}{dx} = 0$$

or

$$[v - g(v)]\,dx + x\,dv = 0.$$

This equation is separable and can be solved as in the previous section.

Example 12. Solve: $\dfrac{dy}{dx} = \dfrac{x^3 + y^3}{x^2 y + xy^2}$

Solution. We have

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

so that the equation is homogeneous. Setting $y = vx$ into (1), we obtain

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

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

or $$\frac{v}{1-v}\,dv = \frac{dx}{x}$$

or $$\left(-1+\frac{1}{1-v}\right)dv = \frac{dx}{x}$$

Integrating, we get

$$-v - \ln|1-v| = \ln|x| + \ln|c|$$

or $$\ln|cx(1-v)| + v = 0.$$

Replacing v by y/x in the above equation, we have

$$\ln|c(x-y)| + \frac{y}{x} = 0$$

which is the required solution.

Example 13. Solve the initial value problem

$$(x^2+3y^2)\,dx - 2xy\,dy = 0, \qquad y(2) = 6$$

Solution. Here

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

which shows that the equation is homogeneous. Putting $y = vx$ into (1), we have

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

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

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

On integrating, we get

$$\ln(1+v^2) = \ln|x| + \ln|c|$$

or $$1+v^2 = |cx|$$

Replacing v by y/x, we obtain

$$1+\frac{y^2}{x^2} = |cx|$$

or $$x^2+y^2 = |cx|\,x^2.$$

If $x \ge 0$, we can write this as

$$x^2+y^2 = cx^3$$

or $$y^2 = cx^3 - x^2$$

or $$y = \pm\sqrt{cx^3-x^2}$$

Applying the initial condition, we get

$$y(2) = 6 = \pm\sqrt{8c-4}$$

or $$8c-4 = 36 \quad \text{i.e.,} \quad c = 5$$

Hence $y = \sqrt{5x^3-x^2}$ is the required solution. We take the plus sign in the radical since $y(2)$ is positive.

D.E. REDUCIBLE TO HOMOGENEOUS FORM

(14) The differential equation

$$(a_1x + b_1y + c_1)\,dx + (a_2x + b_2y + c_2) = 0$$

is non homogeneous. But it can be reduced to a homogeneous form as illustrated below:

Case I. If $\dfrac{a_1}{a_2} \ne \dfrac{b_1}{b_2}$, then make the transformations

$$x = X+h, \; y = Y+k.$$

The given equation becomes

$$(a_1X + b_1Y + a_1h + b_1k + c_1)\,dX + (a_2X + b_2Y + a_2h + b_2k + c_2)\,dY = 0. \qquad (1)$$

Let h and k be the solution of the system of equations

$$\left.\begin{aligned} a_1h + b_1k + c_1 &= 0 \\ a_2h + b_2k + c_2 &= 0 \end{aligned}\right\}$$

Then for these values of h and k, (1) reduces to the homogeneous equation

$$(a_1X + b_1Y)\,dX + (a_2X + b_2Y)\,dY = 0$$

in the variables X and Y.

Case II. If $\dfrac{a_1}{a_2} = \dfrac{b_1}{b_2}$, then put

$z = a_1x + b_1y$ and the given equation will reduce to a separable equation in the variables x and z.

Example 14. Solve $\dfrac{dy}{dx} = \dfrac{2y - x + 5}{2x - y - 4}$ (1)

Solution. Putting

$$x = X + h, \quad y = Y + k$$

the given equation becomes

$$\frac{dY}{dX} = \frac{2Y - X + 2k - h + 5}{2X - Y + 2h - k - 4} \tag{2}$$

The solution of the system of equations

$$\left.\begin{aligned} -h + 2k + 5 &= 0 \\ 2h - k - 4 &= 0 \end{aligned}\right\}$$

is $\quad h = 1, \quad k = -2.$

For these values of h and k, (2) reduces to

$$\frac{dY}{dX} = \frac{2Y - X}{2X - Y} \tag{3}$$

which is homogeneous.

Putting $Y = vX$ into (3), we have

$$v + X\frac{dv}{dX} = \frac{2v - 1}{2 - v}$$

or $\quad X\dfrac{dv}{dX} = \dfrac{2v - 1}{2 - v} - v = \dfrac{v^2 - 1}{2 - v}$ which is separable. Therefore,

$$\left(\frac{2 - v}{v^2 - 1}\right) dv = \frac{dX}{X}$$

Integrating, we obtain

$$\ln\left|\frac{v - 1}{v + 1}\right| - \frac{1}{2}\ln|v^2 - 1| = \ln|X| + \ln|c_0|$$

or $\quad \ln\left|\dfrac{v - 1}{v + 1}\right|^2 = \ln|X|^2 + \ln|c_0|^2 + \ln|v^2 - 1|$

$$= \ln\{|c_0X|^2 . |v^2 - 1|\}$$

or $\quad \left|\dfrac{v - 1}{v + 1}\right|^2 = c|X|^2 . |v - 1| . |v + 1|$, where $c = c_0^2$

or $\quad |v - 1| = c|X|^2 . |v + 1|^3.$

Replacing v by Y/X, this becomes

$$\left|\frac{Y - X}{X}\right| = c|X|^2 \left|\frac{Y + X}{X}\right|^3$$

or $\quad |Y - X| = c|Y + X|^3$ (4)

But $X = x - 1$, $Y = y + 2$. Hence (4) takes the form

$$|y - x + 3| = c|y + x + 1|^3$$

which is the required solution.

Example 15. Solve: $(2x + y + 1)\,dx + (4x + 2y - 1)\,dy = 0.$ (1)

Solution. Since $\dfrac{a_1}{a_2} = \dfrac{2}{4} = \dfrac{b_1}{b_2} = \dfrac{1}{2}$, we put $2x + y = z$. Then (1) becomes

$$(z + 1)\,dx + (2z - 1)(dz - 2\,dx) = 0$$

or $\quad (z + 1 - 4z + 2)\,dx + (2z - 1)\,dz = 0$

i.e., $\quad 3(1 - z)\,dx + (2z - 1)\,dz = 0$ which is separable.

Dividing by $1 - z$, we have

$$3\,dx + \frac{2z - 1}{1 - z}\,dz = 0$$

or $\quad 3\,dx + \left(-2 + \dfrac{1}{1 - z}\right) dz = 0$

Integrating, we obtain

$$3x - 2z - \ln|1 - z| = c_0$$

Replacing z by $2x + y$, it reduces to

$$3x - 2(2x + y) - \ln|1 - 2x - y| = c_0$$

or $\quad -x - 2y - \ln|2x + y - 1| = c_0$

or $\quad \ln|2x + y - 1| + x + 2y = c$, where $c = -c_0$

is the required solution.

EXERCISE 9.3

Solve (Problems 1 – 10):

1. $(x - y)\,dx + (x + y)\,dy = 0$ 2. $(y^2 + 2xy)\,dx + x^2\,dy = 0$
3. $(x^2 - 3y^2)\,dx + 2xy\,dy = 0$
4. $3x\cos(y/x)\,dy = [2x\sin(y/x) + 3y\cos(y/x)]\,dx$
5. $(x^2 + xy + y^2)\,dx - x^2\,dy = 0$ 6. $y\,dy + x\,dx = \sqrt{x^2 + y^2}\,dx$

7. $\dfrac{dy}{dx} = \dfrac{4y - 3x}{2x - y}$

8. $x \sin\left(\dfrac{y}{x}\right) dy = \left[y \sin\left(\dfrac{y}{x}\right) - x\right] dx$

9. $\left(x^3 + y^2\sqrt{x^2 + y^2}\right) dx - xy\sqrt{x^2 + y^2}\, dy = 0$

10. $\left(\sqrt{x+y} + \sqrt{x-y}\right) dx - \left(\sqrt{x+y} - \sqrt{x-y}\right) dy = 0$

Solve the initial value problems (**Problems 11–14**):

11. $\dfrac{dy}{dx} = \dfrac{x+y}{x}$, $\qquad y(1) = 1$

12. $\left(y + \sqrt{x^2 + y^2}\right) dx - x\, dy = 0$, $\qquad y(1) = 0$

13. $(2x - 5y)\, dx + (4x - y)\, dy = 0$, $\qquad y(1) = 4$

14. $(3x^2 + 9xy + 5y^2)\, dx - (6x^2 + 4xy)\, dy = 0$, $\qquad y(2) = -6$.

Solve:

15. $\dfrac{dy}{dx} = \dfrac{x + 3y - 5}{x - y - 1}$

16. $\dfrac{dy}{dx} = -\dfrac{4x + 3y + 15}{2x + y + 7}$

17. $(3y - 7x - 3)\, dx + (7y - 3x - 7)\, dy = 0$

18. $\dfrac{dy}{dx} = \dfrac{3x - 4y - 2}{3x - 4y - 3}$

19. $\dfrac{dy}{dx} = \dfrac{y - x + 1}{y - x + 5}$

20. $\dfrac{dy}{dx} = \dfrac{x - 2y + 5}{2x + y - 1}$

EXACT EQUATIONS

(9.15) Definition. The expression

$$M(x, y)\, dx + N(x, y)\, dy \qquad (1)$$

is called an **exact differential** if there exists a continuously differentiable function $f(x, y)$ of two real variables x and y such that the expression equals the total differential df. We know from calculus that

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$

Thus, if (1) is exact then

$$M(x, y) = \frac{\partial f}{\partial x} = f_x \quad \text{and} \quad N(x, y) = \frac{\partial f}{\partial y} = f_y.$$

If (1) is an exact differential then the differential equation

$$M(x, y)\, dx + N(x, y)\, dy = 0$$

is called an **exact equation**.

(9.16) Theorem. The differential equation

$$M(x, y)\, dx + N(x, y)\, dy = 0 \qquad (1)$$

is an exact equation if and only if

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x},$$

where the functions $M(x, y)$ and $N(x, y)$ have continuous first order partial derivatives.

Proof. Suppose that the equation (1) is exact so that $M\,dx + N\,dy$ is an exact differential. By definition, there exists a function $f(x, y)$ such that

$$M(x, y) = \frac{\partial f}{\partial x} = f_x \quad \text{and} \quad N(x, y) = \frac{\partial f}{\partial y} = f_y.$$

Then $\quad M_y = \dfrac{\partial M}{\partial y} = f_{xy} = \dfrac{\partial^2 f}{\partial y\, \partial x}$

and $\quad N_x = \dfrac{\partial N}{\partial x} = f_{yx} = \dfrac{\partial^2 f}{\partial x\, \partial y}$

Since M and N possess continuous first order partial derivatives, we have $f_{xy} = f_{yx}$ and therefore

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \qquad \text{as desired.}$$

The proof of the converse is omitted since it is beyond our scope.

(9.17) Solution of an Exact Equation.

If $\quad M(x, y)\, dx + N(x, y)\, dy = 0 \qquad (1)$

is an exact equation, then there exists a function $f(x, y)$ such that

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

$$= M(x, y)\, dx + N(x, y)\, dy.$$

Therefore, $\quad \dfrac{\partial f}{\partial x} = M \quad$ and $\quad \dfrac{\partial f}{\partial y} = N.$

Integrating $\quad \dfrac{\partial f}{\partial x} = M \quad$ with respect to x, we have

$$f(x, y) = \int M\, dx + h(y). \qquad (2)$$

The constant of integration $h(y)$ is an arbitrary function of y since it must vanish under differentiation w.r.t. x.

Differentiating (2) w.r.t. y, we get

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\left(\int M\,dx\right) + h'(y)$$

i.e.,
$$\frac{\partial f}{\partial y} = N = \frac{\partial}{\partial y}\left(\int M\,dx\right) + h'(y)$$

or
$$h'(y) = N - \frac{\partial}{\partial y}\left(\int M\,dx\right)$$

Integrating the above equation w.r.t. y, we obtain h and hence $f(x, y) = c$ is the required solution of (1).

Example 16. Solve: $(3x^2y + 2)\,dx + (x^3 + y)\,dy = 0$ (1)

Solution. Here $M = 3x^2y + 2$ and $N = x^3 + y$

$$\frac{\partial M}{\partial y} = 3x^2 \quad , \quad \frac{\partial N}{\partial x} = 3x^2.$$

Thus $\dfrac{\partial M}{\partial y} = \dfrac{\partial N}{\partial x}$ and so the equation is exact.

To find the solution of (1), we note that the left hand side of the equation is an exact differential. Therefore, there exists a function $f(x, y)$ such that

$$\frac{\partial f}{\partial x} = 3x^2y + 2 \qquad (2)$$

and
$$\frac{\partial f}{\partial y} = x^3 + y \qquad (3)$$

Integrating (2) w.r.t. x, we have

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

where $h(y)$ is the constant of integration. Differentiating the above equation w.r.t. y and using (3), we obtain

$$\frac{\partial f}{\partial y} = x^3 + h'(y) = x^3 + y$$

or
$$h'(y) = y.$$

Integrating, we have

$$h(y) = \frac{y^2}{2}.$$

Thus $f(x, y) = x^3y + 2x + \dfrac{y^2}{2}.$

Hence the general solution of (1) is

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

Alternative Method:

Integrating (2) and (3) w.r.t. x and y respectively, we have

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

and
$$f(x, y) = x^3y + \frac{y^2}{2} + g(x)$$

Thus $h(y) = \dfrac{y^2}{2}$ and $g(x) = 2x$

The general solution is

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

Example 17. Solve the initial value problem

$$(2y\sin x\cos x + y^2\sin x)\,dx + (\sin^2 x - 2y\cos x)\,dy = 0, \quad y(0) = 3$$

Solution. Here

$$M = 2y\sin x\cos x + y^2\sin x$$

and
$$N = \sin^2 x - 2y\cos x$$

$$\frac{\partial M}{\partial y} = 2\sin x\cos x + 2y\sin x$$

$$\frac{\partial N}{\partial x} = 2\sin x\cos x + 2y\sin x$$

Thus $\dfrac{\partial M}{\partial y} = \dfrac{\partial N}{\partial x}$

showing that the given equation is exact.

Hence there exists a function $f(x, y)$ such that

$$\frac{\partial f}{\partial x} = 2y\sin x\cos x + y^2\sin x \qquad (1)$$

$$\frac{\partial f}{\partial y} = \sin^2 x - 2y\cos x \qquad (2)$$

Integrating (1) w.r.t. x, we have

$$f(x, y) = y\sin^2 x - y^2\cos x + h(y)$$

Differentiating this equation w.r.t. y and using (2), we get

$$\sin^2 x - 2y\cos x + h'(y) = \sin^2 x - 2y\cos x$$

i.e., $h'(y) = 0$ and so $h(y) = c_1.$

Hence the general solution of the given equation is

$$f(x, y) = c_2$$

i.e., $y \sin^2 x - y^2 \cos x + c_1 = c_2$

or $y \sin^2 x - y^2 \cos x = c_2 - c_1 = c$.

Applying the initial condition that when $x = 0$, $y = 3$, we have

$$-9 = c$$

Hence $\quad y^2 \cos x - y \sin^2 x = 9$

is the required solution

EXERCISE 9.4

Solve (Problems 1–10).

1. $(3x^2 + 4xy)\,dx + (2x^2 + 2y)\,dy = 0$
2. $(2xy + y - \tan y)\,dx + (x^2 - x\tan^2 y + \sec^2 y)\,dy = 0$
3. $\dfrac{x+y}{y-1}dx - \dfrac{1}{2}\left(\dfrac{x+1}{y-1}\right)^2 dy = 0$
4. $\dfrac{dy}{dx} = -\dfrac{ax + hy}{hx + by}$
5. $(1 + \ln xy)\,dx + \left(1 + \dfrac{x}{y}\right)dy = 0$
6. $\dfrac{y\,dx + x\,dy}{1 - x^2y^2} + x\,dx = 0$
7. $(6xy + 2y^2 - 5)\,dx + (3x^2 + 4xy - 6)\,dy = 0$
8. $(y\sec^2 x + \sec x \tan x)\,dx + (\tan x + 2y)\,dy = 0$
9. $(y\cos x + 2xe^y)\,dx + (\sin x + x^2e^y - 1)\,dy = 0$
10. $(ye^{xy}\cos 2x - 2e^{xy}\sin 2x + 2x)\,dx + (xe^{xy}\cos 2x - 3)\,dy = 0$

Solve the initial value problems:

11. $(2xy - 3)\,dx + (x^2 + 4y)\,dy = 0,\quad y(1) = 2$
12. $(2x\cos y + 3x^2y)\,dx + (x^3 - x^2\sin y - y)\,dy = 0,\quad y(0) = 2$
13. $(3x^2y^2 - y^3 + 2x)\,dx + (2x^3y - 3xy^2 + 1)\,dy = 0,\quad y(-2) = 1$
14. $\dfrac{3-y}{x^2}dx + \dfrac{y^2 - 2x}{xy^2}dy = 0,\quad y(-1) = 2$
15. $(4x^3e^{x+y} + x^4e^{x+y} + 2x)\,dx + (x^4e^{x+y} + 2y)\,dy = 0,\quad y(0) = 1.$

INTEGRATING FACTORS

(9.18) Definition. If the differential equation

$$M(x, y)\,dx + N(x, y)\,dy = 0 \qquad (1)$$

is not exact but when it is multiplied by a function $\mu(x, y)$ and the resulting equation

$$\mu(x, y)\,M(x, y)\,dx + \mu(x, y)\,N(x, y)\,dy = 0$$

is exact, then $\mu(x, y)$ is called an **integrating factor (I.F.)** of the differential equation (1). The number of integrating factors of an equation may be infinite.

We list below (without proofs) some rules to find the integrating factors of equations of special types.

Rule I. If $M(x, y)\,dx + N(x, y)\,dy = 0 \qquad (1)$

is not exact and $\dfrac{M_y - N_x}{N} = P$, where P is a function of x only, then (1) has an integrating factor $\mu(x)$ which also depends on x. $\mu(x)$ is solution of the differential equation

$$\frac{d\mu}{dx} = P\mu$$

i.e., $$\mu(x) = \exp\int P\,dx$$

Note that $\quad M_y = \dfrac{\partial M}{\partial y}, \quad N_x = \dfrac{\partial N}{\partial x}$

Rule II. If $\dfrac{N_x - M_y}{M} = Q$, where Q is a function of y only, then the differential equation $M\,dx + N\,dy = 0$ has an integrating factor

$$\mu(y) = \exp\int Q\,dy$$

→ only on 'x' dependent
→ on 'y' dependent

Rule III. If $M\,dx + N\,dy = 0 \qquad (1)$

is homogeneous and $xM + yN \neq 0$, then

$\dfrac{1}{xM + yN}$ is an I.F. of (1).

Rule IV. If $M\,dx + N\,dy = 0$ is of the form

$$yf(xy)\,dx + xg(xy)\,dy = 0$$

and $xM - yN \neq 0$, then

$\dfrac{1}{xM - yN}$ is an I.F. of (1).

The following differential formulas are useful in the calculation of certain exact equations:

(i) $d\left(\dfrac{y}{x}\right) = \dfrac{x\,dy - y\,dx}{x^2}$

(ii) $d\left(\dfrac{x}{y}\right) = \dfrac{y\,dx - x\,dy}{y^2}$

(iii) $d(xy) = x\,dy + y\,dx$

(iv) $d(x^2+y^2) \quad = \quad 2(x\,dx + y\,dy)$

(v) $d\left(\ln\dfrac{x}{y}\right) \quad = \quad \dfrac{y\,dx - x\,dy}{xy}$

(vi) $d\left(\arctan\dfrac{x}{y}\right) \quad = \quad \dfrac{y\,dx - x\,dy}{x^2+y^2}$

Example 18. Solve: $y\,dx + (x^2y - x)\,dy = 0$.

Solution. The equation is not exact. Rearranging the equation, we have

$$y\,dx - x\,dy + x^2y\,dy = 0$$

or $$\frac{y\,dx - x\,dy}{x^2} + y\,dy = 0.$$

Now it is an exact equation and may be written as

$$-d\left(\frac{y}{x}\right) + y\,dy = 0.$$

Integrating, we have

$$-\frac{y}{x} + \frac{y^2}{2} = c$$

or $xy^2 - 2y = cx$ is the general solution.

Example 19. $(x^2 - 2x + 2y^2)\,dx + 2xy\,dy = 0$. (1)

Solution. Here $\dfrac{M_y - N_x}{N} = \dfrac{4y-2y}{2xy} = \dfrac{1}{x}$.

Therefore, I.F. $\mu(x, y)$ is the solution of

$$\frac{d\mu}{dx} = \frac{\mu}{x}$$

or $\mu = x$ is an I.F.

Multiplying the equation by x, we have

$$(x^3 - 2x^2 + 2xy^2)\,dx + 2x^2y\,dy = 0.$$

This equation is exact. The reader can easily find that its solution is

$$\frac{x^4}{4} - \frac{2x^3}{3} + x^2y^2 = c_0$$

or $3x^4 - 8x^3 + 12x^2y^2 = c$ is the solution of (1).

Example 20. Solve: $dx + \left(\dfrac{x}{y} - \sin y\right)dy = 0$

Solution. Here, by Rule II,

$$\frac{N_x - M_y}{M} = \frac{\frac{1}{y} - 0}{1} = \frac{1}{y}$$

function of y only. Therefore,

$$\mu(y) = \exp\int\frac{dy}{y} = e^{\ln y} = y$$

is I.F. Multiplying the equation by y, we have

$$y\,dx + (x - y\sin y)\,dy = 0$$

or $$y\,dx + x\,dy - y\sin y\,dy = 0$$

or $$d(xy) - y\sin y\,dy = 0.$$

Integrating, we get

$$xy + y\cos y - \sin y = c$$

which is the required solution.

Example 21. Solve: $(x^2y - 2xy^2)\,dx - (x^3 - 3x^2y)\,dy = 0$. (1)

Solution. The equation is homogeneous but not exact. We have

$$xM + yN = x^3y - 2x^2y^2 - x^3y + 3x^2y^2 = x^2y^2 \neq 0$$

So, using Rule III,

$\dfrac{1}{xM + yN} = \dfrac{1}{x^2y^2}$ is an I.F. Therefore, multiplying (1) by $\dfrac{1}{x^2y^2}$, we obtain

$$\left(\frac{1}{y} - \frac{2}{x}\right)dx - \left(\frac{x}{y^2} - \frac{3}{y}\right)dy = 0.$$

This equation is exact. Integrating, we get

$$\frac{x}{y} - 2\ln|x| + 3\ln|y| = c$$

the required solution.

Note that, here the term $\dfrac{x}{y}$, involving both x and y, has been taken only once.

Example 22. Solve: $y(xy + 2x^2y^2)\,dx + x(xy - x^2y^2)\,dy = 0$ (1)

Solution. The equation is of the form

$$yf(xy)\,dx + xg(xy)\,dy = 0$$

Now, $xM - yN = x^2y^2 + 2x^3y^3 - x^2y^2 + x^3y^3 = 3x^3y^3 \neq 0$.

Therefore, $\dfrac{1}{x^3y^3}$ is an I.F.

Multiplying (1) by $\frac{1}{x^2y^2}$, we have

$$\left(\frac{1}{x^2y}+\frac{2}{x}\right)dx+\left(\frac{1}{xy^2}-\frac{1}{y}\right)dy = 0$$

This is an exact equation. Integrating, we get

$$\frac{-1}{xy}+2\ln|x|-\ln|y| = c$$

as the required solution.

EXERCISE 9.5

Solve (by finding an I.F.):

1. $(x^9y^3+y)\,dx - x\,dy = 0$
2. $x\,dy - y\,dx = (x^2+y^2)\,dx$
3. $(x^2+x-y)\,dx + x\,dy = 0$
4. $dy + \frac{y-\sin x}{x}\,dx = 0$
5. $y(2xy+e^x)\,dx - e^x\,dy = 0$
6. $(y^4+2y)\,dx + (xy^3+2y^4-4x)\,dy = 0$
7. $(x^2+y^2+2x)\,dx + 2y\,dy = 0$
8. $(x^2+y^2)\,dx - 2xy\,dy = 0$
9. $(3y+4xy^2)\,dx + (2x+3x^2y)\,dy = 0$
10. $(4x+3y^2)\,dx + 2xy\,dy = 0$
11. $(y-xy^2)\,dx + (x+x^2y^2)\,dy = 0$
12. $(3x^2y^4+2xy)\,dx + (2x^3y^3-x^2)\,dy = 0$
13. $y - x\frac{dy}{dx} = x + y\frac{dy}{dx}$
14. $\frac{dy}{dx} = e^{2x}+y-1$
15. $(y^2+xy)\,dx - x^2\,dy = 0$
16. $(3xy+y^2)\,dx + (x^2+xy)\,dy = 0$
17. $(3x^2y+2xy+y^3)\,dx + (x^2+y^2)\,dy = 0$
18. $y\,dx + (2xy-e^{-2y})\,dy = 0$
19. $e^x\,dx + (e^x\cot y + 2y\csc y)\,dy = 0$
20. $(x+2)\sin y\,dx + x\cos y\,dy = 0$

LINEAR EQUATIONS

(9.19) Definition. A first order ordinary differential equation is linear in the dependent variable y and the independent variable x if it is or can be written in the form

$$\frac{dy}{dx}+P(x)y = Q(x), \quad (1)$$

where P and Q are functions of x.

(9.20) Solution of a Linear Equation. The linear equation

$$\frac{dy}{dx}+P(x)y = Q(x)$$

can be written as

$$[P(x)y-Q(x)]\,dx + dy = 0 \quad (2)$$

which is of the form $M\,dx+N\,dy=0$, where

$$M = P(x)y-Q(x) \quad \text{and} \quad N = 1.$$

Now

$$\frac{\partial M}{\partial y} = P(x) \quad \text{and} \quad \frac{\partial N}{\partial x} = 0$$

Thus (2) is not exact unless $P(x)=0$ in which case (1) is separable. However, an integrating factor (depending only on x) of (2) may be easily found. Let $\mu(x)$ be an I.F. of (2). Then multiplying (2) by $\mu(x)$, we get

$$[\mu(x)P(x)y-\mu(x)Q(x)]\,dx+\mu(x)\,dy = 0. \quad (3)$$

Now (3) is an exact equation if and only if

$$\frac{\partial}{\partial y}\left[\mu(x)P(x)y-\mu(x)Q(x)\right] = \frac{\partial}{\partial x}[\mu(x)]$$

This condition reduces to

$$\mu(x)P(x) = \frac{d}{dx}[\mu(x)]$$

i.e.,

$$\mu P(x) = \frac{d\mu}{dx}$$

or

$$\frac{d\mu}{\mu} = P(x)\,dx,$$

Integrating, we obtain

$$\ln|\mu| = \int P(x)\,dx$$

or

$$\mu = \exp\left[\int P(x)\,dx\right] > 0.$$

Thus $e^{\int P\,dx}$ $\left[\text{or } \exp\int P(x)\,dx\right]$ is an I.F. of the linear equation (1). Multiplying (1) by $e^{\int P\,dx}$, we have

$$e^{\int P(x)\,dx}\frac{dy}{dx}+e^{\int P(x)\,dx}P(x)y = Q(x)\,e^{\int P(x)\,dx}$$

or

$$\frac{d}{dx}\left[y\,e^{\int P(x)\,dx}\right] = Q(x)\,e^{\int P(x)\,dx}$$

Integrating this we obtain the solution of (1) in the form

$$y\,e^{\int P(x)\,dx} = \int\left[Q(x)\,e^{\int P(x)\,dx}\right]dx+c.$$

Example 23. Solve: $(x-1)^3\dfrac{dy}{dx} + 4(x-1)^2 y = x+1$ (1)

Solution. We write the equation in the standard form

$$\frac{dy}{dx} + \frac{4}{x-1}y = \frac{x+1}{(x-1)^3}$$

Here $P(x) = \dfrac{4}{x-1}$. Therefore, an I.F. of (1) is

$$\exp\left[\int \frac{4dx}{x-1}\right] = \exp[\ln(x-1)^4] = (x-1)^4$$

Multiplying (1) by this I.F., we get

$$(x-1)^4\frac{dy}{dx} + 4(x-1)^3 y = x^2 - 1$$

or $$\frac{d}{dx}[y(x-1)^4] = x^2 - 1$$

Integrating, we obtain

$$y(x-1)^4 = \frac{x^3}{3} - x + c$$

which is the required solution.

Example 24. Solve: $(x+2y^3)\dfrac{dy}{dx} = y$.

Solution. We have

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

This equation is clearly not linear in y. But in the first order differential equation the roles of x and y are interchangeable in the sense that either variable may be regarded as dependent variable. Let us regard x as dependent variable and y as independent variable. The equation may be written as

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

which is linear in x with

$$\text{I.F.} = \exp\left[\int\left(-\frac{1}{y}\right)dy\right] = \exp\left[\ln\frac{1}{y}\right] = \frac{1}{y}.$$

Multiplying (1) by $\dfrac{1}{y}$, we get

$$\frac{1}{y}\frac{dx}{dy} - \frac{1}{y^2}x = 2y \quad \text{or} \quad \frac{d}{dy}\left(\frac{x}{y}\right) = 2y.$$

Integrating, we have

$$\frac{x}{y} = y^2 + c \quad \text{or} \quad x = y(y^2+c)$$

the required solution

Example 25. Solve the initial value problem

$$\frac{dr}{d\theta} + r\tan\theta = \cos^2\theta, \qquad r\left(\frac{\pi}{4}\right) = 1.$$

Solution. The equation is linear in r with

$$\text{I.F.} = \exp\left[\int \tan\theta\, d\theta\right] = \exp[\ln\sec\theta] = \sec\theta,$$

(taking $\sec\theta$ positive)

Multiplying the given equation by $\sec\theta$, we have

$$\sec\theta\frac{dr}{d\theta} + r\sec\theta\tan\theta = \cos\theta$$

$$\frac{d}{d\theta}[r\sec\theta] = \cos\theta$$

Integrating, we obtain

$$r\sec\theta = \sin\theta + c$$

Applying the initial condition, we have

$$1\,.\,\sec\left(\frac{\pi}{4}\right) = \sin\left(\frac{\pi}{4}\right) + c$$

or $$c = \sqrt{2} - \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}}.$$

Therefore,

$$r\sec\theta = \sin\theta + \frac{1}{\sqrt{2}}$$

or $$r = \sin\theta\cos\theta + \frac{\cos\theta}{\sqrt{2}}$$

or $$2r = \sin 2\theta + \sqrt{2}\cos\theta$$ is the required solution.

THE BERNOULLI EQUATION[1]

(9.21) Definition. An equation of the form

$$\frac{dy}{dx} + P(x)\,y = Q(x)\,y^n \qquad (1)$$

is called the **Bernoulli differential equation**. This equation is linear if $n = 0$ or 1. If n is not zero or 1, then (1) is reducible to a linear equation. Dividing by y^n, (1) becomes

$$y^{-n}\frac{dy}{dx} + P(x)\,y^{1-n} = Q(x) \qquad (2)$$

In (2), put $v = y^{1-n}$ then it reduces to

$$\frac{dv}{dx} + (1-n)\,P(x)\,v = (1-n)\,Q(x)$$

(1) a
b

which is linear in v.

(2) a, b

Note. Consider the equation

$$f'(y)\frac{dy}{dx} + P(x)f(y) = Q(x)$$

Letting $v = f(y)$, this equation becomes

$$\frac{dv}{dx} + P(x)\,v = Q(x)$$

which is linear in v.

Example 26. Solve $\dfrac{dy}{dx} + \dfrac{xy}{1-x^2} = xy^{\frac{1}{2}}$. (1)

Solution. Dividing by $y^{\frac{1}{2}}$, (1) becomes

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

Put $\qquad y^{\frac{1}{2}} = v$

or $\qquad \dfrac{1}{2}y^{-\frac{1}{2}}\dfrac{dy}{dx} = \dfrac{dv}{dx}$

Then (2) reduces to

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

1. After the name of Swiss mathematician Jacques Bernoulli (1654 – 1705) who studied it in 1669 C.E.

This is linear in v.

$$\text{I.F.} = \exp\left[\int \frac{x}{2(1-x^2)}\,dx\right] = \exp\left[\frac{-1}{4}\ln(1-x^2)\right] = (1-x^2)^{-\frac{1}{4}}$$

Multiplying (3) by $(1-x^2)^{-\frac{1}{4}}$, we get

$$(1-x^2)^{-\frac{1}{4}}\frac{dv}{dx} + \frac{x}{2(1-x^2)^{5/4}}\,v = \frac{x}{2(1-x^2)^{1/4}}$$

or $$\frac{d}{dx}\left[(1-x^2)^{-\frac{1}{4}}v\right] = \frac{-1}{4}\left[-2x(1-x^2)^{-\frac{1}{4}}\right]$$

Integrating, we have

$$v(1-x^2)^{-\frac{1}{4}} = \frac{-1}{4}\,\frac{(1-x^2)^{3/4}}{3/4} + c$$

or $$v = c(1-x)^{-\frac{1}{4}} - \frac{1-x^2}{3}$$

or $$y^{\frac{1}{2}} = c(1-x^2)^{-\frac{1}{4}} - \frac{1-x^2}{3}$$

the required solution of (1)

EXERCISE 9.6

Solve (Problems 1–15)

1. $\dfrac{dy}{dx} + \left(\dfrac{2x-1}{x}\right)y = e^{-2x}$
2. $\dfrac{dy}{dx} + \dfrac{3y}{x} = 6x^2$
3. $\dfrac{dy}{dx} + \dfrac{y}{x\ln x} = \dfrac{3x^2}{\ln x}$
4. $\dfrac{dy}{dx} + 3y = 3x^2e^{-3x}$
5. $\cos^3 x\dfrac{dy}{dx} + y\cos x = \sin x$
6. $x\dfrac{dy}{dx} + (1 + x\cot x)\,y = x$
7. $(x+1)\dfrac{dy}{dx} - ny = e^x(x+1)^{n+1}$
8. $(x^2+1)\dfrac{dy}{dx} + 2xy = 4x^2$
9. $x\dfrac{dy}{dx} + 2y = \sin x$
10. $(1+x^2)\dfrac{dy}{dx} + 4xy = \dfrac{1}{(1+x^2)^2}$
11. $\dfrac{dy}{dx} = \dfrac{1}{e^y - x}$
12. $(x+2y^3)\dfrac{dy}{dx} = y$
13. $x\dfrac{dy}{dx} + y = y^2\ln x$
14. $\dfrac{dy}{dx} + y = xy^3$
15. $x\dfrac{dy}{dx} - 2x^2y = y\ln y.$

Solve the initial value problems

16. $(x^2+1)\dfrac{dy}{dx} + 4xy = x,$ $\qquad y(2) = 1$

17. $e^x[y - 3(e^x+1)^2]\,dx + (e^x+1)\,dy = 0,$ $\qquad y(0) = 4$

18. $\dfrac{dy}{dx} + \dfrac{y}{2x} = \dfrac{x}{y^3},$ $\qquad y(1) = 2$

19. $x(2+x)\dfrac{dy}{dx} + 2(1+x)y = 1 + 3x^2,$ $\qquad y(-1) = 1$

20. $x\dfrac{dy}{dx} + 3y = x^3y^2,$ $\qquad y(1) = 2$

ORTHOGONAL TRAJECTORIES

(9.22) (Rectangular Coordinates) Definition. It has been observed that the general solution of a first order differential equation contains one arbitrary constant. When this constant is assigned different values, one obtains a one-parameter family of curves. Each of these curves represents a particular solution of the given differential equation.

On the other hand, given a one-parameter family of curves

$$f(x, y, c) = 0, \qquad (1)$$

c being parameter, then each member of the family is a particular solution of some differential equation. In fact, this differential equation is obtained by elimination of the parameter c between (1) and the relation obtained by differentiating (1) w.r.t. x.

Let $f(x, y, c) = 0$ and $F(x, y, k) = 0$ be two families of curves with parameters c and k. If each curve in either family is intersected orthogonally by every curve in the other family, then each family is said to be **orthogonal trajectory** of the other. Recall that two curves are said to be orthogonal (intersect orthogonally) if their tangents at the point of intersection are perpendicular to each other.

For example, the families of curves given by

$$x^2 + y^2 = c^2 \Rightarrow f(x, y, c) = x^2 + y^2 - c^2 = 0$$

and

$$y = kx \Rightarrow F(x, y, k) = y - kx = 0$$

are orthogonal as illustrated graphically below.

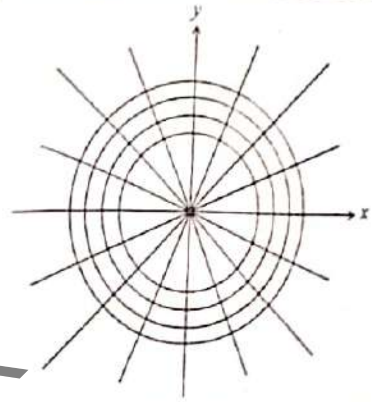


Figure 9.1

Let the given family of curves be (1). By eliminating c, we find the differential equation of this family. Suppose the differential equation of (1) is

$$\frac{dy}{dx} = F(x, y)$$

The differential equation of the orthogonal family is

$$\frac{dy}{dx} = \frac{-1}{F(x, y)} \qquad (2)$$

The solution of (2) is the required family of orthogonal trajectories of (1).

Example 27. Find the orthogonal trajectories of the family of circles

$$x^2 + y^2 = c^2 \qquad (1)$$

Solution. Differentiating (1) w.r.t. x, we have

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

The differential equation of the orthogonal trajectories of (1) is

$$\frac{dy}{dx} = \frac{-1}{-x/y} = \frac{y}{x} \qquad (2)$$

We now solve the differential equation (2). This equation is separable,

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

Integrating, we get

$$y = kx \qquad (3)$$

which is the required equation of the orthogonal trajectories of (1). This equation represents a family of straight line through the origin. The orthogonal trajectories are shown in Figure 9.1 above.

Example 28. Find the orthogonal trajectories of the family of curves

$$y = ce^{-x/4} \qquad (1)$$

Solution. Differentiating (1) w.r.t. x, we obtain

$$\frac{dy}{dx} = -\frac{c}{4}e^{-x/4} = -\frac{1}{4}y$$

which is the differential equation of (1). Differential equation of the orthogonal trajectories is

$$\frac{dy}{dx} = \frac{4}{y}. \qquad (2)$$

Solving (2), we get

$$y^2 = 8x + k$$

as an equation of the orthogonal trajectories of (1).

(9.23) Polar Coordinates

Let $f(r, \theta, c) = 0$ be the equation of a family of curves. The differential equation of this family can be obtained by elimination of c. Suppose the differential equation of this family is

$$P\,dr + Q\,d\theta = 0 \qquad (1)$$

where P and Q are functions of r and θ.

We know from calculus that if ϕ is the angle between the radius vector and the tangent to a curve of the given family at any point (r, θ), then

$$\tan\phi = r\frac{d\theta}{dr}.$$

If ϕ_1 is the angle between the radius vector and the tangent to an orthogonal trajectory at (r, θ), then

$$\phi_1 = \frac{\pi}{2} + \phi$$

or $\tan\phi_1 = -\cot\phi$

i.e., $\tan\phi_1 \tan\phi = -1$

for the two curves to be orthogonal.

From (1), we have

$$\frac{d\theta}{dr} = -\frac{P}{Q}$$

or

$$r\frac{d\theta}{dr} = -\frac{Pr}{Q}$$

Hence the differential equation of the orthogonal trajectories is

$$r\frac{d\theta}{dr} = \frac{Q}{Pr} \qquad (2)$$

Solution of (2) is the required family of orthogonal trajectories of the family $f(r, \theta, c) = 0$.

Example 29. Find the orthogonal trajectories of the family of cardioids

$$r = a(1 + \cos\theta). \qquad (1)$$

Solution. Differentiating (1) w.r.t. r, we have

$$1 = a(-\sin\theta)\frac{d\theta}{dr}$$

or

$$\frac{d\theta}{dr} = \frac{-1}{a\sin\theta} = -\frac{1+\cos\theta}{r\sin\theta}$$

or

$$r\frac{d\theta}{dr} = \frac{-(1+\cos\theta)}{\sin\theta}$$

is the differential equation of (1).

Differential equation of the orthogonal trajectories is

$$r\frac{d\theta}{dr} = \frac{\sin\theta}{1+\cos\theta}. \qquad (2)$$

Separating variables in (2), we get

$$\frac{dr}{r} = \frac{1+\cos\theta}{\sin\theta}d\theta = \csc\theta\,d\theta + \cot\theta\,d\theta.$$

Hence $\ln|r| = \ln|\csc\theta - \cot\theta| + \ln|\sin\theta| + \ln|b|$

or $r = b\sin\theta(\csc\theta - \cot\theta) = b(1-\cos\theta)$

is an equation of the orthogonal trajectories of (1).

EXERCISE 9.7

Find an equation of orthogonal trajectories of the curve of each of the following families (Problems 1–16):

1. $x^2 - y^2 = c$
2. $x = cy^2$
3. $x^2 + y^2 = cx$
4. $y = e^{cx}$
5. $y = x - 1 + ce^{-x}$
6. $xy = c$
7. $x = \frac{y^2}{4} + \frac{c}{y^2}$
8. $y = (x-c)^2$

9. $y^2 = x^2 + cx$

10. $x^2 + y^2 = 1 + 2cy$

11. $r = a(1 + \sin\theta)$

12. $r^2 = a \sin 2\theta$

13. $r^n = a^n \cos n\theta$

14. $r = \dfrac{a}{2 + \cos\theta}$

15. $r = a \sin n\theta$

16. $r^n \cos n\theta = a^n$

17. A family of curves whose family of orthogonal trajectories is the same as the given family is called **self-orthogonal**. Show that

$$y^2 = 4cx + 4c^2$$

is self-orthogonal.

18. Prove that the family of confocal coincs

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

is self-orthogonal.

In this section we shall consider first order differential equations with degree more than one. In what follows, $\dfrac{dy}{dx}$ will be denoted by p.

We have already studied various methods of finding the solution of some special types of the nonlinear first order and first degree differential equations. Such equations were separable, exact, homogenous and so on. We shall briefly discuss techniques to find solutions of special types of first order nonlinear differential equations of higher degree.

EQUATIONS SOLVABLE FOR p

Example 30. Solve: $x^2p^2 + xp - y^2 - y = 0$ (1)

Solution. We factorize the left hand member of (1) to obtain

$$(x^2p^2 - y^2) + (xp - y) = (xp - y)(xp + y + 1) = 0$$

Therefore, either

$$xp - y = 0 \quad (2)$$

or

$$xp + y + 1 = 0 \quad (3)$$

(2) gives $\quad x\dfrac{dy}{dx} = y \quad$ or $\quad \dfrac{dy}{dx} = \dfrac{dx}{x}$

or

$$y = cx \quad (4)$$

From (3), we have $\quad xp = -(y+1)$

or

$$x\frac{dy}{dx} = -(y+1) \quad \text{or} \quad \frac{dy}{y+1} = \frac{-dx}{x}$$

which yields $\quad x(y+1) = c \quad$ (5)

Combining (4) and (5), the required solution of (1) is

$$(y - cx)(xy + x - c) = 0$$

Example 31. Solve:

$$xp^3 - (x^2 + x + y)p^2 + (x^2 + xy + y)p - xy = 0 \quad (1)$$

Solution. By inspection, we find that $p - 1$ is a factor of left hand member of (1). Thus, the given equation is

$$(p-1)[xp^2 - (x^2 + y)p + xy] = 0$$

or

$$(p-1)(xp - y)(p - x) = 0$$

Therefore, either

$$p - 1 = 0 \quad (2)$$

or

$$xp - y = 0 \quad (3)$$

or

$$p - x = 0 \quad (4)$$

From (2), we have $\quad \dfrac{dy}{dx} = 1 \quad$ or $\quad y = x + c \quad$ (5)

From (3), we get $\quad x\dfrac{dy}{dx} = y \quad$ or $\quad y = cx \quad$ (6)

From (4), we obtain $\quad \dfrac{dy}{dx} = x$

or

$$y = \frac{x^2}{2} + c \quad \text{or} \quad x^2 + 2(y - c) = 0 \quad (7)$$

The general solution of (1) is obtained by combining (5), (6) and (7). Thus

$$(y - x - c)(y - cx)(x^2 - 2y + 2c) = 0$$

is the general solution of (1)

EQUATIONS SOLVABLE FOR y

Example 32. Solve: $y = p^2x + p$. (1)

Solution. Differentiating (1) w.r.t. x, we have

$$\frac{dy}{dx} = p = p^2 + 2xp\frac{dp}{dx} + \frac{dp}{dx}$$

or $$\frac{dp}{dx}(2px+1) + p^2 - p = 0$$

or $$\frac{dp}{dx} = \frac{-p(p-1)}{2px+1}$$

or $$\frac{dx}{dp} = \frac{-2px-1}{p(p-1)}$$

i.e., $$\frac{dx}{dp} + \frac{2x}{p-1} = \frac{-1}{p(p-1)} \quad (2)$$

which is linear in x.

$$\text{I.F.} = \exp\left[\int \frac{2}{p-1}\,dp\right] = \exp\left[\ln(p-1)^2\right] = (p-1)^2.$$

Multiplying (2) by $(p-1)^2$ and integrating, we get

$$x(p-1)^2 = c - p + \ln p \quad (3)$$

$$x = \frac{c-p+\ln p}{(p-1)^2}.$$

Substituting this value of x into (1), we have

$$y = p^2 \cdot \frac{c-p+\ln p}{(p-1)^2} + p$$

or $$y(p-1)^2 = p^2(c-p+\ln p) + p \quad (4)$$

Thus, (3) and (4) constitute the required solution of (1) with p as a parameter.

Example 33. Solve: $y + px = p^2x^4$. (1)

Solution. From (1), we get

$$y = p^2x^4 - px.$$

Differentiating the above equation w.r.t. x, we have

$$\frac{dy}{dx} = p = 4x^3p^2 + 2px^4\frac{dp}{dx} - p - x\frac{dp}{dx}$$

or $$2p - 4x^3p^2 - 2x^4p\frac{dp}{dx} + x\frac{dp}{dx} = 0$$

or $$2p(1-2px^3) + x(1-2px^3)\frac{dp}{dx} = 0.$$

or $$(1-2px^2)\left(2p + x\frac{dp}{dx}\right) = 0$$

Hence, either $1 - 2px^3 = 0$. (2)

or $2p + x\dfrac{dp}{dx} = 0$. (3)

The equation (3) gives

$$\frac{dp}{p} + \frac{2dx}{x} = 0$$

or $$\ln p + 2\ln x = \ln c$$

or $$px^2 = c \quad \text{or} \quad p = \frac{c}{x^2}.$$

Substituting this value of p into (1), we obtain

$$y = c^2 - \frac{c}{x}$$

or $xy - c^2x + c = 0$ is the required solution.

We have yet to deal with relation (2). If we eliminate p from (1) and (2), we get

$$y = -\frac{1}{4x^2}. \quad (3)$$

It is easy to check that (3) is also a solution of (1) and this solution does not involve any constant.

EQUATIONS SOLVABLE FOR x

Example 34. Solve: $xp = 1 + p^2$. (1)

Solution. We have

$$x = \frac{1}{p} + p \quad (2)$$

Differentiating (1) w.r.t. y, we have

$$\frac{dx}{dy} = \frac{1}{p} = -\frac{1}{p^2}\frac{dp}{dy} + \frac{dp}{dx}$$

or $$\frac{1}{p} = \left(1 - \frac{1}{p^2}\right)\frac{dp}{dy}$$

or $$1 = \left(p - \frac{1}{p}\right)\frac{dp}{dy}$$

which is a separable equation. Therefore,

$$dy = \left(p - \frac{1}{p}\right)dp.$$

Integrating, we get $$c + y = \frac{p^2}{2} - \ln p$$

i.e., $$2y = p^2 - 2\ln p - 2c \qquad (3)$$

Thus, (2) and (3) constitute the solution of (1).

CLAIRAUT'S EQUATION[1]

(9.24) The equation

$$y = xp + f(p), \qquad (1)$$

is known as **Clairaut's equation**. Differentiating both sides of (1), w.r.t. x, we get

$$p = p + xp' + f'(p)\,.\,p'.$$

Cancelling like terms, we have

$$xp' + p'f'(p) = 0$$

or $$p'\,[x + f'(p)] = 0.$$

Since one of the factors must be zero, two different solutions arise.

(i) If $p' = 0$, then $p = c$ and substitution of this value into (1) yields the general solution

$$y = cx + f(c).$$

(ii) If $x + f'(p) = 0$, then $x = -f'(p)$ and (1) can be rewritten as

$$y = -pf'(p) + f(p)$$

1. Named after the French astronomer and mathematician Alexis Claude Clairaut (1713 – 1765)

Thus x and y are both expressed as functions of p and we obtain the parametric equations

$$\left.\begin{aligned} x &= -f'(p) \\ y &= f(p) - pf'(p) \end{aligned}\right\} \qquad (2)$$

of a curve, representing a solution of (1). The solution (2) is called the **singular solution**. This solution is not deducible from the general solution. p may be eliminated between the two equations in (2) to get a relation in x and y involving no constant.

Example 35. Find the general solution and singular solution of

$$y = xp + \frac{1}{4}p^4. \qquad (1)$$

Solution. The general solution of the equation is

$$y = cx + \frac{1}{4}c^4.$$

Differentiating (1) w.r.t. x, we have

$$p = p + xp' + p^3p'$$

or $$x = -p^3. \qquad (2)$$

We eliminate p from (1) and (2) and get

$$y = x(-x)^{1/3} + \frac{1}{4}(-x)^{4/3} = -\frac{3}{4}x(-x)^{1/3}$$

or $$64y^3 + 27x^4 = 0 \text{ is the singular solution.} \qquad (1)$$

Example 36. Solve: $x^2(y - px) = yp^2$.

Solution. This is not Clairaut's equation. Let us write

$$x^2 = u \quad \text{and} \quad y^2 = v.$$

Then, $$2x\,dx = du \quad \text{and} \quad 2y\,dy = dv.$$

Hence $$\frac{y}{x}\frac{dy}{dx} = \frac{dv}{du}$$

$$p = \frac{dy}{dx} = \frac{x}{y}\frac{dv}{du}.$$

Substituting this value of p into (1), we get

$$x^2\left(y - \frac{x^2}{y}\frac{dv}{du}\right) = \frac{x^2}{y}\left(\frac{dv}{du}\right)^2$$

or $$y^2 - x^2\frac{dv}{du} = \left(\frac{dv}{du}\right)^2$$

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

or $$v = u\frac{dv}{du} + \left(\frac{dv}{du}\right)^2.$$

This is Clairaut's equation. Hence its general solution is

$$v = cu + c^2$$

or $$y^2 = cx^2 + c \qquad \text{is the general solution of (1)}$$

Differentiating (2) w.r.t. u, we get

$$\frac{dv}{du} = u\frac{d^2v}{du^2} + \frac{dv}{du} + 2\,\frac{dv}{du}\,\frac{d^2v}{du^2}$$

or $$\frac{d^2v}{du^2}\left(u + 2\frac{dv}{du}\right) = 0.$$

If $\dfrac{d^2v}{du^2} = 0$, then $v = cu + c^2$, which is the general solution.

If $u + 2\dfrac{dv}{du} = 0$, then

$$u = -2\frac{dv}{du}$$

and $$v = u\frac{dv}{du} + \left(\frac{dv}{du}\right)^2$$

give the singular solution.

Eliminating $\dfrac{dv}{du}$ from (3) and (2), we obtain

$$v = -\frac{u^2}{4}$$

or $$u^2 + 4v = 0 \qquad \text{is the singular solution.}$$

Replacing u, v by x^2 and y^2 respectively, we have

(i) $y^2 = cx^2 + c^2$ as the general solution

(ii) $x^4 + 4y^2 = 0$ as the singular solution.

EXERCISE 9.8

Solve (Problems 1–25)

1. $p^2 + p - 6 = 0$
2. $x^2p^2 + xyp - 6y^2 = 0$
3. $p^2y + (x - y)p - x = 0$
4. $p^3 - (x^2 + xy + y^2)p + xy^2 + x^2y = 0$
5. $xp^2 + (y - 1 - x^2)p - x(y - 1) = 0$
6. $xyp^2 + (x + y)p + 1 = 0$
7. $p^2 - (x^2y + 3)p + 3x^2y = 0$
8. $yp^2 + (x - y^2)p - xy = 0$
9. $(y + x)^2p^2 + (2y^2 + xy - x^2)p + y(y - x) = 0$
10. $xy(x^2 + y^2)(p^2 - 1) = p(x^4 + x^2y^2 + y^4)$
11. $xp^2 - 3yp + 9x^2 = 0$
12. $p^2 + x^3p - 2x^2y = 0$
13. $p^2 + 4x^5p - 12x^4y = 0$
14. $x^8p^2 + 3xp + 9y = 0$
15. $p^2 + 3xp - y = 0$
16. $y = px + x^3p^2$
17. $xp^2 - 2yp + ax = 0$
18. $p = \tan\left(x - \dfrac{p}{1 + p^2}\right)$
19. $p^3 - 4xyp + 8y^2 = 0$
20. $ap^2 + py - x = 0$
21. $e^{4x}(p - 1) + e^{2y}p^2 = 0$
22. $yp^2 - 2xp + y = 0$
23. $p^2\cos^2 y + p\sin x\cos x\cos y - \sin y\cos^2 x = 0$
24. $(px - y)(py + x) = 2p$
25. $y^2(y - xp) = x^4p^2.$

Find the general solution and the singular solution of each of the following differential equations:

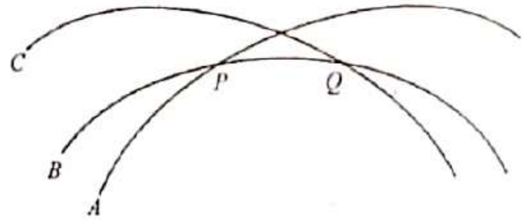
26. $y = xp - \ln p$
27. $y = xp - e^p$
28. $y = xp + a\sqrt{1 + p^2}$
29. $y = xp - \sqrt{p}$
30. $y = xp + p^3.$

ENVELOPE

(9.25) Definition. Let $f(x, y, c) = 0$ be a one-parameter family of curves. Suppose all members of the family of curves are drawn for various values of the parameter c arranged in order of magnitude. The two curves that correspond to two consecutive values of c will be designated as **neighbouring curves**. The locus of the ultimate points of intersection of neighbouring curves is called the **envelope** of the family $f(x, y, c) = 0$.

Let A, B, C represent three neighbouring intersecting members of the family. Let P be the point of intersection of A and B and Q be the point of intersection of B and C. By definition P and Q are points on the envelope. Thus the curve B and the envelope have two contiguous points common, and therefore, have ultimately a common tangent. Hence

B and the envelope touch each other. In the same way, we may show that the envelope touches any other member of the family. Thus a curve E which, at each of its points, touches someone of the curves of the family is the envelope of the family.



For the family of circles with centres on the x-axis and radius 1, that is, the family of circles with equation

$$(x-1)^2+y^2 = 1,$$

the pair of lines $y=\pm 1$ is the envelope. It is easy to see that the lines $y=\pm 1$ touch each member of the family of the circles.

(9.26) Equation of the Envelope. To find an equation of the envelope of the family $f(x, y, c)=0$, consider two neighbouring curves

$$\left.\begin{aligned} &f(x,y,c) = 0\\ \text{and}\quad &f(x,y,c+h) = 0 \end{aligned}\right\} \tag{1}$$

Find the intersection of these two curves and let $h \to 0$. The point of intersection then must approach the point of contact of the curve $f(x, y, c)=0$ with the envelope. At the point of intersection the equation

$$\frac{f(x,y,c+h)-f(x,y,c)}{h} = 0 \tag{2}$$

is true as well as (1). Letting $h\to 0$, we have from (2)

$$f_c(x,y,c) = 0 \tag{3}$$

and from (1) $$f(x,y,c) = 0 \tag{4}$$

If we eliminate c from (3) and (4), we obtain an equation of the envelope. The eliminant is called the **c-discriminant** of the family $f(x, y, c)=0$.

The c-discriminant may contain loci other than the envelope.

Example 37. The family of parabolas

$$(x-c)^2-2y = 0$$

has the x-axis as envelope.

SINGULAR SOLUTIONS

(9.27) Let $$f(x,y,p) = 0 \tag{1}$$

be a nonlinear first order differential equation in which the left hand member is a polynomial in p. The general solution of this differential equation will be a one-parameter family

$$f(x,y,c) = 0. \tag{2}$$

The envelope E of (2) is a curve which, at each of its points, touches some one member of the family (2). At a point of contact P of the envelope and a member of (2), the values of x, y, p are the same. But the values of x, y, p for the curve at P satisfy (1). Hence the values of x, y, p at every point of the envelope also satisfy (1). Thus the envelope of the family (2) is a solution of the differential equation (1).

Example 38. Consider $\quad p^2-xp+y=0.$

This is Clairaut's equation and its general solution (replacing simply p by c) is

$$y = cx+c^2. \tag{1}$$

Now we find the envelope of the family (1).

$$f(x,y,c) \equiv y-cx+c^2 = 0 \tag{2}$$

$$f_c \equiv \frac{\partial f}{\partial c} = -x+2c = 0. \tag{3}$$

Substituting the value of c from (3) into (2), we have

$$y-x\left(\frac{x}{2}\right)+\left(\frac{x}{2}\right)^2 = 0$$

i.e., $$y-\frac{x^2}{2}+\frac{x^2}{4} = 0$$

or $$4y = x^2. \tag{4}$$

(4) is an equation of the envelope of the family (1).

We check whether (4) is a solution of the given differential equation. Differentiating (4) w.r.t. x, we get

$$4\frac{dy}{dx} = 2x \quad\text{or}\quad 2p = x. \tag{5}$$

Substituting from (4) and (5) into p^2-px+y, we have

$$\frac{x^2}{4}-\frac{x^2}{2}+\frac{x^2}{4} = 0$$

Thus (4) is a solution of the given differential equation. This solution does not involve any constant and it cannot be derived from the general solution by giving particular values to c. Such a solution is called a **singular solution (S.S.)**.

(9.28) Definition. A solution of a differential equation $f(x, y, p) = 0$ is called a **singular solution (S.S.)** if

(i) it is not derived from the general solution by giving any particular value to the arbitrary constants

(ii) at each of its points, it is tangent to some member of the one-parameter family of curves represented by the general solution.

We have seen that the envelope, if any, of the one-parameter family of curves represented by the general solution of a differential equation is a singular solution. Thus **the c-discriminant equation may contain singular solution, if any**.

The singular solution may also be obtained from the differential equation directly without finding the general solution. Since at ultimate point of intersection of neighbouring curves the p's for the intersecting curves become equal, and thus the locus of the points where p's have equal roots will include the envelope. If we eliminate p from

$$f(x, y, p) = 0 \quad \text{and} \quad \frac{\partial f}{\partial p} = 0$$

the resulting equation is called the **p-discriminant**. The p-discrminant represents the locus for each point of which $f(x, y, p) = 0$ has equal roots.

If envelope of the general solution of $f(x, y, p) = 0$ exists, it will be contained in the p-discriminant. **Thus the p-discriminant equation may contain (i) singular solution (ii) solutions that are not singular and (iii) such loci which are not solutions at all.**

Example 39. Solve: $xp^2 - 2yp + 4x = 0$ (1)

Solution. $2yp = xp^2 + 4x$ or $2y = xp + \dfrac{4x}{p}$.

Differentiating w.r.t. x, we get

$$2p = x\frac{dp}{dx} + p + \frac{4}{p} - \frac{4x}{p^2}\frac{dp}{dx}$$

or $$p - \frac{4}{p} = \left(x - \frac{4x}{p^2}\right)\frac{dp}{dx}$$

or $$\frac{p^2 - 4}{p} = \frac{x\,(p^2 - 4)}{p^2}\,\frac{dp}{dx}$$

or $$\frac{dp}{p} = \frac{dx}{x}$$

or $$p = cx. \quad (2)$$

Eliminating p from (1) and (2), we obtain

$$x(c^2x^2) - 2ycx + 4x = 0$$

or $$c^2x^2 - 2yc + 4 = 0 \quad (3)$$

is the general solution.

From (3), we obtain the c-discriminant as

$$4y^2 - 16x^2 = 0 \quad \text{or} \quad y^2 = 4x^2.$$

The p-discriminant is [from (1)]

$$4y^2 = 16x^2 \quad \text{or} \quad y^2 = 4x^2$$

Since the c-discrminant and p-discriminant are the same viz., $y^2 = 4x^2$ and it satisfies the given differential equation, it is the singular solution.

Example 40. Solve: $x^2p^2 + yp\,(2x + y) + y^2 = 0$ (1)

by making the substitutions $y = u$, $xy = v$ and find the singular solutions.

Solution. $y = u$, $\quad x = \dfrac{v}{y} = \dfrac{v}{u}$.

Thus $$\frac{dy}{du} = 1, \qquad \frac{dx}{du} = \frac{u\dfrac{dv}{du} - v}{u^2}$$

$$p = \frac{dy}{dx} = \frac{dy}{du}\cdot\frac{du}{dx} = \frac{u^2}{u\dfrac{dv}{du} - v}.$$

Substituting into (1), we have

$$\frac{v^2}{u^2}\left[\frac{u^2}{u\dfrac{dv}{du} - v}\right]^2 + (2v + u^2)\frac{u^2}{u\dfrac{dv}{du} - v} + u^2 = 0$$

or $$v^2 + (2v + u^2)\left(u\frac{dv}{du} - v\right) + \left(u\frac{dv}{du} - v\right)^2 = 0$$

or $$v^2 + 2uv\frac{dv}{du} - 2v^2 + u^3\frac{dv}{du} - u^2v + u^2\left(\frac{dv}{du}\right)^2 - 2uv\frac{dv}{du} + v^2 = 0$$

or $$v = u\frac{dv}{du} + \left(\frac{dv}{du}\right)^2$$

which is a Clairaut's equation. Its solution is

$$v = cu + c^2$$

i.e., $$xy = cy + c^2 \quad (2)$$

is the general solution of (1).

From (2), the c-discriminant is

$$y^2 + 4xy = 0 \quad \text{i.e.,} \quad y(y+4x) = 0$$

Therefore, $y = 0$, $y + 4x = 0$.

Clearly, $y = 0$ satisfies (1) and so it is a singular solution. (3)

We also check whether $y + 4x = 0$ is a solution of (1).

Differentiating (3), we have $p = -4$. Substituting $p = -4$ into the left hand member of (1) and using (3), we obtain

$$16x^2 - 4x(-4)(-2x) + (-4x)^2 = 32x^2 - 32x^2 = 0$$

Thus $y + 4x = 0$ is also a singular solution.

Example 41. Solve

$$(x^2-1)p^2 - 2xyp - x^2 = 0 \tag{1}$$

and find the singular solution, if any.

Solution. Solving the equation for y, we have

$$2y = \frac{(x^2-1)p^2 - x^2}{xp} = xp - \frac{p}{x} - \frac{x}{p}.$$

Differentiating w.r.t. x, we get

$$2p = xp' + p - \frac{xp' - p}{x^2} - \frac{p - xp'}{p^2}, \quad \text{where } p' = \frac{dp}{dx}$$

or $$p = p'\left(x - \frac{1}{x} + \frac{x}{p^2}\right) + \frac{p}{x^2} - \frac{1}{p}$$

or $$p\left(1 - \frac{1}{x^2} + \frac{1}{p^2}\right) = xp'\left(1 - \frac{1}{x^2} + \frac{1}{p^2}\right)$$

Therefore, $\dfrac{dp}{p} = \dfrac{dx}{x}$. or $p = cx$.

Substituting this value of p into (1), we get

$$(x^2-1)c^2x^2 - 2x^2yc - x = 0$$

or $$(x^2-1)c^2 - 2yc - 1 = 0 \tag{2}$$

is the general solution of (1).

From (2), the c-discriminant is

$$4y^2 + 4(x^2-1) = 0$$

i.e., $$x^2 + y^2 = 1$$

which is also envelope of the family (2).

We also note that the p-discriminant is

$$4x^2y^2 + 4x^2(x^2-1) = 0$$

i.e., $$x^2 + y^2 = 1.$$

Since the envelope and the p-discriminant are same, the singular solution is

$$x^2 + y^2 = 1.$$

Example 42. Find the singular solution of

$$x^3p^2 + x^2yp + a^3 = 0 \tag{1}$$

Solution. The p-discriminant of (1) is

$$x^4y^2 - 4x^3a^3 = 0 \quad \text{or} \quad x^3(y^2x - 4a^3) = 0. \tag{2}$$

The parts of (2) that satisfy (1) are singular solutions of the given equation. From (2), we have

$$x = 0 \quad \text{or} \quad y^2x - 4a^3 = 0 \tag{3}$$

We rewrite (1) as

$$x^3 + x^2y\frac{dx}{dy} + a^3\left(\frac{dx}{dy}\right)^2 = 0. \tag{4}$$

Clearly, $x = 0$ satisfies (4).

Thus $x = 0$ is a singular solution of (1).

Differentiating $y^2x - 4a^3 = 0$ w.r.t. x, we have

$$2xyp + y^2 = 0 \quad \text{or} \quad y(2xp + y) = 0$$

i.e., $$p = -\frac{y}{2x}.$$

Putting this value of p into the left hand member of (1) and using (3), we obtain

$$x\left(-\frac{y}{2x}\right)^2 + x^2y\left(-\frac{y}{2x}\right) + a^3 = -\frac{xy^2}{4} + a^3 = -(xy^2 - 4a^3) = 0$$

Thus $xy^2 - 4a^3 = 0$ and its derivatives satisfy (1) and so it is a solution of (1).

Hence $x = 0$ and $xy^2 - 4a^3 = 0$

are singular solutions of (1).

THE RICATTI EQUATION[1]

(9.29) We have already studied first order linear differential equation

$$y' + P(x)\,y = R(x). \tag{1}$$

If we add the term $Q(x)\,y^2$ to the left hand member of (1), we obtain a nonlinear equation

$$y' + P(x)\,y + Q(x)\,y^2 = R(x) \tag{2}$$

(2) is called the Ricatti equation.

In many cases, the solution of (2) cannot be expressed in terms of **elementary functions**. However, the Ricatti equation

$$y' + Py + Qy^2 = R \tag{1}$$

can be reduced to a linear equation by the substitution $y = y_1 + \frac{1}{u}$, where y_1 is a particular solution of (1) and u is an unknown nonzero function of x.

Proof. Let $y = y_1 + \frac{1}{u}$ be as given.

Differentiating w.r.t. x, we have

$$y' = \frac{dy}{dx} = \frac{dy_1}{dx} - \frac{1}{u^2}\frac{du}{dx}$$

Substituting for y and y' into (1), we get

$$\frac{dy_1}{dx} - \frac{1}{u^2}\frac{du}{dx} + P\left(y_1 + \frac{1}{u}\right) + Q\left(y_1^2 + \frac{2y_1}{u} + \frac{1}{u^2}\right) = R$$

or
$$\frac{dy_1}{dx} + Py_1 + Qy_1^2 - R - \frac{1}{u^2}\left(\frac{du}{dx} - Pu - 2Qy_1u - Q\right) = 0 \tag{2}$$

Since y_1 is a solution of (1), we have

$$\frac{dy_1}{dx} + Py_1 + Qy_1^2 - R = 0$$

and so (2) reduces to

$$\frac{du}{dx} - (P + 2Qy_1)\,u = Q \tag{3}$$

which is a linear equation.

Thus if a particular solution of (1) is known then its general solution can be found.

1. After the name of Italian mathematician Count Jacobo Fransceo Ricatti (1676 – 1754).

Example 43. Solve $\frac{dy}{dx} - y^2 = -1$, $y(0) = 3$;

given that $y_1 = 1$ is a particular solution of the given equation

Solution. Here $P = 0,\ Q = -1,\ R = -1$.

Writing $y = 1 + \frac{1}{u}$, the given equation reduces to

$$\frac{du}{dx} - (0 + 2(-1)(1))\,u = -1$$

i.e.,
$$\frac{du}{dx} + 2u = -1, \tag{1}$$

which is a linear equation.

I.F. of (2) is $e^{\int 2\,dx} = e^{2x}$.

Multiplying (1) by the I.F., we get

$$\frac{du}{dx}e^{2x} + 2ue^{2x} = -e^{2x}$$

or
$$\frac{d}{dx}(ue^{2x}) = -e^{2x}$$

Integrating, we obtain

$$ue^{2x} = -\int e^{2x}\,dx + c = \frac{-e^{2x}}{2} + c$$

or
$$u = -\frac{1}{2} + \frac{c}{e^{2x}} \tag{2}$$

From $y = 1 + \frac{1}{u}$, we get by the initial condition,

$$y(0) = 3 = 1 + \frac{1}{u(0)} \quad \text{or} \quad u(0) = \frac{1}{2}.$$

Hence from (2), we have

$$u(0) = \frac{1}{2} = -\frac{1}{2} + c \quad \text{or} \quad c = 1$$

Thus
$$u = -\frac{1}{2} + \frac{1}{e^{2x}} = \frac{2 - e^{2x}}{2e^{2x}}$$

Required solution is

$$y = 1 + \frac{1}{u} = 1 + \frac{2e^{2x}}{2 - e^{2x}} = \frac{2 + e^{2x}}{2 - e^{2x}}$$

Example 44. Solve $\frac{dy}{dx} - \frac{y}{x} - x^3y^2 = -x^5$, (1)

by finding a particular solution.

Solution. It is a Ricatti equation with

$$P = -\frac{1}{x}, \quad Q = -x^3 \quad \text{and} \quad R = -x^5$$

An obvious solution of (1) is $y_1 = x$. Substituting $y = x + \frac{1}{u}$ into (1), we have

$$\frac{du}{dx} - \left[-\frac{1}{x} + 2(-x^3)x\right]u = -x^3$$

or

$$\frac{du}{dx} + \left(\frac{1}{x} + 2x^4\right)u = -x^3, \qquad (2)$$

which is a linear equation.

Its

$$\text{I.F.} = e^{\int\left(\frac{1}{x} + 2x^4\right)dx}$$

$$= e^{\ln x}\, e^{\frac{2}{5}x^5} = xe^{\frac{2}{5}x^5}$$

Multiplying (2) by the I.F. and integrating, we have

$$uxe^{\frac{2}{5}x^5} = -\int x^4 \cdot e^{\frac{2}{5}x^5}dx + c' = -\frac{1}{2}e^{\frac{2}{5}x^5} + c'$$

or

$$u = \frac{c' - \frac{1}{2}e^{\frac{2}{5}x^5}}{xe^{\frac{2}{5}x^5}}$$

Hence the general solution of (1) is

$$y = x + \frac{xe^{\frac{2}{5}x^5}}{c' - \frac{1}{2}e^{\frac{2}{5}x^5}} = \frac{x\left(c' + \frac{1}{2}e^{\frac{2}{5}x^5}\right)}{c' - \frac{1}{2}e^{\frac{2}{5}x^5}}$$

or

$$\frac{y}{x} = \frac{c' + \frac{1}{2}e^{\frac{2}{5}x^5}}{c' - \frac{1}{2}e^{\frac{2}{5}x^5}}$$

or

$$\frac{y-x}{y+x} = \frac{e^{\frac{2}{5}x^5}}{2c'} = ce^{\frac{2}{5}x^5}$$

is the required solution.

EXERCISE 9.9

Solve and find the singular solution, if any (Problems 1–10):

1. $y = px - p^2$
2. $p^2(x^2 - a^2) - 2pxy + y^2 - b^2 = 0$
3. $xp^2 - (x - y)p - 1 - y = 0$
4. $4p^2 = 9x$
5. $4xp^2 = (3x - 1)^2$
6. $p^2 + 2px^3 - 4x^2y = 0$
7. $6p^2y^2 - 3px - y = 0$
8. $x^2p^2 + x^3yp - 1 = 0$
9. $xp^4 - 2yp^3 + 12x^3 = 0$
10. $8p^3x - 12p^2y - 27x = 0$

Investigate each of the following for singular solutions by finding the p-discriminant (Problems 11–16):

11. $e^{y} - p$ [illegible] $= 1 = 0$
12. $4x(x-1)(x-2)p^2 - (3x^2 - 6x + 2)^2 = 0$
13. $p^4 - 2p^2$ [illegible] $- y^4 = 0$
14. $p^3 - 4xyp + 8y^2 = 0$
15. $4p^2 - 3xp - y = 0$
16. $p^3 - 2x^2p - 4xy = 0$

Solve the following Ricatti equations:

17. $\frac{dy}{dx} - y - \frac{2}{x^2}y^2 = -x^2$, given that $y_1 = x^2$ is a particular solution.
18. $\frac{dy}{dx} + \frac{3}{x}y - y^2 = \frac{1}{x^2}$, given that $y_1 = \frac{1}{x}$ is a particular solution.
19. $\frac{dy}{dx} - 4y - y^2 = 4$
20. $\frac{dy}{dx} = 7 - 6y - y^2$
21. $\frac{dy}{dx} + (\cot x)\,y - y^2 = -\csc^2 x$

Mathematics takes us into the region of absolute necessity, for which not only the actual world but also every possible world must conform.

BERTRAND RUSSEL
(1872-1970 C.E.)
English Philosopher, Mathematician and Author

Mathematics is the standard of objective truth for all intellectual endeavours.

HERMANN WEYL
(1885-1955 C.E.)
German Mathematician

Chapter **10**

DIFFERENTIAL EQUATIONS OF HIGHER ORDER

In Chapter 9, we studied methods of solving special types of linear and nonlinear differential equations of the first order. In this chapter, we shall consider systematic methods for the solution of certain classes of differential ~~equations~~ of order more than one.

$\frac{dy}{dx} + P(x)\, y = Q(x)$

LINEAR DIFFERENTIAL EQUATIONS

(10.1) Definition. A linear differential equation of order n in the dependent variable y and the independent variable x is of the form

$$a_0(x)\frac{d^ny}{dx^n} + a_1(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_{n-1}(x)\frac{dy}{dx} + a_n(x)\,y = F(x), \qquad \text{(St. form)} \quad (1)$$

where $a_0(x)$, $a_1(x)$, $\cdots$, $a_{n-1}(x)$, $a_n(x)$ and $F(x)$ are functions of the independent variable x only and $a_0(x)$ is not identically zero. Using primes, (1) is also written as

$$a_0(x)\,y^{(n)} + a_1(x)\,y^{(n-1)} + \cdots + a_{n-1}(x)\,y' + a_n(x)\,y = F(x).$$

We shall first study equations of the type

$$a_0\,y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = F(x) \qquad (2)$$

where a_0, a_1, $\cdots$, a_{n-1}, a_n are real constants.

Equation (1) is with variable coefficients while (2) is with constant coefficients.

In order to solve (2), we shall first consider the equation

$$a_0 y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = 0. \tag{3}$$

The coefficient of $y^{(n)}$ may be made 1 by dividing throughout by a_0. The differential equation (3) is called **homogeneous** linear differential equation of order n [The use of the word **homogeneous** here is quite different from the one already mentioned in **9.12**]. If $F(x)$ is not identically zero then (2) is called **nonhomogeneous** and (3) is called the **associated homogeneous** equation of (2).

(10.2) Definition. If $y_1(x), y_2(x), \cdots, y_m(x)$ are m functions of an independent variable x and $c_1, c_2, \cdots, c_m$ are constants, then the expression

$$c_1 y_1(x) + c_2 y_2(x) + \cdots + c_m y_m(x)$$

is called a **linear combination** of $y_1(x), y_2(x), \cdots, y_m(x)$. We usually write $y_1, y_2, \cdots, y_m$ instead of $y_1(x), y_2(x), \cdots, y_m(x)$ when it is clear from the context that $y_1, y_2, \cdots, y_m$ are functions of x.

As in vector spaces, the m nonzero functions $y_1, y_2, \cdots, y_m$ are called **linearly dependent** if and only if, there exist constants $c_1, c_2, \cdots, c_m$, at least one of which is nonzero, such that

$$c_1 y_1 + c_2 y_2 + \cdots + c_m y_m = 0.$$

The functions $y_1, y_2, \cdots, y_m$ are called **linear independent** if and only if, they are not linearly dependent, i.e., if and only if

$$c_1 y_1 + c_2 y_2 + \cdots + c_m y_m = 0$$

implies $\quad c_1 = c_2 = \cdots = c_m = 0.$

(10.3) Before investigating a solution of (2) of **(10.1)**, we state the following facts:

(i) Every homogeneous linear nth-order differential equation

$$a_0 y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = 0 \tag{3}$$

has n linearly independent solutions $y_1, y_2, \cdots, y_n$.

(ii) If $y_1, y_2, \cdots, y_m$ are n linearly independent solutions of (3), then any linear combination

$$y_c = c_1 y_1 + c_2 y_2 + \cdots + c_n y_n \quad \text{of} \quad y_1, y_2, \cdots, y_n$$

is the general solution of (3); $c_1, c_2, \cdots, c_m$ being arbitrary constants.

(iii) Let y_p be any **particular solution** of (2) of **(10.1)**, i.e., y_p does not contain any constant, then $y_c + y_p$ is the **general solution** of (2).

(ii) and (iii) can be easily checked by actual substitutions into (3) and (2) of (10.1). Proof of (i) is beyond the scope of this book.

Thus to find the general solution of (2) of **(10.1)**, we have to find a linearly ...pendent set of n solutions of (3) so as to determine y_c and a particular solution y_p of ...d then obtain

$$y = y_c + y_p \tag{4}$$

...e general solution of (2). In the general solution (4), y_c is called the **Complementary** ...ction **(C.F.)** and y_p is called the **Particular Integral (P.I.)** of (2).

These statements are also true for the equation (1) of **(10.1)** with variable ...fficients.

HOMOGENEOUS LINEAR EQUATIONS

(10.4) Consider the equation

$$a_0 \frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_{n-1} \frac{dy}{dx} + a_n y = 0 \tag{1}$$

...ere $a_0, a_1, \cdots, a_{n-1}, a_n$ are real constants. To find a solution of (1) we shall try a ...cious guess. The differential equation (1) requires a function y with the property that if ...d its successive derivatives are each multiplied by constants a_j, $j = n, n-1, \cdots, 1, 0$, ...e resulting products are then added, the sum should equal zero. This can only happen if ...function is such that its various derivatives are constant multiples of itself. The ...ponential function $y = e^{mx}$, m being a constant, has such properties. Here we have

$$\frac{dy}{dx} = me^{mx}, \quad \frac{d^2 y}{dx^2} = m^2 e^{mx}, \cdots, \frac{d^n y}{dx^n} = m^n e^{mx}.$$

...bstituting into (1) we have

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

$$e^{mx} (a_0 m^n + a_1 m^{n-1} + \cdots + a_{n-1} m + a_n) = 0$$

...nce $e^{mx} \neq 0$, we have

$$a_0 m^n + a_1 m^{n-1} + \cdots + a_{n-1} m + a_n = 0 \tag{2}$$

...hus, $y = e^{mx}$ is a solution of (1) if and only if m is a solution of (2). Equation (2) is called ...e **characteristics (or auxiliary) equation** of the given differential equation (1). Observe ...at (2) can be obtained from (1) by merely replacing the kth derivative in (1) by ...^{k} ($k = n, n-1, \cdots, 1$). Three cases arise according as the roots of (2) are.

(I) real and distinct

(II) real and repeated

(III) complex.

Case I. Distinct Real Roots

Let $m_1, m_2, \cdots, m_n$

be n distinct real roots of (2). Then $e^{m_1 x}, e^{m_2 x}, \cdots, e^{m_n x}$ are n distinct solutions of (1). These n solutions are linearly independent. Hence the general solution of (2) is

$$y = c_1 e^{m_1 x} + c_2 e^{m_2 x} + \cdots + c_n e^{m_n x},$$

where $c_1, c_2, \cdots, c_n$ are arbitrary constants.

Case II. Repeated Real Roots

In equation (1), writing $D \equiv \dfrac{d}{dx}$, we have

$$(a_0 D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n)\, y = 0$$

or $$\big[f(D)\big]\, y = 0,$$

where $$f(D) = a_0 D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n.$$

Note that if we write D for m in the characteristic equation (2) then it is the same as $f(D) = 0$. If $m_1, m_2, \cdots, m_n$ are the roots of $f(D) = 0$, then (1) may be written as

$$(D - m_1)(D - m_2) \cdots (D - m_n)\, y = 0.$$

Let the root m_1 be repeated twice say $m_2 = m_1$.

The part of the general solution of (1) corresponding to the twice repeated root m_1 of (2) is solution of

$$(D - m_1)^2\, y = 0$$

i.e., of $$(D - m_1)(D - m_1)\, y = 0 \qquad (3)$$

Let $$(D - m_1)\, y = V.$$

Then (3) becomes

$$(D - m_1)\, V = 0$$

or $$\frac{dV}{dx} - m_1 V = 0. \qquad (4)$$

Separating the variables, we have

$$\frac{dV}{V} = m_1\, dx.$$

Therefore, $$\ln V = m_1 x + k, \qquad \text{where } k \text{ is a constant}$$

or $$V = c_2 e^{m_1 x}, \qquad \text{where } c_2 \text{ is a constant.}$$

Replacing V in (4), we obtain

$$(D - m_1)\, y = c_2 e^{m_1 x}$$

or $$\frac{dy}{dx} - m_1 y = c_2 e^{m_1 x} \qquad (5)$$

which is a linear equation of order one

Its $$\text{I.F.} = \exp\left(\int (-m_1)\, dx\right) = e^{-m_1 x}$$

Multiplying (5) by $e^{-m_1 x}$, we get

$$\frac{d}{dx}\left(y e^{-m_1 x}\right) = c_2$$

or $$y e^{-m_1 x} = c_2 x + c_1$$

or $$y = (c_1 + c_2 x)\, e^{m_1 x}$$

is the part of the general solution corresponding to the twice repeated root m_1. The general solution of (1) is

$$y = (c_1 + c_2 x)\, e^{m_1 x} + c_3 e^{m_3 x} + \cdots + c_n e^{m_n x}$$

In the same manner, if the characteristic equation (2) has the triple repeated root m_1, the corresponding part of the general solution of (1) is the solution of

$$(D - m_1)^3\, y = 0.$$

Proceeding as before, we can easily find

$$y = (c_1 + c_2 x + c_3 x^2)\, e^{m_1 x}$$

as the part of the general solution corresponding to this triple repeated root m_1.

If the characteristic equation (2) has the real root m_1 occurring k times then the part of the general solution of (1) corresponding to the k-fold repeated root m_1 is

$$y = (c_1 + c_2 x + c_3 x^2 + \cdots + c_k x^{k-1})\, e^{m_1 x}$$

Case III. Complex Roots

Real Part of

Suppose the characteristic equation has the complex number $a + ib$ as a non-repeated root. Since coefficients of (1) are real, the conjugate complex number $a - ib$ is also a non-repeated root. The corresponding part of the general solution is

$$y = k_1 e^{(a+ib)x} + k_2 e^{(a-ib)x}, \text{ where } k_1 \text{ and } k_2 \text{ are arbitrary constants}$$

i.e.,

$$\begin{aligned} y &= e^{ax}\left[k_1 e^{ibx} + k_2 e^{-ibx})\right] \\ &= e^{ax}\left[k_1(\cos bx + i\sin bx) + k_2(\cos bx - i\sin bx)\right] \\ &= e^{ax}\left[(k_1 + k_2)\cos bx + i(k_1 - k_2)\sin bx\right] \\ &= e^{ax}\left[c_1 \sin bx + c_2 \cos bx\right] \end{aligned}$$

where $c_1 = i(k_1 - k_2)$, $c_2 = k_1 + k_2$ are two arbitray constants.

If $a + ib$ and $a - ib$ are conjugate complex roots, each repeated k times, then the corresponding part of the general solution of (1) may be written as

$$y = \left[(c_1 + c_2 x + c_3 x^2 + \cdots + c_k x^{k-1})\sin bx + (c_{k+1} + c_{k+2} x + \cdots + c_{2k} x^{k-1})\cos bx\right]$$

In the examples given below we use these concepts to solve homogeneous linear differential equations.

Example 1. Solve. $(D^2 + 4D + 3)y = 0$

Solution. The characteristic equation is

$$D^2 + 4D + 3 = 0$$

with roots $D = -1, -3$.

Hence the general solution of the given equation is

$$y = c_1 e^{-x} + c_2 e^{-3x}.$$

Synthetic division

Example 2. Solve: $(D^3 - 5D^2 + 7D - 3)y = 0$

Solution. The characteristic equation is

$$D^3 - 5D^2 + 7D - 3 = 0. \quad (1)$$

$$\begin{array}{r|rrrr} & 1 & -5 & 7 & -3 \\ 1 & \downarrow & 1 & -4 & 3 \\ \hline D-1=0 & 1 & -4 & 3 & \boxed{0} \end{array}$$

$D^2 - 4D + 3 = 0$

By inspection, $D = 1$ is a solution of this equation. The other two roots can be found from the quadratic factor of

$$(D-1)(D^2 - 4D + 3) = 0.$$

Hence all roots of (1) are $D = 1, 1, 3$.

The general solution is

$$y = (c_1 + c_2 x)e^x + c_3 e^{3x}$$

Example 3. Solve: $(D^3 - D^2 + D - 1)y = 0$

Solution. The characteristic equation is

$$D^3 - D^2 + D - 1 = 0$$

By inspection, $D = 1$ is a root of this equation.

Now,

$$D^3 - D^2 + D - 1 = (D-1)(D^2+1) = 0$$

Hence the other two roots are

$$D = \pm i = a + ib \text{ with } a = 0, b = 1$$

The general solution is

$$y = c_1 e^x + e^{0x}(c_2 \sin x + c_3 \cos x)$$

or

$$y = c_1 e^x + c_2 \sin x + c_3 \cos x$$

Example 4. Solve: $(D^2 + D - 12)y = 0$, where $y(2) = 2$, $y'(2) = 0$

Solution. The characteristic equation is

$$D^2 + D - 12 = 0$$

with roots $D = 3, -4$

Hence the general solution of the given equation is

$$y = c_1 e^{3x} + c_2 e^{-4x} \quad (1)$$

We now determine the coefficients (constants) c_1, c_2 in (1) from the initial conditions as follows.

Since the initial conditions are given at $x = 2$, we rewrite the general solution in the form

$$y = k_1 e^{3(x-2)} + k_2 e^{-4(x-2)} \quad (2)$$

where $k_1 = c_1 e^6$, $k_2 = c_2 e^{-8}$.

Applying initial condition, (i.e replacing c_1, c_2 by the new constants $k_1 e^6$, $k_2 e^{-8}$ respectively in (1), we get

$$y(2) = 2 = k_1 + k_2 \quad (3)$$

Differentiating (2) w.r.t. x, we have

$$y' = \frac{dy}{dx} = 3k_1 e^{3(x-2)} - 4k_2 e^{-4(x-2)}$$

Therefore, $y'(2) = 0 = 3k_1 - 4k_2$. (4)

Solving (3) and (4), we obtain

$$k_1 = \frac{8}{7}, \quad k_2 = \frac{6}{7}$$

Hence the solution (2) of the differential equation satisfying the given conditions is

$$y = \frac{8}{7}\, e^{3(x-2)} + \frac{6}{7}\, e^{-4(x-2)}$$

Example 5. Solve $(D^2 + 4D + 5) = 0, \quad y(0) = 1, \quad y'(0) = 0.$

Solution. The characteristic equation is

$$D^2 + 4D + 5 = 0$$

which has roots $D = -2 \pm i.$

The general solution is

$$y = e^{-2x}(c_1 \sin x + c_2 \cos x). \qquad (1)$$

Applying the given conditions, we have, from (1)

$$y(0) = 1 = c_2.$$

Differentiating (1) w.r.t. x, we get

$$y' = -2e^{-2x}(c_1 \sin x + c_2 \cos x) + e^{-2x}(c_1 \cos x - c_2 \sin x)$$

Therefore, $y'(0) = 0 = -2c_2 + c_1$, giving $c_1 = 2.$

Substituting the values of c_1 and c_2 into (1), the required solution is

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

Example 6. Solve $(D^3 - 3D^2 + 4)\, y = 0.$

$$y(0) = 1,\ y'(0) = -8,\ y''(0) = -4$$

Solution. The characteristic equation is

$$D^3 - 3D^2 + 4 = 0$$

or $$(D + 1)(D^2 - 4D + 4) = 0.$$

Therefore, $$D = -1, 2, 2.$$

The general solution is

$$y = c_1e^{-x} + e^{2x}(c_2 + c_3x) \qquad (1)$$

Now, from (1) $$y' = -c_1e^{-x} + 2e^{2x}(c_2 + c_3x) + c_3e^{2x}$$

and $$y'' = c_1e^{-x} + 4e^{2x}(c_2 + c_3x) + 4c_3e^{2x}$$

Applying initial conditions, we get from the above three equation

$$y(0) = 1 = c_1 + c_2 \qquad (2)$$

$$y'(0) = -8 = -c_1 + 2c_2 + c_3 \qquad (3)$$

$$y''(0) = -4 = c_1 + 4c_2 + 4c_3 \qquad (4)$$

Multiplying (3) by -4 and adding to (4), we have

$$28 = 5c_1 - 4c_2 \qquad (5)$$

Multiplying (2) by 4 and adding to (5), we obtain

$$9c_1 = 32 \quad \text{or} \quad c_1 = \frac{32}{9}$$

Therefore, $c_2 = 1 - \frac{32}{9} = -\frac{23}{9}$ and $c_3 = \frac{6}{9}$

The required solution is

$$y = \frac{32}{9}\, e^{-x} + e^{2x}\left(\frac{-23}{9} + \frac{6}{9}x\right) = \frac{1}{9}\left(32\, e^{-x} - 23\, e^{2x} + 6xe^{2x}\right)$$

complete

EXERCISE 10.1

Solve: (Problems 1 – 15)

1. $(9D^2 - 12D + 4)\, y = 0$
2. $(75D^2 + 50D + 12)\, y = 0$
3. $(D^3 - 4D^2 + D + 6)\, y = 0$
4. $(D^3 + D^2 + D + 1)\, y = 0$
5. $(D^3 - 6D^2 + 12D - 8)\, y = 0$
6. $(D^3 - 6D^2 + 3D + 10)\, y = 0$
7. $(D^3 - 27)\, y = 0$
8. $(4D^4 - 4D^3 - 3D^2 + 4D - 1)\, y = 0$
9. $(D^4 + 2D^3 - 2D^2 - 6D + 5)\, y = 0$
10. $(D^4 - 5D^3 + 6D^2 + 4D - 8)\, y = 0$
11. $(D^4 - 4D^3 - 7D^2 + 22D + 24)\, y = 0$
12. $(D^4 + 4)\, y = 0$
13. $(D^4 - D^3 - 3D^2 + D + 2)\, y = 0$
14. $(16D^6 + 8D^4 + D^2)\, y = 0$
15. $(D^4 + 6D^3 + 15D^2 + 20D + 12)\, y = 0$

Solve the following equations with the given conditions:

16. $(D^2 + 8D - 9)\, y = 0;$ $\quad y(1) = 1, \quad y'(1) = 0$
17. $(D^2 + 6D + 9)\, y = 0;$ $\quad y(0) = 2, \quad y'(0) = -3$
18. $(D^2 + 6D + 13)\, y = 0;$ $\quad y(0) = 3, \quad y'(0) = -1$
19. $(D^3 - 6D^2 + 11D - 6)\, y = 0;$ $\quad y(0) = 0 = y'(0),\ y''(0) = 2$
20. $(D^4 - D^3)\, y = 0;$ $\quad y(0) = y'(0) = 1,\ y''(1) = 3e,\ y'''(1) = e$

DIFFERENTIAL OPERATORS

(10.5) Let $T : V \to V$, where V is a vector space over a field F, be a **linear transformation** (or an operator) from V to V.

Linear transformations S, T, U defined on V have the following properties.

(i) $S + T = T + S$ where $(S + T)\,v = S(v) + T(v),\ v \in V$,

(ii) $(S + T) + U = S + (T + U)$

(iii) $S(T + U) = ST + SU$

(iv) $(ST)\,U = S(TU)$

(v) $ST \neq TS$ in general.

Here $ST : V \to V$ is a linear transformation defined by

$$(ST)(v) = S(Tv) \quad \text{for all } v \in V.$$

Let X be the vector space of all real (or complex) valued real (or complex) functions possessing a kth order derivative for every $k = 1, 2, 3, \cdots$. For each $g \in X$, $y = g(x) \in R$ (or $\in C$). D^k which associates with each $y = g(x)$, its kth derivative, is a linear operator on X. We write $D^k y$ as $\dfrac{d^k y}{d^k x}$. Since the sum and product of the linear operators and scalar multiple of a linear operator are also linear operators, the linear combination

$$f(D) = a_0 D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n I \tag{1}$$

of $D^n, D^{n-1}, \cdots, D, I$ is also a linear operator. Here, I is the identity linear operator defined by $I(y) = y$ for all $y = g(x)$ and $a_0, a_1, \cdots, a_n$ are scalars. The image of a $y = g(x)$ under $f(D)$ is written as

$$\begin{aligned} f(D)y &= a_0 D^n y + a_1 D^{n-1} y + \cdots + a_{n-1} Dy + a_n y \\ &= a_0 \frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_{n-1} \frac{dy}{dx} + a_n y. \end{aligned}$$

The equation $f(D)y = 0$

is then a **linear differential equation of order** $\boldsymbol{n}$.

The operators $D, D^2, \cdots, D^n$ and $f(D)$, given by (1), are called **differential operators**.

We have written D for $\dfrac{d}{dx}$, D^2 for $\dfrac{d^2}{dx^2}$ and so on D^n for $\dfrac{d^n}{dx^n}$. Thus

$$Dy = \frac{dy}{dx}, \quad D^2 y = \frac{d^2 y}{dx^2}, \quad \text{and } D^n y = \frac{d^n y}{dx^n}.$$

The expression

$$A = f(D) = a_0 D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n I$$

is called a **differential operator of order** $\boldsymbol{n}$. This may be thought of as an operator which, when applied to any function y, results in

$$Ay = f(D)\,y = a_0 \frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_{n-1} \frac{dy}{dx} + a_n Iy.$$

NONHOMOGENEOUS LINEAR EQUATIONS

In this section we discuss the solution of nonhomogeneous linear equation (2) defined in (10.1).

(10.6) Solution of the equation

$$f(D)\,y = F(x) \tag{1}$$

where $f(D) = a_0 D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n$.

We have already remarked in (10.3) that the general solution of (1) consists of two parts, namely

(i) Complementary Function (C.F.)

(ii) Particular Integral (P.I.)

The C.F. is the solution of the homogeneous equation $f(D)\,y = 0$ and we have described different cases of its solution in (10.4).

To find the P.I. of (1), we write

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

and try to evaluate $y = \dfrac{1}{f(D)}\, F(x)$.

A real number m is said to be a zero of $f(D)$ if $f(m) = 0$.

Suppose $f(D)$ has n distinct zeros $m_1, m_2, \cdots, m_n$.

Then $$\frac{1}{f(D)} F(x) = \frac{1}{(D - m_1)(D - m_2) \ldots (D - m_n)}\, F(x).$$

Suppose $\quad \dfrac{1}{D-m_n} F(x) = y^*$.

Then $\quad (D-m_n)\, y^* = F(x)$

or $\quad \dfrac{dy}{dx} - m_n y^* = F(x)$

which is a linear equation of the first order. Its

$$\text{I.F.} = \exp\left(\int (-m_n)\right) dx = \exp(-m_n x)$$

Hence $\quad \dfrac{d}{dx}\left(y e^{-m_n x}\right) = F(x)\, e^{-m_n x}$

or $\quad y^* = e^{m_n x} \int F(x)\, e^{-m_n x}\, dx,$

omitting the constant of integration, since we are looking for a particular solution.

Therefore,

$$\frac{1}{f(D)} F(x) = \frac{1}{(D-m_1)\ldots(D-m_{n-1})}\, e^{m_n x} \int e^{-m_n x} F(x)\, dx.$$

Next we evaluate $\quad \dfrac{1}{D-m_{n-1}}\, e^{m_n x} \int e^{-m_n x} F(x)\, dx$

as before and continue the process. This is the required P.I. of the equation (1).

Alternatively, we may also write

$$\frac{1}{f(D)} F(x) = \frac{1}{(D-m_1)(D-m_2)\ldots(D-m_n)} F(x)$$

$$= \left[\frac{A_1}{D-m_1} + \frac{A_2}{D-m_2} + \ldots + \frac{A_n}{D-m_n}\right] F(x)$$

after resolving $\dfrac{1}{f(D)}$ into partial fractions

$$= A_1 e^{m_1 x} \int e^{-m_1 x} F(x)\, dx + A_2 e^{m_2 x} \int e^{-m_2 x} F(x)\, dx + \cdots$$

$$+ A_n e^{m_n x} \int e^{-m_n x} F(x)\, dx$$

which is the required P.I. of the equation (1).

Note: Mostly we follow the alternative method.

Example 7. Solve $\quad (D^3 - D)y = e^x$

Solution. The characteristic equation is

$D^3 - D = 0$ with roots $0, 1, -1$

$$\text{C.F.} = y_c = c_1 + c_2' e^x + c_3 e^{-x}$$

$$\text{P.I.} = y_p = \frac{e^x}{D(D+1)(D-1)}$$

Now $\quad \dfrac{1}{D-1}\, e^x = e^x \int e^{-x}\, e^x\, dx = x e^x$, by the method described in (10.6)

Thus $\quad \dfrac{1}{D(D+1)(D-1)}\, e^x = \dfrac{1}{D(D+1)}\, xe^x$

Again,

$$\frac{1}{D+1}(xe^x) = e^{-x} \int e^x\, xe^x\, dx = \frac{xe^x}{2} - \frac{e^x}{4}$$

Hence

$$\frac{1}{D(D+1)}(xe^x) = \frac{1}{D}\left(\frac{xe^x}{2} - \frac{e^x}{4}\right) = \int \frac{xe^x}{2}\, dx - \int \frac{e^x}{4}\, dx$$

$$= \frac{xe^x}{2} - \frac{3}{4} e^x$$

Therefore, the general solution of the equation is

$$y = y_c + y_p = c_1 + c_2' e^x + c_3 e^{-x} + \frac{xe^x}{2} - \frac{3}{4} e^x$$

$$= c_1 + \left(c_2' - \frac{3}{4}\right) e^x + c_3 e^{-x} + \frac{x e^x}{2}$$

$$= c_1 + c_2 e^x + c_3 e^{-x} + \frac{x e^x}{2}$$

Alternative Method, Consider $f(D)^{-1}$ as the inverse of the operator $f(D)$, so that

$$y_p = \frac{1}{D(D+1)(D-1)}\, e^x$$

$$= \left(-\frac{1}{D} + \frac{1}{2}\, \frac{1}{1+D} + \frac{1}{2}\, \frac{1}{D-1}\right) e^x$$

$$= -\int e^x\, dx + \frac{1}{2}\, e^{-x} \int e^x \cdot e^x\, dx + \frac{1}{2}\, e^x \int e^{-x} e^x\, dx$$

$$= -e^x + \frac{1}{4}\, e^x + \frac{1}{2}\, xe^x$$

$= \dfrac{xe^x}{2}$, since e^x already occurs in the C.F. and so the terms involving e^x are omitted

WORKING RULES FOR FINDING P.I.

The method of evaluating $\frac{1}{f(D)} F(x)$ explained in (10.6) involves successive integrations and is quite unwieldy. We list below short methods for finding the P.I. when $F(x)$ is function of a particular type.

(10.7) To Evaluate $\frac{1}{f(D)} e^{ax}$, where

$$f(D) = a_0D^n + a_1D^{n-1} + \cdots + a_{n-1} D + a_n$$

We know that

$$D^k e^{ax} = a^k e^{ax}, \text{ where } k \text{ is a positive integer.}$$

$$\begin{aligned}\text{Hence } f(D)\, e^{ax} &= (a_0D^n + a_1D^{n-1} + \cdots + a_{n-1} D + a_n)\, e^{ax}\\ &= a_0D^n e^{ax} + a_1D^{n-1} e^{ax} + \cdots + a_{n-1} De^{ax} + a_n e^{ax}\\ &= a_0a^n e^{ax} + a_1a^{n-1} e^{ax} + \cdots + a_{n-1}\, a\, e^{ax} + a_n e^{ax}\\ &= f(a)\, e^{ax}.\end{aligned} \qquad (1)$$

Applying $\frac{1}{f(D)}$ to both sides of (1), we obtain

$$\frac{f(D)}{f(D)} e^{ax} = \frac{1}{f(D)} f(a)\, e^{ax}$$

or

$$e^{ax} = f(a) \frac{1}{f(D)} e^{ax}$$

i.e.,

$$\frac{1}{f(D)} e^{ax} = \frac{e^{ax}}{f(a)}, \qquad \text{if } f(a) \neq 0 \qquad (2)$$

If $f(a) = 0$, then a is a root of $f(D) = 0$.

Let a be a k-fold root of $f(D) = 0$ so that $(D-a)^k$ is a factor of $f(D)$. Then

$$f(D) = (D-a)^k\, \phi(D), \text{ where } \phi(a) \neq 0.$$

We first check the effect of $(D-a)^k$ on $e^{ax}\, p(x)$, for a polynomial $p(x)$ in x.

We have

$$(D-a)(e^{ax}\, p(x)) = D(e^{ax}\, p(x)) - a\, e^{ax}\, p(x) = e^{ax}\, D\, p(x)$$

$$(D-a)^2(e^{ax}\, p(x)) = (D-a)\,(e^{ax}\, Dp(x)) = e^{ax}\, D^2 p(x).$$

Continuing in this way, we get

$$(D-a)^k(e^{ax}\, p(x)) = e^{ax}\, D^k p(x) \qquad (3)$$

Setting $p(x) = x^k$ in (3), we are led to

$$(D-a)^k\,(x^k e^{ax}) = e^{ax} D^k x^k = k!\, e^{ax}$$

$$\phi(D)\,(D-a)^k\,(x^k e^{ax}) = \phi(D)\, k!\, e^{ax}$$

or

$$= k!\, \phi(D)\, e^{ax} = k!\, \phi(a)\, e^{ax}, \quad \text{by (1)}$$

Operating on both sides by $\frac{1}{\phi(D)\,(D-a)^k}$, we obtain

$$x^k e^{ax} = \frac{1}{\phi(D)\,(D-a)^k}\, k!\, \phi(a)\, e^{ax}$$

or

$$\frac{1}{\phi(D)\,(D-a)^k}\, e^{ax} = \frac{x^k e^{ax}}{k!\, \phi(a)}. \qquad (5)$$

(10.8) Principle of Superposition. Let

$$f(D)\, y = (a_0D^n + a_1D^{n-1} + \cdots + a_{n-1} D + a_n)\, y = F(x) \qquad (1)$$

be a linear differential equation of order n. If $F(x) = F_1(x) + F_2(x)$, then particular integral of (1) is the sum of particular integrals of

$$F(D)\, y = F_1(x) \qquad (2)$$

and

$$F(D)\, y = F_2(x). \qquad (3)$$

Proof. Let y_1 and y_2 be particular integrals of (2) and (3) respectively.

Then

$$f(D)\, y_1 = F_1(x)$$

and

$$f(D)\, y_2 = F_2(x)$$

Suppose $y = y_1 + y_2$. Setting this value of y into (1), we obtain

$$\begin{aligned} f(D)\, y &= f(D)\,(y_1 + y_2)\\ &= f(D)\, y_1 + f(D)\, y_2\\ &= F_1(x) + F_2(x) = F(x)\end{aligned}$$

showing that y is a solution of (1).

Example 8. Solve. $(D^3 + 1)\, y = 1 + e^{-x} + e^{2x}$.

Solution. The characteristic equation is

$$D^3 + 1 = 0 \text{ with roots } -1, \frac{1 \pm i\sqrt{3}}{2}.$$

Therefore, C.F. is

$$y_c = c_1 e^{-x} + e^{x/2}\left(c_2 \sin \frac{\sqrt{3}}{2} x + c_3 \cos \frac{\sqrt{3}}{2} x\right)$$

The P.I. is given by

$$y_p = \frac{1}{D^3+1}(1 + e^{-x} + e^{2x})$$

$$= \frac{1}{D^3+1} e^{0x} + \frac{1}{D^3+1} e^{2x} + \frac{1}{(D+1)(D^2-D+1)} e^{-x}$$

$$= 1 + \frac{e^{2x}}{9} + \frac{1}{3(D+1)} e^{-x}, \quad \text{using } \frac{1}{f(D)} e^{ax} = \frac{1}{f(a)} e^{ax},$$

for $\frac{1}{D^2-D+1} e^{-x}$ with $D = -1$ when $f(-1) \neq 0$

$$= 1 + \frac{e^{2x}}{9} + \frac{1}{3} x e^{-x}, \text{ by using (4) of (10.7).}$$

Hence the general solution of the equation is

$$y = y_c + y_p$$

$$= c_1 e^{-x} + e^{x/2}\left(c_2 \sin \frac{\sqrt{3}}{2} x + c_3 \cos \frac{\sqrt{3}}{2} x\right) + 1 + \frac{e^{2x}}{9} + \frac{1}{3} x e^{-x}.$$

(10.9) To find the P.I. when $F(x) = \sin ax$ or $\cos ax$.

Here we have to evaluate

$$\frac{1}{f(D)} \sin ax \quad \text{and} \quad \frac{1}{f(D)} \cos ax.$$

From Euler's Theorem, we have

$$e^{iax} = \cos ax + i \sin ax$$

Thus $\frac{1}{f(D)} \sin ax$ and $\frac{1}{f(D)} \cos ax$ are respectively imaginary and real parts of $\frac{1}{f(D)} e^{iax}$.

If $f(D)$ contains only even powers of D, say $f(D) = f(D^2)$, it is easy to see that

$$\frac{1}{f(D^2)} \begin{Bmatrix} \sin ax \\ \cos ax \end{Bmatrix} = \frac{1}{f(-a^2)} \begin{Bmatrix} \sin ax \\ \cos ax \end{Bmatrix}, \text{ provided } -a^2 \text{ is not a zero of } f(D^2).$$

Example 9. Solve: $(D^2 - 5D + 6)y = \sin 3x$

Solution. The roots of the characteristic equation

$$D^2 - 5D + 6 = 0 \text{ are 2 and 3.}$$

$$\text{C.F.} = y_c = c_1 e^{2x} + c_2 e^{3x}$$

$$\text{P.I.} = y_p = \frac{1}{D^2 - 5D + 6} \sin 3x$$

which is imaginary part of $\frac{1}{(D-2)(D-3)} e^{3ix}$

Now $\frac{1}{(D-2)(D-3)} e^{3ix} = \frac{e^{3ix}}{(3i-2)(3i-3)}$, by using $\frac{1}{f(D)} e^{ax} = \frac{e^{ax}}{f(a)}$ and $f(a) \neq 0$

$$= \frac{-1+5i}{78} e^{3ix}$$

$$= \left(-\frac{1}{78} + \frac{5}{78} i\right)(\cos 3x + i \sin 3x)$$

Its imaginary part is $-\frac{1}{78} \sin 3x + \frac{5}{78} \cos 3x = y_p$

Hence the required general solution is

$$y = c_1 e^{2x} + c_2 e^{3x} - \frac{1}{78} \sin 3x + \frac{5}{78} \cos 3x.$$

(10.10) Theorem. If a is not a zero of $f(D)$, then

$$\frac{1}{f(D-a)} e^{ax} F(x) = e^{ax} \frac{1}{f(D)} F(x).$$

This replacing of D by $D + a$ is known as **exponential shift**.

Proof. We have already shown in **(10.7)** that

$$(D-a)^k (e^{ax} p(x)) = e^{ax} D^k p(x)$$

Using linearity of differential operators, we conclude that, when $f(D)$ is a polynomial (with constant coefficients), then

$$f(D-a)(e^{ax} p(x)) = e^{ax} f(D) p(x) \tag{1}$$

Suppose $f(D) p(x) = F(x)$;

Then $p(x) = \frac{1}{f(D)} F(x).$ (2)

From (1) and (2), we have

$$f(D-a)\, e^{ax} \frac{1}{f(D)} F(x) = e^{ax} F(x)$$

Operating on both sides by $\dfrac{1}{f(D-a)}$, we get

$$e^{ax} \frac{1}{f(D)} F(x) = \frac{1}{f(D-a)} e^{ax} F(x)$$

or

$$\frac{1}{f(D-a)} e^{ax} F(x) = e^{ax} \frac{1}{f(D)} F(x),$$

provided that a is not a zero of $f(D)$.

Example 10. Solve: $(D^3 + D^2 - 4D - 4)\, y = e^{2x} \cos 3x$.

Solution. The characteristic equation is

$$D^3 + D^2 - 4D - 4 = 0.$$

By inspection $D = 2$ is a root of this equation. The other roots are $D = -2, -1$

$$\text{C.F.} = y_c = c_1 e^{-x} + c_2 e^{-2x} + c_3 e^{2x}$$

$$\text{P.I.} = y_p = \frac{1}{(D-2)(D+2)(D+1)} e^{2x} \cos 3x$$

$$= e^{2x} \frac{1}{D(D+4)(D+3)} \cos 3x, \text{ by the exponential shift}$$

$$= e^{2x} \frac{1}{(D^3 + 7D^2 + 12D)} \cos 3x$$

$$= e^{2x} \frac{1}{-9D - 63 + 12D} \cos 3x, \quad \text{putting } D^2 = -3^2$$

$$= e^{2x} \frac{1}{3(D-21)} \cos 3x$$

$$= e^{2x} \frac{(D+21)(\cos 3x)}{3(D^2 - 441)}$$

$$= e^{2x} \frac{(D+21)}{3(-9-441)} \cos 3x$$

$$= e^{2x} \frac{-1}{3 \times 450} [-3 \sin 3x + 21 \cos 3x]$$

$$= \frac{e^{2x}}{450} (\sin 3x - 7 \cos 3x).$$

Alternative Method. Here, for $\dfrac{1}{D(D+4)(D+3)} \cos 3x$, we proceed as follow

$$\frac{\cos 3x}{(D^3 + 7D^2 + 12D)} = \text{Re} \frac{1}{(D^3 + 7D^2 + 12D)} e^{3ix}, \text{ by (10.9)}$$

$$= \text{Re} \frac{e^{3ix}}{(-27i - 63 + 36i)}, \text{ by (2) of (10.7)}$$

$$= \text{Re} \frac{e^{3ix}}{9(i-7)}$$

$$= \text{Re} \frac{(i+7)\, e^{3ix}}{-450}$$

$$= \text{Re} \frac{(i+7)(\cos 3x + i \sin 3x)}{-450}$$

$$= -\frac{7}{450} \cos 3x + \frac{1}{450} \sin 3x$$

Therefore,

$$y_p = \frac{e^{2x}}{450} (\sin 3x - 7 \cos 3x) \quad \text{as before}$$

The general solution is

$$y = y_c + y_p$$

$$= c_1 e^{-x} + c_2 e^{-2x} + c_3 e^{2x} + \frac{e^{2x}}{450} (\sin 3x - 7 \cos 3x).$$

(10.11) Series Expansion of Polynomials or Expressions involving Operators.

This method is useful when $F(x)$ is a polynomial in x.

Let $f(D)\, y = F(x)$ be such that $F(x)$ is a polynomial in x. To evaluate the particular integral

$$y_p = \frac{1}{f(D)} F(x)$$

It is often usef[illegible] $\frac{1}{f(D)}$ in a series in D by the Binomial Theorem for negative exponen[illegible]

$$[illegible] \cdots) F(x)$$

The derivatives on the right will vanish after certain stage since

$$D^n x^r = 0 \quad \text{if} \quad n > r.$$

The following binomial expansions will be useful in this connection:

(i) $\dfrac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots$

(ii) $\dfrac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots$

Example 11. Solve $(D^3 - 2D + 1)\,y = 2x^3 - 3x^2 + 4x + 5 \quad (= p(x))$

Solution. The characteristic equation is

$$D^3 - 2D + 1 = 0$$

or

$$(D-1)(D^2 + D - 1) = 0$$

$$D = 1, \ \frac{-1+\sqrt{5}}{2}, \ \frac{-1-\sqrt{5}}{2}$$

$$y_c = c_1 e^x + c_2 e^{\frac{-1+\sqrt{5}}{2}x} + c_3 e^{\frac{-1-\sqrt{5}}{2}x}$$

Next,

$$y_p = \frac{1}{1 - 2D + D^3}\,(2x^3 - 3x^2 + 4x + 5)$$

Now

$$\frac{1}{1-2D+D^3} = [1 - (2D - D^3)]^{-1}$$

$$= 1 + (2D - D^3) + (2D - D^3)^2 + (2D - D^3)^3 + \cdots$$

$$= 1 + 2D - D^3 + 4D^2 + 8D^3,$$

neglecting D^4 and higher powers of D in view of the degree 3 of the polynomial $p(x)$ in x. Therefore

$$y_p = (1 + 2D + 4D^2 + 7D^3)\,(2x^3 - 3x^2 + 4x + 5)$$

$$= 2x^3 - 3x^2 + 4x + 5 + 2\,(6x^2 - 6x + 4) + 4(12x - 6) + 7(12)$$

$$= 2x^3 + 9x^2 + 40x + 73.$$

The general solution is

$$y = y_c + y_p$$

$$= c_1 e^x + c_2 e^{\frac{-1+\sqrt{5}}{2}x} + c_3 e^{\frac{-1-\sqrt{5}}{2}x} + 2x^3 + 9x^2 + 40x + 73.$$

EXERCISE 10.2

Find the general solution of each of the following (Problems 1-15)

1. $(D^2 + 3D - 4)\,y = 15\,e^x$
2. $(D^2 - 3D + 2)\,y = e^x + e^{2x}$
3. $(D^2 - 2D - 3)\,y = 2e^x - 10\sin x$
4. $(D^3 - 2D^2 + D)\,y = x^2 + 3x + 1$
5. $(D^3 - D^2 + D - 1)\,y = 4\sin x$
6. $(D^3 - 2D^2 - 3D + 10)\,y = 40\cos x$
7. $(D^2 + 4)\,y = 4\sin^2 x$ = 2(1 - cos 2x)
8. $(D^3 + D)\,y = 2x^2 + 4\sin x$
9. $(D^4 + D^2)\,y = 3x^2 + 6\sin x - 2\cos x$
10. $(D^2 - 2D + 4)\,y = e^x\cos x$
11. $(D^3 - D^2 + 3D + 5)\,y = e^x \sin 2x$
12. $(D^3 - 7D - 6)\,y = e^{2x}(1 + x)$
13. $(D^2 - 7D + 12)\,y = e^{2x}(x^3 - 5x^2)$
14. $(D^4 + 8D^2 - 9)\,y = 9x^3 + 5\cos 2x$
15. $(D^3 + 3D^2 - 4)\,y = \sinh x - \cos^2 x$

Solve the initial value problems:

16. $y'' - 8y' + 15y = 9x\,e^{2x}$, $\quad y(0) = 5,\ y'(0) = 10$
17. $y'' - 4y' + 13y = 8\sin 3x$, $\quad y(0) = 1,\ y'(0) = 2$
18. $y'' - 4y = 2 - 8x$, $\quad y(0) = 0,\ y'(0) = 5$
19. $y'' + y = x\sin x$, $\quad y(0) = 1,\ y'(0) = 2$
20. $y''' + 3y'' + 7y' + 5y = 16\,e^{-x}\cos 2x$, $\quad y(0) = 2,\ y'(0) = -4,\ y''(0) = -2$

THE METHOD OF UNDETERMINED COEFFICIENTS (U.C.)

(10.12) We have studied some special cases in which particular integral can be evaluated by the inverse operator. Now we consider the method of undetermined coefficients which can prove simpler in finding the particular integral of the equation $f(D)\,y = F(x)$ when $F(x)$ is

(i) an exponential function (e^{ax})

(ii) a polynomial $(b_0 x^n + b_1 x^{n-1} + \cdots + b_n)$

(iii) sinusoidal function ($\sin ax$ or $\cos ax$)

(iv) the more general case in which $F(x)$ is sum of a product of terms of the above types, such as

$$F(x) = e^{ax}\,(b_0 x^n + b_1 x^{n-1} + \cdots + b_n)\begin{Bmatrix}\sin ax \\ \cos ax\end{Bmatrix}$$

The P.I. y_p will be constructed according to the following table:

$F(x)$ is of the form	Take y_p as
1. a	$A x^k$
2. $a x^n$ (n is a +ive integer)	$x^k (A_0 x^n + A_1 x^{n-1} + \cdots + A_{n-1} x + A_n)$
3. $a x^n e^{rx}$ (n is a +ive integer)	$x^k (A_0 x^n + A_1 x^{n-1} + \cdots + A_{n-1} x + A_n) e^{rx}$
4. $C x^n \cos ax$ 5. $C x^n \sin ax$	$x^k (A \cos \alpha x + B \sin \alpha x)$
6. $C x^n e^{rx} \cos \alpha x$ 7. $C x^n e^{rx} \sin \alpha x$	$x^k [(A_0 x^n + A_1 x^{n-1} + \cdots + A_{n-1} x + A_n) e^{rx} \cos \alpha x$ $+ (B_0 x^n + B_1 x^{n-1} + \cdots + B_{n-1} x + B_n) e^{rx} \sin \alpha x)]$

In x^k, k is the smallest nonnegative integer which will ensure that no term in y_p is already in the C.F.

If $F(x)$ is sum of several terms, write y_p for each term individually and then add up all of them.

The y_p and its derivatives will be substituted into the equation $f(D)\, y = F(x)$ and coefficients of like terms on the left hand and right hand sides will be equated to determine the U.C. $A_0, A_1, \cdots, A_n, B_0, B_1, \cdots, B_n$.

The method is illustrated by examples as follows.

Example 12. Solve: $y'' - 3y' + 2y = 2x^3 - 9x^2 + 6x$ (1)

Solution. C.F. is easily found as

$$y_c = c_1 e^x + c_2 e^{2x}$$

For a particular solution, we assume

$$y_p = x^k (Ax^3 + Bx^2 + Cx + D).$$

Since no term of the C.F. is present in y_p, we take $k = 0$ and so

$$y_p = Ax^3 + Bx^2 + Cx + D$$
$$y_p' = 3Ax^2 + 2Bx + C$$
$$y_p'' = 6Ax + 2B.$$

Substituting for y_p, y_p' and y_p'' into (1), we have

$$(6Ax + 2B) - 3(3Ax^2 + 2Bx + C) + 2(Ax^3 + Bx^2 + Cx + D) = 2x^3 - 9x^2 + 6x$$

or
$$2Ax^3 + x^2(2B - 9A) + x(6A - 6B + 2C) + 2B - 3C + 2D = 2x^3 - 9x^2 + 6x.$$

Equating coefficients of like terms, we obtain

Coeff. of x^3:	$2A = 2$	or	$A = 1$
Coeff. of x^2:	$2B - 9A = -9$	or	$2B = 9A - 9$ giving $B = 0$
Coeff. of x:	$6A - 6B + 2C = 6$	or	$C = 0$
Coeff. of x^0:	$2B - 3C + 2D = 0$	or	$D = 0$

So $y_p = x^3$

The required general solution is

$$y = c_1 e^x + c_2 e^{2x} + x^3.$$

Example 13. Solve: $y'' - 3y' + 2y = x^2 e^x$. (1)

Solution. The C.F. is

$$y_c = c_1 e^x + c_2 e^{2x}$$

To construct a particular solution, we use (3) of the table in (10.12).

$$y_p = x^k (Ax^2 + Bx + C) e^x$$

Since e^x is already in C.F., we take $k = 1$ so that no term of C.F. is in y_p. Hence the modified P.I. is

$$y_p = (Ax^3 + Bx^2 + Cx) e^x$$
$$y_p' = (Ax^3 + Bx^2 + Cx) e^x + (3Ax^2 + 2Bx + C) e^x$$
$$y_p'' = (Ax^3 + Bx^2 + Cx) e^x + 2(3Ax^2 + 2Bx + C) e^x + (6Ax + 2B) e^x$$

Substituting for y_p, y_p' and y_p'' into (1), we have

$$(Ax^3 + Bx^2 + Cx) e^x + 2(3Ax^2 + 2Bx + C) e^x + (6Ax + 2B) e^x$$
$$- 3(Ax^3 + Bx^2 + Cx) e^x - 3(3Ax^2 + 2Bx + C) e^x$$
$$+ 2(Ax^3 + Bx^2 + Cx) e^x = x^3 e^x$$

or
$$[-3Ax^2 + (6A - 2B) x + C] e^x = x^3 e^x.$$

Equating coefficients of like terms, we obtain

Coeff. of x^3: $-3A = 1$ or $A = -\frac{1}{3}$

Coeff. of x: $6A - 2B = 0$ or $B = -1$

Coeff. of x^0: $C = 0$

so
$$y_p = -\frac{1}{3}\, x^3 e^x - x^2 e^x$$

The required general solution is

$$y = c_1 e^x + c_2 e^{2x} - \frac{1}{3}\, x^3 e^x - x^2 e^x$$

Example 14. Solve $y'' + 4y = xe^x + x \sin 2x$ (1)

Solution. The C.F., as readily found above, is

$$y_c = c_1 \sin 2x + c_2 \cos 2x.$$

For the P.I. of (1), we find the P.I. of

$$y'' + 4y = x e^x \qquad (2)$$

and
$$y'' + 4y = x \sin 2x \qquad (3)$$

separately. Their sum will be the P.I. of (1) (by the **Principle of Superposition**).

For a particular solution of (2), we have

$$y_p = x^k (Ax + B)\, e^x$$

Since no term of C.F. is in the y_p, we must take $k = 0$. Therefore,

$$y_p = (Ax + B)\, e^x$$
$$y_p' = (Ax + B)\, e^x + Ae^x$$
$$y_p'' = (Ax + B)\, e^x + 2Ae^x.$$

Substituting for y_p, y_p' and y_p'' into (2), we have

$$(Ax + B)\, e^x + 2Ae^x + 4(Ax + B)\, e^x = xe^x$$

or
$$5Ax\, e^x + (2A + 5B)\, e^x = xe^x$$

Equating coefficients of like terms, we obtain

Coeff. of $x e^x$: $5A = 1$ or $A = \frac{1}{5}$

Coeff. of e^x: $2A + 5B = 0$ or $B = -\frac{2}{25}$

So P.I. of (2) is
$$y_p = \frac{1}{5}\, xe^x - \frac{2}{25}\, e^x. \qquad (4)$$

For a particular solution of (3), we assume

$$y_p = x^k \left[(Cx + D) \sin 2x + (Ex + F) \cos 2x\right]$$

Since $\sin 2x$ and $\cos 2x$ are already in the C.F., we must take $k = 1$, so that the P.I. is

$$y_p = (Cx^2 + Dx) \sin 2x + (Ex^2 + Fx) \cos 2x$$
$$y_p' = (Cx^2 + Dx)(2 \cos 2x) + (2Cx + D) \sin 2x$$
$$+ (Ex^2 + Fx)(-2 \sin 2x) + (2Ex + F) \cos 2x$$
$$y_p'' = (Cx^2 + Dx)(-4 \sin 2x) + (2Cx + D)(2 \cos 2x)$$
$$+ (2Cx + D)(2 \cos 2x) + 2C \sin 2x + (Ex^2 + Fx)(-4 \cos 2x)$$
$$+ (2Ex + F)(-2 \sin 2x) + (2Ex + F)(-2 \sin 2x) + 2E \cos 2x$$

Substituting into (3), we have

$$(Cx^2 + Dx)(-4 \sin 2x) + 4(2Cx + D) \cos 2x + 2C \sin 2x$$
$$+ (Ex + Fx)(-4 \cos 2x) - 4(2Ex + F) \sin 2x + 2E \cos 2x$$
$$+ 4(Cx^2 + Dx) \sin 2x + 4(Ex^2 + Fx) \cos 2x = x \sin 2x$$

Equating the coefficients of like terms, we get

Coeff. of $x \sin 2x$: $-8E = 1$ or $E = -\frac{1}{8}$

Coeff. of $x \cos 2x$: $8C = 0$ or $C = 0$

Coeff. of $\sin 2x$: $2C - 4F = 0$ or $F = 0$

Coeff. of $\cos 2x$: $2E + 4D = 0$ or $D = \frac{1}{16}$

Hence P.I. of (3) is

$$y_p = \frac{1}{16}\, x \sin 2x - \frac{1}{8}\, x^2 \cos 2x \qquad (5)$$

Sum of (4) and (5) is the required P.I. of (1), i.e.,

$$y_p = \frac{1}{5}\, xe^x - \frac{2}{25}\, e^x + \frac{1}{16}\, x \sin 2x - \frac{1}{8}\, x^2 \cos 2x$$

is the particular solution of (1)

Hence the general solution of (1) is

$$y = c_1 \sin 2x + c_2 \cos 2x + \frac{1}{5}\, xe^x - \frac{2}{25}\, e^x + \frac{1}{16}\, x \sin 2x - \frac{1}{8}\, x^2 \cos 2x$$

EXERCISE 10.3

Solve by the method of U C (Problems 1–9)

1. $y'' - 4y' + 4y = e^{2x}$
2. $y'' + 2y' + 5y = 6 \sin 2x + 7 \cos 2x$
3. $2y'' + 3y' + y = x^2 + 3 \sin x$
4. $y'' + 2y' + y = e^x \cos x$
5. $y'' + y = 12 \cos^2 x$
6. $y'' - 3y' + 2y = 2x^2 + 2x\,e^x$
7. $y''' + y' = 2x^2 + 4 \sin x$
8. $y''' + y'' + 3y' - 5y = 5 \sin 2x + 10x^2 + 3x + 7$
9. $y^{(4)} + 8y'' + 16y = \sin x$
10. Write the general form of the P I (without evaluating the U C.) for
    (i) $y'' + 2y' + 2y = 4e^{-x} x^2 \sin x + 3e^{-x} + 2e^{-x} \cos x$
    (ii) $y'' + 3y' + 2y = e^x (x^2 + 1) \sin 2x + 4e^x - 3e^{-x} \cos x$

THE CAUCHY-EULER[1] EQUATION

(10.13) An equation of the form

$$a_0 x^n \frac{d^n y}{dx^n} + a_1 x^{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_{n-1} x \frac{dy}{dx} + a_n y = F(x) \quad (1)$$

is called **Cauchy-Euler** (or **equidimensional**) equation $a_0, a_1, \cdots, a_{n-1}, a_n$ are real constants

The equation can be reduced to a linear differential equation with constant coefficients by the transformation

$$x = e^t \quad \text{or} \quad t = \ln x$$

Then

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

$$\frac{d^2 y}{dx^2} = \frac{d}{dx}\left(\frac{1}{x}\,\frac{dy}{dt}\right) = \frac{1}{x}\,\frac{d}{dx}\left(\frac{dy}{dt}\right) + \frac{dy}{dt}\,\frac{d}{dx}\left(\frac{1}{x}\right)$$

$$= \frac{1}{x}\,\frac{d}{dt}\left(\frac{dy}{dt}\right)\frac{dt}{dx} + \frac{dy}{dt}\left(-\frac{1}{x^2}\right) = \frac{1}{x^2}\left(\frac{d^2 y}{dt^2} - \frac{dy}{dt}\right)$$

1. Leonhard Euler (1701 – 1783). A Swiss mathematician.

$$\frac{d^3 y}{dx^3} = \frac{d}{dx}\left[\frac{1}{x^2}\frac{d^2 y}{dt^2} - \frac{1}{x^2}\,\frac{dy}{dt}\right]$$

$$= \frac{1}{x^2}\,\frac{d}{dx}\left(\frac{d^2 y}{dt^2}\right) + \frac{d^2 y}{dt^2}\,\frac{d}{dx}\left(\frac{1}{x^2}\right) - \frac{1}{x^2}\,\frac{d}{dx}\left(\frac{dy}{dt}\right) + \frac{dy}{dt}\,\frac{d}{dx}\left(-\frac{1}{x^2}\right)$$

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

$$= \frac{1}{x^3}\,\frac{d^3 y}{dt^3} - \frac{3}{x^3}\,\frac{d^2 y}{dt^2} + \frac{2}{x^3}\,\frac{dy}{dt}$$

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

If we write $D = \dfrac{d}{dx}$ and $\Delta = \dfrac{d}{dt}$, then

$$xD = \Delta$$

$$x^2 D^2 = \Delta^2 - \Delta = \Delta(\Delta - 1)$$

$$x^3 D^3 = \Delta^3 - 3\Delta^2 + 2\Delta = \Delta(\Delta - 1)(\Delta - 2)$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$x^n D^n = \Delta(\Delta - 1)(\Delta - 2) \cdots (\Delta - n + 1)$$

Substituting these values of $xD, x^2D^2, \cdots$ into (1), we obtain an equation of nth order with constant coefficients having t as the independent variable. This equation can be solved by the previous methods.

Example 15. Solve $x^2 \dfrac{d^2 y}{dx^2} - 2x\dfrac{dy}{dx} + 2y = x^3$. (1)

Solution. Let $x = e^t$. Then we have

$$t = \ln x$$

$$xD = \Delta$$

$$x^2 D^2 = \Delta(\Delta - 1)$$

Substituting into (1), we get

$$[\Delta(\Delta - 1) - 2\Delta + 2]\,y = e^{3t}$$

or $$[\Delta^2 - 3\Delta + 2]\,y = e^{3t}. \quad (2)$$

The characteristic equation has roots 2 and 1. Therefore, C.F. of (2) is

$$y_c = c_1 e^t + c_2 e^{2t}$$

P.I. of equation (2) is

$$y_p = \frac{1}{(\Delta - 1)(\Delta - 2)} e^{3t} = \frac{1}{2} e^{3t}$$

The general solution of (2) is

$$y = c_1 e^t + c_2 e^{2t} + \frac{1}{2} e^{3t}$$

Replacing t by $\ln x$ (or e^t by x), we have

$$y = c_1 x + c_2 x^2 + \frac{1}{2} x^3$$

as the general solution of (1).

Example 16. Solve:

$$x^4 \frac{d^4y}{dx^4} + 6x^3 \frac{d^3y}{dx^3} + 9x^2 \frac{d^2y}{dx^2} + 3x \frac{dy}{dx} + y = (1 + \ln x)^2$$

Solution. We let $x = e^t$ so that $t = \ln x$. By (10.13), we have

$$xD = \Delta$$
$$x^2D^2 = \Delta(\Delta - 1) = \Delta^2 - \Delta$$
$$x^3D^3 = \Delta(\Delta - 1)(\Delta - 2) = \Delta^3 - 3\Delta^2 + 2\Delta$$
$$x^4D^4 = \Delta(\Delta - 1)(\Delta - 2)(\Delta - 3) = \Delta^4 - 6\Delta^3 + 11\Delta^2 - 6\Delta$$

Substituting into (1), we get

$$[\Delta^4 - 6\Delta^3 + 11\Delta^2 - 6\Delta + 6(\Delta^3 - 3\Delta^2 + 2\Delta) + 9(\Delta^2 - \Delta) + 3\Delta + 1]\, y = (1 + t)^2$$

or

$$(\Delta^4 + 2\Delta^2 + 1)\, y = 1 + 2t + t^2 \qquad (2)$$

The roots of the characteristic equation of (2) are $\pm i, \pm i$.

Therefore, C.F. of (2) is

$$y_c = (c_1 + c_2 t)\sin t + (c_3 + c_4 t)\cos t$$

To find a particular integral of (2), we use the method of U.C. Let the particular integral of (2) be

$$y_p = At^2 + Bt + C$$
$$y_p' = 2At + B$$
$$y_p'' = 2A$$

Substituting into (2), we obtain

$$4A + At^2 + Bt + C = 1 + 2t + t^2$$
$$At^2 + Bt + (4A + C) = t^2 + 2t + 1$$

Equating coefficients of like terms, we have

$$A = 1, \quad B = 2, \quad 4A + C = 1 \quad \text{or} \quad C = -3.$$

Therefore, a particular integral of (2) is

$$y_p = t^2 + 2t - 3$$

Hence the general solution of (2) is

$$y = (c_1 + c_2 t)\sin t + (c^3 + c_4 t)\cos t + t^2 + 2t - 3 \qquad (3)$$

Replacing t by $\ln x$ in (3), we get

$$y = (c_1 + c_2 \ln x)\sin(\ln x) + (c_3 + c_4 \ln x)\cos(\ln x) + (\ln x)^2 + 2\ln x - 3$$

as the general solution of (1).

EXERCISE 10.4

Solve

1. $x^2 \dfrac{d^2y}{dx^2} + 7x \dfrac{dy}{dx} + 5y = x^5$

2. $x^2 \dfrac{d^2y}{dx^2} - 3x \dfrac{dy}{dx} + 5y = x^2 \sin(\ln x)$

3. $x^2 \dfrac{d^2y}{dx^2} - (2m - 1)x \dfrac{dy}{dx} + (m^2 + n^2)y = n^2 x^m \ln x$

4. $4x^2 \dfrac{d^2y}{dx^2} - 4x \dfrac{dy}{dx} + 3y = \sin \ln(-x), \quad x < 0$

5. $x^3 \dfrac{d^3y}{dx^3} + 2x^2 \dfrac{d^2y}{dx^2} + 2y = 10x + \dfrac{10}{x}$

6. $x^4 \dfrac{d^3y}{dx^3} + 2x^3 \dfrac{d^2y}{dx^2} - x^2 \dfrac{dy}{dx} + xy = 1$

7. $x^3 \dfrac{d^3y}{dx^3} + 4x^2 \dfrac{d^2y}{dx^2} - 5x \dfrac{dy}{dx} - 15y = x^4$

8. $(x+1)^2\frac{d^2y}{dx^2}+(x+1)\frac{dy}{dx}+y = 4[\cos\ln(x+1)]^2$

9. $(2x+1)^2\frac{d^2y}{dx^2}-6(2x+1)\frac{dy}{dx}+16y = 8(2x+1)^2$

10. $x^2y''+2xy'-6y = 10x^2$, $\quad y(1)=1\ ,\ y'(1)=-6$

11. $x^2y''-2xy'+2y = x\ln x$, $\quad y(1)=1\ ,\ y'(1)=0$

12. $x^3y'''+2x^2y''+xy'-y = 15\cos(2\ln x)$, $\quad y(1)=2\ ,\ y'(1)=-3,\ y''(1)=0$

REDUCTION OF ORDER

(10.14) If one solution of the second order linear equation

$$\frac{d^2y}{dx^2}+P\frac{dy}{dx}+Qy = 0 \tag{1}$$

(where P, Q are not necessarily constants and may be functions of x) is known, then we can use it to find the general solution of

$$\frac{d^2y}{dx^2}+P\frac{dy}{dx}+Qy = F(x) \tag{2}$$

The procedure, due to D' Alembert, is known as the method of reduction of order.

Suppose it is known that $y=y_1$ is a solution of (1). We assume that

$$y = vy_1 \tag{3}$$

is a solution of (2), where v is a function of x to be determined. From (3), we get

$$\frac{dy}{dx} = v\frac{dy_1}{dx}+y_1\frac{dv}{dx},$$

$$\frac{d^2y}{dx^2} = v\frac{d^2y_1}{dx^2}+2\frac{dv}{dx}\frac{dy_1}{dx}+y_1\frac{d^2v}{dx^2}$$

Substituting into (2), we have

$$v\frac{d^2y_1}{dx^2}+2\frac{dv}{dx}\frac{dy_1}{dx}+y_1\frac{d^2v}{dx^2}+Pv\frac{dy_1}{dx}+Py_1\frac{dv}{dx}+Qvy_1 = F(x)$$

or
$$y_1\frac{d^2v}{dx^2}+\left(2\frac{dy_1}{dx}+Py_1\right)\frac{dv}{dx}+\left(\frac{d^2y_1}{dx^2}+P\frac{dy_1}{dx}+Qy_1\right)v = F(x) \tag{4}$$

Since $y=y_1$ is a solution of (1)

$$\frac{d^2y_1}{dx^2}+P\frac{dy_1}{dx}+Qy_1 = 0$$

and, therefore, equation (4) reduces to

$$y_1\frac{d^2v}{dx^2}+\left(2\frac{dy_1}{dx}+Py_1\right)\frac{dv}{dx} = F(x) \tag{5}$$

Setting $\frac{dv}{dx} = u$ in (5), we obtain

$$y_1\frac{du}{dx}+\left(2\frac{dy_1}{dx}+Py_1\right)u = F(x)$$

which is a linear equation of the first order in u and can be solved for u. From $\frac{dv}{dx} = u$, we determine v and hence the solution $y = vy_1$. The method is illustrated by means of examples.

Example 17. Solve: $\frac{d^2y}{dx^2}+y = \csc x$ (1)

Solution. The C.F. of (1) is

$$y_c = c_1\sin x + c_2\cos x.$$

We may take any special value of y_c as y_1. Let us take y_1 to be the value of y_c when $c_1=1$ and $c_2=0$. Then assume that

$$y = v\sin x \tag{2}$$

is a solution of (1). From (2), we get

$$\frac{dy}{dx} = \sin x\frac{dv}{dx}+v\cos x$$

$$\frac{d^2y}{dx^2} = \sin x\frac{d^2v}{dx^2}+2\cos x\frac{dv}{dx}-v\sin x.$$

Substituting into (1), we obtain

$$\sin x\frac{d^2v}{dx^2}+2\cos x\frac{dv}{dx}-v\sin x+v\sin x = \csc x$$

or
$$\frac{d^2v}{dx^2}+2\cot x\frac{dv}{dx} = \csc^2 x.$$

Setting $u = \dfrac{dv}{dx}$, the above equation becomes

$$\frac{du}{dx} + 2\cot x\, u = \csc^2 x \tag{3}$$

which is linear equation of first order. An integrating factor of (3) is

$$\exp\left[\int 2\cot x\, dx\right] = \exp\left[\ln \sin^2 x\right] = \sin^2 x$$

Multiplying (3) by $\sin^2 x$ and integrating, we have

$$u \sin^2 x = x,$$

neglecting constant of integration since we seek only a particular solution

or $$u = x\csc^2 x$$

or $$\frac{dv}{dx} = x\csc^2 x$$

or $$v = \int x\csc^2 x\, dx = -x\cot x + \ln|\sin x|. \qquad \text{(on integrating by parts)}$$

Hence a particular solution of (1) is

$$y_p = v\sin x = -x\cos x + \sin x\left(\ln|\sin x|\right)$$

The general solution of (1) is

$$y = c_1 \sin x + c_2 \cos x - x\cos x + \sin x\left(\ln|\sin x|\right)$$

Example 18. Solve $(x+2)\dfrac{d^2y}{dx^2} - (2x+5)\dfrac{dy}{dx} + 2y = (x+1)e^x$ (1)

Solution. We note that $y = e^{2x}$ makes the left hand side of (1) zero. Therefore, we put

$$y = ve^{2x}$$

$$\frac{dy}{dx} = \left(\frac{dv}{dx} + 2v\right)e^{2x}$$

$$\frac{d^2y}{dx^2} = \left(\frac{d^2v}{dx^2} + 4\frac{dv}{dx} + 4v\right)e^{2x}$$

Substituting into (1), we obtain

$$(x+2)\left(\frac{d^2v}{dx^2} + 4\frac{dv}{dx} + 4v\right)e^{2x} - (2x+5)\left(\frac{dv}{dx} + 2v\right)e^{2x} + 2ve^{2x} = (x+1)e^x$$

or $$(x+2)\frac{d^2v}{dx^2}e^{2x} + \left[4(x+2) - (2x+5)\right]\frac{dv}{dx}e^{2x} + \left[4(x+2) - 2(2x+5) + 2\right]ve^{2x} = (x+1)e^x$$

or $$(x+2)\frac{d^2v}{dx^2} + (2x+3)\frac{dv}{dx} = (x+1)e^{-x}$$

Writing $u = \dfrac{dv}{dx}$, the above equation becomes

$$(x+2)\frac{du}{dx} + (2x+3)u = (x+1)e^{-x}$$

or $$\frac{du}{dx} + \frac{2x+3}{x+2}u = \frac{x+1}{x+2}e^{-x} \tag{2}$$

which is linear equation of first order. An integrating factor of (2) is

$$\exp\left[\int \frac{2x+3}{x+2}dx\right] = \exp\left[\int\left(2 - \frac{1}{x+2}\right)dx\right]$$

$$= \exp\left[2x - \ln(x+2)\right] = \frac{e^{2x}}{x+2}$$

Multiplying (2) by $\dfrac{e^{2x}}{x+2}$, we have

$$\frac{d}{dx}\left(u\frac{e^{2x}}{x+2}\right) = \frac{x+1}{(x+2)^2}\,e^x$$

or $$u\frac{e^{2x}}{x+2} = \int \frac{x+1}{(x+2)^2}\,e^x\,dx + c_1$$

$$= \int \frac{e^x}{x+2}\,dx - \int \frac{e^x}{(x+2)^2}\,dx + c_1$$

$$= \frac{e^x}{x+2} + c_1$$

or $$u = e^{-x} + c_1(x+2)e^{-2x}$$

Therefore, $$\frac{dv}{dx} = e^{-x} + c_1(x+2)e^{-2x}$$

Integrating, we get

$$v = -e^{-x} - \frac{1}{4}c_1(2x+5)e^{-2x} + c_2$$

Hence $$y = ve^{2x} = -e^x - \frac{1}{4}c_1(2x+5) + c_2e^{2x}$$

is the general solution of (1).

Note. It is easy to see that $y'' + Py' + Qy = 0$ is satisfied by $y = e^{\pm x}$ if $1 \pm P + Q = 0$ and by $y = x$ if $P + Qx = 0$. These can be used to find a solution of $y'' + Py' + Q = 0$ by inspection.

EXERCISE 10.5

Solve:

1. $\dfrac{d^2y}{dx^2} + y = \sec^3 x$

2. $\dfrac{d^2y}{dx^2} + 4y = 4\tan 2x$

3. $(1 - x^2)\dfrac{d^2y}{dx^2} - 2x\dfrac{dy}{dx} + 2y = 0$ (**Legendre's**[1] equation of order one)

4. $(x - 1)\dfrac{d^2y}{dx^2} - x\dfrac{dy}{dx} + y = 1$

5. $x^2\dfrac{d^2y}{dx^2} + x\dfrac{dy}{dx} - y = 8x^3$

6. $x^2\dfrac{d^2y}{dx^2} - (x^2 + 2x)\dfrac{dy}{dx} + (x + 2)\,y = x^3e^x$

7. $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} + y = \dfrac{1}{(1 + e^x)^2}$

8. $x\dfrac{d^2y}{dx^2} - (2x + 1)\dfrac{dy}{dx} + (x + 1)\,y = (x^2 + x - 1)\,e^x$

9. $\dfrac{d^2y}{dx^2} - 4\dfrac{dy}{dx} + 4y = (1 + x + x^2 + \dots + x^{25})\,e^{2x}$, given that $y = e^{2x}$ is a solution of the associated homogeneous equation

10. $\dfrac{d^2y}{dx^2} - 2\tan x\dfrac{dy}{dx} + 3y = 2\sec x$, given that $y = \sin x$ is a solution of the associated homogeneous equation

11. $x\dfrac{d^2y}{dx^2} - 2(x + 1)\dfrac{dy}{dx} + (x + 2)\,y = (x - 1)\,e^{2x}$

12. $(1 - x^2)\dfrac{d^2y}{dx^2} + x\dfrac{dy}{dx} - y = x(1 - x^2)^{3/2}$.

1 Legendre, Adrien Marie (1752 – 1833). A French mathematician

THE WRONSKIAN[1]

(10.15) Definition. If y_1, y_2 are two differentiable functions of x on $I = [\,a, b\,]$ then their **Wronskian,** denoted by $W = W[\,y_1, y_2\,]$, is defined by

$$W = W[\,y_1, y_2\,] = y_1y'_2 - y'_1y_2 = \begin{vmatrix} y_1 & y_2 \\ y'_1 & y'_2 \end{vmatrix}$$

Similarly, the Wronskian of three differentiable functions y_1, y_2, y_3 on $I = [\,a, b\,]$ is defined by

$$W = W[\,y_1, y_2, y_3\,] = \begin{vmatrix} y_1 & y_2 & y_3 \\ y'_1 & y'_2 & y'_3 \\ y''_1 & y''_2 & y''_3 \end{vmatrix}.$$

The definition can be extended in a similar manner for the Wronskian of n differentiable functions on $I = [\,a, b\,]$

(10.16) Theorem. Let y_1, y_2 be two solutions of the homogeneous linear equation

$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + Qy = 0, \qquad (1)$$

where P, Q are functions of x and are continuous on $I = [\,a, b\,]$. Then their Wronskian $W = W[\,y_1, y_2\,]$ is either identically zero or is never zero on I.

Proof. We have

$$W = y_1y'_2 - y'_1y_2$$

Differentiating w.r.t. x, we get

$$W' = y_1y''_2 + y_1'y_2' - y_1'y_2' - y''_1y_2$$
$$= y_1y''_2 - y''_1y_2$$

Since y_1, y_2 are solutions of (1), we have

$$y''_1 + Py_1' + Qy_1 = 0 \qquad (2)$$

and

$$y_2'' + Py_2' + Qy_2 = 0 \qquad (3)$$

Multiply (2) by y_2 and (3) by y_1 to have an equivalent system

$$y''_1y_2 + Py'_1y_2 + Qy_1y_2 = 0 \qquad (4)$$
$$y_1y''_2 + Py'_2y_1 + Qy_1y_2 = 0 \qquad (5)$$

1 Named after the Polish mathematician Jozef Maria Hoene Wronski (1778 – 1853).

Subtracting (4) from (5), we obtain

$$(y_1 y_2'' - y_2 y_1'') + P(y_1 y_2' - y_1' y_2) = 0$$

or

$$\frac{dW}{dx} + PW = 0$$

The general solution of the above equation is readily found as

$$W = c\exp\left(-\int P\,dx\right)$$

Since the exponential function is never zero, W is identically zero if $c = 0$ or is never zero when $c \neq 0$.

(10.17) Theorem. If y_1, y_2 are two solutions of

$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + Qy = 0,$$

then their Wronskian $W = W[y_1, y_2] = 0$ if and only if y_1, y_2 are linearly dependent.

Proof. If one of the two solutions is identitically zero, then the theorem is obviously true.

Assume that y_1, y_2 are both nonzero and let y_1, y_2 be linearly dependent. Then $y_1 = cy_2$, where c is a constant.

i.e.,
$$\frac{y_1}{y_2} = c$$

or
$$\frac{d}{dx}\left(\frac{y_1}{y_2}\right) = \frac{y_2 y_1' - y_1 y_2'}{y_2^2} = \frac{dc}{dx} = 0$$

i.e.,
$$y_1 y'_2 - y'_1 y_2 = 0$$

$\Rightarrow$
$$W[y_1, y_2] = 0$$

Conversely, if $W[y_1, y_2] = 0$, then $\frac{d}{dx}\left(\frac{y_1}{y_2}\right) = \frac{W}{y_2^2} = 0$.

$\Rightarrow$
$$\frac{y_1}{y_2} = c, \text{ where } c \text{ is a constant}$$

Thus $y_1 = c\,y_2$ and so y_1, y_2 are linearly dependent.

(10.18) Corollary. $W[y_1, y_2] \neq 0$ if and only if y_1, y_2 are linearly independent.

Note: Theorem 10.17 and Corollary 10.18 can only be applied to check the linear dependence or linear independence of differentiable functions which are solutions of a linear homogeneous O.D.E. Arbitrary differentiable functions exist with their Wronskian zero but they are linearly independent.

Consider the functions $y_1 = x^3$, $y_2 = |x^3|$ on $]-\infty, \infty[$. Here $W(y_1, y_2) = 0$ but y_1 and y_2 are linearly independent. (Verify!).

VARIATION OF PARAMETERS

(10.19) We found the solution of the equation

$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + Qy = F(x), \tag{1}$$

where P, Q are functions of x, in the previous section by reduction of order of (1). The solution of (1) can be determined by a procedure known as the **method of variation of parameters**[1]. This method can be applied even to equations of higher order.

Suppose that linearly independent solutions of

$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + Qy = 0 \tag{2}$$

are given by $y = y_1(x)$ and $y = y_2(x)$. Then the complementary function of (1) is

$$y_c = c_1 y_1 + c_2 y_2$$

where c_1 and c_2 are arbitrary constants. We replace the arbitrary constants c_1 and c_2 by unknown functions $u_1(x)$ and $u_2(x)$ and require that

$$y_p = u_1 y_1 + u_2 y_2 \tag{3}$$

be a particular solution of (1). In order to determine the two functions u_1 and u_2 we need two conditions. One condition is that (3) must satisfy (1). A second condition can be imposed arbitrarily.

Differentiating (3) w.r.t. x, we have

$$y_p' = (u_1' y_1 + u_2' y_2) + u_1 y_1' + u_2 y_2'$$

If we differentiate the above equation, y_p'' will contain u_1'' and u_2''. To avoid second derivatives of u_1 and u_2, we put

$$u_1' y_1 + u_2' y_2 = 0. \tag{4}$$

With this condition, we have

$$y_p' = u_1 y_1' + u_2 y_2'$$

so that

$$y_p'' = u_1' y_1' + u_2' y_2' + u_1 y_1'' + u_2 y_2''.$$

1. Discovered by the French mathematician J.L. Lagrange (1736 – 1813)

Substituting for y_p, y_p', y_p'' into equation (1), we get

$$\left(u_1' y_1' + u_2' y_2' + u_1 y_1'' + u_2 y_2''\right) + P\left(u_1 y_1' + u_2 y_2'\right) + Q\left(u_1 y_1 + u_2 y_2\right) = F(x)$$

or $\quad u_1 (y_1'' + P y_1' + Q y_1) + u_2 (y_2'' + P y_2' + Q y_2) + u_1' y_1' + u_2' y_2' = F(x)$

Expressions within the parentheses are zero since y_1 and y_2 are solutions of (2). Hence

$$u_1' y_1' + u_2' y_2' = F(x) \qquad (5)$$

Taking (4) and (5) together, we have two equations in the two unknowns u_1' and u_2':

$$u_1' y_1 + u_2' y = 0$$

and $\quad u_1' y_1' + u_2' y_2' = F(x).$

Solving these, we have

$$\left.\begin{aligned} u_1' &= \frac{-y_2 F(x)}{y_1 y_2' - y_1' y_2} \\ u_2' &= \frac{y_1 F(x)}{y_1 y_2' - y_1' y_2} \end{aligned}\right\} \qquad (6)$$

In (6), the Wronskian of y_1, y_2, namely, $W(y_1, y_2) = y_1 y_2' - y_1' y_2 \neq 0$, since y_1, y_2 are linearly independent solutions of (2).

Integrating (6), we find u_1 and u_2 as

$$\left.\begin{aligned} u_1 &= \int \frac{-y_2 F(x)}{y_1 y_2' - y_1' y_2}\, dx = \int \frac{-y_2 F(x)}{W}\, dx \\ u_2 &= \int \frac{y_1 F(x)}{y_1 y_2' - y_1' y_2}\, dx = \int \frac{y_1 F(x)}{W}\, dx \end{aligned}\right\} \qquad (7)$$

where W is the Wronskian of y_1, y_2.

Thus y_p is completely determined.

In numerical problems, instead of performing the complete process, formulas (7) will be directly applied to evaluate u_1 and u_2.

Example 19. Find the general solution of

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

Solution. The C.F. of (1) is

$$y_c = c_1 e^x + c_2 x e^x$$

Let $\quad y_p = u_1 e^x + u_2 x e^x.$

Here $\quad y_1 = e^x, \quad y_2 = x e^x, \quad F(x) = x e^x \ln x$

$$W = W(y_1, y_2) = y_1 y_2' - y_1' y_2 = e^x (e^x + x e^x) - x e^x \cdot e^x = e^{2x}.$$

By the formulas (7) of (10.19), we have

$$u_1 = \int \frac{-y_2 F(x)}{W}\, dx = \int \frac{-x e^x \cdot x e^x \ln x}{e^{2x}}\, dx$$

and

$$u_2 = \int \frac{y_1 F(x)}{W}\, dx = \int \frac{e^x \cdot x e^x \ln x}{e^{2x}}\, dx$$

i.e.,

$$u_1 = -\int x^2 \ln x\, dx = -(\ln x)\frac{x^3}{3} + \int \frac{x^2}{3}\, dx = \frac{x^3}{9} - \frac{x^3}{3}\ln x$$

$$u_2 = \int x \ln x\, dx = \frac{x^2}{2}\ln x - \int \frac{x}{2}\, dx = -\frac{x^2}{4} + \frac{x^2}{2}\ln x.$$

$$y_p = \left(\frac{x^3}{9} - \frac{x^3}{3}\ln x\right) e^x + \left(-\frac{x^2}{4} + \frac{x^2}{2}\ln x\right) x e^x$$

$$= e^x\left(\frac{x^3}{9} - \frac{x^3}{3}\ln x - \frac{x^3}{4} + \frac{x^3}{2}\ln x\right) = e^x\left(-\frac{5x^3}{36} + \frac{x^3}{6}\ln x\right).$$

The general solution of (1) is

$$y = y_c + y_p = c_1 e^x + c_2 x e^x + e^x\left(-\frac{5x^3}{36} + \frac{x^3}{6}\ln x\right).$$

Example 20. Find the general solution of

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

given that $y = x e^x$ is a solution of the associated homogeneous equation

$$x^2\frac{d^2y}{dx^2} - x(x+2)\frac{dy}{dx} + (x+2)\, y = 0 \qquad (2)$$

Solution. The given equation in the standard form is

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

Since $P + Qx = -\frac{x+2}{x} + x\frac{x+2}{x^2} = 0$, $y = x$ is also a solution of (2). The two solutions, $y = x$ and $y = x e^x$ are linearly independent. The complementary function of (1) is

$$y_c = c_1 x + c_2 x e^x.$$

We assume that

$$y_p = u_1 x + u_2 x e^x$$

is a particular solution of (1) (3)

Here $\quad y_1 = x, \quad y_2 = x e^x, \quad F(x) = x$

$$W = W(y_1, y_2) = y_1 y_2' - y_1' y_2 = x(x e^x + e^x) - x e^x = x^2 e^x$$

By the formulas (7) of (10.19), we have

$$u_1 = \int \frac{-y_2 F(x)}{W}\, dx \quad \text{and} \quad u_2 = \int \frac{y_1 F(x)}{W}\, dx$$

i.e., $$u_1 = -\int \frac{x^2 e^x}{x^2 e^x}\, dx = -x \quad \text{and} \quad u_2 = \int \frac{x^2}{x^2 e^x}\, dx = -e^{-x}$$

Substituting for u_1, u_2 into (3), a particular solution of (1) is

$$y_p = -x^2 - x$$

The general solution of (1) is

$$y = y_c + y_p = c_1 x + c_2 x e^x - x^2 - x$$

Example 21. Explain how the method of variation of parameters can be applied to find a particular solution of a nonhomogeneous linear third order differential equation whose complementary function is known. Apply the method to find a particular solution of

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

Solution. Suppose the C.F. of a linear third order differential equation

$$\frac{d^3y}{dx^3} + P\frac{d^2y}{dx^2} + Q\frac{dy}{dx} + Ry = F(x) \tag{1}$$

is known to be

$y_c = c_1 y_1 + c_2 y_2 + c_3 y_3$; y_1, y_2, y_3 being linearly independent solutions of the associated homogeneous equation of (1)

We assume that

$$y_p = u_1 y_1 + u_2 y_2 + u_3 y_3, \text{ where } u_1, u_2, u_3 \text{ are functions of } x.$$

$$y_p' = u_1 y_1' + u_2 y_2' + u_3 y_3' + u_1' y_1 + u_2' y_2 + u_3' y_3$$

We set $$u_1' y_1 + u_2' y_2 + u_3' y_3 = 0 \tag{2}$$

Then $$y_p' = u_1 y_1' + u_2 y_2' + u_3 y_3'$$

$$y_p'' = u_1 y_1'' + u_2 y_2'' + u_3 y_3'' + u_1' y_1' + u_2' y_2' + u_3' y_3'$$

Again set $$u_1' y_1' + u_2' y_2' + u_3' y_3' = 0 \tag{3}$$

Then $$y_p'' = u_1 y_1'' + u_2 y_2'' + u_3 y_3''$$

$$y_p''' = u_1 y_1''' + u_2 y_2''' + u_3 y_3''' + u_1' y_1'' + u_2' y_2'' + u_3' y_3''$$

Substituting for y_p, y_p', y_p'', y_p''' into equation (1), we have

$$(u_1 y_1''' + u_2 y_2''' + u_3 y_3''' + u_1' y_1'' + u_2' y_2'' + u_3' y_3'') + P(u_1 y_1'' + u_2 y_2'' + u_3 y_3'')$$
$$+ Q(u_1 y_1' + u_2 y_2' + u_3 y_3') + R(u_1 y_1 + u_2 y_2 + u_3 y_3) = F(x)$$

$$u_1(y_1''' + Py_1'' + Qy_1' + Ry_1) + u_2(y_2''' + Py_2'' + Qy_2' + Ry_2)$$
$$+ u_3(y_3''' + Py_3'' + Qy_3' + Ry_3) + u_1' y_1'' + u_2' y_2'' + u_3' y_3'' = F(x)$$

i.e., $$u_1 . 0 + u_2 . 0 + u_3 . 0 + u_1' y_1'' + u_2' y_2'' + u_3' y_3'' = F(x) \tag{4}$$

From the system of equations (2), (3) and (4), we find u_1', u_2' and u_3' and on integration we get u_1, u_2 and u_3.

Now we apply this method to find P.I. of

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

The C.F. of the equation is obtained by solving the equation $D(D^2 + 1) = 0$. Thus

$$y_c = c_1 + c_2 \cos x + c_3 \sin x \text{ with } y_1 = 1, y_2 = \cos x, y_3 = \sin x.$$

For the particular integral, let

$$y_p = u_1 y_1 + u_2 y_2 + u_3 y_3.$$
$$= u_1 + u_2 \cos x + u_3 \sin x$$

Substituting for y_1, y_2, y_3 and their derivatives into equations (2), (3) and (4) above, we have the following system of linear equations in u_1', u_2', u_3'.

$$u_1' + u_2' \cos x + u_3' \sin x = 0$$
$$u_1' \cdot 0 - u_2' \sin x + u_3' \cos x = 0$$
$$u_1' \cdot 0 - u_2' \cos x - u_3' \sin x = \csc x.$$

We solve this system by the Guassian elimination method

Augmented matrix of the system is

$$\left[\begin{array}{ccc|c} 1 & \cos x & \sin x & 0 \\ 0 & -\sin x & \cos x & 0 \\ 0 & -\cos x & -\sin x & \csc x \end{array}\right]$$

$$\underset{\sim}{R}\left[\begin{array}{ccc|c} 1 & 0 & 0 & \csc x \\ 0 & -\sin x & \cos x & 0 \\ 0 & -\cos x & -\sin x & \csc x \end{array}\right] \quad \text{by } R_1 + R_3$$

$$\underset{\sim}{R}\left[\begin{array}{ccc|c} 1 & 0 & 0 & \csc x \\ 0 & -\sin x\cos x & \cos^2 x & 0 \\ 0 & -\cos x\sin x & -\sin^2 x & 1 \end{array}\right] \quad \text{by } (\cos x)R_2 \text{ and } (\sin x)R_3$$

$$\underset{\sim}{R}\left[\begin{array}{ccc|c} 1 & 0 & 0 & \csc x \\ 0 & -\sin x\cos x & \cos^2 x & 0 \\ 0 & 0 & -1 & 1 \end{array}\right] \quad \text{by } R_3 - R_2$$

Therefore,

$-u'_3 = 1$ or $u'_3 = -1$, (from R_3)

$(-\sin x\cos x)\,u'_2 + (\cos^2 x)\,u'_3 = 0$, (from R_2)

or $u'_2 = -\cot x$.

$u'_1 = \csc x$, (from R_1).

Hence $u_1 = \int \csc x\,dx = \ln|\csc x - \cot x|$

$u_2 = \int \frac{-\cos x}{\sin x}\,dx = -\ln|\sin x|$

and $u_3 = \int -\,dx = -x$.

$y_p = u_1 + u_2\cos x + u_3\sin x$

$= \ln|\csc x - \cot x| - (\cos x)\ln|\sin x| - x\sin x$

is the required particular solution.

EXERCISE 10.6

Find a particular solution of each of the following (**Problems 1–9**)

1. $\frac{d^2y}{dx^2} + 4y = \sec 2x$

2. $\frac{d^2y}{dx^2} + y = \tan x\sec x$

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

4. $\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 5y = e^{-2x}\sec x$

5. $\frac{d^2y}{dx^2} - 4\frac{dy}{dx} + 4y = \frac{e^{2x}}{1+x}$

6. $\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = e^x \arcsin x$

7. $\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + 5y = e^x\tan 2x$

8. $\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y = e^{-x}\ln x$

9. $\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + 2y = 2e^{-x}\tan^2 x$.

10. Find the general solution of

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

given that $y_1 = x^2$ is a solution of the associated homogeneous equation.

11. Find the general solution of

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

given that $y_1 = \frac{1}{x}$ is a solution of the associated homogeneous equation.

12. Find the general solution of

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

given that $y_1 = e^x$ is a solution of the associated homogeneous equation.

13. Find the general solution of

$$(\sin^2 x)\frac{d^2y}{dx^2} - (\sin 2x)\frac{dy}{dx} + (1+\cos^2 x)\,y = \sin^3 x,$$

given that $y_1 = \sin x$ and $y_2 = x\sin x$ are linearly independent solutions of the associated homogeneous equation.

14. Use the method of Example 21 to find a particular solution of

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

15. By the method of variation of parameters, find a particular solution of

$$\frac{d^3y}{dx^3} - 2\frac{dy}{dx} - 4y = e^{-x}\sin x.$$

ONE VARIABLE ABSENT (NONLINEAR D.E.)

(10.20) Consider an equation in which the dependent variable y is missing i.e., an equation of the form

$$f\left(\frac{d^ny}{dx^n}, \frac{d^{n-1}y}{dx^{n-1}}, \ldots, \frac{dy}{dx}, x\right) = 0 \qquad (1)$$

If we set $p = \frac{dy}{dx}$, then (1) takes the form

$$f\left(\frac{d^{n-1}p}{dx^{n-1}}, \frac{d^{n-2}p}{dx^{n-2}}, \ldots, p, x\right) = 0$$

so that the order of (1) has been reduced by one.

If the derivative of lowest order occurring in the equation (1) is $\frac{d^ry}{dx^r}$, then put

$$\frac{d^ry}{dx^r} = p.$$

Similarly, consider the equation

$$f\left(\frac{d^ny}{dx^n}, \frac{d^{n-1}y}{dx^{n-1}}, \ldots, \frac{dy}{dx}, y\right) = 0 \qquad (2)$$

in which x is absent. We write

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

so that
$$\frac{d^2y}{dx^2} = \frac{dp}{dx} = \frac{dp}{dy}\frac{dy}{dx} = p\frac{dp}{dy}$$

$$\frac{d^3y}{dx^3} = \frac{d}{dx}\left(p\frac{dp}{dy}\right) = \frac{d}{dy}\left(p\frac{dp}{dy}\right)\frac{dy}{dx} = p\left(\frac{dp}{dy}\right)^2 + p^2\frac{d^2p}{dy^2}.$$

Substituting for $\frac{dy}{dx}, \frac{d^2y}{dx^2}, \frac{d^3y}{dx^3}, \ldots$ the equation (2) will be transformed into an equation in p and y of order $n-1$. This equation in then solved for p and y. Such an equation may reduce to more than one equation in p and y. The solution of each of these equations must satisfy (2) in order that it is a solution of (2).

Example 22. Solve $(1+x^2)\frac{d^2y}{dx^2} + x\frac{dy}{dx} + ax = 0$ (1)

Solution. Put $\frac{dy}{dx} = p$ and so $\frac{d^2y}{dx^2} = \frac{dp}{dx}$.

Then (1) becomes

$$(1+x^2)\frac{dp}{dx} + xp + ax = 0$$

or
$$\frac{dp}{a+p} = -\frac{x}{1+x^2}\,dx.$$

Integrating, we have

$$\ln|a+p| = \ln(1+x^2)^{-\frac{1}{2}} + \ln|c_1|$$

$$a+p = \frac{c_1}{\sqrt{1+x^2}}$$

Therefore,

$$a + \frac{dy}{dx} = \frac{c_1}{\sqrt{1+x^2}}$$

or
$$dy = \left(\frac{c_1}{\sqrt{1+x^2}} - a\right)dx$$

or
$$y = c_1\sinh^{-1}x - ax + c_2$$

is the general solution of (1).

Example 23. Solve: $y\frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^2 = \frac{dy}{dx}$ (1)

Solution. Here x is absent. We set

$$\frac{dy}{dx} = p \quad \text{and so} \quad \frac{d^2y}{dx^2} = p\frac{dp}{dy}.$$

Hence (1) is exact. Its first integral is

$$R_0 \frac{dy}{dx} + R_1 y = \int (2\cos x - 2x)\, dx + c_1$$

where $R_0 = P_0 = x^2 + 1$

$R_1 = P_1 - P_0' = 4x - 2x = 2x$.

Thus, the first integral becomes

$$(x^2 + 1)\frac{dy}{dx} + 2xy = 2\sin x - x^2 + c_1 \qquad (2)$$

This is again exact, since

$P_0 = x^2 + 1,\ P_1 = 2x$ and $P_1 - P_0' = 2x - 2x = 0$.

First integral of (2) is

$$R_0 y = \int (2\sin x - x^2 + c_1)\, dx + c_2$$

where $R_0 = P_0 = x^2 + 1$.

Thus $(x^2 + 1)y = -2\cos x - \dfrac{x^3}{3} + c_1 x + c_2$

is the general solution of (1).

Example 25. Solve

$$(2x - 1)\frac{d^3y}{dx^3} + (4 + x)\frac{d^2y}{dx^2} + 2\frac{dy}{dx} = 0 \qquad (1)$$

Solution. Here $n = 3$

$P_0 = 2x - 1,\ P_1 = 4 + x, P_2 = 2,\ P_3 = 0$

$P_3 - P_2' + P_1'' - P_0''' = 0 - 0 + 0 - 0 = 0$.

Hence (1) is exact. Its first integral is

$$R_0 \frac{d^2y}{dx^2} + R_1 \frac{dy}{dx} + R_2 y = c_1, \qquad (2)$$

where $R_0 = P_0 = 2x - 1$

$R_1 = P_1 - P_0' = 4 + x - 2 = 2 + x$

$R_2 = P_2 - P_1' + P_0'' = 2 - 1 = 1$

Hence (2) becomes

$$(2x - 1)\frac{d^2y}{dx^2} + (2 + x)\frac{dy}{dx} + y = c_1. \qquad (3)$$

In (3), we have

$P_0 = 2x - 1,\quad P_1 = 2 + x,\quad P_2 = 1$

and $P_2 - P_1' + P_0'' = 1 - 1 + 0 = 0$.

Therefore (3) is exact. First integral of (3) is

$$R_0 \frac{dy}{dx} + R_1 y = \int c_1\, dx + c_2, \qquad (4)$$

where $R_0 = P_0 = 2x - 1$

$R_1 = P_1 - P_0' = 2 + x - 2 = x$

Therefore (4) becomes

$$(2x - 1)\frac{dy}{dx} + xy = c_1 x + c_2 \qquad (5)$$

In (5), we have

$P_0 = 2x - 1,\quad P_1 = x$

and $P_1 - P_0' = x - 2 \neq 0$

Thus, (5) is not exact. We can write it as

$$\frac{dy}{dx} + \frac{x}{2x - 1}\, y = \frac{c_1 x + c_2}{2x - 1}$$

which is a linear equation. Its

$$\text{I F} = \exp\left(\int \frac{x\, dx}{2x - 1}\right) = \exp\left(\int \left(\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2x - 1}\right) dx\right)$$

$$= \exp\left(\frac{1}{2}x + \frac{1}{4}\ln(2x - 1)\right) = e^{\frac{1}{2}x}\,.(2x - 1)^{\frac{1}{4}}.$$

So $$\frac{d}{dx}\left[y e^{\frac{1}{2}x}\cdot(2x - 1)^{\frac{1}{4}}\right] = \frac{e^{\frac{1}{2}x}(2x - 1)^{\frac{1}{4}}(c_1 x + c_2)}{(2x - 1)}$$

or $$y e^{\frac{1}{2}x}\cdot(2x - 1)^{\frac{1}{4}} = \int \frac{e^{\frac{1}{2}x}(c_1 x + c_2)}{(2x - 1)^{\frac{3}{4}}}\, dx$$

TWO SPECIAL TYPES

(10.22) In this section we consider differential equations of the following types

(i) $\dfrac{d^n y}{dx^n} = f(x)$ (ii) $\dfrac{d^2 y}{dx^2} = f(y)$

The differential equation in (i) is exact and its solution can be found by successive integration.

The differential equation in (ii) is not exact. In order to solve it, multiply both sides of (ii) by $2\dfrac{dy}{dx}$ to obtain

$$2\frac{dy}{dx}\frac{d^2y}{dx^2} = 2\frac{dy}{dx}f(y) = 2f(y)\frac{dy}{dx}$$

or
$$\frac{d}{dx}\left(\frac{dy}{dx}\right)^2 = 2\frac{dy}{dx}f(y)$$

Integrating, we get

$$\left(\frac{dy}{dx}\right)^2 = 2\int f(y)\,dy + c$$

or
$$\frac{dy}{dx} = \left[2\int f(y)\,dy + c\right]^{1/2}$$

which is a separable equation and can be solved.

Example 26. Solve: $x^4\dfrac{d^4y}{dx^4} + 1 = 0.$ (1)

Solution. $\dfrac{d^4y}{dx^4} = -\dfrac{1}{x^4} = -x^{-4}.$

Integrating successively, we have

$$\frac{d^3y}{dx^3} = -\frac{x^{-3}}{-3} + c_1$$

$$\frac{d^2y}{dx^2} = -\frac{x^{-3+1}}{3\cdot 2} + c_1x + c_2$$

$$\frac{dy}{dx} = -\frac{x^{-2+1}}{3\cdot 2\,(-1)} + c_1\frac{x^2}{2} + c_2x + c_3$$

$$y = \frac{1}{6}\ln|x| + c_1\frac{x^3}{6} + c_2\frac{x^2}{2} + c_3x + c_4$$

is the general solution of (1).

Example 27. Solve: $\dfrac{d^2y}{dx^2} = a^2 y$ (1)

Solution. Multiplying both the sides of (1) by $2\dfrac{dy}{dx}$, we have

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

Integrating, we get

$$\left(\frac{dy}{dx}\right)^2 = a^2y^2 + c = a^2(y^2 + c_1)$$

$$\frac{dy}{dx} = a\sqrt{y^2 + c_1}$$

or
$$\frac{dy}{\sqrt{y^2 + c_1}} = a\,dx$$

or
$$\ln\left(y + \sqrt{y^2 + c_1}\right) = ax + c_2$$

is the general solution of the given equation.

EXERCISE 10.7

Solve:

1. $2\dfrac{d^2y}{dx^2} - \left(\dfrac{dy}{dx}\right)^2 + 4 = 0$
2. $2x\dfrac{d^3y}{dx^3}\dfrac{d^2y}{dx^2} = \left(\dfrac{d^2y}{dx^2}\right)^2 - a^2$
3. $x\dfrac{d^2y}{dx^2} - \left(\dfrac{dy}{dx}\right)^3 - \dfrac{dy}{dx} = 0$
4. $x\dfrac{d^3y}{dx^3} - 2\dfrac{d^2y}{dx^2} = 12x^3,$ $y(1) = 0,\ y'(1) = 1,\ y''(1) = 0$
5. $2y\dfrac{d^2y}{dx^2} - \left(\dfrac{dy}{dx}\right)^2 = 1$
6. $y\dfrac{d^2y}{dx^2} - \left(\dfrac{dy}{dx}\right)^2 = 4y^2\ln x,$ $y(1) = e,\ y'(1) = 2e$
7. $(1 + y^2)\dfrac{d^2y}{dx^2} + \left(\dfrac{dy}{dx}\right)^3 + \dfrac{dy}{dx} = 0$
8. $y\dfrac{d^2y}{dx^2} + 4y^2 - \dfrac{1}{2}\left(\dfrac{dy}{dx}\right)^2 = 0,$ $y(0) = 1,\ y'(0) = \sqrt{8}$

9. $(2x^2 + 3x)\dfrac{d^2y}{dx^2} + (6x + 3)\dfrac{dy}{dx} + 2y = (x + 1)\,e^x$

10. $\sin x\dfrac{d^2y}{dx^2} - \cos x\dfrac{dy}{dx} + 2y\sin x = 0$

11. $(x + \sin x)\dfrac{d^3y}{dx^3} + 3(1 + \cos x)\dfrac{d^2y}{dx^2} - 3\sin x\dfrac{dy}{dx} - y\cos x = -\sin x$

12. $(e^x + 2x)\dfrac{d^4y}{dx^4} + 4(e^x + 2)\dfrac{d^3y}{dx^3} + 6e^x\dfrac{d^2y}{dx^2} + 4e^x\dfrac{dy}{dx} + e^x y = \dfrac{1}{x^5}$

13. $x^5\dfrac{d^2y}{dx^2} + 3x^3\dfrac{dy}{dx} + (3 - 6x)\,x^2y = x^4 + 2x - 5$

14. (i) $\dfrac{d^3y}{dx^3} = \ln x$ (ii) $\dfrac{d^2y}{dx^2} = x^2\sin x$

15. (i) $\dfrac{d^2y}{dx^2} = -\cot y\csc^2 y;\ y(0) = \dfrac{\pi}{2},\ y'(0) = 1$ (ii) $\dfrac{d^2y}{dx^2} = -\dfrac{a^2}{}$

LINEAR SYSTEMS OF D.E.

(10.23) The differential equations studied in previous sections involved one independent variable x and a dependent variable y. In many problems in applied mathematics, there occur several dependent variables which are functions of a single independent variable. The mathematical formulation of such problems gives rise to a system of simultaneous differential equations. In this section, we briefly discuss such a system. We shall restrict our discussion to three dependent variables x, y and z which are all functions of an independent variable t. The theory can be easily extended to n dependent variables of a single independent variable.

Let x, y, z be functions of t and let D denote $\dfrac{d}{dt}$. The system of three equations in three unknowns x, y, z such as

$$\left.\begin{aligned} L_{11}(D)\,x + L_{12}(D)\,y + L_{13}(D)\,z &= f_1(t)\\ L_{21}(D)\,x + L_{22}(D)\,y + L_{23}(D)\,z &= f_2(t)\\ L_{31}(D)\,x + L_{32}(D)\,y + L_{33}(D)\,z &= f_3(t)\end{aligned}\right\}\qquad (1)$$

where $L_{ij}(D)$, $(1 \le i, j \le 3)$ are linear operators, is called a [1]**linear system of differential equations.** The system (1) is said to be a **linear system with constant coefficients** if $L_{ij}(D)$ are polynomials in D with coefficients not involving the independent variable. For example,

$$\begin{aligned}(D^2 + D)\,x \qquad - D^2y + 3Dz &= 0\\ (D^2 - 1)\,x + (2D + 3)\,y + Dz &= t\\ (D + 1)^2\,x \qquad + 5\,Dy - z &= t^2\end{aligned}$$

is a linear system of differential equations with constant coefficients.

Here we shall study only linear system of differential equations with constant coefficients.

(10.24) Linear Systems with Constant Coefficients. Let

$$\left.\begin{aligned} L_{11}(D)\,x + L_{12}(D)\,y + L_{13}(D)\,z &= f_1(t)\\ L_{21}(D)\,x + L_{22}(D)\,y + L_{23}(D)\,z &= f_2(t)\\ L_{31}(D)\,x + L_{32}(D)\,y + L_{33}(D)\,z &= f_3(t)\end{aligned}\right\}\qquad (1)$$

be a linear system of differential equations with constant coefficients, i.e. $L_{ij}(D)$ are polynomials in D, $1 \le i, j \le 3$. A set of functions $x = f(t)$, $y = g(t)$, $z = h(t)$ is called a **solution** of the system (1), if each of the three equations of the system is satisfied on substitution of these values of the variables x, y, z and their derivatives with respect to t.

The solution of (1) can be found only if

$$\det[\,L\,] \neq 0,$$

where $[\,L\,] = [L_{ij}(D)]$ is the matrix of coefficients in (1).

As in the case of linear algebraic systems discussed in Chapter 4, the following elementary operations can be performed on a linear differential system with constant coefficients to produce another equivalent system:

(1) Interchanging two equations.

(2) Multiplying an equation by a nonzero constant.

(3) Operating on both sides of an equation by a polynomial operator and adding (subtracting) the result to (from) another equation.

These elementary operations will be used to find the general solution of the system (1).

Before solving a linear system, it should be changed into operator notation.

1. Recall that
$L \equiv a_0(t)\,D^n + a_1(t)\,D^{n-1} + \dots + a_{n-1}(t)\,D + a_n(t)$ is an nth order linear operator. $a_0(t) \neq 0$

Example 28. Solve the system

$$\frac{dx}{dt} + \frac{dy}{dt} + y - x = e^{2t}$$

$$\frac{d^2x}{dt^2} + \frac{dy}{dt} = 3e^{2t}$$

Solution. Using the operator notation with $D = \frac{d}{dt}$, we can write the given system as

$$(D-1)\,x + (D+1)\,y = e^{2t} \qquad (1)$$

$$D^2x + Dy = 3e^{2t} \qquad (2)$$

We shall eliminate one of the dependent variables from (1) and (2) to get an equation in one dependent variable.

Operating on (1) by D and on (2) by $D + 1$, we have

$$D(D-1)\,x + D(D+1)\,y = De^{2t} = 2e^{2t} \qquad (3)$$

$$(D+1)\,D^2x + (D+1)\,Dy = (D+1)\,3e^{2t} = 9e^{2t} \qquad (4)$$

Subtracting (3) from (4), we get

$$(D^3 + D^2 - D^2 + D)\,x + 0 = 7e^{2t}$$

or $$(D^3 + D)\,x = 7e^{2t}$$

Its solution is easily found as

$$x = c_1 + c_2 \sin t + c_3 \cos t + \frac{7}{10}\, e^{2t} \qquad (5)$$

Substituting for x from (5) into (2), we obtain

$$Dy = c_2 \sin t + c_3 \cos t + \frac{1}{5}\, e^{2t}$$

Integrating, we have

$$y = -c_2 \cos t + c_3 \sin t + \frac{1}{10}\, e^{2t} + c_4 \qquad (6)$$

To verify, substitute x from (5) and y from (6) into both (1) and (2) and obtain after simplification

$$c_4 - c_1 = 0, \quad 0 = 0$$

Thus $c_4 = c_1$ and (5) and (6) constitute the required general solution of the given system provided that we replace c_4 in (6) by c_1.

Note. It is proved in advanced books that the number of independent constants in the general solution of a linear system of differential equations is the same as the degree in D of the determinant of the operational coefficients of the equations. If the determinant of operational coefficients is identically zero the system may have no solution or it may have solutions with any number of constants.

We note that the coefficient determinant of the system of Example 28 is

$$\begin{vmatrix} D-1 & D+1 \\ D^2 & D \end{vmatrix}$$

which is of degree three in D. Hence the number of independent constants in the general solution should be three as we have already obtained.

Example 29. Solve

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

$$\frac{d^2y}{dt^2} + \frac{dx}{dt} - x + y = 0$$

Solution. Writing $D = \frac{d}{dt}$, the given system in operator notation is

$$(D^2-1)\,x + (D+1)\,y = 1 \qquad (1)$$

$$(D-1)\,x + (D^2+1)\,y = 0 \qquad (2)$$

Operating on (1) by $(D^2 + 1)$ and on (2) by $(D + 1)$, we have

$$(D^2+1)\,(D^2-1)\,x + (D^2+1)\,(D+1)\,y = 1 \qquad (3)$$

$$(D+1)\,(D-1)\,x + (D+1)\,(D^2+1)\,y = 0 \qquad (4)$$

Subtracting (4) from (3), we get

$$(D^4 - 1 - D^2 + 1)\,x = 1$$

or $$D^2(D^2-1)\,x = 1.$$

Its solution is

$$x = c_1 + c_2 t + c_3 e^{t} + c_4 e^{-t} - \frac{t^2}{2}. \qquad (A)$$

Again applying the operator $(D - 1)$ on (1) and $(D^2 - 1)$ on (2), we obtain

$$(D-1)\,(D^2-1)\,x + (D-1)\,(D+1)\,y = (D-1)\,1 = -1 \qquad (5)$$

$$(D^2-1)\,(D-1)\,x + (D^2-1)\,(D^2+1)\,y = 0. \qquad (6)$$

Subtracting (5) from (6), we have

$$(D^4 - 1 - D^2 + 1)\,y = 1$$

or $$D^2(D^2 - 1)\,y = 1$$

Its general solution is

$$y = k_1 + k_2 t + k_3 e^t + k_4 e^{-t} - \frac{t^2}{2} \qquad \text{(B)}$$

Thus the solution of the system must be of the forms given by (A) and (B) for some choice of constants $c_1, c_2, c_3, c_4, k_1, k_2, k_3, k_4$. The determinant of operational coefficients for the system is

$$\begin{vmatrix} D^2 - 1 & D + 1 \\ D - 1 & D^2 + 1 \end{vmatrix}$$

which is of degree 4 in D. Hence the number of independent constants in the general solution should be four.

Substituting for x from (A) and for y from (B) into the second equation of the system, we get

$$\left[c_2 + c_3 e^t - c_4 e^{-t} - t\right] + \left[-c_1 - c_2 t - c_3 e^t - c_4 e^{-t} + \frac{t^2}{2}\right]$$

$$+ \left[k_3 e^t + k_4 e^{-t} - 1\right] + \left[k_1 + k_2 t + k_3 e^t + k_4 e^{-t} - \frac{t^2}{2}\right] = 0$$

or $$(c_2 - c_1 - 1 + k_1) + (-1 - c_2 + k_2)\,t + 2k_3 e^t + 2(-c_4 + k_4)\,e^{-t} = 0$$

In order that the pair (A) and (B) satisfy the second equation of the system, we must have

$$\begin{aligned} c_2 - c_1 - 1 + k_1 &= 0 & \text{or} \quad k_1 &= 1 + c_1 - c_2 \\ -1 - c_2 + k_2 &= 0 & \text{or} \quad k_2 &= 1 + c_2 \\ k_3 &= 0 & \text{or} \quad k_3 &= 0 \\ -c_4 + k_4 &= 0 & \text{or} \quad k_4 &= c_4 \end{aligned}$$

Hence

$$x = c_1 + c_2 t + c_3 e^t + c_4 e^{-t} - \frac{t^2}{2}$$

$$y = (1 + c_1 - c_2) + (c_2 + 1)\,t + c_4 e^{-t} - \frac{t^2}{2}$$

is the general solution of the given system.

EXERCISE 10.8

Find the general solution of each of the following linear systems:

1. $\dfrac{dx}{dt} = y$, $\quad \dfrac{dy}{dt} = -4x + 4y$

2. $\dfrac{dx}{dt} = x + y$, $\quad \dfrac{dy}{dt} = 4x - 2y$

3. $2\dfrac{dx}{dt} + \dfrac{dy}{dt} - x - y = e^{-t}$, $\quad \dfrac{dx}{dt} + \dfrac{dy}{dt} + 2x + y = e^t$

4. $2\dfrac{dx}{dt} - 2\dfrac{dy}{dt} - 3x = t$, $\quad 2\dfrac{dx}{dt} + 2\dfrac{dy}{dt} + 3x + 8y = 2$

5. $\dfrac{d^2x}{dt^2} + \dfrac{dy}{dt} + 6x = 0$, $\quad \dfrac{dx}{dt} + \dfrac{dy}{dt} + 2x - 2y = 2$

6. $\dfrac{dx}{dt} - 2x - y = 3e^{2t}$, $\quad \dfrac{dy}{dt} + 4x - 2y = te^{2t}$

7. $\dfrac{dx}{dt} + \dfrac{dy}{dt} + 5x + 4y = e^{-t}$, $\quad \dfrac{dx}{dt} + \dfrac{dy}{dt} + 2x + y = 3$

8. $2\dfrac{dx}{dt} + \dfrac{dy}{dt} + 3x + 4y = -\cos t$, $\quad \dfrac{dx}{dt} + \dfrac{dy}{dt} + x + 2y = 2\sin t$

9. $\dfrac{dx}{dt} + \dfrac{dy}{dt} + 2x + 6y = 2e^t$, $\quad 2\dfrac{dx}{dt} + 3\dfrac{dy}{dt} + 3x + 8y = -1$

10. $\dfrac{dx}{dt} - x - y = 0, \quad x(0) = 1$; $\quad \dfrac{dy}{dt} - y = 0, \quad y(0) = 0$

11. $\dfrac{dx}{dt} - x - 2y = t - 1, \quad x(0) = 0$; $\quad \dfrac{dy}{dt} - 3x - 2y = -5t - 2, \quad y(0) = 4$

12. $\dfrac{dx}{dt} - 3y = 0$, $\quad \dfrac{dy}{dt} + z = 0$, $\quad \dfrac{dz}{dt} + y = 0$

13. $\dfrac{d^2x}{dt^2} + \dfrac{dy}{dt} - y = 0$, $\quad 2\dfrac{dx}{dt} + \dfrac{dz}{dt} - x - z = 0$, $\quad \dfrac{dx}{dt} + \dfrac{dy}{dt} + 3x - 4y + 3z = 0$

There are many differential equations whose solutions cannot be obtained in terms of **elementary functions** by the methods already discussed. However, their solutions can be obtained in the form of power series. The procedure is similar to the method of undetermined coefficients for polynomials but the number of coefficients will not be finite in this case.

We begin our study with examples of first order differential equations.

POWER SERIES SOLUTIONS OF FIRST ORDER D.E.

Example 30. Consider the equation

$$y' = -2y \tag{1}$$

We assume that (1) has a power series solution

$$y = \sum_{n=0}^{\infty} c_n x^n$$

The series can be differentiated term by term so that

$$y' = \sum_{n=0}^{\infty} nc_n x^{n-1} = \sum_{n=1}^{\infty} nc_n x^{n-1}$$

Substituting for y and y' into (1), we have

$$\sum_{n=1}^{\infty} nc_n x^{n-1} = -2\sum_{n=0}^{\infty} c_n x^n \tag{2}$$

In (2), we reindex the series on the left by setting $n-1 = k$ so that

$$\sum_{n=1}^{\infty} n c_n x^{n-1} = \sum_{k=0}^{\infty} (k+1) c_{k+1} x^k$$

$$= \sum_{n=0}^{\infty} (n+1) c_{n+1} x^n \text{ , since } k \text{ is dummy index.}$$

Now (2) becomes

$$\sum_{n=0}^{\infty} (n+1) c_{n+1} x^n = -2\sum_{n=0}^{\infty} c_n x^n.$$

Equating coefficients of like terms, we have the recursion formula

$$(n+1)c_{n+1} = -2c_n \tag{3}$$

Setting $n = 0, 1, 2, 3, \cdots$ into (3), we get the successive coefficients in terms of c_0 as below

$$c_1 = -2c_0$$

$$c_2 = -\frac{2}{2}c_1 = 2c_0$$

$$c_3 = -\frac{2}{3}c_2 = -\frac{2\cdot 2}{3}c_0$$

The solution of the given equation becomes

$$y = c_0 - 2c_0 x + 2c_0 x^2 - \frac{2\cdot 2}{3}c_0 x^3 + \cdots$$

$$= c_0\left(1 + (-2x) + \frac{(-2x)^2}{2!} + \frac{(-2x)^3}{3!} + \cdots\right) \tag{4}$$

The power series (4) can be easily seen as a convergent series with interval of convergence $]-\infty, \infty[$. Thus (4) is indeed a solution of (1).

If we solve (1) directly, we find that

$$y = c_0 e^{-2x}$$

a solution of (1) which is clearly the same as (4) with e^{-2x} expressed as a series.

Note: When an exact solution of an equation can be easily found, series solution is not beneficial.

Example 31. Find a series solution of the initial value problem

$$(1+x)y' - my = 0, \qquad y(0) = 1,$$

m being any real number.

Solution. We assume a series solution as

$$y = \sum_{n=0}^{\infty} c_n x^n. \tag{1}$$

Differentiation of both sides of (1) gives

$$y' = \sum_{n=0}^{\infty} nc_n x^{n-1}.$$

Substituting the values of y and y' into the given equation, we have

$$(1+x)\sum_{n=0}^{\infty} nc_n x^{n-1} - m\sum_{n=0}^{\infty} c_n x^n = 0$$

or $$\sum_{n=1}^{\infty} nc_n x^{n-1} + \sum_{n=0}^{\infty} nc_n x^n - m\sum_{n=0}^{\infty} c_n x^n = 0$$

or $$\sum_{n=0}^{\infty} (n+1)\, c_{n+1} x^n + \sum_{n=0}^{\infty} (n-m)\, c_n x^n = 0.$$

Equating coefficients of x^n, we get

$$(n+1)\, c_{n+1} = (m-n)\, c_n$$

Since $y(0) = 1$, we obtain $c_0 = 1$ from (1).

Substituting $n = 0, 1, 2, 3, \cdots$, into (2), we have

$$c_1 = mc_0 = m$$

$$c_2 = \frac{m-1}{2}\, c_1 = \frac{m(m-1)}{2!}$$

$$c_3 = \frac{m-2}{3}\, c_2 = \frac{m(m-1)(m-2)}{3!}$$

$$\vdots \quad \vdots \quad \vdots$$

Inserting the values of $c_0, c_1, c_2, \cdots$ into (1), we obtain

$$y = 1 + mx + \frac{m(m-1)}{2!}\, x^2 + \cdots$$

It may be easily checked that this series converges for $|x| < 1$. Thus (3) is a series solution of the given initial value problem.

If we solve the given equation directly, we have by separation of variables

$$\frac{y'}{y} = \frac{m}{1+x}.$$

Integrating, we get

$$\ln y = m \ln(1+x) + \ln c$$

or $$y = c\,(1+x)^m.$$

Applying the initial condition, we have

$$c = 1. \quad \text{Thus} \quad y = (1+x)^m$$

is solution of the given problem.

Comparing (3) and (4), we obtain

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{2!}\, x^2 + \cdots + \frac{m(m-1)(m-2)\cdots(m-r+1)}{r!}\, x^r + \cdots$$

which is the Binomial Theorem for exponent m, m real.

If m is a positive integer, then all the terms $c_r x^r$ for $r > m$ are zero and the series terminates at x^m. The right hand side of (3) becomes a polynomial of degree m which converges for all values of x.

EXERCISE 10.9

Apply the power series method to solve the following differential equations **(Problems 1 – 9)**

1. $y' = y\left(1 + \frac{1}{x}\right)$
2. $(x^2 + x)\, y' = (2x+1)\, y$
3. $y' - k\,y = 0$
4. $y' + y - 1 = 0$
5. $x(1-x)\, y' = y$
6. $xy' = (2x^2 + 1)\, y$
7. $(1 - x^2)\, y' = y$
8. $xy' - (x+2)\, y = -2x^2 - 2x$
9. $(x+1)\, y' - (2x+3)\, y = 0.$
10. Express $\arcsin x$ in the form of a power series $\sum_{n=0}^{\infty} c_n x^n$ by solving

$$y' = \frac{1}{\sqrt{1-x^2}}, \quad y(0) = 0$$

in two different ways. Hence deduce the formula

$$\frac{\pi}{6} = \frac{1}{2} + \frac{1}{2}\cdot\frac{1}{3\cdot 2^3} + \frac{1\cdot 3}{2\cdot 4}\cdot\frac{1}{5\cdot 2^5} + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\frac{1}{7\cdot 2^7} + \cdots$$

SECOND-ORDER LINEAR EQUATIONS

Before we take up the problem of solving second order linear differential equations with variable coefficients by the power series method, we need some definitions.

(10.25) Definition. A real valued function $f : R \to R$ is said to be analytic at a point x_0 if it has a power series representation

$$f(x) = \sum_{n=0}^{\infty} c_n (x - x_0)^n$$

that is convergent in some **neighbourhood**[1] of x_0. For example the functions e^x, $\sin x$, $\cos x$ and polynomials in x are all analytic at all points of R. The function represented by the sum of a power series is analytic at all points inside its interval of convergence. The powers series

$$1 + x + x^2 + \cdots$$

has the sum function $\frac{1}{1-x}$, $(|x| < 1)$ and it is analytic at all points within its interval of convergence i.e. at all points of $]-1, 1[$.

(10.26) Definition. Let

$$a_0(x)\,y'' + a_1(x)\,y' + a_2(x)\,y = 0 \qquad (1)$$

be a linear homogeneous differential equation with variable coefficients $a_0(x)$, $a_1(x)$ and $a_2(x)$ that are continuous functions of x in some interval $]\alpha, \beta[$. If $a_0(x) \neq 0$ in its domain, we divide by $a_0(x)$ so that the equation (1) becomes

$$y'' + \frac{a_1(x)}{a_0(x)}\,y' + \frac{a_2(x)}{a_0(x)}\,y = 0$$

or $\quad y'' + P(x)\,y' + Q(x)\,y = 0 \qquad (2)$

where $\quad P(x) = \frac{a_1(x)}{a_0(x)}, \quad Q(x) = \frac{a_2(x)}{a_0(x)}$

(2) is called the **standard** (or **normalized**) form of (1).

(10.27) Definition. Let $\quad y'' + P(x)\,y' + Q(x)\,y = 0 \qquad (1)$

be in the standard form. If the coefficients $P(x)$ and $Q(x)$ are analytic at a point $x = x_0$, then the point $x = x_0$ is called an **ordinary point** of the differential equation (1). Any point, that is not an ordinary point of (1), is called a **singular point.**

Thus for a point $x = x_0$ to be an ordinary point of (1), $P(x)$ and $Q(x)$ must be represented by power series in $x - x_0$ that are covergent in some neighbourhood of x_0.

If $P(x)$ and $Q(x)$ are algebraic functions then they are analytic at all points except where their denominators vanish. In this case all points, except where the denominators of $P(x)$ and $Q(x)$ vanish, are ordinary points of the differential equation.

1. A neighbourhood of x_0 is any open interval $]-a, a[$ containing x_0.

Example 32. Find the ordinary points of

$$(x^2 - 1)\,y'' + 3xy' + (x + 1)\,y = 0 \qquad (1)$$

Solution. Writing (1) in the standard form, we have

$$y'' + \frac{3x}{x^2-1}\,y' + \frac{x+1}{x^2-1}\,y = 0$$

Here $\quad P(x) = \frac{3x}{x^2-1}, \quad Q(x) = \frac{x+1}{x^2-1} = \frac{1}{x-1}$

$P(x)$ and $Q(x)$ are algebraic functions whose denominators vanish at $x = 1$, $x = -1$. Except for these points, all points on the real line R, are ordinary points of (1).

The points $x = \pm 1$ are singular points of (1).

Example 33. Find ordinary points of

$$y'' + \cos x\, y' + e^x\, y = 0 \qquad (1)$$

Solution. Here $\quad P(x) = \cos x, \qquad Q(x) = e^x$

and these are transcendental functions. Since $\cos x$ and e^x have power series representations

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

and the series converge for all x, every point x of R is an ordinary point of (1). There are no singular points of (1).

SERIES SOLUTIONS AROUND ORDINARY POINTS

A sufficient condition for the existence of a power series solution of a linear differential equation

$$y'' + P(x)\,y' + Q(x)\,y = 0 \qquad (1)$$

around an ordinary point is stated in the following theorem whose proof is omitted.

(10.28) Theorem. If $x = x_0$ is an ordinary point of (1) then it is always possible to find its two independent power series solutions of the form $y = \sum_{n=0}^{\infty} c_n\,(x - x_0)^n$. The power series will converge for $|x - x_0| < R$, where R is a positive number.

(10.29) The following procedure will be adopted to find a power series solution of

$$y'' + P(x)\,y' + Q(x) = 0 \qquad (1)$$

around an ordinary point $x = x_0$ of (1)

I. If $P(x)$ and $Q(x)$ are algebraic functions then multiply (1) by the L.C.M of denominators of $P(x)$ and $Q(x)$ so that (1) becomes

$$a_0(x)\,y'' + a_1(x)\,y' + a_2(x)\,y = 0$$

and $a_0(x)$, $a_1(x)$ and $a_2(x)$ are polynomials in x.

If $P(x)$ and $Q(x)$ involve **transcendental functions**[1], replace them by their power series about the ordinary point $x = x_0$.

II. Assume $y = \sum_{n=0}^{\infty} c_n(x-x_0)^n$ is a power series solution of (1)

III. Differentiate $y = \sum_{n=0}^{\infty} c_n(x-x_0)^n$ twice and substitute into (1) for y, y' and y''.

IV. Equate coefficients of like powers of x to have a recursion relation from which coefficients c_n will be obtained step by step.

The method is illustrated by the following examples.

Example 34. Find the power series solution of

$$y'' - xy' - y = 0 \qquad (1)$$

around the ordinary point $x = 0$.

Solution. Let

$$y = \sum_{n=0}^{\infty} c_n x^n$$

be a solution of (1). Then

$$y' = \sum_{n=0}^{\infty} nc_n x^{n-1}, \quad y'' = \sum_{n=0}^{\infty} n(n-1)\,c_n x^{n-2}.$$

Substituting into (1) for y, y′ and y″, we obtain

$$\sum_{n=0}^{\infty} n(n-1)\,c_n x^{n-2} - x\sum_{n=0}^{\infty} nc_n x^{n-1} - \sum_{n=0}^{\infty} c_n x^n = 0$$

1. Functions which are not algebraic functions.

$$\sum_{n=2}^{\infty} n(n-1)\,c_n x^{n-2} - \sum_{n=0}^{\infty} (n+1)\,c_n x^n = 0 \qquad (2)$$

We reindex the first sum by setting $n - 2 = k$ so that

$$\sum_{n=2}^{\infty} n(n-1)\,c_n x^{n-2} = \sum_{k=0}^{\infty} (k+2)(k+1)\,c_{k+2}\,x^k$$

$$= \sum_{n=0}^{\infty} (n+2)(n+1)\,c_{n+2}\,x^n, \text{ since } k \text{ is dummy variable.}$$

Now (2) becomes

$$\sum_{n=0}^{\infty} \left[(n+2)(n+1)\,c_{n+2} - (n+1)\,c_n\right] x^n = 0$$

Equating the coefficients of x^n, we get, for all natural numbers n,

$$(n+2)\,c_{n+2} - c_n = 0 \quad \text{or} \quad c_{n+2} = \frac{c_n}{n+2} \qquad (3)$$

(3) is the recursion formula from which the coefficients c_n will be obtained step by step.

Putting $n = 0$ and 1, we note that

$$c_2 = \frac{c_0}{2} \quad \text{and} \quad c_3 = \frac{c_1}{3}$$

Thus the coefficients c_n can be expressed in terms of c_0 and c_1 according as n is even or odd. For even and odd n, (3) may be written as

$$c_{2n+2} = \frac{c_{2n}}{2n+2} \qquad (4)$$

$$\text{and} \quad c_{2n+3} = \frac{c_{2n+1}}{2n+3} \qquad (5)$$

Setting $n = 0, 1, 2, 3 \cdots$, into (4) and (5), we have

$$c_2 = \frac{c_0}{2}$$

$$c_4 = \frac{c_2}{4} = \frac{c_0}{2^3} = \frac{c_0}{2^2 \cdot 2!}$$

$$c_6 = \frac{c_4}{2 \cdot 3} = \frac{c_0}{2^4 \cdot 3} = \frac{c_0}{2^3 \cdot 3!}$$

$$c_8 = \frac{c_6}{8} = \frac{c_0}{2^3\cdot 3} = \frac{c_0}{2^4\cdot 4!}$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

and $$c_3 = \frac{c_1}{3}$$

$$c_5 = \frac{c_3}{5} = \frac{c_1}{3\cdot 5}$$

$$c_7 = \frac{c_5}{7} = \frac{c_1}{3\cdot 5\cdot 7}$$

$$c_9 = \frac{c_7}{9} = \frac{c_1}{3\cdot 5\cdot 7\cdot 9}$$

$$\vdots \quad \vdots \quad \vdots$$

Thus $$u(x) = c_0 + \frac{c_0}{2}x^2 + \frac{c_0}{2^2\cdot 2!}x^4 + \frac{c_0}{2^3\cdot 3!}x^6 + \cdots$$

and $$v(x) = c_1 x + \frac{c_1}{3}x^3 + \frac{c_1}{3\cdot 5}x^5 + \frac{c_1}{3\cdot 5\cdot 7}x^7 + \cdots$$

are two linearly independent solutions of (1).

The general power series solution is

$$\begin{aligned} y &= u(x) + v(x) \\ &= c_0\left(1 + \frac{x^2}{2} + \frac{x^4}{2^2\cdot 2!} + \frac{x^6}{2^3\cdot 3!} + \frac{x^8}{2^4\cdot 4!} + \cdots\right) + c_1\left(x + \frac{x^3}{3} + \frac{x^5}{3\cdot 5} + \frac{x^7}{3\cdot 5\cdot 7} + \cdots\right) \\ &= c_0 e^{\frac{x^2}{2}} + c_1\left(x + \frac{x^3}{3} + \frac{x^5}{3\cdot 5} + \frac{x^7}{3\cdot 5\cdot 7} + \cdots\right) \end{aligned}$$

Example 35. Find the series solution of **Airy's[1] equation**

$$y'' - xy = 0 \qquad (1)$$

Solution. Here $x = 0$ is an ordinary point of (1).

Let $y = \sum_{n=0}^{\infty} c_n x^n$ be a series solution of (1).

Then $y' = \sum_{n=0}^{\infty} n c_n x^{n-1}$ and $y'' = \sum_{n=0}^{\infty} n(n-1)c_n x^{n-2}$.

1. Named after the English astronomer Sir George Biddel Airy (1801 – 1892).

Substituting for y and y'' into (1), we have

$$\sum_{n=0}^{\infty} n(n-1)c_n x^{n-2} - \sum_{n=0}^{\infty} c_n x^{n+1} = 0$$

or $$\sum_{n=2}^{\infty} n(n-1)c_n x^{n-2} - \sum_{n=3}^{\infty} c_{n-3} x^{n-2} = 0$$

or $$2c_2 + \sum_{n=3}^{\infty} n(n-1)c_n x^{n-2} - \sum_{n=3}^{\infty} c_{n-3} x^{n-2} = 0$$

Equating coefficients of like terms, we get

$$2c_2 = 0$$

and $$n(n-1)c_n - c_{n-3} = 0, \qquad n \geq 3$$

or $$c_n = \frac{c_{n-3}}{n(n-1)} \qquad (2)$$

Setting $n = 3, 4, 5, 6, \cdots$ into (2), we have

$$c_3 = \frac{c_0}{6}$$

$$c_4 = \frac{c_1}{3\cdot 4}$$

$$c_5 = \frac{c_2}{5\cdot 4} = 0 = c_8 = c_{11} = \cdots = c_{3n+2}$$

$$c_6 = \frac{c_3}{6\cdot 5} = \frac{c_0}{6\cdot 5\cdot 3\cdot 2}$$

$$c_7 = \frac{c_4}{7\cdot 6} = \frac{c_1}{7\cdot 6\cdot 4\cdot 3}$$

Thus

$c_2, c_5, c_8, \cdots, c_{3n+2}, \cdots$ are all zero.

$c_3, c_6, c_9, \cdots, c_{3n}, \cdots$ are all dependent on c_0.

$c_4, c_7, c_{10}, \cdots, c_{3n+1}, \cdots$ are all dependent on c_1.

$c_n = \dfrac{c_{n-3}}{n(n-1)}$ implies that $\quad c_{3n} = \dfrac{c_{3n-3}}{3n(3n-1)}$

Therefore, $c_3 = \dfrac{c_0}{3\cdot 2} = \dfrac{c_0}{6}$

$$c_6 = \frac{c_3}{6\cdot 5} = \frac{c_0}{6\cdot 5\cdot 3\cdot 2} = \frac{c_0}{180}$$

$$c_9 = \frac{c_6}{9\cdot 8} = \frac{c_0}{9\cdot 8\cdot 6\cdot 5\cdot 3\cdot 2} = \frac{c_0}{12960}$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$c_{3n+1} = \frac{c_{3n-2}}{3n(3n+1)} \quad \text{and so}$$

$$c_4 = \frac{c_1}{3\cdot 4} = \frac{c_1}{12}$$

$$c_7 = \frac{c_4}{6\cdot 7} = \frac{c_1}{3\cdot 4\cdot 6\cdot 7} = \frac{c_1}{504}$$

$$c_{10} = \frac{c_7}{9\cdot 10} = \frac{c_1}{504\times 90} = \frac{c_1}{45360}$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

Two independent solutions are

$$u(x) = c_0\left[\frac{1}{6}x^3 + \frac{1}{180}x^6 + \frac{1}{12960}x^9 + \cdots\right]$$

and $\quad v(x) = c_1\left[x + \dfrac{1}{12}x^4 + \dfrac{1}{504}x^7 + \dfrac{1}{45360}x^{10} + \cdots\right]$

So the general power series solution is

$$y = u(x) + v(x)$$

Example 36. Find the series solution of

$$(x^2 - 2x)\,y'' + 5(x-1)\,y' + 3y = 0 \tag{1}$$

around the ordinary point $x = 1$

Solution. Let $y = \sum_{n=0}^{\infty} c_n (x-1)^n$ be a solution of (1). Then

$$y' = \sum_{n=0}^{\infty} n\,c_n (x-1)^{n-1},$$

$$y'' = \sum_{n=0}^{\infty} n(n-1)\,c_n (x-1)^{n-2}$$

Substituting these values of y, y' and y'' into (1), we have

$$(x^2-2x)\sum_{n=0}^{\infty} n(n-1)\,c_n (x-1)^{n-2} + 5(x-1)\sum_{n=0}^{\infty} n\,c_n (x-1)^{n-1} + 3\sum_{n=0}^{\infty} c_n (x-1)^n = 0$$

or

$$[(x-1)^2 - 1]\sum_{n=0}^{\infty} n(n-1)\,c_n (x-1)^{n-2} + 5(x-1)\sum_{n=0}^{\infty} n c_n (x-1)^{n-1}$$

$$+\,3\sum_{n=0}^{\infty} c_n (x-1)^n = 0$$

or

$$\sum_{n=0}^{\infty} n(n-1)\,c_n (x-1)^n + 5\sum_{n=0}^{\infty} n c_n (x-1)^n + 3\sum_{n=0}^{\infty} c_n (x-1)^n$$

$$-\sum_{n=2}^{\infty} n(n-1)\,c_n (x-1)^{n-2} = 0$$

or

$$\sum_{n=0}^{\infty} [n(n-1) + 5n + 3]\,c_n (x-1)^n - \sum_{n=0}^{\infty} (n+2)(n+1)\,c_{n+2} (x-1)^n = 0$$

Equating the coefficients of $(x-1)^n$, we have

$$(n^2 + 4n + 3)\,c_n = (n+2)(n+1)\,c_{n+2}$$

or

$$c_{n+2} = \frac{n+3}{n+2}\,c_n \tag{2}$$

is the recursion relation for finding the coefficients. From (2), we get

$$c_{2n} = \frac{2n+1}{2n}\,c_{2n-2} \tag{3}$$

by taking $n + 2 = 2k$ so that $n = 2k - 2$ and replacing k by n in the resulting expression involving k.

and $\quad c_{2n+1} = \dfrac{2n+2}{2n+1}\,c_{2n-1} \qquad (4)$

Setting $n = 1, 2, 3, \cdots$ into (3) and (4), we obtain

$$c_2 = \frac{3}{2}\,c_0$$

$$c_4 = \frac{5}{4}\,c_2 = \frac{5\cdot 3}{4\cdot 2}\,c_0$$

$$c_6 = \frac{7}{6}\, c_4 = \frac{7\cdot 5\cdot 3}{6\cdot 4\cdot 2}\, c_0$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$c_{2n} = \frac{3\cdot 5\cdot 7 \cdots (2n+1)}{2\cdot 4\cdot 6 \cdots 2n}\, c_0$$

and $c_3 = \frac{4}{3}\, c_1$

$$c_5 = \frac{6}{5}\, c_3 = \frac{6\cdot 4}{5\cdot 3}\, c_1$$

$$c_7 = \frac{8}{7}\, c_5 = \frac{8\cdot 6\cdot 4}{7\cdot 5\cdot 3}\, c_1$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$c_{2n+1} = \frac{4\cdot 6\cdot 8 \cdots (2n+2)}{3\cdot 5\cdot 7 \cdots (2n+1)}\, c_1$$

Required solution is

$$y = c_0 + c_0\left[\frac{3}{2}(x-1)^2 + \frac{3\cdot 5}{2\cdot 4}(x-1)^4 + \cdots\right] + c_1(x-1) + c_1\left[\frac{4}{3}(x-1)^3 + \frac{6\cdot 4}{5\cdot 3}(x-1)^5 + \cdots\right]$$

$$= c_0 + c_0 \sum_{n=1}^{\infty} \frac{3\cdot 5\cdot 7 \cdots (2n+1)}{2\cdot 4\cdot 6 \cdots 2n}(x-1)^{2n} + c_1(x-1) + c_1 \sum_{n=1}^{\infty} \frac{4\cdot 6\cdot 8 \cdots (2n+2)}{3\cdot 5\cdot 7 \cdots 2n+1}(x-1)^{2n+1}$$

Example 37. Find the series solution of

$$xy'' + y\sin x = 0$$

around the ordinary point $x = 0$.

Solution. The given equation in the standard form is

$$y'' + \frac{\sin x}{x}\, y = 0 \qquad (1)$$

Here
$$Q(x) = \frac{\sin x}{x} = \frac{1}{x}\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right) = 1 - \frac{x^2}{6} + \frac{x^4}{120} - \cdots$$

and this series converges for all x. Thus every value of x is an ordinary point and in particular $x = 0$ is an ordinary point.

Let $y = \sum_{n=0}^{\infty} c_n x^n$ be a solution of (1)

Then, on substitution of the values of y and y'' into (1), we have

$$\sum_{n=2}^{\infty} n(n-1)\, c_n x^{n-2} + \left(1 - \frac{x^2}{6} + \frac{x^4}{120} - \frac{x^6}{720} + \cdots\right)\left(c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 + \cdots\right) = 0$$

or
$$2c_2 + 6c_3 x + 12c_4 x^2 + 20c_5 x^3 + 30c_6 x^4 + \cdots + c_0 + c_1 x + \left(c_2 - \frac{c_0}{6}\right)x^2 + \left(-\frac{c_1}{6} + c_3\right)x^3 + \left(c_4 + \frac{c_0}{120} - \frac{c_2}{6}\right)x^4 + \cdots = 0$$

Equating coefficients of like terms, we get

$$2c_2 + c_0 = 0 \quad \text{or} \quad c_2 = -\frac{1}{2}\, c_0$$

$$6c_3 + c_1 = 0 \quad \text{or} \quad c_3 = -\frac{1}{6}\, c_1$$

$$12c_4 + c_2 - \frac{c_0}{6} = 0$$

or $$12c_4 = -c_2 + \frac{c_0}{6} = \frac{1}{2}\, c_0 + \frac{c_0}{6} = \frac{2}{3}\, c_0 \quad \text{or} \quad c_4 = \frac{1}{18}\, c_0$$

$$20c_5 - \frac{c_1}{6} + c_3 = 0$$

or $$20c_5 = \frac{c_1}{6} - c_3 = \frac{c_1}{6} + \frac{c_1}{6} = \frac{c_1}{3} \quad \text{or} \quad c_5 = \frac{c_1}{60}$$

$$30c_6 + c_4 + \frac{c_0}{120} - \frac{c_2}{6} = 0$$

or $$30c_6 = -c_4 - \frac{c_0}{120} + \frac{c_2}{6} = -\frac{1}{18}\, c_0 - \frac{1}{120}\, c_0 - \frac{1}{12}\, c_0 = -\frac{53}{360}\, c_0$$

Required solution is

$$y = c_0 - \frac{1}{2} c_0 x^2 + \frac{1}{18} c_0 x^4 - \frac{53}{360} c_0 x^6 + \cdots$$
$$+ c_1 x - \frac{1}{6} c_1 x^3 + \frac{1}{60} c_1 x^5 + \cdots$$
$$= c_0 \left(1 - \frac{1}{2} x^2 + \frac{1}{18} x^4 - \frac{53}{360} x^6 + \cdots\right) + c_1 \left(x - \frac{1}{6} x^3 + \frac{1}{60} x^5 + \cdots\right)$$

Example 38. Find the power series solution of the initial value problem

$$x(2 - x)\, y'' - 6(x - 1)\, y' - 4y = 0, \quad y(1) = 1, \quad y'(1) = 0$$

Solution. Since the initial conditions are given at $x = 1$, we seek a power series solution of the form $\quad y = \sum_{n=0}^{\infty} c_n (x - 1)^n$.

Substituting into the given equation for y, y' and y'', we have

$$[1 - (x - 1)^2] \sum_{n=0}^{\infty} n(n - 1) c_n (x - 1)^{n-2} - 6(x - 1) \sum_{n=0}^{\infty} nc_n (x - 1)^{n-1}$$
$$- 4 \sum_{n=0}^{\infty} c_n (x - 1)^n = 0$$

or
$$\sum_{n=2}^{\infty} n(n - 1) c_n (x - 1)^{n-2} - \sum_{n=2}^{\infty} n(n - 1) c_n (x - 1)^n - 6 \sum_{n=1}^{\infty} nc_n (x - 1)^n$$
$$- 4 \sum_{n=0}^{\infty} c_n (x - 1)^n = 0$$

or
$$\sum_{n=0}^{\infty} (n + 1)(n + 2) c_{n+2} (x - 1)^n - \sum_{n=2}^{\infty} n(n - 1) c_n (x - 1)^n - 6 \sum_{n=1}^{\infty} nc_n (x - 1)^n$$
$$- 4 \sum_{n=0}^{\infty} c_n (x - 1)^n = 0$$

or
$$2c_2 + 6c_3(x - 1) + \sum_{n=2}^{\infty} (n + 1)(n + 2) c_{n+2} (x - 1)^n - \sum_{n=2}^{\infty} n(n - 1) c_n (x - 1)^n$$
$$- 6c_1 (x - 1) - 6 \sum_{n=2}^{\infty} nc_n (x - 1)^n - 4c_0 - 4c_1 (x - 1) - 4 \sum_{n=2}^{\infty} c_n (x - 1)^n = 0$$

or
$$(2c_2 - 4c_0) + (6c_3 - 10c_1)(x - 1) + \sum_{n=2}^{\infty} \Big[(n + 2)(n + 1) c_{n+2} - n(n - 1) c_n$$
$$- 6nc_n - 4c_n\Big] (x - 1)^n = 0$$

Equating coefficients of like powers of $x - 1$, we have

Coeff of $(x - 1)^0$: $\quad 2c_2 - 4c_0 = 0 \quad$ or $\quad c_2 = 2c_0 = 2$,

since from initial condition, $y(1) = c_0 = 1$

Coeff of $(x - 1)$: $\quad 6c_3 - 10c_1 = 0 \quad$ or $\quad c_3 = \frac{10}{6} c_1 = 0$, since $c_1 = y'(1) = 0$

Coeff of $(x - 1)^n$: $\quad (n + 2)(n + 1)c_{n+2} + (-n^2 + n - 6n - 4) c_n = 0, \; n \geq 2$

i.e.,
$$c_{n+2} = \frac{n^2 + 5n + 4}{(n + 2)(n + 1)} c_n = \frac{n + 4}{n + 2} c_n$$

Therefore,
$$c_4 = \frac{6}{4} c_2 = 3$$
$$c_5 = \frac{7}{6} c_3 = 0 = c_7 = \cdots = c_{2n+1}$$
$$c_{2n+2} = \frac{2n + 4}{2n + 2} c_{2n} = \frac{n + 2}{n + 1} c_{2n} \qquad n = 0, 1, 2, \cdots$$
$$c_6 = \frac{4}{3} c_4 = \frac{4}{3} \cdot 3 = 4$$
$$c_8 = \frac{5}{4} c_6 = 5$$
$$c_{10} = \frac{6}{5} c_8 = 6$$
$$\vdots \quad \vdots \qquad \vdots$$
$$c_{2n} = \frac{n + 1}{n} c_{2n-2} = n + 1.$$

Required solution is

$$y = 1 + 2(x - 1)^2 + 3(x - 1)^4 + 4(x - 1)^6 + \cdots + (n + 1)(x - 1)^{2n} + \cdots$$
$$= \sum_{n=0}^{\infty} (n + 1 (x - 1)^{2n}.$$

EXERCISE 10.10

Find the series solution of each of the following differential equations around the indicated ordinary point (**Problems 1 – 14**)

1. $(x^2-1)y''+4xy'+2y=0$, around $x=0$
2. $y''-x^2y=0$, around $x=0$
3. $y''-x^3y=0$, around $x=0$
4. $y''+\dfrac{3x}{1+x^2}\,y'+\dfrac{1}{1+x^2}\,y=0$, around $x=0$
5. $y''+xy'+(x^2+2)y=0$, around $x=0$
6. $y''-4xy'-4y=4+6x$, around $x=0$
7. $y''-2x^2y'+4xy=x^2+2x+4$, around $x=0$
8. $(1-x^2)y''-2xy'+m(m+1)y=0$, (Legendre's equation) around $x=0$
9. $y''+(\alpha+\beta\cos 2x)y=0$, (Mathieu's equation) around $x=0$
10. $y''+x^3y'+3x^2y=e^x$, around $x=0$
11. $y''+(x^2-1)y=0$, around $x=1$
12. $y''+(x-3)y'+y=0$, around $x=2$
13. $y''+xy'+(\ln x)y=0$, around $x=1$
14. $y''+(\sin x)y=0$, around $x=\frac{\pi}{2}$

Solve the initial value problems using power series method:

15. $y''+xy'+2y=0$, $y(0)=4$, $y'(0)=-1$
16. $y''+e^xy=0$, $y(0)=1$, $y'(0)=-1$
17. $(x^2+1)y''+xy'+2xy=0$, $y(0)=2$, $y'(0)=3$
18. $y''+(x^2+2x+1)y'-(4+4x)y=0$, $y(-1)=0$, $y'(-1)=1$
19. $y''+xy=e^{2x}$, $y(0)=0$, $y'(0)=1$
20. $xy''+y'+2y=0$, $y(1)=2$, $y'(1)=4$.

APPLICATIONS OF D.E.

Differential equations can be used in the solutions of many problems in physical, biological and social sciences. In this section, we shall study such problems which can be formulated in terms of first and second order differential equations. The solution of the differential equation will lead to the solution of the given problem.

The following examples illustrate the techniques.

Example 39. A stone weighting 8 lbs falls from rest toward the earth from a great height. As it falls it is acted upon by air resistance that is numerically $\frac{1}{2}v$ (in pounds) where v is the velocity (in feet per second). Find the velocity and distance fallen at time t.

Solution. We use Newton's second law $F=ma$ to formulate mathematical model of this problem.

The forces acting on the body are:

(i) F_1, the weight of the stone acting downward and hence is positive.

(ii) F_2, the air resistance, numerically equal to $\frac{1}{2}v$, which acts upward and is therefore negative.

Newton's second law becomes

$$m\frac{dv}{dt}=F_1+F_2$$

i.e., $\quad \frac{1}{4}\frac{dv}{dt}=8-\frac{1}{2}v,\quad$ taking $g=32$.

or $\quad \frac{dv}{dt}=32-2v$

or $\quad \frac{dv}{dt}+2v=32$

which is a linear equation. We now solve this differential equation.

Here

$$\text{I.F.}=e^{\int 2dt}=e^{2t}$$

and so, $\quad \frac{d}{dt}(ve^{2t})=32e^{2t}$

Integrating, we get

$ve^{2t}=16e^{2t}+c \quad$ or $\quad v=16+ce^{-2t}$.

Now $\quad v(0)=0$.

Hence $\quad 0=16+c \quad$ or $\quad c=-16$

Therefore, $v = 16\{1 - e^{-2t}\}$ (1)

is the velocity after time t.

(1) may be written as

$$\frac{dx}{dt} = 16(1 - e^{-2t}). \qquad (2)$$

From (2), we have,

$$dx = 16(1 - e^{-2t})\, dt$$

or $\quad x = 16t + 8e^{-2t} + k.$

Applying the initial condition $x(0) = 0$, we find $k = -8$.

Hence

$$x = 16t + 8e^{-2t} - 8$$

is the distance fallen after time t.

Example 40. A body of constant mass is projected upward from the earth's surface with an initial velocity v_0. Assuming there is no air resistance, but taking into consideration the variation of the earth's gravitational field with altitude, find the smallest initial velocity for which the body will not return to the earth (This is the so called **escape velocity**).

Solution. The general expression for the weight $w(x)$ of a body of mass m is obtained from Newton's inverse-square law of gravitational attraction. If R is the radius of the earth and x is the altitude above sea level, then $w(x) = \frac{k}{(R+x)^2}$, k being constant. At $x = 0$, $w = mg$, hence $k = mgR^2$ and $w(x) = \frac{mgR^2}{(R+x)^2}$. The only force acting on the body is its weight which acts downward. Thus the equation of motion is

$$m\frac{dv}{dt} = \frac{-mgR^2}{(R+x)^2}$$

or $$v\frac{dv}{dx} = \frac{-mgR^2}{(R+x)^2}.$$

Separating the variables and integrating, we have

$$\frac{1}{2}v^2 = \frac{gR^2}{R+x} + c.$$

But $v = v_0$ at $t = 0$ i.e., at $x = 0$.

Therefore, $c = \frac{1}{2}v_0^2 - gR$

and so $$v^2 = v_0^2 - 2gR + \frac{2gR^2}{R+x}. \qquad (1)$$

The escape velocity is found by requiring that v given by (1) remains positive for all (positive) value of x. Thus we must have

$$v_0^2 \geq 2gR.$$

Hence the escape velocity is

$$v_0 = \sqrt{2gR} = 6.9 \text{ miles/sec. taking } R = 4000 \text{ miles.}$$

Example 41. The radioactive isotopes thorium 234 disintegrates at a rate proportional to the amount present. It is found that in one week 17.06 % of this material has disintegrated. Find an expression for the amount of material at any time. Also determine how long will it take for one half of this material to disintegrate?

Solution. Let y be the amount of thorium 234 present at any time t (t in days). Let $y = y_0$ at $t = 0$. We have

$$\frac{dy}{dt} = -ky \qquad (1)$$

$$y(0) = y_0, \quad y(7) = 0.8204\, y_0,$$

k being constant of proportionality.

Solution of (1) is

$$y = ce^{-kt}$$

At $\quad t = 0, \quad c = y_0.$

Therefore, $\quad y = y_0 e^{-kt}.$

Now $\quad y(7) = 0.8204\, y_0 = y_0 e^{-7k}$

or $\quad e^{-7k} = 0.8204$

or $\quad -7k = \ln 0.8204$

or $\quad k = \frac{-\ln 0.8204}{7} = -0.02828.$

Hence, $\quad y = y_0 e^{-0.02828t} \qquad (2)$

gives the value of y at any time.

Now we want to find t when $y = 0.5\, y_0$. This is obtained from (2) by

$$0.5\, y_0 = y_0 e^{-0.02828t}$$

or $\quad e^{0.02828t} = 0.5$

or $\quad -0.02828t = \ln 0.5$

$$= -\ln 2 = -0.6951$$

or $\quad t = \frac{0.66951}{0.02828} = 24.5$ days.

Example 42. A tank contains x_0 kg of salt dissolved in 200 litres of water. Starting at time $t = 0$ water containing 1/2 kg of salt per litre enters the tank at the rate of 4 litres/min and the well-stirred solution leaves the tank at the same rate. Find the concentration of salt in the tank at any time $t > 0$.

Solution. Let x denote the amount of salt in the tank at time t. Then the rate of change of the salt in the tank at time t is equal to the rate at which salt enters the tank minus the rate at which it leaves the tank.

The rate at which salt enters the tank is

$$\left(\frac{1}{2}\text{ kg/litre}\right)(4 \text{ litres/min}) = 2 \text{ kg/min}$$

The rate at which salt leaves the tank is

$$\left(\frac{x}{200}\text{ kg/litre}\right)(4 \text{ litres/min}) = \frac{x}{50}\text{ kg/min}$$

Hence we have

$$\frac{dx}{dt} = 2 - \frac{x}{50}$$

or

$$\frac{dx}{dt} + \frac{x}{50} = 2$$

$$\text{I.F.} = e^{\int \frac{1}{50} dt} = e^{0.02\,t}$$

Therefore,

$$\frac{d}{dt}\left(xe^{0.02\,t}\right) = 2e^{0.02\,t}$$

Integrating, we get

$$xe^{0.02\,t} = \int 2e^{0.02\,t}\,dt + c$$

$$= \frac{2}{0.02}\,e^{0.02\,t} + c$$

or $\quad x = 100 + c\,e^{-0.02\,t}$

When $\quad t = 0, \quad x = x_0$. Thus

$x_0 = 100 + c \quad$ i.e., $\quad c = x_0 - 100.$

Hence $\quad x = 100 + (x_0 - 100)\,e^{-0.02\,t}$

$$= x_0 e^{-0.02\,t} + 100\left(1 - e^{-0.02\,t}\right).$$

The concentration $c(t)$ of salt in the tank is given by

$$c(t) = \frac{x}{200} = \frac{x_0\,e^{-0.02\,t}}{200} + \frac{1}{2}\left(1 - e^{-0.02\,t}\right)\text{ kg/litre.}$$

Example 43. A litre of ice cream at a temperature of $-15°C$ is removed from the deep freezer and placed in a room where the temperature is $20°C$. If after 15 minutes the temperature of the ice cream is $-10°C$, how long will it take the ice cream to reach a temperature of $0°C$?

Solution. Newton's law of cooling states that the temperature of an object changes at a rate proportional to the difference between the temperatures of the object and its surroundings. Thus if $\theta(t)$ is temperature of the ice cream at time t and T is the temperature of the room, then

$$\frac{d\theta}{dt} = -k(\theta - T), \qquad (1)$$

where $k > 0$ is constant of proportionality.

From (1), we have

$$\frac{d\theta}{\theta - T} = -k\,dt.$$

Integrating, we get

$$\ln(\theta - T) = -kt + \ln c$$

or $\quad \theta - T = ce^{-kt}$

At $t = 0$, let $\theta = \theta(0)$, then

$$\theta(0) - T = c.$$

Therefore,

$$\theta - T = (\theta(0) - T)\,e^{-kt}. \qquad (2)$$

To determine k we need one more condition. At $t = t_1$, let $\theta = \theta(t_1)$. Then from (2),

$$\theta(t_1) = T + (\theta(0) - T)\,e^{-kt_1}$$

or

$$\frac{\theta(t_1) - T}{\theta(0) - T} = e^{-kt_1}$$

or

$$k = -\frac{1}{t_1}\ln\frac{\theta(t_1) - T}{\theta(0) - T}$$

At $\quad t = 0, \qquad \theta(0) = -15.$

At $\quad t = t_1 = 20, \qquad \theta(20) = -10.$

Therefore,

$$k = -\frac{1}{20}\ln\frac{-10 - 20}{-15 - 20}$$

$$= -\frac{1}{20}\ln\frac{6}{7} \text{ which is the value of the constant of proportionality.}$$

We need t such that $\theta(t) = 0$.

From (2), we have

$$0 = T + (\theta(0) - T)\,e^{-kt}$$

or $$0 = 20 + (-35)\,e^{-kt}, \text{ since } T = 20,\ \theta(0) = -15$$

or $$e^{-kt} = \frac{20}{35}$$

or $$-kt = \ln\frac{4}{7}$$

i.e., $$t = -\frac{1}{k}\ln\frac{4}{7} = \frac{20}{\ln\frac{6}{7}}\ln\frac{4}{7}$$

$$= \frac{20(\ln 4 - \ln 7)}{\ln 6 - \ln 7} = \frac{20(1.3863 - 1.9459)}{1.7928 - 1.9459}$$

$$= \frac{20(-0.5596)}{-0.1541} = 72.63 \text{ min.}$$

Thus the ice cream reaches a temperature of 0°C after 72.63 min. of its removal from the deep freezer.

Example 44. A car, moving at a certain velocity and with constant acceleration, is applied brakes to make it stop. The car stops 10 seconds after the brakes are applied and travels 300 meters during this time. Find the law of motion of the car during this 10 second interval. Also find the distance covered, the speed and acceleration of the car at the moment the brakes are applied.

Solution. Suppose that the car is represented by a particle moving in a straight line at the time brakes are applied. We have the following initial conditions.

$$s = 0 \text{ at } t = 0,\quad s = 300 \text{ at } t = 10,\quad v = 0 \text{ at } t = 10$$

Now the particle (car) is moving at a constant acceleration. So

$\dfrac{dv}{dt} = a =$ constant. This is a first order separable differential equation. Hence

$$v = at + c_1 \tag{1}$$

Also $\dfrac{ds}{dt} = v = at + c_1$ which, again, is a first order separable differential equation. The solution of this differential equation is

$$s = \frac{1}{2}at^2 + c_1 t + c_2 \tag{2}$$

Now $s = 0$ at $t = 0$. So $c_2 = 0$. Next, at $t = 10$, $v = 0$, so, from (1), $c_1 = -10a$. Hence (2) becomes

$$s = \frac{1}{2}at^2 - 10at \tag{3}$$

At $s = 300$, $t = 10$, so from equation (3), we have

$$300 = \frac{1}{2}a \times 100 - 10a \times 10$$

$$= -50a$$

So the initial acceleration is $a = -6$ m/sec^2

Putting this value of a into (3), we get

$$s = 60t - 3t^2 \tag{4}$$

as the law of motion.

Initial speed of the car $= v = -10a = 60$ m/sec.

The distance covered by the car before coming to rest is (with $t = 10$ seconds)

$$s = 600 - 300 \text{ m} = 300 \text{ m}$$

Example 45. A projectile is fired from a platform 5 m above the ground, with an initial velocity of 250 m/sec. The only force affecting the projectile, during its flight, is taken to be the gravity which is equivalent to a force having a downward acceleration of 9.8 m/sec^2. Find the equation which gives the projectile's height above the ground as a function of the time t with $t = 0$ when the projectile is fired. Find also the height of the projectile from the ground after 5 seconds. After how many seconds the velocity will be zero and what will be height at that time?

Solution. Let s denote the height of the projectile after t seconds of its flight, v its velocity at time t and a its acceleration. The quantities a, v and s are related by the following derivatives

$$\frac{ds}{dt} = v,\quad \frac{dv}{dt} = a,\quad \frac{d^2s}{dt^2} = a$$

Since the gravitational acceleration acts in the direction of decreasing s, we have the following differential equation

$$\frac{dv}{dt} = -9.8$$

or $$\frac{d^2s}{dt^2} = -9.8 \tag{1}$$

with initial conditions

$$v = \frac{ds}{dt} = 0,\quad s = 5,\quad t = 0.$$

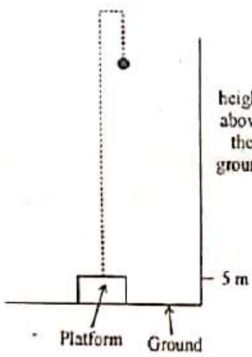


Figure 10.1

The second order differential equation (1) can to be integrated twice to give v and s as :

$$v = \frac{ds}{dt} = -9.8\,t + c_1 \qquad (2)$$

$$s = -9.8\,\frac{t^2}{2} + c_1 t + c_2 \qquad (3)$$

As $v = 250$ m/sec when $t = 0$, $c_1 = 250$ from (2). Also, at $t = 0$, $s = 5$, so that, from (3), we have

$$5 = s = 0 + c_2 \qquad \text{so that} \qquad c_2 = 5$$

Hence (3) becomes

$$s = -44.9\,t^2 + 250\,t + 5 \qquad (4)$$

To find the height of the projectile after 5 seconds, we have, from (4), when $t = 0$,

$$s = (-4.9 \times 25 + 250 \times 5 + 5)\text{ m}$$

$$= (-122.5 + 1250 + 5)\text{ m} = 1132.5\text{ m}$$

Also from (2), the velocity after 5 seconds is

$$v = (-9.8 \times 5 + 250)\text{ m/sec}$$

$$= (-49.0 + 250)\text{ m/sec} = 201\text{ m/sec.}$$

To find the time taken for the velocity to be zero, we use equation (2) with $c_1 = 250$. This equation gives

$$0 = -9.8\,t + 250$$

so that $$t = \frac{250}{9.8} = 25.51 \text{ seconds.}$$

At $t = 25.5$, the distance covered is (from (4))

$$s = (-4.9) \times (25.5)^2 + 250 \times (25.5) + 5 = 3193\text{ m.}$$

Thus the height is 3193 m when $v = 0$.

(10.30) The Harmonic Oscillator Equation. In the real world, there are a number of quantities that oscillate or vibrate in a uniform manner, repeating themselves periodically in definite intervals of time. Some of the examples are alternating electric currents, sound waves, light waves, radio waves, electromagnetic waves, pendulums, mass-spring systems, human heartbeat, periodic variation of the population of a plant or animal species.

The simples mathematical model for such quantities is

$$y = A\cos(\omega t - \phi) \qquad (1)$$

where y is the oscillating quantity, t is time, usually measured in seconds, A is a positive constant, called **amplitude** of the oscillation, ω is a positive constant called the **angular frequency** of the oscillation and ϕ is a constant called the **phase angle**.

For each equation of the type (1), the constant ν given by

$$\nu = \frac{w}{2p}$$

is called the **frequency** of the oscillation. If t is measured in seconds, then ν represents the number of complete oscillations, called **cycles** per second. One cycle per second is called hertz (or Hz). The quantity T defined by

$$T = \frac{1}{\nu}$$

is called the **period of the oscillation.**

A quantity y oscillating in accordance with the equation (1) is said to be undergoing a **simple harmonic oscillation.** Also, any physical device which produces a quantity that undergoes simple harmonic oscillation is called a **harmonic oscillator.**

The **harmonic oscillator equation**

$$\frac{d^2y}{dt^2} + \omega^2 y = 0 \qquad (2)$$

is obtained from (1) by differentiating the equation (1) twice. Hence (1) is a solution of the differential equation (2).

The differential equation (2) is a second order linear differential equation and can be solved by the usual techniques of solving such equations.

Example 46. (Mass-Spring System). Suppose that a mass m is suspended by a perfectly elastic spring with spring constant k. The mass of the spring and the air friction are neglected. A vertical axis y is set up so that when the mass and the spring are hanging in equilibrium, the mass is lifted to the position with y-coordinate $y = A_0$ and is released at time $t = 0$ with the initial velocity $v_0 = 0$. Discuss the motion of the particle of mass m and find

(i) the equation of motion and (ii) the frequency of the oscillation.

Solution. The mass moves up and down between A_0 and $-A_0$. Suppose that, at time t, the velocity of m is given by

$$v = \frac{dy}{dt}$$

and its acceleration is

$$a = \frac{dv}{dt} = \frac{d^2y}{dt^2}.$$

By Hooke's law, the unbalanced force F applied on a mass m by the spring is given by $F = -ky$

So, by Newton's second law of motion, we have

$$F = ma = -ky,$$

so that $$a + \frac{k}{m}y = 0$$

i.e., $$\frac{d^2y}{dt^2} + \frac{k}{m}y = 0 \qquad (1)$$

is the differential equation of the vibrating mass

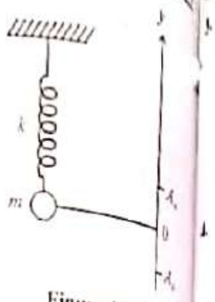


Figure 10.2

Let $\frac{k}{m} = \omega^2$ so that $\omega = \sqrt{\frac{k}{m}}$ Then (1) can be written as

$$\frac{d^2y}{dt^2} + \omega^2 y = 0.$$

The general solution of this differential equation is

$$y = A_1 \cos \omega t + B_1 \sin \omega t = A \cos(\omega t - \phi).$$

When $t = 0, y(0) = A_0, A = \sqrt{A_0^2 + 0^2} = A_0, \phi = 0$ so that the equation of motion is

$$y = A_0 \cos \omega t = A_0 \cos \sqrt{\frac{k}{m}}\, t$$

The frequency of the oscillation is

$$v = \frac{w}{2p} = \frac{1}{2p}\sqrt{\frac{k}{m}}.$$

EXERCISE 10.11

1. A ball weighting $\frac{3}{4}$ lb is thrown vertically upward from a point 6 ft above the surface of the earth with an initial velocity of 20 ft/sec. As it rises it is acted upon by air resistance that is numerically equal to $\frac{1}{64}$ v, where v is the velocity in feet per sec. How high will the ball rise?

2. A bullet weighing 1 oz is fired vertically downward from a stationary helicopter with a muzzle velocity of 1200 ft/sec. The air resistance (in pounds) is numerically equal to $10^{-4} v^2$, where v is the velocity (in feet per second). Find the velocity of the bullet as a function of the time.

3. A parachutist weighting 192 lbs (including equipment) falls from rest toward the earth. Before the parachute opens, the air resistance equals $\frac{3}{4}$ v. The parachute opens 10 sec. after the fall begins and the air resistance is then $\frac{3}{4}v^2$, where v is the velocity in feet per second. Find the velocity of the parachutist (i) when the parachute opens (ii) after the parachute opens

4. The population of Pakistan increases at a rate proportional to the number of inhabitants present at any time t. If the population doubles in 40 years, in how many years will it triple?

5. Assume that the population of the earth changes at a rate proportional to the current population. It is estimated that at time $t = 0$ (1650 C.E.) the earth's population was 600 million, at $t = 300$(1950 C.E.) its population was 2.8 billion. Find an expression giving the population of the earth at any time. Assuming that the maximum population the earth can support is 25 billion, when will this limit be reached?

6. A newly built fish farm is stocked with 400 fish at time $t = 0$ (month). Thereafter the population increases at the rate of $\sqrt{P}$ per month when there are P fish in the farm. What is fish population at time t ?

7. A 1000/litre tank contains a mixture of water and chlorine. Fresh water is pumped in at a rate of 6 litres per second. The fluid is well stirred and pumped out at a rate of 8 litres per second. If the initial concentration of chlorine is 0.02 grams per litre, find the amount of chlorine in the tank as a function of t

8. A tank contains 200 litres of toxic solution containing 20 kg of pollutant. Fresh water is poured into the tank at the rate of 8 litres/min and the well stirred mixture leaves the tank at the same rate. Find an expression for the amount of pollutant in the tank at any time t

9. A coffee cup has a temperature of 100°C when freshly poured and one minute later has cooled to 95°C in a room at 25°C. Determine when the coffee reaches a temperature of 70°C?

10. A steel bar at a temperature of 110°C is moved to a room where the constant temperature is 10°C. After one hour, the temperature of the bar is 60°C. How much time is required for it to reach a temperature of 30°C?

11. The marginal cost for producing x units of a product is given by

$$\frac{dC}{dt} = \frac{360}{\sqrt{x}} \text{ rupees per unit of the product}$$

Find the cost C of manufacturing x units of the product if the fixed cost is Rs. 48,000 when $x = 1600$ units

12. Suppose that a certain disease in Pakistan spreads in such a way that the rate of change of the infected people varies as the number of infected people. If the number of cases of the disease in a given year is reduced by 25 %, in how many years the present 75000 infected cases will reduce to 3000 cases only?

13. Bacteria grown in a certain culture increase at a rate proportional to the number of bacteria present. If there are 4000 bacteria present initially and if the number of bacteria triples in half an hour, how many bacteria are present after t hours? How many are present after 2 hours? After how many hours bacteria will grow to one million?

14. It is observed that the rate at which a solid substance dissolves in fluid varies directly as the product of the amount of undissolved solid present in the solvent and the difference between the saturation concentration and the instantaneous concentration of the substance. If 20 kg of soluble substance is put into a tank containing 120 kg of solvent and, at the end of 12 minutes the concentration is observed to be 1 part in 30, find the amount of soluble substance in the solution at any time t if the saturation concentration is 1 part of soluble substance in 3 parts of solvent.

15. The temperature of a machine, when it is first shut down after operating, is 220°C and temperature of the surrounding air is 30°C. After 20 minutes, the temperature of the machine is 160°C. Find a function that gives the temperature of the machine at any time t and then find the temperature of the machine 30 minutes after it is shut down.

16. A mass of 10 kg attached to a spring stretches it 0.5 m from its natural length. The system is set in motion by displacing the mass 0.1 m above the equilibrium position. Find the spring constant k and solve the initial value problem.

17. Establish the mass-spring equation

$$m\frac{d^2y}{dt^2} + c\frac{dy}{dt} + ky = F(t)$$

by taking into account damping due to air resistance on a mass m by a force proportional to the velocity and also a time dependent force $F(t)$ acting on the mass, where k is the spring constant and c is the damping coefficient.

18. A mass of 4 kg attached to a spring stretches it 0.2 m from its natural length. The damping coefficient is 10 kg/sec. and the system is set in motion from equilibrium with a downward initial velocity of 2 m/sec. Find the differential equation and the initial conditions satisfied by the system.

10

Chapter 11

THE LAPLACE TRANSFORM

The Laplace transform[1] is an efficient technique for solving linear differential equations with constant coefficients. We shall study its basic properties and will apply them to solve initial value problems. As the name suggests, Laplace Transform is an operator which transforms a function f of the variable t into a function F of the variable s.

It will be seen later that the Laplace transform of a function is a convergent improper integral. A necessary pre-requisite for the study of Laplace transform is familiarity with the convergence of improper integrals.

THE LAPLACE TRANSFORM

(11.1) Definition. **(Piecewise Continents Function)** A real-valued function f defined on an interval $[a, b]$ is said to be **piecewise continuous** in $[a, b]$ if there exists a partition

$$P = \{a = x_0, x_1, x_2, \cdots, x_{n-1}, x_n = b\}$$

of $[a, b]$ such that f is continuous in the interior of each subinterval $[x_i, x_{i+1}]$ and has finite one-sided limits $\lim\limits_{x \to x_i^+} f(x)$ and $\lim\limits_{x \to x_{i+1}^-} f(x)$ at the end points of each subinterval $(i = 0, 1, 2, \cdots, n-1)$. The one-sided limits $\lim\limits_{x \to x_i^+} f(x)$ and $\lim\limits_{x \to x_{i+1}^-} f(x)$ are finite but unequal so that f has a finite number of **jumps (ordinary) discontinuities at the points of**

1 Named for the eminent French astronomer, mathematician and physicist Pierre Simon Marquis de Laplace (1749 – 1827)

subdivision of $[a, b]$. Clearly, a continuous function is piecewise continuous and a piecewise continuous function is bounded. It is also known from calculus that if f is piecewise continuous on $[a, b]$, then $\int_a^b f(x)\,dx$ exists.

(11.2) Definition. Let f be a real-valued piecewise continuous function defined on $[0, \infty[$. The **Laplace transform** of f, denoted by $\mathscr{L}(f)$, is the function F defined by

$$F(s) = \int_0^\infty e^{-st} f(t)\,dt \tag{1}$$

provided the improper integral in (1) converges.

The domain of F is the set of all real numbers s for which the above integral converges.

Note that the operation transforms the given function f of the variable t into a new function F of the variable s and is written symbolically $F(s) = \mathscr{L}\{f(t)\}$ or simply $F = \mathscr{L}(f)$.

(11.3) Definition. If $F = \mathscr{L}\{f\}$ as in **11.2**, then the original function f is called the **inverse Laplace transform** of F and is denoted by $f = \mathscr{L}^{-1}\{F\}$. Clearly,

$$\mathscr{L}^{-1}\{F\} = \mathscr{L}^{-1}\{\mathscr{L}(f)\} = f.$$

Thus, if $\mathscr{L}\{f(t)\} = F(s)$ then $\mathscr{L}^{-1}\{F(s)\} = f(t)$

The inverse Laplace transforms will be useful for solving initial value problems.

Example 1. Let $f(t) = 1$ on $[0, \infty[$. Then

$$\mathscr{L}\{f\} = \mathscr{L}\{1\} = \int_0^\infty e^{-st}\,dt = \lim_{h\to\infty}\int_0^h e^{-st}\,dt$$

$$= \lim_{h\to\infty}\left[-\frac{e^{-st}}{s}\right]_0^h = \lim_{h\to\infty}\left[-\frac{e^{-hs}}{s} + \frac{1}{s}\right]$$

$$= \frac{1}{s} = F(s), \quad \text{provided } s > 0.$$

Example 2. Let $f(t) = t^n$, n being a positive integer. Evaluate $\mathscr{L}\{f(t)\}$.

Solution. Here $\mathscr{L}\{f(t)\} = \mathscr{L}\{t^n\} = \int_0^\infty e^{-st}\, t^n\,dt$

Integrating by parts, taking t^n as first function, we have

$$\mathscr{L}\{t^n\} = \left[t^n \cdot \frac{e^{-st}}{-s}\right]_0^\infty + \int_0^\infty nt^{n-1}\cdot\frac{e^{-st}}{s}\,dt$$

$$= -\left[\frac{t^n}{se^{st}}\right]_0^\infty + \frac{n}{s}\int_0^\infty t^{n-1}e^{-st}\,dt = \frac{n}{s}\cdot\mathscr{L}\{t^{n-1}\}.$$

since successive application of L' Hospital's rule to $\lim_{t\to\infty}\frac{t^n}{e^{st}}$ gives its value zero.

Hence
$$\mathscr{L}\{t^n\} = \frac{n}{s}\,\mathscr{L}\{t^{n-1}\}$$

$$= \frac{n}{s}\cdot\frac{n-1}{s}\,\mathscr{L}\{t^{n-2}\}$$

$$\vdots \quad \vdots \quad \vdots$$

$$= \frac{n}{s}\cdot\frac{n-1}{s}\cdot\ldots\cdot\frac{1}{s}\,\mathscr{L}\{1\}.$$

$$= \frac{n!}{s^n}\cdot\frac{1}{s}, \qquad \text{as } \mathscr{L}\{1\} = \frac{1}{s} \text{ by Example 1.}$$

$$= \frac{n!}{s^{n+1}} = F(s).$$

Example 3. Compute $\mathscr{L}\{e^{at}\}$, where a is a constant and $s \neq a$.

Solution. Here $f(t) = e^{at}$

$$\mathscr{L}\{f(t)\} = \mathscr{L}\{e^{at}\} = \int_0^\infty e^{-st}e^{at}\,dt$$

$$= \int_0^\infty e^{(a-s)t}\,dt = \lim_{h\to\infty}\int_0^h e^{(a-s)t}\,dt$$

$$= \lim_{h\to\infty}\left[\frac{e^{(a-s)t}}{a-s}\right]_0^h$$

$$= \lim_{h\to\infty}\left[\frac{e^{(a-s)h}}{a-s} - \frac{1}{a-s}\right]$$

$$= \begin{cases}\dfrac{1}{s-a} & \text{if } s > a\\ \infty & \text{if } s = a\end{cases}$$

Here $e^{-(s-a)h} \to 0$ as $h \to \infty$ and $s > a$, while $e^{-(s-a)h} \to \infty$ as $h \to \infty$ and $s < a$

When $a = s$, $f(t) = e^{st}$ and

$$\mathscr{L}\{f(t)\} = \int_0^\infty e^{-st}\, e^{st}\, dt = \int_0^\infty dt = (t)_0^\infty = \infty$$

Therefore, $\mathscr{L}\{e^{at}\} = \dfrac{1}{s-a}$, $s > a$

Example 4. Find the Laplace transforms of

(i) $\cos at$ (ii) $\sin at$

Solution. By definition

$$\mathscr{L}\{\cos at\} = \int_0^\infty e^{-st} \cos at\, dt$$

and $\mathscr{L}\{\sin at\} = \displaystyle\int_0^\infty e^{-st} \sin at\, dt.$

Therefore, for $i = \sqrt{-1}$,

$$\begin{aligned}
\mathscr{L}\{\cos at\} + i\,\mathscr{L}\{\sin at\} &= \int_0^\infty e^{-st} \cos at\, dt + i\int_0^\infty e^{-st} \sin at\, dt \\
&= \int_0^\infty e^{-st} (\cos at + i \sin at)\, dt \\
&= \int_0^\infty e^{-st} e^{iat}\, dt \\
&= \lim_{h \to \infty} \left[\frac{e^{(ia-s)t}}{ia-s}\right]_0^h \\
&= \lim_{h \to \infty} \left[\frac{e^{(ia-s)h}}{ia-s} - \frac{1}{ia-s}\right] \\
&= \begin{cases} \dfrac{1}{s-ia} & \text{if } s > 0 \\ \text{undefined} & \text{if } s < 0 \end{cases} \\
&= \frac{s+ia}{s^2+a^2} \qquad \text{if } s > 0
\end{aligned}$$

Here $\displaystyle\lim_{h \to \infty} \frac{e^{(ia-s)h}}{ia-s} = 0$ for $s > 0$ and is undefined for $s < 0$.

Equating real and imaginary parts, we get

(i) $\mathscr{L}\{\cos at\} = \dfrac{s}{s^2+a^2}$, $s > 0$

(ii) $\mathscr{L}\{\sin at\} = \dfrac{a}{s^2+a^2}$, $s > 0$

Note. $\mathscr{L}\{\cos at\}$ and $\mathscr{L}\{\sin at\}$ could also be evaluated directly by definition

$$\mathscr{L}\{\cos at\} = \int_0^\infty e^{-st} \cos at\, dt$$

Integrate twice by parts and the result will follow

Example 5. Consider the function f defined by $f(t) = \dfrac{1}{t}$. For the Laplace transform of $\dfrac{1}{t}$, we first check the convergence of $\displaystyle\int_0^\infty \frac{e^{-st}}{t}\, dt$.

$$\int_0^\infty \frac{e^{-st}}{t}\, dt = \int_0^1 \frac{e^{-st}}{t}\, dt + \int_1^\infty \frac{e^{-st}}{t}\, dt.$$

For $0 \le t \le 1$, we have $e^{-st} \ge e^{-s}$ if $s > 0$.

Therefore, $\displaystyle\int_0^\infty \frac{e^{-st}}{t}\, dt \ge \int_0^1 \frac{e^{-s}}{t}\, dt + \int_1^\infty \frac{e^{-st}}{t}\, dt.$

But

$$\begin{aligned}
\int_0^1 \frac{e^{-s}}{t}\, dt &= \frac{1}{e^s} \lim_{h \to 0} [\ln t]_h^1 \\
&= \frac{1}{e^s} \lim_{h \to 0} (\ln 1 - \ln h) \\
&= \infty, \text{ since } \ln h \to -\infty \text{ as } h \to 0
\end{aligned}$$

Hence $\displaystyle\int_0^1 \frac{e^{-st}}{t}\, dt$ also diverges to ∞. Consequently, $\displaystyle\int_0^\infty \frac{e^{-st}}{t}\, dt$ diverges and so by definition, $\mathscr{L}\left\{\dfrac{1}{t}\right\}$ does not exist.

It is obvious from this example that Laplace transforms exist only for certain class of functions. We shall later state and prove a theorem that guarantees the existence of Laplace transform of functions satisfying certain conditions. For that we need

(11.4) Definition. A function f defined $[0, \infty[$ is said to be of **exponential order** a as $t \to \infty$ if there exist real constants a, $M > 0$ and $T > 0$ such that

$$|f(t)| \leq Me^{at} \qquad \text{for } t \geq T$$

(11.5) Theorem. Let f be a piecewise continuous function defined on $[0, \infty[$. If f is of exponential order a as $t \to \infty$ then $\mathscr{L}\{f(t)\}$ exists for all $s > a$.

Proof. Since f is of exponential order a, there exist positive real numbers M and T such that

$$|f(t)| \leq Me^{at} \qquad , \qquad t \geq T \tag{1}$$

The theorem will be proved if we show that $\int_0^\infty e^{-st} f(t)\,dt$ converges.

Now

$$\int_0^\infty e^{-st} f(t)\,dt = \int_0^T e^{-st} f(t)\,dt + \int_T^\infty e^{-st} f(t)\,dt \tag{2}$$

Since f is piecewise continuous, the first integral on the right of (2) exists. Thus the convergence of $\int_0^\infty e^{-st} f(t)\,dt$ depends on the convergence of $\int_T^\infty e^{-st} f(t)\,dt$. But

$$\begin{aligned}
\left|\int_T^\infty e^{-st} f(t)\,dt\right| &\leq \int_T^\infty \left|e^{-st} f(t)\right| dt \\
&\leq \int_T^\infty e^{-st} Me^{at}\,dt, \qquad \text{by (1)} \\
&= M\int_T^\infty e^{(a-s)t}\,dt \\
&= M \lim_{h\to\infty}\left[\frac{e^{(a-s)t}}{a-s}\right]_T^h \\
&= \lim_{h\to\infty} M\left[\frac{e^{(a-s)h}}{a-s} - \frac{e^{(a-s)T}}{a-s}\right] \\
&= \begin{cases} M\left(0 - \dfrac{e^{(a-s)T}}{a-s}\right) & \text{if } a-s<0 \\ \infty & \text{if } a-s \geq 0 \end{cases}
\end{aligned}$$

Here $\lim\limits_{h\to\infty} \dfrac{e^{(a-s)h}}{a-s} = 0$ if $a-s<0$ and $\lim\limits_{h\to\infty} \dfrac{e^{(a-s)h}}{a-s} = \infty$ if $a-s \geq 0$.

Thus $\mathscr{L}\{f(t)\}$ exists for all $s > a$.

Note. The conditions stated in Theorem 11.5 are sufficient but not necessary. There exist functions which do not satisfy the hypothesis of **(11.5)** but still possess Laplace transforms.

Example 6. Consider the function f defined by

$$f(t) = t^{-\frac{1}{2}}$$

Clearly, f is not defined at $t = 0$, but it will be shown that $\mathscr{L}\left\{t^{-\frac{1}{2}}\right\}$ exists. By definition, we have

$$\mathscr{L}\left\{t^{-\frac{1}{2}}\right\} = \int_0^\infty e^{-st}\, t^{-\frac{1}{2}}\,dt \tag{1}$$

Let $st = x$. Then $s\,dt = dx$ or $dt = \frac{1}{s}\,dx$. So

$$t^{-\frac{1}{2}} = \left(\frac{x}{s}\right)^{-\frac{1}{2}} = \sqrt{\frac{s}{x}}$$

Substituting these values of t and dt into (1), we have

$$\begin{aligned}
\mathscr{L}\left\{t^{-\frac{1}{2}}\right\} &= \frac{1}{\sqrt{s}}\int_0^\infty e^{-x} x^{-\frac{1}{2}}\,dx \\
&= \frac{1}{\sqrt{s}}\,\Gamma\left(\frac{1}{2}\right) = \sqrt{\frac{\pi}{s}} \qquad \text{(from calculus)}
\end{aligned}$$

Here $\Gamma(t)$ (the **gamma function** of t) is defined by

$$\Gamma(t) = \int_0^\infty e^{-x}\, x^{t-1}\,dx \qquad \text{with} \qquad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Thus $\mathscr{L}\left\{t^{-\frac{1}{2}}\right\}$ exists.

PROPERTIES OF THE LAPLACE TRANSFORM

(11.6) Theorem. (The Linearity Property).

Let $f(t) = a\,g(t) + b\,h(t)$, where a, b are constants and $\mathscr{L}\{g(t)\}$ and $\mathscr{L}\{h(t)\}$ exist. Then $\mathscr{L}\{f(t)\}$ exists and

$$\mathscr{L}\{f(t)\} = \mathscr{L}\{a\,g(t) + b\,h(t)\} = a\,\mathscr{L}\{g(t)\} + b\,\mathscr{L}\{h(t)\}.$$

Proof. By definition,

$$\mathscr{L}\{a\,g(t)+b\,h(t)\} = \int_0^\infty e^{-st}\{a\,g(t)+b\,h(t)\}\,dt$$

$$= a\int_0^\infty e^{-st}g(t)\,dt + b\int_0^\infty e^{-st}h(t)\,dt$$

$$= a\,\mathscr{L}\{g(t)\} + b\,\mathscr{L}\{h(t)\}$$

(11.7) Theorem. (**The Differentiation Formula**). Let f be continuous on $[\,0,\infty\,[$ and of exponential order a. Let $f\,'$ be piecewise continuous on every finite closed interval $0 \le t \le b$. Then $\mathscr{L}\{f\,'(t)\}$ exists for $s > a$ and

$$\mathscr{L}\{f\,'(t)\} = s\,\mathscr{L}\{f(t)\} - f(0)$$

Proof. By Theorem 11.5, $\mathscr{L}\{f\,'(t)\}$ exists.

Let $\quad F(s) = \mathscr{L}\{f(t)\}$

Then, by definition,

$$\mathscr{L}\{f\,'(t)\} = \int_0^\infty e^{-st}f\,'(t)\,dt$$

$$= \left[e^{-st}f(t)\right]_0^\infty + s\int_0^\infty e^{-st}f(t)\,dt$$

$$= 0 - f(0) + s\,\mathscr{L}\{f(t)\} = s\,\mathscr{L}\{f(t)\} - f(0) = s\,F(s) - f(0)$$

Here $\quad |e^{-st}f(t)| \le Me^{(a-s)t} \to 0 \quad$ as $t\to\infty$ and $s > a$.

(11.8) Corollary. If $F(s) = \mathscr{L}\{f(t)\}$, then

$$\mathscr{L}\{f\,''(t)\} = s^2F(s) - s\,f(0) - f\,'(0)$$

Proof. $\quad \mathscr{L}\{f\,''(t)\} = s\,\mathscr{L}\{f\,'(t)\} - f\,'(0),\quad$ by (**11.7**)

$$= s\left[s\,\mathscr{L}\{f(t)\} - f(0)\right] - f\,'(0)$$

$$= s^2\,\mathscr{L}\{f(t)\} - s\,f(0) - f\,'(0)$$

$$= s^2F(s) - s\,f(0) - f\,'(0).$$

(11.9) Corollary. $\mathscr{L}\{f^{(n)}(t)\} = s^n\,\mathscr{L}\{f(t)\} - s^{n-1}f(0) - s^{n-2}f\,'(0) - s^{n-3}f\,''(0) - \cdots - f^{(n-1)}(0)$

Proof. $\quad \mathscr{L}\{f^{(n)}(t)\} = s\,\mathscr{L}\{f^{(n-1)}(t)\} - f^{(n-1)}(0)$

$$= s\left[s\,f^{(n-2)}(t)\} - f^{(n-2)}(0)\right] - f^{(n-1)}(0)$$

$$= s^2\,\mathscr{L}\{f^{(n-2)}(t)\} - s\,f^{(n-2)}(0) - f^{(n-1)}(0)$$

$$= s^2\left[s\,\mathscr{L}\{f^{(n-3)}(t)\} - f^{(n-3)}(0)\right] - s\,f^{(n-2)}(0) - f^{(n-1)}(0)$$

$$= s^3\,\mathscr{L}\{f^{(n-3)}(t)\} - s^2f^{(n-3)}(0) - s\,f^{(n-2)}(0) - f^{(n-1)}(0)$$

Continuing in this way, we get the required result.

(11.10) Theorem. (**First Shifting or Translation Property**).

Let $F(s) = \mathscr{L}\{f(t)\}$ exist for $s > b$. For any constant a such that $s > a + b$, $\mathscr{L}\{e^{at}f(t)\}$ exists and $\quad \mathscr{L}\{e^{at}f(t)\} = F(s-a)$

Proof. By definition, $\quad F(s) = \int_0^\infty e^{-st}f(t)\,dt$

So $\quad \mathscr{L}\{e^{at}f(t)\} = \int_0^\infty e^{-st}e^{at}f(t)\,dt = \int_0^\infty e^{-(s-a)t}f(t)\,dt\quad$ exists

provided $\;(s-a) > b\;$ i.e. $\;s > a+b$, and

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

Example 7. Compute $\mathscr{L}\{\sinh at\}$ and $\mathscr{L}(\cosh at)$.

Solution. Here $\sinh at = \dfrac{e^{at}-e^{-at}}{2}$

We have

$$\mathscr{L}\{\sinh at\} = \mathscr{L}\left\{\frac{e^{at}-e^{-at}}{2}\right\}$$

$$= \frac{1}{2}\,\mathscr{L}\{e^{at}\} - \frac{1}{2}\,\mathscr{L}\{e^{-at}\},\ \text{by } (\mathbf{11.6})$$

$$= \frac{1}{2}\cdot\frac{1}{s-a} - \frac{1}{2}\cdot\frac{1}{s+a} = \frac{a}{s^2-a^2}.$$

Similarly, $\quad \mathscr{L}\{\cosh at\} = \dfrac{s}{s^2-a^2}.$

Example 8. Compute $\mathcal{L}\{\cos^2 at\}$

Solution. Let
$$f(t) = \cos^2 at$$
$$f'(t) = -2a\cos at \sin at = -a\sin 2at$$

By (11.7), $\mathcal{L}\{f'(t)\} = s\,\mathcal{L}\{f(t)\} - f(0)$

or $s\,\mathcal{L}\{f(t)\} = \mathcal{L}\{f'(t)\} + f(0)$

i.e., $s\,\mathcal{L}\{\cos^2 at\} = -a\,\mathcal{L}\{\sin 2at\} + f(0)$

$$= -a \cdot \frac{2a}{s^2+4a^2} + 1, \text{ as in Example 4}$$
$$= \frac{s^2+2a^2}{s^2+4a^2}$$

Therefore, $\mathcal{L}\{f(t)\} = \mathcal{L}\{\cos^2 at\} = \dfrac{s^2+2a^2}{s(s^2+4a^2)}$

Alternative Method:
$$\cos^2 at = \frac{\cos 2at + 1}{2}$$

So, by the Linearity Property of the transform, we have
$$\mathcal{L}\{\cos^2 at\} = \frac{1}{2}\,[\,\mathcal{L}\{\cos 2at\} + \mathcal{L}\{1\}]$$
$$= \frac{1}{2}\left\{\frac{s}{s^2+4a^2} + \frac{1}{s}\right\} = \frac{s^2+2a^2}{s(s^2+4a^2)}$$

Example 9. Evaluate $\mathcal{L}\{e^{3t}(t^3 + \sin 2t)\}$

Solution.
$$\mathcal{L}\{e^{3t}(t^3+\sin 2t)\} = \mathcal{L}\{e^{3t}t^3 + e^{3t}\sin 2t\}$$
$$= \mathcal{L}\{e^{3t}t^3\} + \mathcal{L}\{e^{3t}\sin 2t\}, \text{ by (11.6)}$$

Now $\mathcal{L}\{t^3\} = \dfrac{3!}{s^4}$

$$\mathcal{L}\{\sin 2t\} = \frac{2}{s^2+4}$$

By (11.10), $\mathcal{L}\{e^{3t}t^3\} = \dfrac{3!}{(s-3)^4}$ and $\mathcal{L}\{e^{3t}\sin 2t\} = \dfrac{2}{(s-3)^2+4}$

Therefore,
$$\mathcal{L}\{e^{3t}(t^3+\sin 2t)\} = \frac{3!}{(s-3)^4} + \frac{2}{(s-3)^2+4}.$$

Example 10. Compute $\mathcal{L}\{te^{at}\cos bt\}$.

Solution. Consider $te^{at}e^{ibt} = te^{(a+ib)t}$

Let $f(t) = t$, then $\mathcal{L}\{f(t)\} = \mathcal{L}\{t\} = \dfrac{1}{s^2}$.

By (11.10), $\mathcal{L}\{te^{(a+ib)t}\} = \dfrac{1}{[s-(a+ib)]^2} = \dfrac{1}{[(s-a)-ib]^2}$

$$= \frac{[(s-a)+ib]^2}{[\{(s-a)-ib\}\{(s-a)+ib\}]^2}$$
$$= \frac{(s-a)^2 - b^2 + 2ib(s-a)}{[(s-a)^2+b^2]^2}$$

Equating real parts, we have
$$\mathcal{L}\{t\,e^{at}\cos bt\} = \frac{(s-a)^2-b^2}{[(s-a)^2+b^2]^2}$$

(11.11) Theorem. Suppose $\mathcal{L}\{f(t)\} = F(s)$ exists for $s > a$. Then
$$\mathcal{L}\{t f(t)\} = -F'(s).$$

Proof. Consider
$$\frac{d}{ds}\mathcal{L}\{f(t)\} = \frac{d}{ds}F(s) = \frac{d}{ds}\int_0^\infty e^{-st} f(t)\,dt$$
$$= \int_0^\infty \frac{d}{ds}\,e^{-st} f(t)\,dt = \int_0^\infty -t\,e^{-st} f(t)\,dt = -\mathcal{L}\{t f(t)\}$$

Thus $F'(s) = -\mathcal{L}\{t f(t)\}$

or $\mathcal{L}\{t f(t)\} = -F'(s)$, as desired.

Using the above result repeatedly, we have
$$\mathcal{L}\{t^n f(t)\} = -\frac{d}{ds}\,[\mathcal{L}\{t^{n-1} f(t)\}]$$
$$= (-1)^2\frac{d^2}{ds^2}\,[\mathcal{L}\{t^{n-2} f(t)\}]$$
$$\vdots \quad \vdots \quad \vdots$$
$$= (-1)^n\frac{d^n}{ds^n}\,[\mathcal{L}\{f(t)\}].$$

Example 11. Compute $\mathscr{L}\{t^3 e^{-t}\}$

Solution.

$$f(t) = e^{-t}$$

$$\mathscr{L}\{f(t)\} = \mathscr{L}\{e^{-t}\} = \frac{1}{s+1}$$

By **(11.11)**, $$\mathscr{L}\{t^3 e^{-t}\} = (-1)^3 \frac{d^3}{ds^3}\left[\mathscr{L}\{e^{-t}\}\right] = -\frac{d^3}{ds^3}\left(\frac{1}{s+1}\right)$$

$$= -\frac{(-1)^3\, 3!}{(s+1)^4} = \frac{6}{(s+1)^4}$$

(11.12) Theorem. If $\mathscr{L}\{f(t)\} = F(s)$ then

$$\mathscr{L}\left\{\frac{f(t)}{t}\right\} = \int_s^\infty F(u)\,du, \text{ provided } \lim_{t \to 0^+} \frac{f(t)}{t} \text{ exists}$$

Proof. Let $\frac{f(t)}{t} = g(t)$ Then $f(t) = t\,g(t)$

$$F(s) = \mathscr{L}\{f(t)\} = \mathscr{L}\{t\,g(t)\}$$

$$= -\frac{d}{ds}\left(\mathscr{L}\{g(t)\}\right), \text{ by (11.11)}$$

Integrating, we have

$$\mathscr{L}\{g(t)\} = -\int_\infty^s F(u)\,du = \int_s^\infty F(u)\,du$$

(11.13) Theorem. If f is piecewise continuous and is of exponential order a, then

$$\mathscr{L}\left\{\int_0^t f(u)\,du\right\} = \frac{1}{s}\,\mathscr{L}\{f(t)\}$$

Proof. The integral

$$g(t) = \int_0^t f(u)\,du$$

is a continuous function of t. Since $f(t)$ is of exponential order a, $|f(t)| \le M e^{at}$. Therefore,

$$|g(t)| = \left|\int_0^t f(u)\,du\right| \le M\int_0^t e^{au}\,du = \frac{M}{a}\left\{e^{at} - 1\right\}.$$

By the Fundamental Theorem of Integral Calculus, $g'(t) = f(t)$ except at points where f is discontinuous. Hence $g'(t)$ is piecewise continuous. By (11.7), we have

$$\mathscr{L}\{f(t)\} = \mathscr{L}\{g'(t)\} = s\,\mathscr{L}\{g(t)\} - g(0), \quad s > a$$

$$= s\,\mathscr{L}\{g(t)\}, \text{ since } g(0) = 0$$

Thus $$\mathscr{L}\{g(t)\} = \frac{1}{s}\,\mathscr{L}\{f(t)\}$$

Example 12. Compute $\mathscr{L}\left\{\frac{\sin t}{t}\right\}$

Solution. Let $f(t) = \sin t$ Then

$$\mathscr{L}\{f(t)\} = \mathscr{L}\{\sin t\} = \frac{1}{s^2+1} = F(s)$$

Set $g(t) = \frac{\sin t}{t}$ By **(11.12)**, we have

$$\mathscr{L}\left\{\frac{\sin t}{t}\right\} = \int_s^\infty F(u)\,du$$

$$= \int_s^\infty \frac{1}{1+u^2}\,du = \left[\arctan u\right]_s^\infty = \frac{\pi}{2} - \arctan s$$

Example 13. Evaluate $\mathscr{L}\left\{\int_0^t \frac{1-\cosh au}{u}\,du\right\}$

Solution. Let $f(t) = 1 - \cosh at$ Then

$$\mathscr{L}\{f(t)\} = \frac{1}{s} - \frac{s}{s^2 - a^2} = F(s), \text{ as in Example 7.}$$

Set $g(t) = \frac{1-\cosh at}{t}$ Then by **(11.12)**, we get

$$\mathscr{L}\left\{\frac{1-\cosh at}{t}\right\} = \int_s^\infty \left(\frac{1}{u} - \frac{u}{u^2 - a^2}\right) du$$

$$= \left[\ln u - \frac{1}{2}\ln(u^2 - a^2)\right]_s^\infty$$

$$= \lim_{u \to \infty} \left[\frac{1}{2}\ln u^2 - \frac{1}{2}\ln(u^2 - a^2)\right] + \frac{1}{2}\ln(s^2 - a^2) - \ln s$$

$$= \frac{1}{2}\lim_{u \to \infty} \ln\left(\frac{u^2}{u^2 - a^2}\right) + \frac{1}{2}\ln\left(\frac{s^2 - a^2}{s^2}\right)$$

$$= \frac{1}{2}\ln \lim_{u \to \infty}\left(\frac{u^2}{u^2 - a^2}\right) + \frac{1}{2}\ln\left(\frac{s^2 - a^2}{s^2}\right)$$

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

By **(11.13)**, we have

$$\mathscr{L}\left\{\int_0^t \frac{1 - \cosh au}{u}\, du\right\} = \frac{1}{s}\,\mathscr{L}\left\{\frac{1 - \cosh at}{t}\right\} = \frac{1}{2s}\ln\left(\frac{s^2 - a^2}{s^2}\right)$$

(11.14) Definition. **(Unit Step Function).** Let $a \ge 0$. The function u_a defined on $]0, \infty[$ by

$$u_a(t) = \begin{cases} 0 & \text{if } t < a \\ 1 & \text{if } t > a \end{cases}$$

is called the **unit step function**. If $a = 0$ then

$$u_0(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t > 0 \end{cases}$$

Since u_a is defined on $]0, \infty[$, we have $u_0(t) = 1$ for $t > 0$.

Clearly, the unit step function is of exponential order.

(11.15) Theorem. Let u_a be the unit step function. Then

$$\mathscr{L}\{u_a(t)\} = \frac{e^{-as}}{s}$$

Proof. By definition,

$$\mathscr{L}\{u_a(t)\} = \int_0^\infty e^{-st}\, u_a(t)\, dt$$

$$= \int_0^a 0\, e^{-st}\, dt + \int_a^\infty e^{-st}\, dt = \int_a^\infty e^{-st}\, dt$$

$$= \lim_{h \to \infty}\left[\frac{e^{-st}}{-s}\right]_a^h = \lim_{h \to \infty}\left[\frac{-e^{-hs}}{s} + \frac{e^{-as}}{s}\right]$$

2nd shifting thm

$$= \frac{e^{-as}}{s},\ s > 0, \text{ because } \lim_{h \to \infty}\frac{e^{-hs}}{s} = 0$$

(11.16) Theorem. Let f be a function of exponential order α and $\mathscr{L}\{f(t)\} = F(s)$. For the function

$$u_a(t)f(t-a) = \begin{cases} 0 & \text{if } 0 < t < a \\ f(t-a) & \text{if } t > a, \end{cases}$$

$$\mathscr{L}\{u_a(t)f(t-a)\} = e^{-as}F(s)$$

Proof.

$$\mathscr{L}\{u_a(t)f(t-a)\} = \int_0^\infty e^{-st}\, u_a(t)f(t-a)\, dt$$

$$= \int_0^a e^{-st}\ 0\, dt + \int_a^\infty e^{-st} f(t-a)\, dt$$

$$= \int_a^\infty e^{-st} f(t-a)\, dt \qquad (1)$$

Putting $t - a = \tau$ into (1), we get

$$\mathscr{L}\{u_a(t)f(t-a)\} = e^{-as}\int_0^\infty e^{-s\tau} f(\tau)\, d\tau = e^{-as}F(s)$$

This is known as the **Second Translation Property.**

Example 14. Find the Laplace transform of

$$f(t) = \begin{cases} 0 & \text{if } 0 < t < \frac{\pi}{2} \\ \cos t & \text{if } t > \frac{\pi}{2} \end{cases}$$

Solution. We have to express $\cos t$ in terms of $t - \frac{\pi}{2}$ so as to apply **(11.16)**.

As $\cos t = -\sin\left(t - \frac{\pi}{2}\right)$, let

$$g(t) = \begin{cases} 0 & \text{if } 0 < t < \frac{\pi}{2} \\ \sin\left(t - \frac{\pi}{2}\right) & \text{if } t > \frac{\pi}{2} \end{cases}$$

Then $f(t) = -u_{\pi/2}(t)\, g(t)$

$$\mathscr{L}\{f(t)\} = -\mathscr{L}\left\{u_{\pi/2}(t)\sin\left(t-\frac{\pi}{2}\right)\right\}$$

$$= -e^{\left(-\frac{\pi}{2}\right)s}\ \mathscr{L}\{\sin t\} \text{ by (11.16)}$$

$$= -e^{-\frac{\pi}{2}s}\ \frac{1}{s^2+1}$$

Alternative Method

By definition,

$$F(s) = \mathscr{L}\{f(t)\} = \int_0^\infty e^{-st} f(t)\, dt$$

$$= \int_0^{\pi/2} e^{-st} f(t)\, dt + \int_{\pi/2}^\infty e^{-st} f(t)\, dt$$

$$= 0 + \int_{\pi/2}^\infty e^{-st}\cos t\, dt$$

$$= \left[\frac{e^{-st}}{-s}\cos t\right]_{\pi/2}^\infty + \frac{1}{s}\int_{\pi/2}^\infty e^{-st}(-\sin t)\, dt$$

$$= \frac{-1}{s}\int_{\pi/2}^\infty e^{-st}\sin t\, dt$$

$$= \frac{1}{s^2}\left[e^{-st}\sin t\right]_{\frac{\pi}{2}}^\infty - \frac{1}{s^2}\int_{\pi/2}^\infty e^{-st}\cos t\, dt$$

$$= -\frac{e^{-\frac{\pi}{2}s}}{s^2} - \frac{1}{s^2}F(s)$$

Therefore, $\left(1+\frac{1}{s^2}\right)F(s) = -\frac{e^{-\frac{\pi}{2}s}}{s^2}$

or $F(s) = -\frac{e^{-\frac{\pi}{2}s}}{s^2+1}$

TABLE OF SOME LAPLACE TRANSFORMS

$f(t)$	$\mathscr{L}\{f(t)\} = F(s)$	$f(t)$	$\mathscr{L}\{f(t)\} = F(s)$
1. 1	$\frac{1}{s}$, $s>0$	2. t	$\frac{1}{s^2}$
3. t^n	$\frac{n!}{s^{n+1}}$, $s>0$	4. t^α, $\alpha > -1$	$\frac{\Gamma(\alpha+1)}{s^{\alpha+1}}$
5. $\frac{1}{\sqrt{\pi t}}$	$\frac{1}{\sqrt{s}}$	6. e^{at}	$\frac{1}{s-a}$, $s>a$
7. te^{at}	$\frac{1}{(s-a)^2}$	8. $\sin at$	$\frac{a}{s^2+a^2}$, $s>0$
9. $\cos at$	$\frac{s}{s^2+a^2}$, $s>0$	10. $\sinh at$	$\frac{a}{s^2-a^2}$, $s>\|a\|$
11. $\cosh at$	$\frac{s}{s^2-a^2}$, $s>\|a\|$	12. $t^n e^{at}$	$\frac{n!}{(s-a)^{n+1}}$, $s>a$
13. $e^{at}\sin bt$	$\frac{b}{(s-a)^2+b^2}$, $s>a$	14. $e^{at}\cos bt$	$\frac{s-a}{(s-a)^2+b^2}$, $s>a$
15. $t\sin at$	$\frac{2as}{(s^2+a^2)^2}$, $s>0$	16. $t\cos at$	$\frac{s^2-a^2}{(s^2+a^2)^2}$, $s>0$
17. $f(ct)$	$\frac{1}{c}F\left(\frac{s}{c}\right)$, $c>0$	18. $\int_0^t f(u)\, du$	$\frac{1}{s}F(s)$
19. $t^n f(t)$	$(-1)^n \frac{d^n}{ds^n}F(s)$	20. $\frac{f(t)}{t}$	$\int_s^\infty F(u)\, du$
21. $u_a(t)$	$\frac{e^{-as}}{s}$	22. $u_a(t) f(t-a)$ u(t-a)	$e^{-as}F(s)$
23. $\sin at - at\cos at$	$\frac{2a^3}{(s^2+a^2)^2}$	24. $f'(t)$	$sF(s)-f(0)$
25. $1-\cos at$	$\frac{a^2}{s(s^2+a^2)}$	26. $at-\sin at$	$\frac{a^3}{s^2(s^2+a^2)}$
27. $\sinh at - \sin at$	$\frac{2a^3}{s^4-a^4}$	28. $\cosh at - \cos at$	$\frac{2a^2 s}{s^4-a^4}$

EXERCISE 11.1

Compute the Laplace transform of each of the following (**Problems 1 – 28**):

1.	$t^2 + 6t - 17$	**2.**	e^{2t+3}
3.	$\sin(7t+4)$	**4.**	$\cos(at+b)$
5.	$\cosh(5t-3)$	**6.**	$(t^3-1)\,e^{-2t}$
7.	$e^{-t}\sin 2t$	**8.**	$e^{3t}\cosh 4t$
9.	$\cos t\cos 2t$	**10.**	$\sin^3 t$
11.	$te^{-3t}\sin at$	**12.**	$\sinh^2 at$
13.	$\cosh at\sin at$	**14.**	$\sinh at\cos at$
15.	$\cosh at\cos bt$	**16.**	$[\,t\,]$, the bracket function
17.	t^{α}, $\alpha > -1$. Hence find $\mathscr{L}\{t^{1/2}\}$	**18.**	$t^2\sin at$
19.	$t^2\cos at$	**20.**	$t\sin^2 at$
21.	$t^2\cos^2 2t$	**22.**	$\dfrac{\sin at}{t}$
23.	$\dfrac{1-\cos at}{t}$	**24.**	$\displaystyle\int_0^t \frac{\sin au}{u}\,du$
25.	$\displaystyle\int_0^t \frac{1-\cos au}{u}\,du$	**26.**	$\dfrac{\sinh at}{t}$
27.	$\ln t$	**28.**	$f(t) = \begin{cases} 0 & \text{if } t<3 \\ (t-3)^3 & \text{if } t>3 \end{cases}$

29. If $\mathscr{L}\{f(t)\} = F(s)$ for $s > a$, show that $\mathscr{L}\{f(ct)\} = \dfrac{1}{c}\,F\!\left(\dfrac{s}{c}\right)$, $c > 0$ and $s > ca$.

30. Compute $\mathscr{L}\{\sin\sqrt{t}\}$. Deduce $\mathscr{L}\left\{\dfrac{\cos\sqrt{t}}{\sqrt{t}}\right\}$.

INVERSE LAPLACE TRANSFORM

Recall Definition 11.3 that if $F(s)$ is the Laplace transform of a function $f(t)$, then $f(t)$ is called the **inverse Laplace transform** of $F(s)$ and is denoted by $\mathscr{L}^{-1}\{F(s)\} = f(t)$. In this section, we shall compute inverse Laplace transforms of certain functions. This will be done with the aid of Table of Laplace Transforms of elementary functions and properties of the inverse transform which follow from the theorems and properties of the Laplace transform.

Eg

(11.17) Theorem. (Linearity Property). If $\mathscr{L}^{-1}\{F_1(s)\} = f_1(t)$ and $\mathscr{L}^{-1}\{F_2(s)\} = f_2(t)$, then for any constants a, b

$$\mathscr{L}^{-1}\{a\,F_1(s) + b\,F_2(s)\} = a\,\mathscr{L}^{-1}\{F_1(s)\} + b\,\mathscr{L}^{-1}\{F_2(s)\} = a\,f_1(t) + b\,f_2(t)$$

It is a direct consequence of Theorem 11.6.

(11.18) Theorem. If $\mathscr{L}^{-1}\{F(s)\} = f(t)$, then

(i) $\mathscr{L}^{-1}\{F(s-a)\} = e^{at}f(t)$

(ii) $\mathscr{L}^{-1}\{F(cs)\} = \dfrac{1}{c}\,f\!\left(\dfrac{t}{c}\right)$, $c > 0$

(iii) $\mathscr{L}^{-1}\{F^{(n)}(s)\} = (-1)^n t^n f(t)$

(iv) $\mathscr{L}^{-1}\left\{\dfrac{F(s)}{s}\right\} = \displaystyle\int_0^t f(u)\,du$

(v) $\mathscr{L}^{-1}\{e^{-as}F(s)\} = u_a(t)\,f(t-a)$, where $u_a(t)$ is the unit step function.

These results follow from the relevant theorems on Laplace transform.

Example 15. Compute $\mathscr{L}^{-1}\left\{\dfrac{5s}{s^2+5}\right\}$

Solution. Here $\dfrac{5s}{s^2+5} = 5\,\dfrac{s}{s^2+(\sqrt{5})^2}$.

So, from the Table of Laplace Transforms, we have

$$\mathscr{L}^{-1}\left\{\frac{5s}{s^2+5}\right\} = 5\,\mathscr{L}^{-1}\left\{\frac{s}{s^2+(\sqrt{5})^2}\right\} = 5\cos\sqrt{5}\,t.$$

Example 16. Find $\mathscr{L}^{-1}\left\{\dfrac{1}{s^2+2s-15}\right\}$

Solution. Here

$$\frac{1}{s^2+2s-15} = \frac{1}{(s+5)(s-3)} = \frac{A}{s+5} + \frac{B}{s-3}$$

$$= \frac{1}{8}\left[\frac{1}{s-3} - \frac{1}{s+5}\right]$$

Therefore, $\mathscr{L}^{-1}\left\{\dfrac{1}{s^2+2s-15}\right\} = \dfrac{1}{8}\,\mathscr{L}^{-1}\left\{\dfrac{1}{s-3}\right\} - \dfrac{1}{8}\,\mathscr{L}^{-1}\left\{\dfrac{1}{s+5}\right\}$

$$= \frac{1}{8}\,e^{3t} - \frac{1}{8}\,e^{-5t}$$

Example 17. Compute $\mathscr{L}^{-1}\left\{\dfrac{3s+17}{s^2+8s+25}\right\}$ ✓

Solution. Here

$$\frac{3s+17}{s^2+8s+25} = \frac{3(s+4)+5}{(s+4)^2+3^2}$$

So $\mathscr{L}^{-1}\left\{\dfrac{3s+17}{s^2+8s+25}\right\} = 3\,\mathscr{L}^{-1}\left\{\dfrac{s+4}{(s+4)^2+3^2}\right\} + 5\,\mathscr{L}^{-1}\left\{\dfrac{1}{(s+4)^2+3^2}\right\}$

$$= 3e^{-4t}\,\mathscr{L}^{-1}\left\{\frac{s}{s^2+3^2}\right\} + \frac{5}{3}\,e^{-4t}\,\mathscr{L}^{-1}\left\{\frac{3}{s^2+3^2}\right\}$$

$$= 3e^{-4t}\cos 3t + \frac{5}{3}\,e^{-4t}\sin 3t$$

Example 18. Evaluate $\mathscr{L}^{-1}\left\{\dfrac{s^2}{(s^2+a^2)^2}\right\}$ ✓

Solution. Here $\dfrac{s^2}{(s^2+a^2)^2} = \dfrac{1}{s^2+a^2} - \dfrac{a^2}{(s^2+a^2)^2}$

So $\mathscr{L}^{-1}\left\{\dfrac{s^2}{(s^2+a^2)^2}\right\} = \mathscr{L}^{-1}\left\{\dfrac{1}{s^2+a^2}\right\} - a^2\,\mathscr{L}^{-1}\left\{\dfrac{1}{(s^2+a^2)^2}\right\}$

$$= \frac{1}{a}\sin at - a^2\,\mathscr{L}^{-1}\left\{\frac{1}{(s^2+a^2)^2}\right\} \qquad (1)$$

Now $\mathscr{L}\{t\sin at\} = -\dfrac{d}{ds}\,\mathscr{L}\{\sin at\}$

$$= -\frac{d}{ds}\left(\frac{a}{s^2+a^2}\right)$$

$$= \frac{2as}{(s^2+a^2)^2}$$

Therefore, $2a\,\mathscr{L}^{-1}\left\{\dfrac{s}{(s^2+a^2)^2}\right\} = t\sin at$

By the integration property (11.18) (iv), we get

$$2a\,\mathscr{L}^{-1}\left\{\frac{s}{(s^2+a^2)^2}\cdot\frac{1}{s}\right\} = 2a\,\mathscr{L}^{-1}\left\{\frac{1}{(s^2+a^2)^2}\right\}$$

$$= \int_0^t u\sin au\,du$$

$$= \left[\frac{-u\cos au}{a}\right]_0^t - \frac{1}{a}\int_0^t \cos au\,du$$

$$= -\frac{t\cos at}{a} + \frac{1}{a^2}\sin at$$

Substituting into (1), we have

$$\mathscr{L}^{-1}\left\{\frac{s^2}{(s^2+a^2)^2}\right\} = \frac{1}{a}\sin at + \frac{t\cos at}{2} - \frac{\sin at}{2a}$$

$$= \frac{1}{2a}\left(\sin at + at\cos at\right).$$

Example 19. Compute $\mathscr{L}^{-1}\left\{\dfrac{e^{-2s}}{s^2}\right\}$

Solution. Here $\mathscr{L}\{u_2(t)\} = \dfrac{e^{-2s}}{s}$

Therefore, $\mathscr{L}^{-1}\left\{\dfrac{e^{-2s}}{s^2}\right\} = \displaystyle\int_0^t u_2(\tau)\,d\tau$

$$= (t-2)\,u_2(t)$$

Again, $\mathscr{L}^{-1}\left\{\dfrac{e^{-2s}}{s^3}\right\} = \displaystyle\int_0^t (t-2)\,u_2(\tau)\,d\tau$

$$= u_2(t)\,\frac{(t-2)^2}{2}.$$

Alternative Method:

$$\mathscr{L}^{-1}\left\{\frac{e^{-2s}}{s^3}\right\} = u_2(t)f(t-2), \quad \text{where } f(t) = \mathscr{L}^{-1}\left\{\frac{1}{s^3}\right\} = \frac{t^2}{2}$$

$$= u_2(t)\,\frac{(t-2)^2}{2}$$

Example 20. Find $\mathscr{L}^{-1}\left\{\frac{1}{s^2(s^2-a^2)}\right\}$

Solution. Here $\mathscr{L}^{-1}\left\{\frac{1}{s^2-a^2}\right\} = \frac{1}{a}\sinh at$

Using the integration property **(11.18)(iv)**, we get

$$\mathscr{L}^{-1}\left\{\frac{1}{s(s^2-a^2)}\right\} = \frac{1}{a}\int_0^t \sinh au\,du$$

$$= \frac{1}{a^2}\left[\cosh au\right]_0^t = \frac{1}{a^2}\{\cosh at - 1\}.$$

Applying the same property again, we obtain

$$\mathscr{L}^{-1}\left\{\frac{1}{s^2(s^2-a^2)}\right\} = \frac{1}{a^2}\int_0^t (\cosh au - 1)\,du$$

$$= \frac{1}{a^3}\sinh at - \frac{1}{a^2}t$$

Alternative Method:

Here $$\frac{1}{s^2(s^2-a^2)} = \frac{1}{a^2}\left(\frac{1}{s^2-a^2} - \frac{1}{s^2}\right)$$

Therefore, $$\mathscr{L}^{-1}\left\{\frac{1}{s^2(s^2-a^2)}\right\} = \frac{1}{a^2}\mathscr{L}^{-1}\left\{\frac{1}{s^2-a^2}\right\} - \frac{1}{a^2}\mathscr{L}^{-1}\left\{\frac{1}{s^2}\right\}$$

$$= \frac{1}{a^3}\sinh at - \frac{1}{a^2}t.$$

(11.19) Definition. (**Convolution**). Let $f(t)$ and $g(t)$ be piecewise continuous functions on $[0, \infty[$. The **convolution** of f and g, denoted by $f \bullet g$, is defined by

$$(f \bullet g)t = \int_0^t f(t-u)\,g(u)\,du$$

It can be easily verified that

(i) $f \bullet g = g \bullet f$

(ii) $f \bullet (g+h) = f \bullet g + f \bullet h$

(iii) $f \bullet 0 = 0$

The following result is useful in evaluating the Laplace transform and its inverse.

(11.20) Theorem. Let $f(t)$ and $g(t)$ be piecewise continuous on $[0, \infty[$ and of exponential order and $\mathscr{L}(f) = F(s)$, $\mathscr{L}(g) = G(s)$. Then

(i) $\mathscr{L}(f \bullet g)(s) = F(s)\,G(s)$

(ii) $\mathscr{L}^{-1}\{F(s)\,G(s)\}(t) = (f \bullet g)t$

$$= \int_0^t f(t-u)\,g(u)\,du$$

The proof of this theorem is omitted.

Example 21. Let $f(t) = t$ and $g(t) = \cos t$. Compute $f \bullet g$ and verify

Solution. $(f \bullet g)\,t = \int_0^t (t-u)\cos u\,du$

$$= \int_0^t t\cos u\,du - \int_0^t u\cos u\,du$$

$$= t\left[\sin u\right]_0^t - \left\{\left[u\sin u\right]_0^t - \int_0^t \sin u\,du\right\}$$

$$= t\sin t - t\sin t + \left[-\cos u\right]_0^t$$

$$= 1 - \cos t$$

$$\mathscr{L}(f) = \frac{1}{s^2} = F(s), \quad \mathscr{L}(g) = \frac{s}{1+s^2} = G(s)$$

(i) $$\mathscr{L}(f \bullet g) = \int_0^\infty e^{-st}\left\{\int_0^t (t-u)\cos u\,du\right\}dt$$

$$= \int_0^\infty e^{-st}(1-\cos t)\,dt$$

$$= \int_0^\infty e^{-st}\,dt - \int_0^\infty e^{-st}\cos t\,dt$$

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

$$= F(s)\, G(s)$$

(ii) $$\mathscr{L}^{-1}\{F(s)\, G(s)\} = \mathscr{L}^{-1}\left(\frac{1}{s} - \frac{s}{s^2+1}\right)$$

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

$$= 1 - \cos t = (f * g).$$

Example 22. Compute $\mathscr{L}^{-1}\left\{ \dfrac{2a^2}{(s^2+a^2)^2} \right\}$.

Solution. We have $\mathscr{L}^{-1}\left\{ \dfrac{a}{s^2+a^2} \right\} = \sin at.$

By **(11.20) (ii)**, we get

$$\mathscr{L}^{-1}\left\{ \frac{2a^2}{(s^2+a^2)^2} \right\} = 2(\sin at * \sin at)$$

$$= 2\int_0^t \sin a(t-u) \sin au \, du$$

$$= 2\int_0^t \{\sin at \cos au - \cos at \sin au\} \sin au \, du$$

$$= \sin at \int_0^t 2\cos au \sin au \, du - 2\cos at \int_0^t \sin^2 au \, du$$

$$= \sin at \int_0^t \sin 2au \, du - \cos at \int_0^t 2\sin^2 au \, du$$

$$= \sin at \left[-\frac{\cos 2at}{2a} \right]_0^t - \cos at \int_0^t (1 - \cos 2au)\, du$$

$$= -\frac{\sin at(\cos 2at - 1)}{2a} - \cos at\left[t - \frac{\sin 2at}{2a}\right]$$

$$= \frac{1}{2a}(\sin 2at \cos at - \cos 2at \sin at) - t\cos at + \frac{1}{2a}\sin at$$

$$= \frac{1}{2a}\sin(2at - at) + \frac{1}{2a}\sin at - t\cos at$$

$$= \frac{1}{a}\sin at - t\cos at.$$

EXERCISE 11.2

Compute the inverse Laplace transform of each of the following (Problems 1 – 20).

1. $\dfrac{s-2}{s^2-2}$

2. $\dfrac{3s+1}{s^2-6s+18}$

3. $\dfrac{9s-67}{s^2-16s+\cancel{49}\,35}$

4. $\dfrac{as+b}{s^2+2cs+d}$, $d > c^2 > 0$

5. $\dfrac{s}{(s+a)^2+b^2}$

6. $\dfrac{1}{(s^2+a^2)(s^2+b^2)}$

7. $\dfrac{1}{(s-1)(s^2+4)}$

8. $\dfrac{7s+5}{(3s-8)^2}$

9. $\dfrac{5s+3}{(s+7)^3}$

10. $\dfrac{2s-3}{2s^3+3s^2-2s-3}$

11. $\dfrac{2s^3+6s^2+21s+52}{s(s+2)(s^2+4s+13)}$

12. $\dfrac{1}{(s^2+4)(s^2+6s-5)}$

13. $\dfrac{s^3+3s^2-s-3}{(s^2+2s+5)^2}$

14. $\arctan \dfrac{a}{s}$

15. $\ln \dfrac{s^2+1}{(s-1)^2}$

16. $\ln \dfrac{s^2+a^2}{s^2+b^2}$

17. $\dfrac{e^{-3s}}{s^2(s^2+9)}$

18. $e^{-\pi s}\, \dfrac{s}{s^2-4s+5}$

19. $e^{-2s} \cdot \dfrac{s+6}{s^3-5s^2+6s}$

20. $e^{-3s} \cdot \dfrac{3s-7}{s^2-10s+26}$

In each of **Problems 21 – 23**, use the Convolution Theorem 11.20 to evaluate the inverse Laplace transform:

21. $\dfrac{1}{s^2(s+5)}$

22. $\dfrac{s}{(s+1)(s^2+4)}$

23. $\dfrac{1}{(s^2+1)(s^2+4s+5)}$

24. Show that $\mathscr{L}^{-1}\left\{ \dfrac{s^3}{s^4+4a^4} \right\} = \cosh at \cos at$

25. Show that $\mathscr{L}^{-1}\left\{ \dfrac{s}{s^4+4a^4} \right\} = \dfrac{1}{2a^2} \sinh at \sin at.$

SOLUTIONS OF INITIAL VALUE PROBLEMS

(11.21) In this section we shall use the powerful tool of Laplace transform to solve constant coefficients linear initial value problems. It will be seen that by the methods of Laplace transform, a differential equation can be converted into an algebraic equation. The independent variable will be t instead of x and the dependent variable will remain y as before in the differential equations to be considered here.

If $y(t)$ is a solution of a differential equation, then the Laplace transform of $y(t)$ will be denoted by $\mathscr{L}\{y(t)\} = Y(s)$ in conformity with our earlier notation. The differentiation formula **(Theorem 11.7)** and its corollaries are the most important properties to be employed in the solutions of the problems.

The following procedure will be adopted:

I. Given an initial value problem, take Laplace transform of both sides. Use Theorem 11.7 and initial conditions to convert the differential equation into an algebraic equation in $Y(s)$.

II. Solve the algebraic equation for $Y(s)$.

III. The inverse transform $\mathscr{L}^{-1}\{Y(s)\} = y(t)$ is the required solution of the given problem.

The method is applicable to both homogeneous and nonhomogeneous equations. However, for nonhomogeneous equations, the function $f(t)$ on the right hand side of an equation in standard form must possess Laplace transform.

This method is now illustrated by examples.

Example 23. Use the Laplace transform to solve

$$\frac{dy}{dt} + y = -e^t, \qquad (1)$$

$$y(0) = 2.$$

Solution. Taking Laplace transform of both sides of (1), we have

$$\mathscr{L}\left\{\frac{dy}{dt}\right\} + \mathscr{L}\{y(t)\} = -\mathscr{L}\{e^t\}. \qquad (2)$$

Now $\mathscr{L}\left\{\frac{dy}{dt}\right\} = s\,Y(s) - y(0)$, by (11.7)

$= s\,Y(s) - 2$, on applying the initial condition.

Substituting into (2), we get

$$s\,Y(s) - 2 + Y(s) = -\frac{1}{s-1}$$

or $$(s+1)\,Y(s) = 2 - \frac{1}{s-1} = \frac{2s-3}{s-1}$$

or $$Y(s) = \frac{2s-3}{(s-1)(s+1)} = \frac{5}{2}\cdot\frac{1}{s+1} - \frac{1}{2}\cdot\frac{1}{s-1}$$

Applying $\mathscr{L}^{-1}$ on both sides, we have

$$\mathscr{L}^{-1}\{Y(s)\} = y(t) = \frac{5}{2}\,\mathscr{L}^{-1}\left\{\frac{1}{s+1}\right\} - \frac{1}{2}\,\mathscr{L}^{-1}\left\{\frac{1}{s-1}\right\}$$

$$= \frac{5}{2}\,e^{-t} - \frac{1}{2}\,e^{t}$$

is the required solution.

Example 24. Solve $\frac{d^2y}{dt^2} - 2\frac{dy}{dt} + y = t\,e^t, \qquad (1)$

$$y(0) = 0, \quad y'(0) = 0$$

by the method of Laplace transform.

Solution. Taking Laplace transform of both sides of (1), we have

$$\mathscr{L}\left\{\frac{d^2y}{dt^2}\right\} - 2\,\mathscr{L}\left\{\frac{dy}{dt}\right\} + \mathscr{L}\{y(t)\} = \mathscr{L}\{t\,e^t\}$$

or $$s^2\,Y(s) - s\,y(0) - y'(0) - 2s\,Y(s) - 2y(0) + Y(s) = \frac{1}{(s-1)^2}$$

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

i.e., $$Y(s) = \frac{1}{(s-1)^4}$$

Therefore, $$\mathscr{L}^{-1}\{Y(s)\} = y(t) = \mathscr{L}^{-1}\left\{\frac{1}{(s-1)^4}\right\} = \frac{1}{6}\,t^3 e^t$$

is the required solution.

Example 25. Solve $\frac{d^2y}{dt^2} + y = 2u_4(t)\sin\pi t, \qquad (1)$

$$y(0) = 1, \; y'(0) = 0.$$

Solution. Taking Laplace transform of both sides of (1), we get

$$s^2 Y(s) - s \cdot 1 - 0 + Y(s) = 2e^{-4s} \frac{\pi}{s^2+\pi^2}, \quad \sin \pi t = \sin \pi(t-4)$$

or
$$Y(s) = \frac{s + \dfrac{2\pi e^{-4s}}{s^2+\pi^2}}{s^2+1}$$

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

$$= \frac{s}{s^2+1} + 2\pi e^{-4s}\left[\frac{1}{\pi^2-1} \cdot \frac{1}{s^2+1} - \frac{1}{\pi^2-1} \cdot \frac{1}{s^2+\pi^2}\right]$$

$$\mathcal{L}^{-1}\{Y(s)\} = y(t) = \cos t + \frac{2\pi}{\pi^2-1} u_4(t) \sin(t-4) - \frac{2}{\pi^2-1} u_4(t) \sin \pi(t-4)$$

$$= \cos t + \frac{2}{\pi^2-1} u_4(t)\left[\pi \sin(t-4) - \sin \pi t\right]$$

is the required solution.

Example 26. Solve

$$\frac{d^3y}{dt^3} - 6\frac{d^2y}{dt^2} + 11\frac{dy}{dt} - 6y = e^{4t}, \quad (1)$$

$$y(0) = y'(0) = y''(0) = 0.$$

Solution. Taking Laplace transform of both sides of (1), we have

$$s^3 Y(s) - s^2 y(0) - s y'(0) - y''(0) - 6s^2 Y(s) + 6s y(0) + 6y'(0)$$
$$+ 11s Y(s) - 11 y(0) - 6Y(s) = \frac{1}{s-4}$$

or
$$(s^3 - 6s^2 + 11s - 6)\, Y(s) = \frac{1}{s-4}$$

or
$$Y(s) = \frac{1}{(s^3 - 6s^2 + 11s - 6)(s-4)}$$

$$= \frac{1}{(s-1)(s-2)(s-3)(s-4)}$$

$$= \frac{A}{s-1} + \frac{B}{s-2} + \frac{C}{s-3} + \frac{D}{s-4}$$

$$= -\frac{1}{6}\,\frac{1}{s-1} + \frac{1}{2}\,\frac{1}{s-2} - \frac{1}{2} \cdot \frac{1}{s-3} + \frac{1}{6} \cdot \frac{1}{s-4}$$

Therefore,
$$\mathcal{L}^{-1}\{Y(s)\} = y(t) = -\frac{1}{6} e^{t} + \frac{1}{2} e^{2t} - \frac{1}{2} e^{3t} + \frac{1}{6} e^{4t}$$

is the solution of (1).

Example 27. Solve
$$\frac{d^4y}{dt^4} - y = u_1(t) - u_2(t), \quad (1)$$

$$y(0) = y'(0) = y''(0) = y'''(0) = 0$$

Solution. Taking Laplace transform of (1), we have

$$s^4 Y(s) - s^3 y(0) - s^2 y'(0) - s y''(0) - y'''(0) - Y(s) = \frac{e^{-s}}{s} - \frac{e^{-2s}}{s}$$

or
$$(s^4 - 1) Y(s) = \frac{e^{-s} - e^{-2s}}{s}$$

or
$$Y(s) = (e^{-s} - e^{-2s}) \frac{1}{s(s-1)(s+1)(s^2+1)}$$

Now,
$$\frac{1}{s(s-1)(s+1)(s^2+1)} = \frac{A}{s} + \frac{B}{s-1} + \frac{C}{s+1} + \frac{Ds+E}{s^2+1}$$

or
$$1 = A(s^4-1) + Bs(s+1)(s^2+1) + Cs(s-1)(s^2+1) + (Ds+E)\, s(s^2-1) \quad (2)$$

Setting $s = 0, 1, -1$ into (2), we find

$$A = -1, \qquad B = \frac{1}{4}, \qquad C = \frac{1}{4}.$$

Equate coefficients of like terms in (2).

Coeff. of s^4 : $0 = A + B + C + D$ or $D = \frac{1}{2}$

Coeff. of s^3 : $0 = B - C + E$ or $E = 0.$

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

Taking inverse Laplace transform, we have

$$\mathcal{L}^{-1}\{Y(s)\} = y(t) = u_1(t)\left[-1 + \frac{1}{4} e^{t-1} + \frac{1}{4} e^{-(t-1)} + \frac{1}{2}\cos(t-1)\right]$$
$$- u_2(t)\left[-1 + \frac{1}{4} e^{t-2} + \frac{1}{4} e^{-(t-2)} + \frac{1}{2}\cos(t-2)\right]$$

as the desired solution.

Example 28. Solve $t\dfrac{d^2y}{dt^2} - t\dfrac{dy}{dt} - y = 0,$ (1)

$$y(0) = 0,\; y'(0) = 3.$$

Solution. We have

$t = -\dfrac{d}{ds}$

$$\mathcal{L}\left\{t\frac{d^2y}{dt^2}\right\} = -\frac{d}{ds}\,\mathcal{L}\left\{\frac{d^2y}{dt^2}\right\}$$

$$= -\frac{d}{ds}\,\{s^2 Y(s) - s\,y(0) - y'(0)\}$$

$$= -s^2 Y'(s) - 2s\,Y(s) + y(0)$$

$$= -s^2 Y'(s) - 2s\,Y(s). \qquad (2)$$

$$\mathcal{L}\left\{t\frac{dy}{dt}\right\} = -\frac{d}{ds}\,\{s\,Y(s) - y(0)\}$$

$$= -s\,Y'(s) - Y(s) \qquad (3)$$

Taking Laplace transform of (1) and using (2) and (3), we get

$$-s^2 Y'(s) - 2s\,Y(s) + s\,Y'(s) + Y(s) - Y(s) = 0$$

or $\quad (s^2 - s)\,Y'(s) + 2s\,Y(s) = 0$

i.e., $\quad Y'(s) + \dfrac{2s}{s^2 - s}\,Y(s) = 0$

or $\quad \dfrac{Y'(s)}{Y(s)} = \dfrac{-2}{s-1}$

Integrating, we have

$$\ln|Y(s)| = -2\ln|s-1| + \ln c$$

or $\quad Y(s) = \dfrac{c}{(s-1)^2}$

Therefore, $\mathcal{L}^{-1}\{\,Y(s)\,\} = y(t) = c\,t\,e^t$ (4)

Differentiate (4) w.r.t t and apply the initial conditions.

$$y'(t) = c\,e^t + c\,t\,e^t$$

$$3 = y'(0) = \left[c\,(t+1)\,e^t\right]_{t=0} = c.$$

Thus the solution is $y(t) = 3t\,e^t$.

Example 29. Find the solution $(x(t), y(t))$ of the system

$$\frac{dx}{dt} - x + y = 2e^t, \qquad x(0) = 0$$

$$\frac{dy}{dt} + x - y = e^t, \qquad y(0) = 0$$

Solution. Taking Laplace transform, we have

$$s\,X(s) - x(0) - X(s) + Y(s) = \frac{2}{s-1}$$

$$s\,Y(s) - y(0) + X(s) - Y(s) = \frac{1}{s-1}$$

or $\quad (s-1)\,X(s) + Y(s) = \dfrac{2}{s-1}$ (1)

and $\quad X(s) + (s-1)\,Y(s) = \dfrac{1}{s-1}$ (2)

Multiply (2) by $s-1$ and subtract (1) from the resulting equation to have

$$[(s-1)^2 - 1]\,Y(s) = 1 - \frac{2}{s-1} = \frac{s-3}{s-1}$$

or $\quad Y(s) = \dfrac{s-3}{(s-1)\,s(s-2)}$

$$= \frac{A}{s} + \frac{B}{s-1} + \frac{C}{s-2}$$

$$= -\frac{3}{2}\cdot\frac{1}{s} + \frac{2}{s-1} - \frac{1}{2}\cdot\frac{1}{s-2}$$

Therefore, $\mathcal{L}^{-1}\{\,Y(s)\,\} = y(t) = -\dfrac{3}{2} + 2e^t - \dfrac{1}{2}\,e^{2t}.$

Differentiating w.r.t. t, we obtain

$$\frac{dy}{dt} = 2e^t - e^{2t}.$$

Substituting for y and $\dfrac{dy}{dt}$ into second equation of the system, we get

$$x = e^t + y - \frac{dy}{dt} = e^t - \frac{3}{2} + 2e^t - \frac{1}{2}\,e^{2t} - 2e^t + e^{2t}$$

$$= -\frac{3}{2} + e^t + \frac{1}{2}\,e^{2t}.$$

Solution of the system is

$$x(t) = -\frac{3}{2} + e^t + \frac{1}{2}\,e^{2t},$$

$$y(t) = -\frac{3}{2} + 2e^t - \frac{1}{2}\,e^{2t}.$$

EXERCISE 11.3

Use the Laplace transform method to solve the following initial value problems.

1. $\frac{dy}{dt} - ky = ce^{at}$, $\quad y(0) = 0$

2. $\frac{dy}{dt} + 4y = 2e^{t} - 4e^{-t}$, $\quad y(0) = 0$

3. $\frac{dy}{dt} + y = f(t)$ where $f(t) = \begin{cases} 0 & \text{if } 0 \le t < 1 \\ 5 & \text{if } t \ge 1 \end{cases}$ $\quad y(0) = 0$

4. $\frac{dy}{dt} + 2y = f(t)$ where $f(t) = \begin{cases} t & \text{if } 0 \le t < 1 \\ 0 & \text{if } t \ge 1 \end{cases}$ $\quad y(0) = 0$

5. $\frac{dy}{dt} = \cos t + \int_0^t y(u)\cos(t-u)\,du$, $\quad y(0) = 1$

6. $\frac{d^2y}{dt^2} + 2\frac{dy}{dt} - 3y = e^{t}$, $\quad y(0) = 1,\ y'(0) = 0$

7. $\frac{d^2y}{dt^2} + y = \cos t$, $\quad y(0) = 0,\ y'(0) = -1$

8. $\frac{d^2y}{dt^2} + y = 4t\sin t$, $\quad y(0) = 0 = y'(0)$

9. $\frac{d^2y}{dt^2} - 2\frac{dy}{dt} = 20\,e^{-t}\cos t$, $\quad y(0) = 0 = y'(0)$

10. $\frac{d^2y}{dt^2} - 3\frac{dy}{dt} - 4y = 12e^{-3t}\sin 2t$, $\quad y(0) = 1,\ y'(0) = 0$

11. $\frac{d^2y}{dt^2} - 4\frac{dy}{dt} + 4y = u_1(t)$, $\quad y(0) = 0,\ y'(0) = 1$

12. $\frac{d^2y}{dt^2} - 3\frac{dy}{dt} + 2y = f(t)$ where $f(t) = \begin{cases} 0 & \text{if } 0 < t < 2 \\ 3 & \text{if } 2 < t < 5 \\ 0 & \text{if } 5 > t \end{cases}$ $\quad y(0) = 0 = y'(0)$

13. $\frac{d^2y}{dt^2} + 2\frac{dy}{dt} + y = 2(t-3)\,u_3(t)$, $\quad y(0) = 2,\ y'(0) = 1$

14. $\frac{d^2y}{dt^2} + y = \begin{cases} \cos t & \text{if } 0 \le t < \frac{\pi}{2} \\ 0 & \text{if } \frac{\pi}{2} < t < \infty \end{cases}$, $\quad y(0) = 3,\ y'(0) = -1$

15. $\frac{d^2y}{dt^2} + 4y = \sin t - u_{2\pi}(t)\sin(t - 2\pi)$, $\quad y(0) = 0 = y'(0) = 0$

16. $\frac{d^3y}{dt^3} - 4\frac{d^2y}{dt^2} + \frac{dy}{dt} + 6y = t\,e^{t}$, $\quad y(0) = 0 = y'(0),\ y''(0) = 1$

17. $\frac{d^3y}{dt^3} - 5\frac{d^2y}{dt^2} + 7\frac{dy}{dt} - 3y = 20\sin t$, $\quad y(0) = 0 = y'(0),\ y''(0) = -2$

18. $\left(\frac{d^2}{dt^2} + 6\frac{d}{dt} + 7\right)^2 y = 0$, $\quad y(0) = 0 = y'(0) = y''(0),\ y'''(0) = 4\sqrt{2}$

19. $\frac{d^4y}{dt^4} + 5\frac{d^2y}{dt^2} + 4y = 1 - u_\pi(t)$, $\quad y(0) = 0 = y'(0) = y''(0) = y'''(0)$

20. $t\frac{d^2y}{dt^2} + (t-1)\frac{dy}{dt} - y = 0$, $\quad y(0) = 5,\ y(\infty) = 0$

21. $\frac{dx}{dt} - x - 3y = 0$, $\quad x(0) = 2$
 $\frac{dy}{dt} - 5x - 3y = 0$, $\quad y(0) = 1$

22. $\frac{dx}{dt} - 4x - 5y = e^{-4t}$, $\quad x(0) = 0$
 $\frac{dy}{dt} + 4x + 4y = e^{4t}$, $\quad y(0) = 0$

23. $2\frac{dx}{dt} + \frac{dy}{dt} - x - y = 1$, $\quad x(0) = 0$
 $\frac{dx}{dt} + \frac{dy}{dt} + 2x - y = t$, $\quad y(0) = 0$

24. $\frac{dx}{dt} + \frac{dy}{dt} = t$, $\quad x(0) = 3,\ x'(0) = -2$
 $\frac{d^2x}{dt^2} - y = e^{-t}$, $\quad y(0) = 0$

25. $\frac{dx}{dt} + 2\frac{d^2y}{dt^2} = e^{-t}$, $\quad x(0) = 0$
 $\frac{dx}{dt} + 2x - y = 1$, $\quad y(0) = 0 = y'(0)$

SALMAN.

ANSWERS

EXERCISE 1.1 (Page 20)

1. $2 \text{ cis } \frac{5\pi}{6}$ 2. $\text{cis}\left(-\frac{\pi}{2}\right)$ 3. $\text{cis}\left(-\frac{2\pi}{3}\right)$

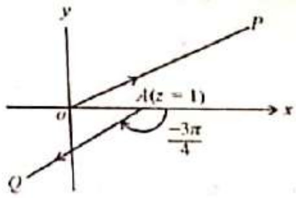
4. $\sqrt{2} \text{ cis}\left(\frac{-3\pi}{4}\right)$ 5. $4 \text{ cis } \frac{\pi}{2}$ 6. $\sqrt{34} \text{ cis}\left(\tan^{-1}\left(-\frac{5}{3}\right)\right)$

7. $\sqrt{3} + i$ 8. $\frac{-5}{\sqrt{2}}(1-i)$ 9. $\frac{-3}{2} - \frac{\sqrt{3}}{2}i$

10. $\frac{5\sqrt{3}}{4} - \frac{5}{4}i$ 11. (i) 40 (ii) $\frac{5}{2}$

14. Greatest value = 31. Least value = 19.

17. (i) Circle with centre (5, 0) and radius 6.
(ii) Set of all points on or outside the circle with centre (0, 2) and radius 1.
(iii) $x = -3$ (iv) $y = 3$ (v) x-axis
(vi) $3x^2 + 4y^2 + 12x = 0$
(vii) $-1 \le x < 1$ (viii) $y < 0$.
(ix) The half-line $\overrightarrow{oP}$
(x) The half-line $\overrightarrow{AQ}$.



EXERCISE 1.2 (Page 35)

1. (i) $2(1 - \sqrt{3}\, i)$ (ii) 81 (iii) 1 2. (i) $\text{cis}(-20\theta)$
(ii) $\text{cis}[-(11\alpha + 9\beta)]$ (iii) $\text{cis}(\alpha + \beta - \gamma - \delta)$ (iv) $\frac{3^7}{4^3} \text{ cis}\left(-\frac{\pi}{2}\right)$

5. (i) $\sqrt{3}+i,\ -\sqrt{3}+i,\ -2i$

(ii) $2\,\text{cis}\left(-\frac{\pi}{8}\right),\ 2\,\text{cis}\left(\frac{3\pi}{8}\right),\ 2\,\text{cis}\left(\frac{7\pi}{8}\right),\ 2\,\text{cis}\left(\frac{11\pi}{8}\right),\ 2\sqrt{2}\ \text{cis}\left(\frac{k\pi}{2}\right),$ $k=0,1,2,3$

$\sqrt{2}\ \text{cis}\left(\frac{k\pi}{2}+\frac{5\pi}{24}\right),\ k=0,1,2,3$

6. $\pm i,\ \frac{\pm\sqrt{3}\pm i}{2},\ \pm 2^{1/12}\,\text{cis}\left(\frac{\pi}{24}\right),\ \pm 2^{1/12}\,\text{cis}\left(\frac{3\pi}{8}\right),\ \pm 2^{1/12}\,\text{cis}\left(\frac{17\pi}{24}\right)$

7. $\text{cis}\ 2(6k+1)\frac{\pi}{15},\ k=0,1,2,3,4$

8. (i) $-1,\ \text{cis}\left(\pm\frac{n\pi}{7}\right),\ n=1,3,5$ (ii) $-1,\ \frac{1\pm i\sqrt{3}}{2},\ \frac{1\pm i}{\sqrt{2}},\ \frac{-1\pm i}{\sqrt{2}}$

(iii) $2^{1/6}\,\text{cis}\left(\frac{k\pi}{3}+\frac{\pi}{9}\right),\ k=0,1,2,3,4,5$

9. $\pm 1,\ \pm i,\ \frac{\pm\sqrt{3}\pm i}{2},\ \frac{\pm 1\pm i\sqrt{3}}{2}$

The last four values are the roots of $x^4+x^2+1=0$

10. (i) $\frac{1}{8}(\cos 4\theta+4\cos 2\theta+3)$ (ii) $\frac{1}{8}[\cos 4\theta-4\cos 2\theta+3]$

(iii) $-\frac{1}{2^5}(\cos 6\theta-6\cos 4\theta+15\cos 2\theta-10)$

(iv) $\frac{1}{2^6}(\cos 7\theta+7\cos 5\theta+21\cos 3\theta+35\cos\theta)$

(v) $\frac{1}{2^8}(\sin 9\theta-9\sin 7\theta+36\sin 5\theta-84\sin 3\theta+126\sin\theta)$

(vi) $-\frac{1}{2^7}(\cos 8\theta-4\cos 6\theta+4\cos 4\theta+4\cos 2\theta-5)$

(vii) $-\frac{1}{2^6}(\sin 7\theta+\sin 5\theta-3\sin 3\theta-3\sin\theta)$

(viii) $-\frac{1}{2^{11}}(\sin 12\theta-2\sin 10\theta-4\sin 8\theta+10\sin 6\theta+5\sin 4\theta-20\sin 2\theta)$

15. $t=-1,\ 2\pm\sqrt{3}$

EXERCISE 1.5 (Page 59)

1. (i) $\dfrac{\cos\left(\frac{n+1}{2}A\right)\sin\frac{n}{2}A}{\sin\frac{A}{2}}$ (ii) $\dfrac{\sin\left(\frac{n+1}{2}A\right)\sin\frac{n}{2}A}{\sin\frac{A}{2}}$ 2. $\dfrac{\sin 2n\theta}{2\sin\theta}$

3. $\dfrac{1-x\cos\theta-x^{n+1}\cos(n+1)\theta+x^{n+2}\cos n\theta}{1-2x\cos\theta+x^2}$

4. $\dfrac{\sin\alpha+(2n+3)\sin n\alpha-(2n+1)\sin(n+1)\alpha}{2(1-\cos\alpha)}$ 5. $\frac{n}{2}+\frac{1}{2}\cos(n+1)\theta\cdot\frac{\sin n\theta}{\sin\theta}$

6. $(2\sin\theta)^{-1/2}\sin\left(\frac{\pi}{4}+\frac{\theta}{2}\right),\ \theta\neq n\pi$ 7. $e^{\cosh\theta}\sinh(\sinh\theta)$ 8. $\dfrac{\sin^2\alpha}{1-\sin 2\alpha+\sin^2\alpha}$

9. $\dfrac{1}{\sqrt{2\cos\theta/2}}\cdot\cos\frac{\theta}{4}$ 10. $\left(2\sin\frac{\theta}{2}\right)^{-n}\sin\left(\frac{n\pi}{2}-\frac{n\theta}{2}\right)$

11. $\dfrac{1}{\sqrt{2\sin\theta/2}}\cdot\cos\left(\frac{\pi-\theta}{4}\right)$

12. $\cos(\alpha-\beta)\sin(\cos\beta)\cosh(\sin\beta)-\sin(\alpha-\beta)\cos(\cos\beta)\sinh(\sin\beta)$

13. $e^{c\cos\theta}\cos(c\sin\theta)$ 14. $\frac{-1}{2}\log(1-2c\cos\theta+c^2),\ |c|<1$

15. $\theta\cos\theta-\sin\theta\ln(2\cos\theta)$

EXERCISE 2.1 (Page 69)

1. (i) False (ii) False (iii) False (iv) False (v) False
(vi) False (vii) False (viii) True (ix) False

3. (i) Group (ii) Group (iii) Not a group (iv) Not a group
(v) Group (vi) Not a group (vii) Group

5. Not a group

EXERCISE 2.2 (Page 80)

1. The set $G=\{\bar{1},\bar{3},\bar{5},\bar{7}\}$ of residue classes with the binary operation in G as multiplication modulo 8 is an abelian group which is not a cyclic group.

2. Since the order of G is a prime number. 3. (i) No (ii) No (iii) No

5. No 6. (i) 1 (ii) 8 (iii) 12

7. Subgroups of order 60, 15, 12, 10, 6, 5, 4, 3, 2 and 1 having generators $a, a^4,$ $a^5, a^6, a^{10}, a^{12}, a^{15}, a^{20}, a^{30}$ and $a^{60} = e$ respectively

8. Subgroups are $C_0 = \{a^{18} = e\}, C_1 = \{e, a^9\}, C_2 = \{e, a^6, a^{12}\}$
$C_3 = \{e, a^3, a^6, a^9, a^{12}, a^{15}\}, C_4 = \{e, a^2, a^4, a^6, a^8, a^{10}, a^{12}, a^{14}, a^{16}\}$
$C_5 = \{a, a^2, a^3, \cdots, a^{17}, a^{18} = e\}$

13. $H, 1+H, 2+H, \cdots, (n-1)+H$, n in all

18. H_1, H_2 and H_3 are not subgroups. H_4 is a subgroup

EXERCISE 2.3 (Page 96)

1. (i) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 5 & 2 & 3 \end{pmatrix}$ (ii) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 4 & 5 & 6 \end{pmatrix}$

(iii) $\begin{pmatrix} 2 & 3 & 6 & 7 & 8 & 9 \\ 9 & 7 & 3 & 2 & 8 & 6 \end{pmatrix}$ 2. $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}$

3. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 6 & 1 & 3 & 2 & 7 & 5 \end{pmatrix}$ 5. $f o g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 3 & 2 & 4 & 6 \end{pmatrix}$

$g o f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 5 & 2 & 4 & 6 \end{pmatrix}$, $g o h = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 2 & 5 & 4 & 1 \end{pmatrix}$

$h o g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 6 & 5 & 3 & 4 \end{pmatrix}$, $f^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 1 & 6 & 3 & 2 \end{pmatrix}$

$g^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 6 & 5 & 4 \end{pmatrix}$, $h^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 2 & 5 & 4 & 3 & 1 \end{pmatrix}$

$|f| = 6, |g| = 4 = |h|$

7. Let $I = (1)$, $a = (1\ 2)(3\ 4)$, $b = (1\ 3)(2\ 4)$, $c = (1\ 4)(3\ 2)$

	I	a	b	c
I	I	a	b	c
a	a	b	c	I
b	b	c	I	a
c	c	I	a	b

8. $a = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 2 & 6 & 1 \end{pmatrix}$ $a^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 6 & 4 & 1 & 3 \end{pmatrix}$

$a^3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 1 & 2 & 3 & 5 \end{pmatrix}$ $a^4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} = I$

9. (124) (13654) = (12) (3654), (1234) (25341) = (153) (24),
(14) (235) (35) (45) = (1524)

10. (i) (18) (364) (57) (ii) (134) (26) (587)

11. (i) (12) (14) (16) (13) (78) (79) (ii) (12) (13) (14)

12. (i) I, (231), (312) (ii) (132), (213), (321)

13. (123) (256) (4351) is an odd permutation
(147) (345) (87) (8345) is an even permutation

14. (i) 2 (ii) 3 (iii) 6 (iv) 3

15. (i) Even (ii) Even (iii) Odd (iv) Odd

EXERCISE 2.4 (Page 105)

1. (i) True (ii) False (iii) False (iv) True
(v) True (vi) True (vii) True

2. (i) No (ii) No (iii) Yes, but not field (iv) No
(v) Ring, but not field (vi) Ring as well as field (vii) Ring, but not field
(viii) Ring as well as field (ix) Ring and field

EXERCISE 3.1 (Page 126)

1. (i) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ (ii) $\begin{bmatrix} 3 & -6 & -10 \\ -2 & 7 & 10 \\ 2 & -6 & -7 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 3 & 5 \\ 1 & -1 & -5 \\ -1 & 3 & 1 \end{bmatrix}$

(iv) $\begin{bmatrix} 11 & -24 & -40 \\ -8 & 27 & 40 \\ 8 & -24 & -27 \end{bmatrix}$ (v) $\begin{bmatrix} 0 & 0 & 10 \\ 0 & 0 & -10 \\ 0 & 0 & 8 \end{bmatrix}$ (vi) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 6 & 8 \end{bmatrix}$

2. (i) $\begin{bmatrix} 2 & 8 \\ 4 & -12 \end{bmatrix}$ (ii) $[ax^2 + by^2 + cz^2 + 2hxy + 2fyz + 2gzx]$

(iii) $\begin{bmatrix} 9 & -4 \\ -8 & 17 \end{bmatrix}$ (iv) $\begin{bmatrix} -7 & 30 \\ 60 & -67 \end{bmatrix}$ (v) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(vi) $\begin{bmatrix} -1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

5. Because $AB \neq BA$ in general. Equality holds if $AB = BA$

8. Nilpotency index is 2

22. $\left[\begin{array}{ccc|c} 9 & 8 & 15 & 4 \\ 19 & 18 & 33 & 7 \\ \hline 0 & 0 & 0 & 2 \end{array}\right]$ 23. $\begin{bmatrix} 0 & 0 & 0 & 4 & -3 \\ 0 & 0 & 0 & -3 & -4 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$

EXERCISE 3.2 (Page 144)

5. (i) $\begin{bmatrix} 1 & 0 & -k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$ (iii) $\begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix}$

(iv) $\begin{bmatrix} 8 & -1 & -3 \\ -5 & 1 & 2 \\ 10 & -1 & -4 \end{bmatrix}$ (v) $\begin{bmatrix} -\frac{i}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{i}{2} \end{bmatrix}$ (vi) $\begin{bmatrix} 0 & \frac{1}{4} & -\frac{1}{2} \\ -1 & \frac{3}{4}i & \frac{1}{2}i \\ 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$

(vii) $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (viii) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 7 & -2 \\ 0 & 0 & 0 & -3 & 1 \end{bmatrix}$

(i) $\begin{bmatrix} 1 & 0 & 0 & \frac{15}{7} \\ 0 & 1 & 0 & -\frac{4}{7} \\ 0 & 0 & 1 & -\frac{10}{7} \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & -\frac{7}{2} & \frac{5}{2} \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 2 & -1 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{5}{6} \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 0 & \frac{4}{11} & \frac{13}{11} \\ 0 & 1 & -\frac{5}{11} & \frac{3}{11} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

8. (i) 2 (ii) 3 (iii) 2 (iv) 3 (v) 1

9. (i) I_2, $P = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix}$, $Q = \begin{bmatrix} 1 & \frac{2}{3} \\ 0 & -\frac{1}{3} \end{bmatrix}$

(ii) $[I_2 \quad 0_{2\times1}]$, $P = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$, $Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

(iii) I_3, $P = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 1 \\ 6 & -3 & -2 \end{bmatrix}$, $Q = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & \frac{3}{11} \\ 0 & 1 & \frac{1}{11} \end{bmatrix}$

(iv) $[I_3 \quad 0_{3\times1}]$, $P = \begin{bmatrix} 1 & 0 & 0 \\ 8 & -2 & 1 \\ -3 & 1 & 0 \end{bmatrix}$, $Q = \begin{bmatrix} 1 & -\frac{2}{11} & 0 & -\frac{3}{77} \\ 0 & \frac{1}{11} & 0 & \frac{1}{11} \\ 0 & \frac{2}{11} & 1 & -\frac{41}{77} \\ 0 & 0 & 0 & \frac{1}{7} \end{bmatrix}$

Note: P and Q are not unique

EXERCISE 4 (Page 174)

1. $x_1 = 1/5, x_2 = -9/20, x_3 = 3/5$ 2. $x_1 = a-3, x_2 = 1, x_3 = 2$
3. $x_1 = 1/2 - 5x_5, x_2 = 1/2 - 3x_5, x_3 = 1/2 + 3x_5, x_4 = -1/2 + 2x_5$ and value of x_5 in arbitrary
4. $x_1 = 1, x_2 = 1, x_3 = 1, x_4 = 1, x_5 = 1$ 5. $x_1 = 3, x_2 = -9/2, x_3 = 0, x_4 = -1$
6. $x_1 = -1, x_2 = -2, x_3 = 4$. The system has unique solution.
7. The system has no solution.
8. (i) System inconsistent (ii) No nontrivial solution
   (iii) System inconsistent (iv) $x_1 = 3, x_2 = 2, x_3 = 0$
9. $\lambda = 0, x_2 = -x_1$; $\lambda = 2, x_2 = x_1$
10. $\lambda = 0, x_1 = 0, x_2 = x_3$, $\lambda = 1, x_1 = x_2 = -x_3$
    $\lambda = 4, x_1 = 2x_3, x_3 = -x_2$
11. $\lambda = 0, x_1 = -2x_2, x_2 = -x_3$ 12. $x_1 = 5, x_2 = 11/5, x_1 = -9/5$
13. $x_1 = -\frac{3}{5}x_3, \quad x_2 = \frac{4}{5}x_3 \quad x_3$ arbitrary.
14. The sytem has infinite number of solutions. $x_1 = 1 - x_3, x_2 = 1 - 2x_3, x_3$ arbitrary
15. The system has no solution.
16. $x_1 = 1, x_2 = 2, x_3 = 1$
17. $x_1 = -19, x_2 = -10, x_3 = 10, x_4 = 0, x_5 = 28$
18. $x_1 = 2, x_2 = -1/2, x_3 = 1/2$ 19. $x_1 = 1, x_2 = -1, x_3 = -1, x_4 = 1$
21. $\begin{bmatrix} 2-s & -1-t \\ -1-s & 1-t \\ s & t \end{bmatrix}$, $s, t \in R$. No.
22. The amounts spent on radio, magazine and T.V. are Rs. 100,000, Rs. 200,000 and Rs. 300,000 respectively.
23. (iii) 850
24. (ii) Along $AB = 100, BC = 100, CE = 400 - s, s - 200, CF = s, 200 \le s \le 400$
    (iii) Along $EA = 100, AB = 200, BC = 200$
    $CE = 300 - s, 0 \le s \le 300$
    $CF = s - 200, 0 \le s \le 200$

$x_1 = t - 10, x_2 = 30 \quad t, x_3 = t - 10, \quad x_4 = 10 + t, \quad x_5 = t$, t being arbitrary

25. If $x_1 = x_3 = 0$, then $x_2 = 20, x_4 = 20, x_5 = 10$
26. $x_1 = t - 200, \quad x_2 = t - 100, \quad x_3 = t + 100, \quad x_2 = 200, \quad x_3 = 400$

(a positive integer ≥ 200)

When $x_4 = 300$, then $x_1 = 100, x_2 = 200, x_3 = 400$

EXERCISE 5.1 (Page 209)

1. The range of the mapping is R if M_2 are real matrices. This mapping is not one-to-one
2. None 3. (i) -6 (ii) 27 (iii) 60
4. (i) -286 (ii) 92 (iii) 0 (iv) 652,992 (v) 0
   (vi) $(a+b+c+d)(a-b+c-d)(a+ib-c-id)(a-ib-c+id)$, $i = \sqrt{-1}$.
20. The other factor is $\begin{vmatrix} 0 & 3 & 2 \\ 2 & 0 & 3 \\ 1 & 1 & 0 \end{vmatrix}$
24. (i) $\begin{bmatrix} 3/4 & -1/4 & -1/4 \\ -1/4 & 3/4 & -1/4 \\ -1/4 & -1/4 & 3/4 \end{bmatrix}$ (ii) $\frac{1}{\Delta}\begin{bmatrix} A & H & G \\ H & B & F \\ G & F & C \end{bmatrix}$
    where A, B, C, F, G, H are cofactors of a, b, c, f, g, h respectively and
    $\Delta = abc + 2fgh + af^2 - bg^2 - ch^2 \neq 0$
    (iii) $\begin{bmatrix} 0 & 2 & -1 \\ 1 & -8 & 5 \\ 1 & -5 & 3 \end{bmatrix}$

EXERCISE 5.2 (Page 219)

1. (i) $x = -1, -6$ (ii) $x = \pm 1, \pm \frac{2}{\sqrt{3}}$ (iii) $x = 2, 3, 4$ (iv) $x = 1, 2, 3, 4$
   (iv) $x = a, a, a, a, -4a$ (vi) $x = 0, 0, 0, -2$

2. (i) $(a-b)^{n-1}[a+(n-1)b]$ (ii) 0 (iii) $\sum n(-1)^{\frac{n(n-1)}{2}} n^{n-2}$

(iv) $xn!\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}+\frac{1}{x}\right)$ 3. $\det 2A = 16$, $\det(B)^{-1} = \frac{1}{3}$

4. For all values of α. 8. $\det P(n) = 1$

9. (i) $1, -1$ (ii) $e^{\pm i\theta}$ (iii) $1, 3, 4$ (iv) $0, 1 \pm 1$

EXERCISE 6.1 (Page 236)

2. (i) Not a vector space. Axiom (iii) (ii) Not a vector space. Axioms (i) and (ii)
   (iii) Not a vector space. Axiom (iii) (iv) Not a vector space. Axiom (v)
3. (i) Yes (ii) Yes (iii) Yes (iv) No, axiom (i) fails to hold
4. (i) No (ii) No (iii) Yes
5. In R^3, let $X = \{(x_1, 0, 0) : x_1 \in R\}$ and $Y = \{(0, x_2, 0) : x_2 \in R\}$. Then X and Y are both subspaces of R^3, but $X \cup Y$ is not a subsapce of R^3.
6. (i) Yes (ii) No (iii) No (iv) No (v) Yes (vi) No
7. (i) Yes (ii) Yes (iii) Yes (iv) Yes
8. (i) No (ii) Yes (iii) Yes
9. Not expressible as a linear combination of the given vectors.
10. $k = -8$ 13. No
15. (i) The straight line $x = t$, $y = -3t$, $z = 5t$ (ii) The plane $9x + 7y + 6z = 0$
    (iii) The plane $6x + 5y + 4z = 0$
17. (i) Let $S = \{(1, 0)\}$, $T = \{(0, 1)\}$ in R^2, $<S> \cup <T> \neq <S \cup T> = R^2$
    (ii) Let $S = \{(0, 0), (0, 1)\}$, $T = \{(0, 0), (0, 3)\}$ in R^2.
    $<S \cap T> = \{(0, 0)\}$, $<S> \cap <T> = \{(0, y) : y \in R\}$.

EXERCISE 6.2 (Page 251)

1. (i) Linearly dependent (ii) Linearly independent
2. (i) Linearly independent (ii) Linearly independent 6. $k = -2$ or 3
7. (i) $\{(1, -3, 1), (2, 1, -4)\}$ (ii) $\{1, \sin^2 x\}$ (iii) $\{1, 4x+3, x^2+2\}$
8. (i) $1 + x = 3(2 - x^2) - (x^3 - x) - (2 - 3x^2) - (3 - x^2)$
   (ii) $x + x^2 = \frac{11}{4}(2 - x^2) - (x^3 - x) - \frac{5}{4}(2 - 3x^2) - (3 - x^3)$

(i) Basis (ii) Basis 10. (i) Basis (ii) Basis

(i) Linearly independent (ii) Linearly dependent

(i) Linearly independent (iv) Linearly dependent (v) Linearly independent

(iii) Linearly independent

(i) $(2, 1, 0), (-5, 0, 1)$ (ii) $(-2, 1, 6)$ (iii) $(1, 0, -3), (0, 1, 2)$

Dimension = 3, Basis = $\{(1, 0, 0, 0), (0, 1, 1, 0), (0, 0, 0, 1)\}$

$\dim U = 2$, $\dim W = 3$ 15. $\dim(U \cap W) = 2$ or 3

(i) $\dim W = 3$. The given set of vectors is a basis for W

(ii) $\dim W = 2$. First two vectors form a basis for W

EXERCISE 6.3 (Page 272)

(i) Linear (ii) Not linear (iii) Not linear

(iv) Linear (v) Not linear (vi) Linear

(a)(i) Linear (ii) Not linear (b)(i) Linear (ii) Not linear

(i) One-to-one (ii) Not one-to-one (iii) One-to-one

$T(x_1, x_2, x_3) = (4x_1 - 2x_2 - x_3, 3x_1 - 4x_2 + x_3)$ 10. $5x_1 - 2x_2$

11. All constant polynomials.
12. $N(T) = \{(0, x_2, 0) : x_2 \in R\}$. T is not one-to-one.
15. (i) Basis for $R(T) = \{(1, 0, 1), (-1, 1, -2)\}$. $\dim R(T) = 2$
    Basis for $N(T) = \{(-3, 1, -1)\}$. $\dim N(T) = 1$
    (ii) Basis for $R(T) = \{(2, 4, 1, 0), (0, 1, 0, -4), (1, 0, 1, 1)\}$
    $\dim R(T) = 3$, $\dim N(T) = 0$, T is one-to-one.

EXERCISE 6.4 (Page 279)

1. (i) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

   (iii) $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

2. (i) $\begin{bmatrix} 3 \\ 5 \end{bmatrix}$ (ii) $\begin{bmatrix} 3 & -5 & 9 \\ 5 & 3 & -2 \end{bmatrix}$ (iii) $\begin{bmatrix} 3 & 4 \\ 5 & -2 \\ 1 & 7 \\ 4 & 0 \end{bmatrix}$ (iv) $[2 \quad 3 \quad -7 \quad 1]$

3. (i) $m = 3,\ n = 5$

$T(x_1, x_2, x_3, x_4, x_5) = (3x_1 + x_2 + 2x_4 + x_5,\ x_1 + x_4 + x_5,\ -x_2 + x_3 + x_4 + x_5)$

(ii) $m = 3, n = 2, T(x_1, x_2) = (6x_1 - x_2,\ x_1 + 2x_2,\ x_1 + 3x_2)$

(iii) $m = 3, n = 3, T(x_1, x_2, x_3) = (x_1 + x_2 + 2x_3,\ 2x_1 + 5x_2 + 6x_3,\ -2x_1 + 3x_2 + x_3)$

4. $T(x_1, x_2, x_3) = (x_2 + x_3,\ x_1 - x_3,\ -x_2 - x_3)$

$$\begin{bmatrix} -11 & -2 & 14 \\ 9 & 2 & -11 \\ -6 & 0 & 9 \end{bmatrix}$$

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

6. $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

EXERCISE 7.1 (Page 294)

1. (ii) $k \geq 92$. (i) $\sqrt{13}$ (ii) 5 3. $\frac{\sqrt{65}}{12}$ 4. $\sqrt{30}$

6. (i) $\begin{cases} < u_1, u_2 > = 6, \\ < u_1, u_3 > = 0 \\ < u_2, u_3 > = -3 \end{cases}$ (ii) $\sqrt{6},\ 3,\ \sqrt{21}$ (iii) $[1 \quad 0 \quad -1]^T$ (iv) $[3 \quad -2 \quad 1]^T$

7. $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ 8. (i) (0, 0, 1), (0, 1, 0), (1, 0, 0)

(ii) $\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right), \left(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$

9. $\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0\right), \left(\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{-2\sqrt{2}}{3}\right), \left(-\frac{2}{3}, -\frac{2}{3}, -\frac{1}{3}\right)$

EXERCISE 7.2 (Page 299)

2. $\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix}$

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

4. $\begin{bmatrix} 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \\ \frac{-5}{\sqrt{30}} & \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{30}} \end{bmatrix}$

EXERCISE 7.3 (Page 306)

1. (i) Characteristic polynomial: $\lambda^2 - 3\lambda - 4$, Eigenvalues: $-1, 4$

Eigenvectors: for $\lambda_1 = -1$, $[-2 \quad 3]^T$

for $\lambda_2 = 4$, $[1 \quad 1]^T$

(ii) Characteristic polynomial: $\lambda^2 - 4\lambda - 1$, Eigenvalues: $2 \pm \sqrt{5}$

Eigenvectors: for $\lambda_1 = 2 + \sqrt{5}$, $[2 \quad 1 + \sqrt{5}]^T$

for $\lambda_2 = 2 - \sqrt{5}$, $[2 \quad 1 - \sqrt{5}]^T$

(iii) Characteristic polynomial: $\lambda^3 - 10\lambda^2 - 28\lambda - 24$, Eigenvalues: 2, 2, 6;

Eigenvectors: for $\lambda_1 = 2$, $[1 \quad 1 \quad 0]^T, [1 \quad 0 \quad 1]^T$

for $\lambda_2 = 6$, $[1 \quad 2 \quad 1]^T$

(iv) Characteristic polynomial: $(\lambda - 1)(\lambda^2 - 6\lambda + 9)$, Eigenvalues: 1, 3, 3.

Eigenvectors: for $\lambda_1 = 1$, $[2 \quad -1 \quad 1]^T$

for $\lambda_2 = 3$, $[1 \quad 1 \quad 0]^T, [1 \quad 0 \quad 1]^T$

(v) Characteristic polynomial: $\lambda^3 - 12\lambda - 16$, Eigenvalues: 4, −2, −2

Eigenvectors: for $\lambda_1 = -2$, $[1 \quad 1 \quad 0]^T, [1 \quad 0 \quad -1]^T$

for $\lambda_2 = 4$, $[1 \quad 1 \quad 2]^T$

(vi) Characteristic polynomial: $(\lambda - 1)(\lambda + 3)^2$, Eigenvalues: −3, −3, 1

Eigenvectors: for $\lambda_1 = -3$, $[1 \quad -1 \quad 0]^T, [2 \quad 0 \quad -1]^T$

for $\lambda_2 = 1$, $[2 \quad 1 \quad 1]^T$.

10. (i) Characteristic polynomial $\lambda^2 - 8\lambda + 12$, Eigenvalues 2, 6

Eigenvectors for $\lambda_1 = 2$, $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$

for $\lambda_2 = 6$, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(ii) Characteristic polynomial $\lambda^2 - 1$, Eigenvalues 1, −1

Eigenvectors for $\lambda_1 = 1$, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

for $\lambda_2 = -1$, $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

(iii) Characteristic polynomial $\lambda^2 + 1$ No real eigenvalues

11. (i) Characteristic polynomial $(\lambda^3 - 1)(\lambda^2 - 5\lambda + 4)$. Eigenvalues 1, 1, 4

Eigenvectors for $\lambda_1 = 1$, $[1 \ 0 \ 0]^T$, $[0 \ -1 \ 1]^T$

for $\lambda_2 = 4$, $[1 \ 1 \ 2]^T$

(ii) Characteristic polynomial $(\lambda - 1)(\lambda^2 + 1)$, Real eigenvalue 1

For $\lambda = 1$, eigenvector $[1 \ 0 \ 0]^T$

(iii) Characteristic polynomial $\lambda^3 - 6\lambda^2 + 11\lambda - 6$, Eigenvalues 1, 2, 3

Eigenvectors for $\lambda_1 = 1$, $[-1 \ 0 \ 1]^T$, $[-2 \ 2 \ 1]^T$

for $\lambda_2 = 3$, $[-1 \ 2 \ 1]^T$

EXERCISE 7.4 (Page 323)

1. (i) Eigenvalues −3, 2 Eigenvectors $[-2 \ 1]^T$, $[1 \ 2]^T$

$$P = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \text{ and } P^T AP = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$$

(ii) Eigenvalues 7, −3 Eigenvectors $[2 \ 1]^T$, $[1 \ -2]^T$

$$P = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}, \quad P^T AP = \begin{bmatrix} 7 & 0 \\ 0 & -3 \end{bmatrix}$$

(iii) Eigenvalues 8, −2 Eigenvectors $[3 \ 1]^T$, $[1 \ -3]^T$

$$P = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{bmatrix}, \quad P^T AP = \begin{bmatrix} 8 & 0 \\ 0 & -2 \end{bmatrix}$$

2. Eigenvalues 1, 1, 4 Eigenvectors $[-1 \ 1 \ 0]^T$, $[-1 \ 0 \ 1]^T$, $[1 \ 1 \ 1]^T$

Orthonormal vectors are

$$v_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}^T, \quad v_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix}^T$$

and $v_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}^T$

$$P = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}, \quad P^T AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

3. (i) Eigenvalues 1, 2, 3 Eigenvectors $[1 \ 0 \ 1]^T$, $[0 \ 1 \ 0]^T$, $[-1 \ 0 \ 1]^T$

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad P^T AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

(ii) Eigenvalues 1, 3, 3. Eigenvectors $[-1 \ 0 \ 1]^T$, $[0 \ 1 \ 0]^T$, $[1 \ 0 \ 1]^T$

$$P = \begin{bmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad P^T AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

(iii) Eigenvalues: $-2, -2, 7$ Eigenvectors: $[1 \;\; 0 \;\; -2]^T, [0 \;\; 1 \;\; 2]^T, [-2 \;\; 2 \;\; -1]^T$

$$P = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{4}{\sqrt{45}} & \frac{-2}{3} \\ 0 & \frac{5}{\sqrt{45}} & \frac{2}{3} \\ \frac{-2}{\sqrt{5}} & \frac{2}{\sqrt{45}} & \frac{-1}{3} \end{bmatrix}, P^T AP = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

(iv) Eigenvalues: $0, 3, 6$. Eigenvectors: $[-2 \;\; 2 \;\; 1]^T, [1 \;\; 2 \;\; -2]^T, [2 \;\; 1 \;\; 2]^T$

$$P = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}, \quad P^T AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

4. (i) Eigenvalues: $0, 2a$ Eigenvectors: $[1 \;\; 1]^T, [-1 \;\; 1]^T$

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}, \quad P^T AP = \begin{bmatrix} 2a & 0 \\ 0 & 0 \end{bmatrix}$$

(ii) Eigenvalues: $3a, 0, 0$. Eigenvectors: $[1 \;\; 1 \;\; 1]^T, [-1 \;\; 1 \;\; 0]^T, [-1 \;\; 0 \;\; 1]^T$

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix}, P^T AP = \begin{bmatrix} 3a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A vector orthogonal to $[-1 \;\; 1 \;\; 0]^T$ is $[1 \;\; 1 \;\; -2]^T$

EXERCISE 8.1 (Page 330)

Note: C = Sequence converges, D = Sequence diverges

1. C, 0	**2.** C, 0	**3.** C, 1	**4.** C, 0	**5.** C, 1
6. C, 0	**7.** D	**8.** C, 0	**9.** D	**10.** C, 0
11. C, 0	**12.** D	**13.** D	**14.** C, 0	**15.** C, $\frac{1}{2}$
16. C, $\frac{1}{5}$	**17.** C, d	**18.** D	**19.** C, $e-1$	**20.** C, 2

EXERCISE 8.2 (Page 347)

Note: C = Series converges, D = Series diverges

1. D	**2.** C, $\frac{x+2}{x+1}$	**3.** C, $\frac{3}{7}(3+\sqrt{2})$	**4.** C, $\frac{1}{2}$
5. C, $\frac{1}{6}$	**6.** C, $\frac{3}{4}$	**7.** D	**8.** C, $\frac{1}{2}$

10. Take $\sum_{1}^{\infty} a_n = \sum_{1}^{\infty}\left(\frac{1}{n^2}+\frac{1}{n}\right)$ and $\sum_{1}^{\infty} b_n = \sum_{1}^{\infty}\left(\frac{1}{n^2}-\frac{1}{n}\right)$

11. C	**12.** D	**13.** D	**14.** C	**15.** C
16. D	**17.** D	**18.** D	**19.** D	**20.** C
21. C	**22.** D	**23.** C	**24.** D	**25.** D
26. C	**27.** C	**28.** C	**29.** D	
30. C if $p > 1$, D if $p \le 1$		**31.** C	**32.** D	**33.** C
34. C	**35.** C	**36.** D	**37.** D	**38.** C
39. C	**40.** D			

EXERCISE 8.3 (Page 354)

Note: C = Series converges, D = Series diverges

1. C	**2.** C	**3.** D	**4.** D	**5.** D
6. D	**7.** C	**8.** C	**9.** C	**10.** C
11. D	**12.** C	**13.** D	**14.** C	**15.** D
16. C	**17.** C	**18.** C	**19.** C	**20.** C
21. C	**22.** C	**23.** D	**24.** C	**25.** D
26. C	**27.** C	**28.** D	**29.** D	**30.** C
31. C	**32.** C	**33.** D	**34.** C	**35.** C

36. C if $a < b$, D if $a \ge b$

40. C for $x < 1$ and D for $x \ge 1$

EXERCISE 8.4 (Page 362)

Note: C = Series converges, D = Series diverges

CC = Series converges conditionally, CA = Series converges absolutely

1. C	2. C	3. D	4. C	5. C
6. C	7. CA	8. CA	9. CC	10. CA
11. CA	12. D	13. CA	14. CC	15. CC
16. D	17. CA	18. CC	19. D	20. CA
21. CC	22. D	23. CA	24. CC	25. CC
26. D	27. D	28. D	29. CA	

30. CA for $|x| < 3$, D for $|x| \geq 3$
31. CA for $|x| < 1$, CC for $x = -1$ D for $|x| \geq 1$
32. CA for $|x| > 4$, D for $0 < |x| \leq 4$
33. CA for all x
34. CA for $|x| \leq 1$ and D for $|x| > 1$
35. CA for all x

EXERCISE 8.5 (Page 375)

1. Radius of convergence $= \infty$ Interval of convergence $=]-\infty, \infty[$

2. $\frac{1}{2}, \left[-\frac{1}{2}, \frac{1}{2}\right[$	3. $1, [4, 6, [$	4. $\infty,]-\infty, \infty[$
5. $1,]1, 3[$	6. $\infty,]-\infty, \infty[$	7. $\frac{1}{2}, \left]\frac{1}{2}, \frac{3}{2}\right[$
8. $2, [0, 4[$	9. $1,]-1, 1[$	10. $\infty,]-\infty, \infty[$
11. $0, [-1, -1]$	12. $\frac{1}{2}, \left[\frac{5}{2}, \frac{7}{2}\right]$	13. $1,]-1, 1[$
14. $3,]-3, 3[$	15. $1, [-1, 1]$	16. $2, [-2, 2[$
17. $\frac{\pi}{6}, \left]-\frac{\pi}{6}, \frac{\pi}{6}\right[$	18. $e,]0, 2e[$	19. $\frac{c}{a}, \left[\frac{-c-b}{a}, \frac{c-b}{a}\right]$
20. $2,]-1, 3[$	21. $\frac{1}{e}$	22. 1
23. 1	24. 3	25. $\sum_{n=0}^{\infty} (-1)^n (n+1) x^{2n+1}$

26. $\frac{2}{(1-x)^3}$ 28. $x + \frac{x^3}{3} + \frac{2x^5}{15} + \cdots, \; -\frac{\pi}{2} < x < \frac{\pi}{2}$

29. 0.494 30. 0.187

EXERCISE 9.1 (Page 384)

1. (i) Ordinary, third order, first degree, linear
   (ii) Ordinary, first order, first degree, nonlinear
   (iii) Partial, third order, first degree, nonlinear
   (iv) Partial, second order, first degree, nonlinear
   (v) Ordinary, second order, third degree, nonlinear

2. (i) $\frac{dy}{dx} + y = x + 1$ (ii) $\frac{dy}{dx} + 3y = 3x^2 e^{-3x}$
   (iii) $(y^2 - y) \frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^2 = 0$
   (iv) $x(2+x) \frac{d^4y}{dx^4} = (2x+6) \frac{d^2y}{dx^2} - (6 - x^2) \frac{d^3y}{dx^3}$
   (v) $\left[1 + \left(\frac{dy}{dx}\right)^2\right] \frac{d^3y}{dx^3} - 3 \frac{dy}{dx} \left(\frac{d^2y}{dx^2}\right)^2$
   (vi) $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$ (vii) $a^2 \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial y^2}$

3. (i) $\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{3/2} = a \frac{d^2y}{dx^2}$
   (ii) $(x^2 + y^2) \frac{d^2y}{dx^2} = 2\left(x \frac{dy}{dx} - y\right)\left[1 + \left(\frac{dy}{dx}\right)^2\right]$
   (iii) $xy \frac{d^2y}{dx^2} + x\left(\frac{dy}{dx}\right)^2 - y \frac{dy}{dx} = 0$ (iv) $2a \frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^3 = 0$
   (v) $xy \frac{d^2y}{dx^2} + x\left(\frac{dy}{dx}\right)^2 - y \frac{dy}{dx} = 0$ (vi) $xy \frac{d^2y}{dx^2} + x\left(\frac{dy}{dx}\right)^2 - y \frac{dy}{dx} = 0$

4. (i) $y = \sqrt{25 - x^2}$ (ii) $y = (x^2 + 3/e) e^{-x}$ (iii) $y = -2e^{-3x}$
   (iv) $y = x^2 + \frac{1}{x^2}$ (v) $y = 2x - 3x^2 + x^3$

5. (i) $y = \sin x + \cos x$ (ii) $y = \frac{1}{e - e^2} (e^x - e^{2x})$

EXERCISE 9.2 (Page 388)

1. $3y^2 - 2\ln|1+x| = c, \quad x \neq -1, \ y \neq 0$
2. $\frac{1}{y} + \cos x = c$
3. $\arctan y - x - \frac{x^2}{2} = c$
4. $\ln\sqrt{x^2+2x} + \ln|y+2| = c$
5. $(2-y)^{-6} = c\,e^{2x^3-3x^2-6x}$
6. $\sin x - \cos y = c$
7. $x + y + \ln|xy| = c$
8. $\sqrt{1+x^2} + \sqrt{1+y^2} + \ln\left[\frac{(\sqrt{1+x^2}-1)(\sqrt{1+y^2}-1)}{|xy|}\right]$

$\tan^{-1}$

9. $y = \sin(c - \arcsin x)$ if $|x| < 1$ and $|y| < 1$,
   $y = \cosh(c - \cosh^{-1} x)$ if $|x| > 1$ and $|y| > 1$
10. $(1+y)(1+e^x) = c\,e^y$
11. $\frac{\sqrt{y}}{(y^2+2)^{1/4}} = \left(\frac{cx}{x+3}\right)^{1/3}$
12. $e^y(\sin x + \cos x) = c$
13. $y = \ln\left|\frac{x^2}{2} + c\right| - x$
14. $e^{x^2} = -2e^{-y}(y+1) + c$
15. $2\ln|x| = \ln|\sec y| - \tan y + c$
16. $y = 1 - \sqrt{x^3 + 2x^2 + 2x + 4}$
17. $16(x+3)(x+2)^2 = 9(y^2+4)^2$
18. $\ln|y| + y^2 = 1 + \sin x$
19. $4x - 2\sin 2x + \tan y = \frac{\pi}{3}$
20. $y = -\sqrt{\frac{x^2+1}{2}}$

EXERCISE 9.3 (Page 393)

1. $\frac{1}{4}\ln(x^2+y^2) + \arctan\left(\frac{y}{x}\right) = c$
2. $|yx^3| = c|y+3x|$
3. $|y^2 - x^2| = |cx|x^2$
4. $|\sin(y/x)|^3 = cx^2$
5. $\arctan(y/x) - \ln|x| = c$
6. $y^2 - 2cx + c^2 = 0$
7. $|y - x| = c|y + 3x|^5$
8. $\ln|x| = \cos\left(\frac{y}{x}\right) + c$
9. $(x^2+y^2)^{3/2} = x^3\ln|cx^3|$
10. $x + \sqrt{x^2 - y^2} = c$
11. $y = x + x\ln|x|$
12. $y + \sqrt{x^2+y^2} = x^2$

$\int \frac{1}{\sqrt{1-y^2}}$

13. $(y+2x)^2 = 12(y-x)$
14. $2(y^2+3xy+3x^2)^2 = 9x^5$
15. $\frac{-2(x-2)}{x+y-3} = \ln c|x+y-3|$
16. $|y+x+4|(y+4x+13)^2 = c$
17. $(x+y-1)^5(x-y+1)^2 = c$
18. $x - y + c = \ln|3x - 4y + 1|$
19. $(y-x)^2 + 10y - 2x = c$
20. $x^2 - y^2 - 4xy + 10x + 2y = c$

EXERCISE 9.4 (Page 398)

1. $x^3 + 2x^2y + y^2 = c$
2. $x^2y + xy + (1-x)\tan y = c$
3. $x^2 + 2xy + 1 = c(y-1)$
4. $ax^2 + 2hxy + y^2 = c$
5. $x\ln xy + y = c,\ x > 0,\ y > 0$
6. $\ln\left|\frac{1+xy}{1-xy}\right| + x^2 = c$
7. $3x^2y + 2y^2x - 5x - 6y = c$
8. $y\tan x + \sec x + y^2 = c$
9. $y\sin x + x^2e^y - y = c$
10. $e^{3y}\cos 2x + x^2 - 3y = c$
11. $x^2y - 3x + 2y^2 = 7$
12. $x^2\cos y + x^3y - \frac{y^2}{2} = -2$
13. $x^3y^2 - y^3x + x^2 + y + 1 = 0$
14. $2x - 3y + y^2 = 2xy$
15. $x^4e^{x+y} + x^2 + y^2 = 1$

EXERCISE 9.5 (Page 402)

1. $\frac{x}{y} + \frac{x^2}{2} = c$
2. $y = x\tan(x+c)$
3. $y = x(c - x - \ln|x|)$
4. $-xy + \cos x = c$
5. $x^2 + \frac{e^x}{y} = c$
6. $xy + y^2 + \frac{2y}{y^2} = c$
7. $e^x(x^2+y^2) = c$
8. $x^2 - y^2 = cx$
9. $x^3y^2(1+xy) = c$
10. $x^4 + x^3y^2 = c$
11. $y = c + \frac{1}{xy} + \ln|x|$
12. $x^3y^3 + x^2 = cy$
13. $\ln(x^2+y^2)^{1/2} + \arctan\frac{y}{x} = c$
14. $y = c\,e^x + e^{2x} + 1$

15. $\ln|x| + \frac{x}{y} = c$

16. $x^3y + \frac{1}{2}x^2y^2 = c,\ x > 0,\ y > 0$

17. $(3x^2y + 2xy + y^3)e^{3x} = c$

18. $xe^{2y} - \ln|y| = c$

19. $e^x \sin y + y^2 = c$

20. $x^2 e^y \sin y = c$

EXERCISE 9.6 (Page 407)

1. $y = \frac{1}{2}xe^{-2x} + \frac{c}{x}$

2. $y = x^3 + cx^{-3}$

3. $y = \frac{x^3 + c}{\ln x}$

4. $y = (x^3 + c)e^{-3x}$

5. $1 + y = \tan x + ce^{-\tan x}$

6. $y = -\cot x + \frac{1}{x} + \frac{c}{x}\csc x$

7. $y = (e^x + c)(x + 1)^n$

8. $3(x^2 + 1)y = 4x^3 + c$

9. $y = \frac{1}{x^2}(c - x\cos x + \sin x)$

10. $y = \frac{1}{(1 + x^2)^2}\left(c + \arctan x\right)$

11. $x = \frac{e^y}{2} + ce^{-y}$

12. $x = cy + y^3$

13. $\frac{1}{y} = cx + 1 + \ln x$

14. $\frac{1}{y^2} = x + \frac{1}{2} + ce^{2x}$

15. $\ln y = 2x^2 + cx$

16. $(x^2 + 1)^2 y = \frac{x^4}{4} + \frac{x^2}{2} + 19$

17. $y = (e^x + 1)^2$

18. $x^2y^4 = x^4 + 15$

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

20. $y = \frac{1}{x^3\left(\frac{1}{2} - \ln x\right)},\ 0 < x < \sqrt{e}$

EXERCISE 9.7 (Page 411)

1. $x = \frac{k}{y}$

2. $2x^2 + y^2 = k$

3. $x^2 + y^2 = ky$

4. $y^2(\ln y^2 - 1) = 2(k - x^2)$

5. $x = y - 1 + ke^{-y}$

6. $y^2 - x^2 = k$

7. $y^2 = 2x - 1 + ke^{-2x}$

8. $16y^3 = 9(k - x)^2$

9. $3x^2y + y^3 = k$

10. $(x - k)^2 + y^2 = k^2 - 1,\ k^2 > 1$

11. $r = b(1 - \sin\theta)$

12. $r^2 = b\cos 2\theta$

13. $r^n = b^n \sin n\theta$

14. $r\sin^3\theta = b(1 + \cos\theta)^2$

15. $r^{n^2} = b\cos n\theta$

16. $r^n \sin n\theta = b^n$

EXERCISE 9.8 (Page 419)

1. $(y - 2x - c)(y + 3x - c) = 0$

2. $(y - cx^2)(yx^3 - c) = 0$

3. $(x^2 + y^2 - c)(y - x - c) = 0$

4. $(2y - x^2 - c)(y - ce^x)(y + x - 1 - ce^{-x}) = 0$

5. $(2y - x^2 + c)(xy - x + c) = 0$

6. $(y^2 + 2x - c)(y + \ln|cx|) = 0$

7. $(y - 3x - c)(x^3 - 3\ln|cy|) = 0$

8. $(x^2 + y^2 - c)(y - ce^x) = 0$

9. $(y^2 + 2xy - c)(y^2 + 2xy - x^2 - c) = 0$

10. $[y^2(y^2 + 2x^2) - c][y^2 - 2x^2\ln|cx|] = 0$

11. $c^2x^3 - 3cy + 9 = 0$

12. $2y = c^2 + cx^2$

13. $12y = c(c + 4x^3)$

14. $x^3(9y + c^2) + 3c = 0$

15. $p^3(5x + 2p)^2 = c$ and the given equation

16. $x^2 = cp^{-4/3} - 2/p$ and the given equation.

17. $c^2x^2 - 2cy + a = 0$

18. $y = \frac{-1}{1 + p^2} + c$ and the given equation

19. $64y = c(4x - c)^2$

20. $y = \frac{c - a\cosh^{-1} p}{\sqrt{p^2 - 1}} - ap$ and the given equation

21. $e^{2y} = ce^{2x} + c^2$

22. $y^2 - 2cx + c^2 = 0$

23. $\sin y = c\sin x + c^2$

24. $y^2 - cx^2 = -\frac{2c}{1 + c}$

25. $x = cy + c^2xy$

26. G.S. $y = cx - \ln c$, S.S $y = 1 + \ln x,\ x > 0$

27. G.S. $y = cx - e^c$. S.S. $y = x\ln x - x$

28. G.S. $y = cx + a\sqrt{1+c^2}$ S.S. $x^2 + y^2 = a^2$

29. G.S. $y = cx - \sqrt{c}$ S.S. $y = -\frac{1}{4x}$

30. G.S. $y = cx + c^3$. S.S. $27y^2 = -4x^3$

EXERCISE 9.9 (Page 429)

Note: G.S. = General Solution, S.S. = Singular Solution

1. G.S. $y = cx - c^n$. S.S. $\left(\frac{x}{n-1}\right)^{n-1} - \left(\frac{x}{n}\right)^{n} = 0$

2. G.S. $(y - cx)^2 = a^2c^2 - b^2$. S.S. $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

3. G.S. $y = cx - \frac{1}{c-1}$. S.S. $(x-y)^2 - 4x = 0$

4. G.S. $(y - cx)^2 = c^2$ No S.S.

5. G.S. $(y - cx)^2 = c(x-1)^2$. S.S. $x = 0$

6. G.S. $c^2 - 2cx^2 - 4y = 0$. S.S. $x^4 - 4y = 0$

7. G.S. $3c^2 - 3cx - y^3 = 0$. S.S. $3x^2 - 8y^3 = 0$

8. G.S. $c^2 - cxy + x = 0$. S.S. $xy^2 = 4,\ x = 0$

9. G.S. $c^2x^2 - 2c^2y - 12 = 0$. S.S. $9y^4 = 64x^3$

10. G.S. $c[x^2 - c(y-c)^2] = 0$. S.S. $27x^3 - [illegible] = 0$

11. $y - 1 = 0$
12. $x = 0,\ x = 1,\ x = 2$

13. $y = \pm 1$
14. $4x^3 - 27y = 0$

15. No singular solution.
16. $2x^4 - 27y^2 = 0$

17. $y = x^2 + \frac{x^4 e^x}{c - 2e^x(x-1)}$
18. $y = \frac{1 + e^{cx}}{x e^{cx}}$

19. $y = \frac{2x - 2c + 1}{c - x}$
20. $y = \frac{e^{16(x+c)} + 15}{e^{16(x+c)} - 1}$

21. $y = \frac{1 + c\cos x}{(c + \cos x)\sin x}$

EXERCISE 10.1 (Page 439)

1. $y = (c_1 + c_2x)e^{\frac{2x}{3}}$
2. $y = \left(c_1\cos\frac{\sqrt{11}\,x}{18} + c_2\sin\frac{\sqrt{11}\,x}{18}\right)e^{-\frac{x}{18}}$

3. $y = c_1e^{-x} + c_2e^{2x} + c_3e^{3x}$
4. $y = c_1e^{-x} + c_2\sin x + c_3\cos x$

5. $y = (c_1 + c_2x + c_3x^2)e^{2x}$
6. $y = c_1e^{-x} + c_2e^{-2x} + c_3e^{3x}$

7. $y = c_1e^{3x} + e^{-\frac{3x}{2}}\left(c_2\cos\frac{3\sqrt{3}}{2}x + c_3\sin\frac{3\sqrt{3}}{2}x\right)$

8. $y = c_1e^{x} + c_2e^{-x} + (c_3 + c_4x)e^{\frac{x}{2}}$

9. $y = (c_1 + c_2x)e^{x} + (c_3\cos x + c_4\sin x)e^{-2x}$

10. $y = c_1e^{-4x} + (c_2 + c_3x + c_4x^2)e^{2x}$
11. $y = c_1e^{-x} + c_2e^{-3x} + c_3e^{3x} + c_4e^{4x}$

12. $y = e^{x}(c_1\cos x + c_2\sin x) + e^{-x}(c_3\cos x + c_4\sin x)$

13. $y = c_1e^{x} + c_2e^{2x} + (c_3 + c_4x)e^{-x}$

14. $y = c_1 + c_2x + (c_3 + c_4x)\cos\frac{x}{2} + (c_5 + c_6x)\sin\frac{x}{2}$

15. $y = (c_1 + c_2x)e^{-2x} + (c_3\cos\sqrt{2}x + c_4\sin\sqrt{2}x)e^{-x}$

16. $y = \frac{9}{10}e^{x-1} + \frac{1}{10}e^{-9(x-1)}$
17. $y = (2 + 3x)e^{-3x}$

18. $y = (3\cos 2x + 4\sin 2x)e^{-3x}$
19. $y = e^{x} - 2e^{2x} + e^{3x}$

20. $y = x^2 + e^x$

EXERCISE 10.2 (Page 451)

1. $y = c_1e^{-4x} + c_2e^{x} + 3xe^{x}$
2. $y = c_1e^{x} + c_2e^{3x} + xe^{x} + xe^{3x}$

3. $y = c_1e^{-x} + c_2e^{3x} - \frac{1}{2}e^{x} + 2\sin x - \cos x$

4. $y = c_1 + c_2e^{x} + c_3e^{\frac{1+\sqrt{5}}{2}x} + c_4e^{\frac{1-\sqrt{5}}{2}x} + \frac{x^5}{5} + 8x^3 - \frac{21x^2}{2} + 97x$

5. $y = c_1e^{x} + c_2\sin x + c_3\cos x - x\sin x + x\cos x$

6. $y = c_1e^{-2x} + (c_2 \cos x + c_3 \sin x)\, e^{2x} - \sin x + 3 \cos x$

7. $y = c_1 \sin 2x + c_2 \cos 2x + \frac{1}{2}(1 - x \sin 2x)$

8. $y = c_1 + c_2 \sin x + c_3 \cos x + \frac{2}{3}x^3 - 4x - \frac{3}{2}x \sin x$

9. $y = c_1 + c_2x + c_3 \sin x + c_4 \cos x + \frac{1}{4}x^4 - 3x^2 + x \sin x + 3x \cos x$

10. $y = e^x \left(c_1 \sin\sqrt{3}\, x + c_2 \cos\sqrt{3}\, x \right) + \frac{1}{2}e^x \cos x$

11. $y = c_1e^{-x} + e^x (c_2 \sin 2x + c_3 \cos 2x) - \frac{1}{16}xe^x (\cos 2x + \sin 2x)$

12. $y = c_1e^{-x} + c_2e^{-2x} + c_3e^{3x} - \frac{1}{12}\left(\frac{17}{12} + x\right)e^{2x}$

13. $y = c_1e^{3x} + c_2e^{4x} + \frac{e^{x}}{8}(4x^3 - 2x^2 - 18x + 45)$

14. $y = c_1e^x + c_2e^{-x} + c_3 \cos 3x + c_4 \sin 3x - \frac{\cos 2x}{5} - x^3 - \frac{16}{3}x$

15. $y = c_1e^x + c_2e^{-x} + c_3 \cos 2x + c_4 \sin 2x + \frac{x}{10}\cosh x + \frac{x \sin 2x}{40} + \frac{1}{8}$

16. $y = 3e^{3x} - 2e^{5x} + 3xe^x + 4e^{2x}$

17. $y = \frac{1}{5}\left[e^{2x}(\sin 3x + 2 \cos 3x) + \sin 3x + 3 \cos 3x\right]$

18. $y = e^{2x} - \frac{1}{2}e^{-2x} + 2x - \frac{1}{2}$

19. $y = 2 \sin x + \cos x - \frac{1}{4}x^2 \cos x + \frac{1}{4}x \sin x$

20. $y = 2e^{-x}(1 - x) \cos 2x.$

EXERCISE 10.3 (Page 456)

$y = (c_1 + c_2x)\, e^{2x} + \frac{1}{2}x^2 e^{2x}$

$y = e^{-x}(c_1 \sin 2x + c_2 \cos 2x) + 2 \sin 2x - \cos 2x$

$y = c_1e^{-x} + c_2e^{-x/2} + (x^2 - 6x + 14) - \frac{3}{10}\sin x - \frac{9}{10}\cos x$

4. $y = (c_1 + c_2x)\, e^{-x} + \frac{1}{25}e^x (3 \cos x + 4 \sin x)$

5. $y = c_1 \sin x + c_2 \cos x - 2 \cos 2x + 6$

6. $y = c_1e^x + c_2e^{2x} - xe^x (x + 2) + x^2 + 3x + \frac{7}{2}$

7. $y = c_1 + c_2 \sin x + c_3 \cos x + \frac{2x^3}{3} - 4x - 2x \sin x$

8. $y = c_1e^x + e^{-x}(c_2 \sin 2x + c_3 \cos 2x) - \frac{9 \sin 2x}{17} + \frac{2 \cos 2x}{17} - 2x^2 - 3x - 4$

9. $y = c_1 \cos 2x + c_2 \sin 2x + c_3 x \cos 2x + c_4 x \sin 2x + \frac{1}{9}\sin x$

10. (i) $y_p = Ae^{-x} + x(Bx^2 + Cx + D)\, e^{-x} \cos x + x(Ex^2 + Fx + G)\, e^{-x} \sin x$

(ii) $y_p = (Ax^2 + Bx + C)\, e^x \sin 2x + (Dx^2 + Ex + F)\, e^x \cos 2x$
$+ e^{-x}(G \sin x + H \cos x) + Ke^x.$

EXERCISE 10.4 (Page 459)

1. $y = c_1x^{-1} + c_2x^{-5} + \frac{x^5}{60}$

2. $y = x^2 (c_1 \sin(\ln x) + c_2 \cos(\ln x)) - \frac{x^2 \ln x}{2}\cos(\ln x)$

3. $y = x^m (c_1 \sin(\ln x^n) + c_2 \cos(\ln x^n) + \ln x)$

4. $y = c_1(-x)^{3/2} + c_2(-x)^{1/2} + \frac{8}{65}\cos \ln(-x) - \frac{1}{65}\sin \ln(-x)$

5. $y = x(c_1 \cos(\ln x) + c_2 \sin(\ln x) + 5) + x^{-1}(c_3 + 2 \ln x)$

6. $y = x(c_1 + c_2 \ln x) + c_3 x^{-1} + \frac{\ln x}{4x}$

7. $y = c_1x^3 + x^{-2}[c_2 \cos(\ln x) + c_3 \sin(\ln x)] + \frac{x^4}{37}$

8. $y = c_1 \cos \ln(x + 1) + c_2 \sin \ln(x + 1) + 2 - \frac{2}{3}\cos \ln(x + 1)^2$

9. $y = (2x + 1)^2 [c_1 + c_2 (\ln(2x + 1) + \{\ln(2x + 1)\}^3]$

10. $y = -x^2 + 2x^{-3} + 2x^2 \ln x$ 11. $y = x - x \ln x - \frac{1}{2}x(\ln x)^2$

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

EXERCISE 10.5 (Page 464)

1. $y = c_1 \sin x + c_2 \cos x + \frac{1}{2} \sec x$

2. $y = \left[c_1 - \ln \left| \tan\left(\frac{\pi}{4} + x\right) \right| \right] \cos 2x + c_2 \sin 2x$

3. $y = c_1 x + c_2 \left(1 - \frac{x}{2} \ln \left| \frac{1+x}{1-x} \right| \right)$ 4. $y = c_1 x + c_2 e^x + 1$

5. $y = x^3 + c_1 x - c_2 x^{-1}$ 6. $y = (x^2 + c_1 x) e^x + c_2 x$

7. $y = e^{-x} [c_1 + c_2 x + \ln (1 + e^{-x})]$ 8. $y = \left(c_1 x^2 + c_2 + x + \frac{x^3}{3} + \frac{x^2}{2} \ln x \right) e$

9. $y = c_1 e^{2x} + c_2 x e^{2x} + e^{2x} \left[\frac{x^2}{1 \cdot 2} + \frac{x^3}{2 \cdot 3} + \cdots + \frac{x^{27}}{26 \cdot 27} \right]$

10. $y = c_1 (2 \cos x - \sec x) + c_2 \sin x + \frac{1}{2} \sec x$ 11. $y = e^{2x} + c_1 x^3 e^x + c_2 e^x$

12. $y = \frac{-x(1-x^2)^{3/2}}{9} + c_1 \left(\sqrt{1-x^2} + x \arcsin x \right) + c_2 x$

EXERCISE 10.6 (Page 473)

1. $y_p = \frac{1}{4} (\cos 2x) \ln |\cos 2x| + \frac{1}{2} x \sin 2x$

2. $y_p = -\sin x + x \cos x + (\sin x) \ln |\sec x|$

3. $y_p = e^x \ln (1 + e^{-x}) - e^x + e^{2x} (\ln 1 + e^{-x})$

4. $y_p = x e^{2x} \sin x + (\ln |\cos x|) e^{2x} \cos x$

5. $y_p = e^{2x} (-x + \ln |1 + x|) + x e^{2x} \ln |1 + x|$

6. $y_p = \frac{1}{4} e^x \left(2x^2 \arcsin x + 3x \sqrt{1-x^2} + \arcsin x \right)$

7. $y_p = \frac{1}{8} e^x \cos 2x \left(2 \sin 2x - \ln \left| x + \frac{\pi}{4} \right| \right) - \frac{e^x}{4} \sin 2x \cos x$

8. $y_p = \frac{e^{-x}}{4} (2x^2 \ln x - 3x^2)$

9. $y_p = -2e^{-x}(1 + \cos^2 x) + e^{-x} \sin x \left[\ln \tan \left(\frac{x}{2} + \frac{\pi}{4} \right) - 2 \sin x \right]$

10. $y = c_1 x + c_2 x^2 + x e^x$

11. $y = \frac{c_1}{x} + c_2 x - \frac{1}{2x} \ln x$

12. $y = c_1 e^x + c_2 x^2 + x^2 e^x - 3x^2$

13. $y = c_1 \sin x + c_2 x \sin x + \frac{x^2}{2} \sin x$ 14. $y_p = 2x e^x \ln |x|$

15. $y_p = \frac{e^{2x}}{10} \int e^{3x} \tan x \, dx + \frac{e^{-x}}{10} \left[(3 \cos x - \sin x) \ln |\sec x + \tan x| + 1 \right]$

EXERCISE 10.7 (Page 481)

1. $y = 2x - 2 \ln |1 - c_1 e^{2x}| + c_2$ 2. $15 c_1^2 y = 4(c_1 x + a)^{5/2} + c_2 x + c_3$

3. $y = c_2 \pm (c - x^2)^{1/2}$ or $x^2 + (y - c)^2 = c_1^2$

4. $y = \frac{1}{5} (3x^5 - 5x^4 + 10x - 8)$ 5. $\pm c_1 x = c_2 + 2(c_1 y - 1)^{1/2}$

6. $2x - 2 = \ln |\ln y|$ 7. $x = c_2 - c_1 y - (1 + c_1^2) \ln |y - c_1|$

8. $y = 1 + \sin \left(\sqrt{8}\, x \right)$ 9. $y(2x + 3) = e^x + c_1 \ln |x| + c_2$

10. $y = -\frac{1}{2} \left(c_1 \cos x - \sin^2 x \ln \left| \tan \frac{x}{2} \right| \right) + c_2 \sin^2 x$

11. $(x + \sin x) y = \frac{c_1 x^2}{2} + c_2 x + c_3 - \cos x$

12. $(e^x + 2x) y = \frac{c_1 x^3}{6} + \frac{c_2 x^2}{2} + c_3 x + c_4 + \frac{1}{24x}$

13. $y = x^3 e^{3x} \int \frac{1}{x^3 e^{3x}} \left(\frac{1}{3} + \frac{2 \ln x}{x^3} + \frac{5}{x^2} + \frac{c_1}{x} \right) dx + c_2$

14. (i) $36y = 6x^3 \ln x - 11x^3 + c_1 x^2 + c_2 x + c_3$

(ii) $y = -x^2 \sin x - 4x \cos x + 6 \sin x + c_1 x + c_2$

15. (i) $\cos y = -x$ (ii) $\sqrt{y + c_1 y^2} - \frac{1}{\sqrt{c_1}} \ln \left(\sqrt{c_1 y} + \sqrt{1 + c_1 y} \right) = 2\sqrt{a}\, c_1 x +$

EXERCISE 10.8 (Page 487)

1. $x = c_1e^{2t} + c_2te^{2t}$, $y = (2c_1 - c_2)e^{2t} + 2c_2te^{2t}$
2. $x = c_1e^{-3t} + c_2e^{2t}$, $y = -4c_1e^{-3t} + c_2e^{2t}$
3. $x = c_1 \cos t + c_2 \sin t$, $y = \left(\frac{-3c_1 + c_2}{2}\right)\cot t - \left(\frac{c_1 + 3c_2}{2}\right)\sin t - \frac{e^t}{2} - \frac{e^{-t}}{2}$
4. $x = c_1e^t + c_2e^{-3t} - \frac{1}{3}t - \frac{11}{36}$, $y = -\frac{1}{2}c_1e^t + \frac{3}{2}c_2e^{-3t} + \frac{1}{8}t + \frac{5}{12}$
5. $x = c_1e^{3t} + c_2\cos 2t + c_3\sin 2t$, $y = -1 - 5c_1e^{3t} - c_2\sin 2t + c_3\cos 2t$
6. $x = e^{2t}(c_1\cos 2t + c_2\sin 2t) + \frac{1}{4}te^{2t}$, $y = 2e^{2t}(c_2\cos 2t - c_1\sin 2t) - \frac{11}{4}te^{2t}$
7. $x = 4$, $y = -5 + \frac{1}{3}e^{t}$
8. $x = e^{t}[(c_2 - c_1)\cos t - (c_1 + c_2)\sin t] + 2\cos t - 3\sin t$
   $y = e^{t}(c_1\cos t + c_2\sin t) - \cos t + 3\sin t$
9. $x = -5c_1e^{-t} - 2c_2e^{2t} - 11e^t - 3$, $y = c_1e^{-t} + c_2e^{2t} + 5e^t + 1$
10. $x = e^t$, $y = 0$
11. $x = \frac{6}{5}e^{4t} + \frac{4}{5}e^{-t} + 3t - 2$, $y = \frac{9}{5}e^{4t} - \frac{4}{5}e^{-t} - 2t + 3$
12. $y = c_1e^t + c_2e^{-t}$, $z = -c_1e^t + c_2e^{-t}$, $x = 3c_1e^t - 3c_2e^{-t} + c_3$
13. $x = c_1 + c_2e^t + c_3e^{4t}$,
    $y = c_4e^t - c_2te^t - \frac{16}{3}c_3e^{4t}$ $\quad z = -c_1 + (c_4 - c_2)e^t - c_2te^t - \frac{7}{3}c_3e^{4t}$

EXERCISE 10.9 (Page 491)

1. $y = c_1x\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\right) = c_1xe^x$ 2. $y = c_1(x + x^2)$
3. $y = c_0\left(1 + kx + \frac{(kx^2)}{2!} + \frac{(kx^3)}{3!} + \cdots\right) = c_0e^{kx}$
4. $y = c_0 - (c_0 - 1)x + (c_0 - 1)\frac{x^2}{2!} - \frac{(c_0 - 1)}{3!}x^3 + \cdots$
5. $y = c_1x(1 + x + x_2 + \cdots)$ 6. $y = c_1x\left(1 + x + \frac{x^4}{2!} + \frac{x^6}{3!} + \cdots\right)$

7. $y = c_0\left(1 + x + \frac{x^2}{2} + \frac{x^3}{8} + \frac{3}{8}x^4 + \cdots\right)$
8. $y = 2x + c_2x^2\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\right) = 2x + c_2x^2e^x$
9. $y = c_0\left(1 + 3x + 4x^2 + \frac{10}{3}x^3 + 2x^4 + \frac{14}{15}x^5 + \cdots\right)$
10. $\arcsin x = x + \sum_{n=1}^{\infty} \frac{1\cdot3\cdot5\cdot\cdots\cdot(2n-1)}{2\cdot4\cdot6\cdot\cdots\cdot2n(2n+1)}x^{2n+1}$

EXERCISE 10.10 (Page 504)

1. $y = (c_0 + c_1x)(1 + x^2 + x^4 + x^6 + \cdots) = \frac{c_0 + c_1x}{1 - x^2}$
2. $y = c_0\left(1 + \frac{x^4}{3\cdot4} + \frac{x^8}{3\cdot4\cdot7\cdot8} + \cdots\right) + c_1\left(x + \frac{x^5}{4\cdot5} + \frac{x^9}{4\cdot5\cdot8\cdot9} + \cdots\right)$
3. $y = c_0\left(1 + \frac{x^5}{4\cdot5} + \frac{x^{10}}{4\cdot5\cdot9\cdot10} + \cdots\right) + c_1\left(x + \frac{x^6}{5\cdot6} + \frac{x^{11}}{5\cdot6\cdot10\cdot11} + \cdots\right)$
4. $y = c_0\sum_{n=0}^{\infty}(-1)^n\frac{1\cdot3\cdot5\cdot\cdots\cdot(2n-1)}{2\cdot4\cdot6\cdot\cdots\cdot2n}x^{2n} + c_1\sum_{n=0}^{\infty}(-1)^n\frac{2\cdot4\cdot6\cdot\cdots\cdot2n}{3\cdot5\cdot7\cdot\cdots\cdot(2n+1)}$
5. $y = c_0\left(1 - x^2 + \frac{x^4}{4} + \cdots\right) + c_1\left(1 - \frac{x^3}{2} + \frac{3}{40}x^5 + \cdots\right)$
6. $y = c_0 + 2(c_0 + 1)x^2 + 2(c_0 + 1)x^4 + \frac{4}{3}(c_0 + 1)x^6 + \cdots + c_1x + \left(\frac{4}{3}c_1 + 1\right)x^3$
   $+ \frac{4}{5}\left(\frac{4}{3}c_1 + 1\right)x^5 + \cdots$
7. $y = c_0 + 2x^2 + \left(\frac{1}{3} - \frac{2}{3}c_0\right)x^3 + \left(\frac{1}{45} - \frac{2}{45}c_0\right)x^6 + \left(\frac{1}{405} - \frac{2}{405}c_0\right)x^9$
   $+ \cdots + c_1x + \left(-\frac{1}{12} - \frac{1}{6}c_1\right)x^4 + \left(-\frac{1}{126} - \frac{1}{63}c_1\right)x^7 + \left(-\frac{1}{1134} - \frac{1}{567}c_1\right)x^{10} + \cdots$
8. $y = c_0\left[1 - \frac{m(m+1)}{2!}x^2 + \frac{(m-2)m(m+1)(m+3)}{4!}x^4 + \cdots\right]$
   $+ c_1\left[x - \frac{(m-1)(m+2)}{3!}x^3 + \frac{(m-3)(m-1)(m+2)(m+4)}{5!}x^5 - \cdots\right]$

9. $y = c_0 + c_1 x - \frac{1}{2}(\alpha+\beta)c_0 x^2 - \frac{1}{6}(\alpha+\beta)c_1 x^3$

$$+\frac{1}{12}\left[\frac{(\alpha+\beta)^2}{2!}c_0 + 2\beta c_0\right]x^4 + \frac{1}{20}\left[\frac{(\alpha+\beta)^2}{3!}c_1 + 2\beta c_1\right]x^5 \ldots$$

10. $y = c_0 + c_1 x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \left(\frac{1}{24} - \frac{1}{4}c_0\right)x^4 + \left(\frac{1}{120} - \frac{3}{10}c_1\right)x^5 - \frac{83}{61}x^6 + \ldots$

11. $y = c_0\left[1 - \frac{1}{3}(x-1)^3 - \frac{1}{12}(x-1)^4 - \frac{1}{45}(x-1)^6 + \frac{1}{84}(x-1)^7 + \ldots\right]$

$$+ c_1\left[(x-1) - \frac{1}{6}(x-1)^4 - \frac{1}{20}(x-1)^5 + \frac{1}{126}(x-1)^7 + \ldots\right]$$

12. $y = c_0\left[1 - \frac{1}{2}(x-2)^2 - \frac{1}{6}(x-2)^3 + \frac{1}{12}(x-2)^4 + \cdots\right]$

$$+ c_1\left[(x-2) + \frac{1}{2}(x-2)^2 - \frac{1}{6}(x-2)^3 - \frac{1}{6}(x-2)^4 + \cdots\right]$$

13. $y = c_0\left[1 - \frac{1}{6}(x-1)^3 + \frac{1}{12}(x-1)^4 - \frac{1}{6}(x-1)^5 + \cdots\right]$

$$+ c_1\left[(x-1) - \frac{1}{2}(x-1)^4 + (x-1)^5 + \ldots\right]$$

14. $y = c_0\left[1 - \frac{1}{2}\left(x-\frac{\pi}{2}\right)^2 + \frac{1}{12}\left(x-\frac{\pi}{2}\right)^4 - \cdots\right]$

$$+ c_1\left[\left(x-\frac{\pi}{2}\right) - \frac{1}{6}\left(x-\frac{\pi}{2}\right)^3 + \frac{1}{30}\left(x-\frac{\pi}{2}\right)^5 - \ldots\right]$$

15. $y = 4 - x - 4x^2 + \frac{1}{2}x^3 + \frac{4}{3}x^4 - \frac{1}{12}x^5 - \frac{4}{15}x^6 + \cdots$

16. $y = 1 - x - \frac{1}{2}x^2 + \frac{1}{12}x^4 + \frac{1}{24}x^5 + \cdots$ 17. $y = 2 + 3x - \frac{7}{6}x^3 - \frac{1}{2}x^4 + \frac{21}{40}x^5 + \ldots$

18. $y = (x+1) + \frac{1}{4}(x+1)^4$ 19. $y = x + \frac{x^2}{2!} + \frac{2x^3}{3!} + \frac{2x^4}{4!} + \frac{5x^5}{5!} + \cdots$

20. $y = 2 + 4(x-1) - 4(x-1)^2 + \frac{4}{3}(x-1)^3 - \frac{1}{3}(x-1)^4 + \frac{2}{15}(x-1)^5 + \cdots$

EXERCISE 10.11 (Page 514)

1. 10.92 ft. 2. $25\,\frac{49e^{2.36t}+47}{49e^{2.36t}-47}$

3. (a) 182.65 ft/sec. (b) $v = \frac{16\left(1+\frac{83}{99}e^{10-t}\right)}{-1+\frac{83}{99}e^{10-t}}$

4. 63.40 years. 5. $P = (6\times 108)\exp.(0.005\,t)$, 2396 C.E.

6. $P = \left(\frac{1}{2}t+20\right)^2$ 7. $20\left[\frac{500-t}{500}\right]^t$, $0 < t < 500$

8. $P = 20e^{\ldots}$ 9. 7.4 minutes

10. 2.32 hours 11. $C = 720\sqrt{x} + 19200$ Rupees

12. t is given by: $3000 = 75000\,e^{\ln(0.75)t}$ i.e., $t = 4.83$ years.

13. 324,000; 2.51 hours 14. $u = \frac{40\left(1-e^{-0.05889t}\right)}{2-e^{-0.05889t}}$

15. 137.45°C

16. $10\frac{d^2y}{dt^2} + 196y = 0$; $y(0) = -0.1$, $y'(0) = 0$, $y = -(0.1)\cos\sqrt{\frac{98t}{5}}$

18. $4\frac{d^2y}{dt^2} + 10\frac{dy}{dt} + 196y = 0$; $y(0) = 0$, $y'(0) = 2$

EXERCISE 11.1 (Page 534)

1. $\frac{2}{s^3} + \frac{6}{s^2} - \frac{7}{s}$, $s > 0$ 2. $\frac{e^5}{s-3}$, $s > 3$

3. $\frac{7\cos 4}{s^2+49} + \frac{s\sin 4}{s^2+49}$, $s > 0$ 4. $\frac{s\cos b}{s^2+a^2} - \frac{a\sin b}{s^2+a^2}$, $s > 0$

5. $\frac{s\cosh 3}{s^2-25} - \frac{5\sinh 3}{s^2-25}$ 6. $\frac{3!}{(s+2)^4} - \frac{1}{s+2}$, $s > -2$

7. $\frac{2}{(s+1)^2+4}$, $s > -1$ 8. $\frac{s-3}{(s-3)^2-16}$, $s > 3$

9. $\dfrac{s}{2(s^2+9)} + \dfrac{s}{2(s^2+1)}$

10. $\dfrac{3}{4}\dfrac{1}{s^2+1} - \dfrac{1}{4}\dfrac{3}{s^2+9}$

11. $\dfrac{2a\,(s+3)}{[(s+3)^2+a^2]^2}$

12. $\dfrac{2a^2}{s(s^2-4a^2)}$

13. $\dfrac{a(s^2+2a^2)}{s^4+4a^4}$

14. $\dfrac{a\,(s^2-2a^2)}{s^4+4a^4}$

15. $\dfrac{s\,(s^2-a^2+b^2)}{(s^2+a^2+b^2)^2-4a^2s^2}$

16. $\dfrac{e^{-s}}{s\,(1-e^{-s})}$

17. $\dfrac{1}{s^{\alpha+1}}\,\Gamma(\alpha+1),\quad s>0,\quad \dfrac{15\sqrt{\pi}}{8\,s^{7/2}}$

18. $\dfrac{2a(3s^2-a^2)}{(s^2+a^2)^3}$

19. $\dfrac{2s(s^2-3a^2)}{(s^2+a^2)^3}$

20. $\dfrac{2a^2(3s^2+4a^2)}{s^2(s^2+4a^2)}$

21. $\dfrac{(s^2+16)^3+s^4(s^2+16)-64\,s^4}{s^2\,(s^2+16)^3}$

22. $\arctan\dfrac{a}{s}$

23. $\dfrac{1}{2}\ln\dfrac{s^2+a^2}{s^2}$

24. $\dfrac{1}{s}\arctan\dfrac{a}{s}$

25. $\dfrac{1}{2s}\ln\dfrac{s^2+a^2}{s^2}$

26. $\dfrac{1}{2}\ln\dfrac{s+a}{s-a}$

27. $\dfrac{1}{s}\Gamma'(1) - \dfrac{\ln s}{s}$

28. $\dfrac{6e^{-3s}}{s^4}$

30. $\dfrac{\sqrt{\pi}}{2\,s^{3/2}}\,e^{-1/4s},\quad \sqrt{\dfrac{\pi}{s}}\,e^{-1/4s}$

EXERCISE 11.2 (Page 541)

1. $\cosh\sqrt{2}\,t - \sqrt{2}\sinh\sqrt{2}\,t$

2. $3\,e^{3t}\cos 3t + \dfrac{10}{3}\,e^{3t}\sin 3t$

3. $9\,e^{3t}\cosh 5t + e^{3t}\sinh 5t$

4. $ae^{-ct}\cos\sqrt{d-c^2}\,t + \dfrac{b-ac}{\sqrt{d-c^2}}\,e^{-ct}\sin\sqrt{d-c^2}\,t$

5. $e^{-at}\cos bt - \dfrac{a}{b}e^{-at}\sin bt$

6. $\dfrac{1}{b^2-a^2}\left(\dfrac{1}{a}\sin at - \dfrac{1}{b}\sin bt\right)$

7. $\dfrac{e^t}{5} - \dfrac{\cos 2t}{5} - \dfrac{\sin 2t}{10}$

8. $\dfrac{7}{9}\,e^{\frac{8t}{3}} + \dfrac{71}{27}\,te^{\frac{8t}{3}}$

9. $t^3 e^{-7t}\left(\dfrac{5}{6} - \dfrac{4}{3}t\right)$

10. $-\dfrac{12}{5}e^{-3t/2} - \dfrac{1}{10}e^t + \dfrac{5}{2}e^{-t}$

11. $2 - e^{-2t} + e^{-2t}\cos 3t - 2e^{-2t}\sin 3t$

12. $\dfrac{-2}{75}\cos 2t - \dfrac{1}{50}\sin 2t + \dfrac{2}{75}e^{-3t}\cosh\sqrt{14}\,t + \dfrac{3}{25\sqrt{14}}e^{-3t}\sinh\sqrt{14}\,t$

13. $e^{-t}\cos 2t - 2t\,e^{-t}\sin 2t$

14. $\dfrac{\sin at}{t}$

15. $\dfrac{2e^t}{t} - \dfrac{2\cos t}{t}$

16. $-\dfrac{2}{t}(\cos at - \cos bt)$

17. $\dfrac{1}{9}u_3(t)\cdot(t-3) - \dfrac{1}{27}u_3(t)\sin 3(t-3)$

18. $-u_\pi(t)\,e^{2(t-\pi)}(\cos t + 2\sin t)$

19. $u_2(t)\,(1-4e^{2t-4}+3\,e^{6t-6})$

20. $3u_3(t)\,e^{5(t-3)}\cos(t-3) + 8u_3(t)\,e^{5(t-3)}\sin(t-3)$

21. $\dfrac{1}{25}(e^{-5t}+5t-1)$

22. $\dfrac{1}{5}(\cos 2t + 2\sin 2t - e^{-t})$

23. $\dfrac{1}{8}(\sin t - \cos t) + \dfrac{e^{-2t}}{8}(\sin t + \cos t)$

EXERCISE 11.3 (Page 548)

1. $y(t) = c\,t\,e^{kt}$

2. $y(t) = \dfrac{2}{5}e^t - \dfrac{4}{3}e^{-t} + \dfrac{14}{15}e^{-4t}$

3. $y(t) = 5u_1(t)\left(1-e^{-(t-1)}\right)$

4. $y(t) = -\dfrac{1}{4} + \dfrac{1}{2}t + \dfrac{1}{4}e^{-2t} + \dfrac{1}{4}u_1(t)\left(1-2t+e^{-2(t-1)}\right)$

5. $y(t) = 1 + t + \dfrac{1}{2}t^2$

6. $y(t) = \dfrac{11}{16}e^t + \dfrac{1}{4}t\,e^t + \dfrac{5}{16}e^{-3t}$

7. $y(t) = \dfrac{1}{2}t\sin t - \sin t$

8. $y(t) = t \sin t - t^2 \cos t$

9. $y(t) = -5 + 3e^{2t} + 2e^{-t} \cos t - 4e^{-t} \sin t$

10. $y(t) = \frac{77}{265} e^{4t} + \frac{1}{5} e^{-t} + \frac{27}{53} e^{-3t} \cos 2t + \frac{111}{106} e^{-3t} \sin 2t$

11. $y(t) = te^{2t} + \frac{1}{4} u_3(t) \left[1 - e^{2(t-3)} + 2(t-3)\, e^{2(t-3)}\right]$

12. $y(t) = \frac{3}{2} u_2(t) \left[1 - 2e^{t-2} + e^{2(t-2)}\right] - \frac{3}{2} u_5(t) \left[1 - 2e^{t-5} + e^{2(t-5)}\right]$

13. $y(t) = 2e^{-t} + 3t\, e^{-t} + u_3(t) \left[-4 + 2t + 4e^{-(t-3)} + 2(t-3)\, e^{-(t-3)}\right]$

14. $y(t) = 3 \cos t - \sin t + \frac{1}{2} t \sin t - \frac{1}{2} u_{\pi/2}(t) \left[\cos t + \left(t - \frac{\pi}{2}\right) \sin t\right]$

15. $y(t) = \frac{1}{6} (1 - u_{2\pi}(t))\, (2 \sin t - \sin 2t)$

16. $y(t) = \frac{1}{4} e^{t} + \frac{1}{4} t\, e^{t} + \frac{5}{48} e^{-t} - \frac{2}{3} e^{2t} + \frac{5}{16} e^{3t}$

17. $y(t) = 3e^{t} - 4t\, e^{t} - 3 \cos t + \sin t$

18. $y(t) = e^{-3t} \left(\sqrt{2}\, t \cosh \sqrt{2}\, t - \sinh \sqrt{2}\, t\right)$

19. $y(t) = f(t) - u_{\pi}(t) f(t - \pi)$, where $f(t) = \frac{1}{4} - \frac{1}{3} \cos t + \frac{1}{12} \cos 2t$

20. $y(t) = 5e^{-t}$

21. $x(t) = \frac{7}{8} e^{-2t} + \frac{9}{8} e^{6t}$, $\quad y(t) = -\frac{7}{8} e^{-2t} + \frac{15}{8} e^{6t}$

22. $x(t) = \frac{1}{4} e^{4t} - \frac{1}{4} \cos 2t$, $\quad y(t) = -\frac{1}{5} e^{-4t} + \frac{1}{5} \cos 2t + \frac{1}{10} \sin 2t$

23. $x(t) = -\frac{2}{9} + \frac{1}{3} t + \frac{2}{9} e^{3t}$, $\quad y(t) = -\frac{4}{9} - \frac{1}{3} t + e^{t} - \frac{5}{9} e^{3t}$

24. $x(t) = 2 + \frac{t^2}{2} + \frac{1}{2} e^{-t} + \frac{1}{2} \cos t - \frac{3}{2} \sin t$, $y(t) = 1 - \frac{1}{2} e^{-t} - \frac{1}{2} \cos t + \frac{3}{2} \sin t$

25. $x(t) = 1 + e^{-t} - 2e^{-t} \cosh \frac{t}{\sqrt{2}}$, $y(t) = 1 + e^{-t} - 2e^{-t} \cosh \frac{t}{\sqrt{2}} - \sqrt{2}\, e^{-t} \sinh \frac{t}{\sqrt{2}}$

INDEX

A

B

C

D

E

F

G

H

I

J K

L

Kashif Ali

T

U V

W Z