PEARSON

Engineering Mathematics I

For UPTU

Semester I

Babu Ram

Engineering Mathematics-I

B.Tech. Semester-I, U.P. Technical University, Lucknow

BABU RAM
Formerly Dean, Faculty of Physical Sciences,
Maharshi Dayanand University, Rohtak

PEARSON
Delhi • Chennai • Chandigarh

Associate Acquisitions Editor: Anita Yadav
Associate Production Editor: Jennifer Sargunar
Composition: White Lotus Infotech Pvt. Ltd, Pondicherry
Printer:

ISBN: 978-81-317-3335-6

10 9 8 7 6 5 4 3 2 1

Published by Dorling Kindersley (India) Pvt. Ltd, licensees of Pearson Education in South Asia.

Head Office: 7th Floor, Knowledge Boulevard, A-8 (A), Sector 62, Noida 201309, UP, India.
Registered Office: 11 Community Centre, Panchsheel Park, New Delhi 110 017, India.

In memory of

my parents

Smt. Manohari Devi and Sri Makhan Lal

Contents

Preface

All branches of Engineering, Technology and Science require mathematics as a tool for the description of their contents. Therefore, a thorough knowledge of various topics in mathematics is essential to pursue courses in Engineering, Technology and Science. The aim of this book is to provide students with sound mathematics skills and their applications. Although the book is designed primarily for use by engineering students, it is also suitable for students pursuing bachelor degrees with mathematics as one of the subjects and also for those who prepare for various competitive examinations. The material has been arranged to ensure the suitability of the book for class use and for individual self study. Accordingly, the contents of the book have been divided into seven chapters covering the complete syllabus prescribed for B.Tech. Semester-I of Uttar Pradesh Technical University, Lucknow. A number of examples, figures, tables and exercises have been provided to enable students to develop problem-solving skills. The language used is simple and lucid. Suggestions and feedback on this book are welcome.

Acknowledgements

I am extremely grateful to the reviewers for their valuable comments. My family members provided moral support during the preparation of this book. My son, Aman Kumar, software engineer, Adobe India Ltd, offered wise comments on some of the contents of the book. I am thankful to Sushma S. Pradeep for excellently typing the manuscript. Special thanks are due to Thomas Mathew Rajesh, Anita Yadav, M. E. Sethurajan and Jennifer Sargunar at Pearson Education for their constructive support.

Babu Ram

Symbols and Basic Formulae

1 Greek Letters

α	alpha	ϕ	phi
β	beta	Φ	capital phi
γ	gamma	ψ	psi
Γ	capital gamma	Ψ	capital psi
δ	delta	ξ	xi
Δ	capital delta	η	eta
ε	epsilon	ζ	zeta
ι	iota	χ	chi
θ	theta	π	pi
λ	lambda	σ	sigma
μ	mu	Σ	capital sigma
ν	nu	τ	tau
ω	omega	ρ	rho
Ω	capital omega	κ	kapha

2 Algebraic Formulae

(i) Arithmetic progression $a, a+d, a+2d, \ldots,$
nth term $T_n = a + (n-1)d$
Sum of n terms $= \frac{n}{2}[2a + (n-1)d]$

(ii) Geometrical progression: $a, ar, ar^2, \ldots,$
nth term $T_n = ar^{n-1}$
Sum of n terms $= \frac{a(1-r^n)}{1-r}$

(iii) Arithmetic mean of two numbers a and b is $\frac{1}{2}(a+b)$

(iv) Geometric mean of two numbers a and b is $\sqrt{ab}$

(v) Harmonic mean of two numbers a and b is $\frac{2ab}{a+b}$

(vi) If $ax^2 + bx + c = 0$ is quadratic, then

- (a) its roots are given by $\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$
- (b) the sum of the roots is equal to $-\frac{b}{a}$
- (c) product of the roots is equal to $\frac{c}{a}$
- (d) $b^2 - 4ac = 0 \Rightarrow$ the roots are equal
- (e) $b^2 - 4ac > 0 \Rightarrow$ the roots are real and distinct
- (f) $b^2 - 4ac < 0 \Rightarrow$ the roots are complex
- (g) if $b^2 - 4ac$ is a perfect square, the roots are rational

3 Properties of Logarithm

(i) $\log_a 1 = 0,\ \log_a 0 = -\infty$ for $a > 1$, $\log_a a = 1$
$\log_e 2 = 0.6931, \log_e 10 = 2.3026, \log_{10} e = 0.4343$

(ii) $\log_a p + \log_a q = \log_a pq$

(iii) $\log_a p - \log_a q = \log_a \frac{p}{q}$

(iv) $\log_a p^q = q \log_a p$

(v) $\log_a n = \log_a b \,.\, \log_b n = \frac{\log_b n}{\log_b a}$

4 Angles Relations

(i) 1 radian $= \frac{180°}{\pi}$

(ii) $1° = 0.0174$ radian

5 Algebraic Signs of Trigonometrical Ratios

(a) First quadrant: All trig. ratios are positive

(b) Second quadrant: sin θ and cosec θ are positive, all others negative

(c) Third quadrant: tan θ and cot θ are positive, all others negative

(d) Fourth quadrant: cos θ and sec θ are positive, all others negative

6 Commonly Used Values of Triganometrical Ratios

$$\sin\frac{\pi}{2} = 1,\ \cos\frac{\pi}{2} = 0,\ \tan\frac{\pi}{2} = \infty$$
$$\text{cosec}\frac{\pi}{2} = 1,\ \sec\frac{\pi}{2} = \infty,\ \cos\frac{\pi}{2} = 0$$
$$\sin\frac{\pi}{6} = \frac{1}{2},\ \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2},\ \tan\frac{\pi}{6} = \frac{1}{\sqrt{3}}$$
$$\text{cosec}\frac{\pi}{6} = 2,\ \sec\frac{\pi}{6} = \frac{2}{\sqrt{3}},\ \cot\frac{\pi}{6} = \sqrt{3}$$
$$\sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}, \cos\frac{\pi}{3} = \frac{1}{2},\ \tan\frac{\pi}{3} = \sqrt{3}$$
$$\text{cosec}\frac{\pi}{3} = \frac{2}{\sqrt{3}},\ \sec\frac{\pi}{3} = 2,\ \cot\frac{\pi}{3} = \frac{1}{\sqrt{3}}$$
$$\sin\frac{\pi}{4} = \frac{1}{\sqrt{2}},\ \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}},\ \tan\frac{\pi}{4} = 1$$
$$\text{cosec}\frac{\pi}{4} = \frac{1}{\sqrt{2}},\ \sec\frac{\pi}{4} = \sqrt{2},\ \cot\frac{\pi}{4} = 1$$

7 Trig. Ratios of Allied Angles

(a) $\sin(-\theta) = -\sin\theta,\ \cos(-\theta) = \cos\theta$
$\tan(-\theta) = -\tan\theta$
$\text{cosec}(-\theta) = -\text{cosec}\,\theta,\ \sec(-\theta) = \sec\theta$
$\cot(-\theta) = -\cot\theta$

(b) Any trig. ratio of
$(n.90 \pm \theta) =$
$\begin{cases} \pm \text{ same trig. ratio of } \theta \text{ when } n \text{ is even} \\ \pm \text{ co-ratio of } \theta \text{ when } n \text{ is odd} \end{cases}$

For example: $\sin(4620) = \sin[90°(52) - 60°] = \sin(-60°) = -\sin 60° = -\frac{\sqrt{3}}{2}$.

Similarly, $\text{cosec}(270° - \theta) = \text{cosec}(90°(3) - \theta) = -\sec\theta$.

8 Transformations of Products and Sums

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

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

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

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

(e) $\tan(A+B) = \frac{\tan A + \tan B}{1 - \tan A\tan B}$

(f) $\tan(A-B) = \frac{\tan A - \tan B}{1 + \tan A\tan B}$

(g) $\sin 2A = 2\sin A\cos A = \frac{2\tan A}{1+\tan^2 A}$

(h) $\cos 2A = \cos^2 A - \sin^2 A = 1 - 2\sin^2 A = 2\cos^2 A - 1 = \frac{1-\tan^2 A}{1+\tan^2 A}$

(i) $\tan 2A = \frac{\sin 2A}{\cos 2A} = \frac{2\tan A}{1-\tan^2 A}$

(j) $\sin 3A = 3\sin A - 4\sin^3 A$

(k) $\cos 3A = 4\cos^3 A - 3\cos A$

(l) $\tan 3A = \frac{3\tan A - \tan^3 A}{1 - 3\tan^2 A}$

(m) $\sin A + \sin B = 2\sin\frac{A+B}{2}\cos\frac{A-B}{2}$

(n) $\sin A - \sin B = 2\cos\frac{A+B}{2}\sin\frac{A-B}{2}$

(o) $\cos A + \cos B = 2\cos\frac{A+B}{2}\cos\frac{A-B}{2}$

(p) $\cos A - \cos B = 2\sin\frac{A+B}{2}\sin\frac{B-A}{2}$

(q) $\sin A\cos B = \frac{1}{2}[\sin(A+B) + \sin(A-B)]$

(r) $\cos A\sin B = \frac{1}{2}[\sin(A+B) - \sin(A-B)]$

(s) $\cos A\cos B = \frac{1}{2}[\cos(A+B) + \cos(A-B)]$

(t) $\sin A\sin B = \frac{1}{2}[\cos(A-B) - \cos(A+B)]$

9 Expressions for $\sin\frac{A}{2}$; $\cos\frac{A}{2}$ and $\tan\frac{A}{2}$

(a) $\sin\frac{A}{2} = \pm\sqrt{\frac{1-\cos A}{2}}$

(b) $\cos\frac{A}{2} = \pm\sqrt{\frac{1+\cos A}{2}}$

(c) $\tan\frac{A}{2} = \pm\sqrt{\frac{1-\cos A}{1+\cos A}}$

(d) $\sin\frac{A}{2} + \cos\frac{A}{2} = \pm\sqrt{1+\sin A}$

(e) $\sin\frac{A}{2} - \cos\frac{A}{2} = \pm\sqrt{1-\sin A}$

10 Relations Between Sides and Angles of a Triangle

(a) $\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}$ (sine formulae)

(b) $\left.\begin{aligned}\cos A &= \frac{b^2+c^2-a^2}{2bc}\\ \cos B &= \frac{c^2+a^2-b^2}{2ca}\\ \cos C &= \frac{a^2+b^2-c^2}{2ab}\end{aligned}\right\}$ cosine formulae

(c) $\left.\begin{aligned}a &= b\cos C + c\cos B\\ b &= c\cos A + a\cos C\\ c &= a\cos B + b\cos A\end{aligned}\right\}$ Projection formulae.

11 Permutations and Combinations Formulae

$${}^nP_r = \frac{n!}{(n-r)!},$$

$${}^nC_r = \frac{n!}{r!\,(n-r)!} = {}^nC_{n-r},$$

$${}^nC_0 = {}^nC_n = 1$$

12 Differentiation Formulae

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

(c) $\frac{d}{dx}(\tan x) = \sec^2 x$ (d) $\frac{d}{dx}(\cot x) = -\text{cosec}^2 x$

(e) $\frac{d}{dx}(\sec x) = \sec x\tan x$ (f) $\frac{d}{dx}(\text{cosec}\, x) = -\text{cosec}\, x\cot x$

(g) $\frac{d}{dx}(e^x) = e^x$ (h) $\frac{d}{dx}(a^x) = a^2\log_e a$

(i) $\frac{d}{dx}(\log_a x) = \frac{1}{x\log a}$ (j) $\frac{d}{dx}(\log_e x) = \frac{1}{x}$

(k) $\frac{d}{dx}(ax+b)^n = na(ax+b)^{n-1}$

(l) $\frac{d^n}{dx^n}(ax+b)^m = m(m-1)(m-2)\ldots(m-n+1)(ax+b)^{m-n}$

(m) $\frac{d}{dx}(\sin^{-1}x) = \frac{1}{\sqrt{1-x^2}}$ (n) $\frac{d}{dx}(\cos^{-1}x) = -\frac{1}{\sqrt{1-x^2}}$

(o) $\frac{d}{dx}(\tan^{-1}x) = \frac{1}{1+x^2}$ (p) $\frac{d}{dx}(\cot^{-1}x) = -\frac{1}{1+x^2}$

(q) $\frac{d}{dx}(\sec^{-1}x) = \frac{1}{x\sqrt{x^2-1}}$ (r) $\frac{d}{dx}(\text{cosec}^{-1}x) = -\frac{1}{x\sqrt{x^2-1}}$

(s) $\frac{d}{dx}(\sinh x) = \cosh x$ (t) $\frac{d}{dx}(\cosh x) = \sinh x$

(u) $D^n(uv) = D^n u + nc_1 D^{n-1}uDv + nc_2 D^{n-2}uD^2 v + \cdots + {}^nC_r D^{n-r}uD^r v + \cdots + {}^nC_n uD^n v$

(Leibnitz's Formula)

13 Integration Formulae

(a) $\int \sin x\,dx = -\cos x$ (b) $\int \cos x\,dx = \sin x$

(c) $\int \tan x\,dx = -\log\cos x$ (d) $\int \cot x\,dx = \log\sin x$

(e) $\int \sec x\,dx = \log(\sec x + \tan x)$

(f) $\int \text{cosec}\, x\,dx = \log(\text{cosec}\, x - \cot x)$

(g) $\int \sec^2 x\,dx = \tan x$ (h) $\int \text{cosec}^2 x\,dx = -\cot x$

(i) $\int e^x dx = e^x$ (j) $\int a^x dx = \frac{a^x}{\log_e a}$

(k) $\int \frac{1}{x}dx = \log_e x$ (l) $\int x^n dx = \frac{x^{n+1}}{n+1}$, $n \neq -1$

(m) $\int \frac{dx}{a^2+x^2} = \frac{1}{a}\tan^{-1}\frac{x}{a}$ (n) $\int \frac{dx}{a^2-x^2} = \frac{1}{2a}\log_e\frac{a+x}{a-x}$

(o) $\int \frac{dx}{x^2-a^2} = \frac{1}{2a}\log_e\frac{x-a}{x+a}$ (p) $\frac{dx}{\sqrt{a^2-x^2}} = \sin^{-1}\frac{x}{a}$

(q) $\int \frac{dx}{\sqrt{a^2+x^2}} = \sinh^{-1}\frac{x}{a}$ (r) $\int \frac{dx}{\sqrt{x^2-a^2}} = \cosh^{-1}\frac{x}{a}$

(s) $\int \sqrt{a^2+x^2}\,dx = \dfrac{x\sqrt{a^2+x^2}}{2} + \dfrac{a^2}{2}\sinh^{-1}\dfrac{x}{a}$

(t) $\int \sqrt{x^2-a^2}\,dx = \dfrac{x\sqrt{x^2-a^2}}{2} - \dfrac{a^2}{2}\cosh^{-1}\dfrac{x}{a}$

(u) $\int \sqrt{a^2-x^2}\,dx = \dfrac{x\sqrt{a^2-x^2}}{2} + \dfrac{a^2}{2}\sin^{-1}\dfrac{x}{a}$

(v) $\int e^{ax}\sin bx\,dx = \dfrac{e^{ax}}{a^2+b^2}(a\sin bx - b\cos bx)$

(w) $\int e^{ax}\cos bx\,dx = \dfrac{e^{ax}}{a^2+b^2}(a\cos bx + b\sin bx)$

(x) $\int_0^{\frac{\pi}{2}} \sin^n x\,dx = \int_0^{\frac{\pi}{2}} \cos^n x\,dx$

$$= \begin{cases} \dfrac{(n-1)\,(n-3)\,(n-5)\ldots}{n(n-2)\,(n-4)\ldots}\cdots \text{ if } n \text{ is odd} \\ \dfrac{(n-1)\,(n-3)\,(n-5)\ldots}{n(n-2)\,(n-4)\ldots}\,\dfrac{\pi}{2} \text{ if } n \text{ is even} \end{cases}$$

(y) $\int_0^{\frac{\pi}{2}} \sin^m x\cos^n x\,dx$

$$= \begin{cases} \dfrac{(m-1)(m-3)\ldots(n-1)(n-3)\ldots}{(m+n)(m+n-2)(m+n-4)\ldots} & \text{if } m \text{ and } n \text{ are not simultaneously even} \\ \dfrac{(m-1)(m-3)\ldots(n-1)(n-3)\ldots}{(m+n)(m+n-2)(m+n-4)\ldots}\dfrac{\pi}{2} & \text{if both } m \text{ and } n \text{ are even} \end{cases}$$

14 Beta and Gamma Functions

(a) $\beta(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx$ converges for $m, n > 0$

(b) $\Gamma\,(n) = \int_0^\infty e^{-x}x^{n-1}dx$ converges for $n > 0$

(c) $\Gamma\,(n+1) = n\,\Gamma(n)$ and $\Gamma\,(n+1) = n\,!$ if n is positive integer

(d) $\Gamma\,(1) = 1 = \Gamma(2)$ and $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

(e) $\beta(m,n) = \dfrac{\Gamma\,(m)\Gamma\,(n)}{\Gamma\,(m+n)}$

(f) $\displaystyle\int_0^{\frac{\pi}{2}} \sin^p x\cos^q x\,dx = \frac{1}{2}\beta\left(\frac{p+1}{2}, \frac{q+1}{2}\right) = \frac{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{q+1}{2}\right)}{2\,\Gamma\left(\frac{p+q+2}{2}\right)}$

(g) $\displaystyle\int_0^{\frac{\pi}{2}} \sqrt{\tan\theta}\,d\theta = \int_0^{\frac{\pi}{2}} \sin^{1/2}\theta\cos^{-1/2}\theta\,d\theta = \frac{\Gamma\left(\frac{3}{4}\right)\Gamma\left(\frac{1}{4}\right)}{2\,\Gamma\,(1)}$

$= \frac{1}{2}\,\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right)$

UNIT I

Differential Calculus

1 Successive Differentiation and Leibnitz`s Theorem

2 Asymptotes and Curve Tracing

3 Partial Differentiation

1 Successive Differentiation and Leibnitz`s Theorem

1.1 SUCCESSIVE DIFFERENTIATION

Let f be a real-valued function defined on an interval $[a, b]$. Then, it is said to be derivable at an interior point c if $\lim_{x\to c}\frac{f(x)-f(c)}{x-c}$ exists. This limit, if exists, is called the *derivative* or the *differential coefficient* of the function at $x = c$ and is denoted by $f'(c)$.

The above limit exists, if both the following limits exist and are equal:

(i) $\lim_{x\to c-0}\frac{f(x)-f(c)}{x-c}$, called the *left-hand derivative* and denoted by $f'(c-0)$,

(ii) $\lim_{x\to c+0}\frac{f(x)-f(c)}{x-c}$, called the *right-hand derivative* and denoted by $f'(c+0)$.

The derivative $f'(c)$ exists when $f'(c-0) = f'(c+0)$. If $y = f(x)$ be a derivable (differentiable) function of x, then $f'(x) = \frac{dy}{dx}$ is called the *first differential coefficient* of y with respect to x. If $f'(x)$ is again derivable, then $\frac{d}{dx}\left(\frac{dy}{dx}\right)$ is called the *second differential coefficient* of y with respect to x and denoted by $f''(x)$ *or* $\frac{d^2y}{dx^2}$. If $f''(x)$ is further derivable, then $\frac{d}{dx}\left(\frac{d^2y}{dx^2}\right)$, denoted by $\frac{d^3y}{dx^3}$ is called the *third differential coefficient* of y with respect to x. In general, the nth differential coefficient of y with respect to x is $\frac{d^ny}{dx^n} = \frac{d}{dx}\left(\frac{d^{n-1}y}{dx^{n-1}}\right)$. The differential operator $\frac{d}{dx}$ is also denoted by D. As such, the nth derivative of y is denoted by D^ny. For the sake of convenience, the derivatives $Dy, D^2y, \ldots, D^ny$ are generally represented by $y_1, y_2, \ldots, y_n$.

EXAMPLE 1.1

Find the nth derivative of

(i) $y = e^{ax+b}$, (ii) $y = (ax+b)^m$

(iii) $y = \frac{1}{ax+b}$, (iv) $y = a^x$

(v) $y = \sin(ax+b)$, and $y = \cos(ax+b)$

(vi) $y = e^{ax}\sin(bx+c)$ and $y = e^{ax}\cos(bx+c)$

(vii) $y = \log(ax+b)$.

Solution. (i) We are given that $y = e^{ax+b}$. Therefore

$$y_1 = \frac{dy}{dx} = a\,e^{ax+b},$$
$$y_2 = \frac{d^2y}{dx^2} = a^2e^{ax+b},$$
$$y_3 = a^3e^{ax+b}$$

and so on. Therefore the nth derivative D^ny is given by

$$y_n = a^ne^{ax+b}.$$

(ii) We have $y = (ax+b)^m$. Therefore

$$y_1 = ma(ax+b)^{m-1},$$
$$y_2 = m(m-1)a^2(ax+b)^{m-2},$$
$$y_3 = m(m-1)(m-2)a^3(ax+b)^{m-3},$$

and so on. Hence, in general,

$$y_n = m(m-1)(m-2)\ldots(m-n+1)a^n(ax+b)^{m-n}.$$

Further, if m is a positive integer, then

$$y_n = \frac{m(m-1)(m-2)\ldots(m-n+1)\,(m-n)(m-n-1)\ldots 2.1}{(m-n)\ (m-n-1)\ \ldots\ 1} \times a^n(ax+b)^{m-n}$$
$$= \frac{m\,!}{(m-n)!}a^n(ax+b)^{m-n} \qquad (1)$$

From (1), it follows that the mth derivative of the given function is

$$y_m = \frac{m\,!}{o\,!} a^m (ax+b)^o = m\,!\, a^m.$$

In particular, taking $a = 1$, $b = 0$, we get

$$y_m = m\,!.$$

Thus, if m is a positive integer, the mth differential coefficient of $(ax+b)^m$ is constant.
In case m is negative, then $m = -p$, where p is a positive integer and so

$$\begin{aligned} y_n &= (-p)(-p-1)(-p-2)\ldots(-p-n+1)a^n \\ &\quad \times (ax+b)^{-p-n} \\ &= (-1)^n p(p+1)(p+2)\ldots.(p+n-1)a^n \\ &\quad \times (ax+b)^{-p-n} \\ &= (-1)^n \frac{(p+n-1)\,!}{(p-1)\,!} a^n (ax+b)^{-p-n}. \end{aligned}$$

(iii) We have $y = (ax+b)^{-1}$ and so

$$\begin{aligned} y_1 &= (-1)a(ax+b)^{-2}, \\ y_2 &= (-1)(-2)a^2(ax+b)^{-3}, \\ y_3 &= (-1)(-2)(-3)a^3(ax+b)^{-4} \end{aligned}$$

and so on. Hence, in general,

$$\begin{aligned} y_n &= (-1)(-2)(-3)\ldots(-n)a^n(ax+b)^{-(n+1)} \\ &= (-1)^n n\,!\, a^n (ax+b)^{-(n+1)}. \end{aligned}$$

(iv) If is given that $y = a^x$. Therefore

$$y_1 = a^x \log a,$$

$$y_2 = a^x (\log a)^2,$$

$$y_3 = a^2 (\log a)^2,$$

and so on. Hence, in general,

$$y_n = a^x (\log a)^n.$$

(v) We have $y = \sin(ax+b)$. Therefore

$$\begin{aligned} y_1 &= a\cos(ax+b) = a\sin\left(ax+b+\frac{\pi}{2}\right), \\ y_2 &= a^2\cos\left(ax+b+\frac{\pi}{2}\right) = a^2\sin(ax+b+\pi), \\ y_3 &= a^3\cos(ax+b+\pi) = a^3\sin\left(ax+b+\frac{3\pi}{2}\right), \end{aligned}$$

and so on. Hence, in general,

$$y_n = a^n \sin\left(ax+b+\frac{n\pi}{2}\right).$$

In a similar fashion, we can show that if $y = \cos(ax+b)$, then

$$y_n = a^n \cos\left(ax+b+\frac{n\pi}{2}\right).$$

(vi) The given function is $y = e^{ax}\sin(bx+c)$. Therefore

$$\begin{aligned} y_1 &= ae^{ax}\sin(bx+c) + b\,e^{ax}\cos(bx+c) \\ &= e^{ax}[a\ \sin(bx+c) + b\ \cos(bx+c)]. \quad (2) \end{aligned}$$

Let us choose $a = r\cos\phi$ and $b = r\sin\phi$. Then

$$r = \sqrt{a^2+b^2} \quad \text{and} \quad \phi = \tan^{-1}\left(\frac{b}{a}\right),$$

and (2) reduces to

$$y_1 = re^{ax}\sin(bx+c+\phi).$$

Similarly, repeating the above argument, we have

$$\begin{aligned} y_2 &= r^2 e^{ax}\sin(bx+c+2\phi), \\ y_3 &= r^3 e^{ax}\sin(bx+c+3\phi), \end{aligned}$$

and so on. Hence, in general,

$$y_n = r^n e^{ax} \sin(bx + c + n\,\phi),$$

where

$$r = \sqrt{a^2 + b^2} \quad \text{and} \quad \phi = \tan^{-1}\left(\frac{b}{a}\right).$$

In a similar fashion, we can show that if $y = e^{ax}\cos(bx + c)$, then

$$y_n = r^n e^{ax} \cos(bx + c + n\phi).$$

(vii) When $y = \log(ax + b)$, then

$$y_1 = \frac{a}{ax + b},$$
$$y_2 = \frac{dy_1}{dx} = (-1)a^2(ax + b)^{-2},$$
$$y_3 = (-1)(-2)a^2(ax + b)^{-3},$$

and so on. Therefore, in general,

$$y_n = (-1)(-2)\ldots(-(n-1))a^n(ax + b)^{-n}$$
$$= \frac{(-1)^{n-1}(n-1)!\,a^n}{(ax + b)^n}.$$

EXAMPLE 1.2

Find the nth derivative of

(i) $y = e^{2x}\sin^3 x$, (ii) $\cos^2 x \sin^3 x$
(iii) $y = \cos x \cos 2x \cos 3x$,
(iv) $y = e^{ax}\sin bx \cos cx$

Solution. (i) We have

$$y = e^{2x}\sin^3 x = \frac{1}{4}e^{2x}[3\sin x - \sin 3x].$$
$$= \frac{3}{4}e^{2x}\sin x - \frac{1}{4}e^{2x}\sin 3x.$$

Therefore, using Example 1.1 (vi), we have

$$y_n = \frac{3}{4}\left[\sqrt{2^2 + 1^2}\right]^n e^{2x}\sin\left[x + n\tan^{-1}\frac{1}{2}\right]$$
$$-\frac{1}{4}\left[\sqrt{2^2 + 3^2}\right]^n e^{2x}\sin\left[2x + n\tan^{-1}\frac{3}{2}\right].$$

(ii) We have

$$y = \cos^2 x \sin^3 x$$
$$= \frac{1}{2}(1 + \cos 2x)\left[\frac{1}{4}(3\sin x - \sin 3x)\right]$$
$$= \frac{1}{8}\left[3\sin x - \sin 3x + \frac{3}{2}(2\sin x\cos 2x)\right.$$
$$\left. -\frac{1}{2}(2\sin 3x\cos 2x)\right]$$
$$= \frac{1}{8}\left[3\sin x - \sin 3x + \frac{3}{2}(\sin 3x\right.$$
$$\left. -\sin x) - \frac{1}{2}(\sin 5x + \sin x)\right]$$
$$= \frac{1}{16}[2\sin x + \sin 3x - \sin 5x].$$

Therefore, using Example 1.1 (v), we have

$$y_n = \frac{1}{16}\left[2\sin\left(x + \frac{n\pi}{2}\right) + 3^n\sin\left(3x + \frac{n\pi}{2}\right)\right.$$
$$\left.-5^n\sin\left(5x + \frac{n\pi}{2}\right)\right].$$

(iii) We are given that

$$y = \cos x\cos 2x\cos 3x.$$
$$= \cos x\left[\frac{1}{2}(\cos 5x + \cos x)\right]$$
$$= \frac{1}{4}(2\cos x\cos 5x + 2\cos^2 x)$$
$$= \frac{1}{4}[(\cos 6x + \cos 4x) + (1 + \cos 2x)].$$

Therefore, using Example 1.1 (v), we have

$$y_n = \frac{1}{4}D^n(1) + \frac{1}{4}D^n[\cos 2x + \cos 4x + \cos 6x]$$
$$= 0 + \frac{1}{4}\left[2^n\cos\left(2x + \frac{n\pi}{2}\right) + 4^n\cos\left(4x + \frac{n\pi}{2}\right)\right.$$
$$\left.+6^n\cos\left(6x + \frac{n\pi}{2}\right)\right]$$
$$= \frac{1}{4}\left[2^n\cos\left(2x + \frac{n\pi}{2}\right) + 4^n\cos\left(4x + \frac{n\pi}{2}\right)\right.$$
$$\left.+6^n\cos\left(6x + \frac{n\pi}{2}\right)\right].$$

(iv) We have

$$\begin{aligned} y &= e^{2x}\sin bx\cos cx \\ &= e^{ax}\left[\frac{1}{2}(2\sin bx\,\cos cx)\right] \\ &= \frac{1}{2}e^{ax}[\sin(bx+cx)+\sin(bx-cx)] \\ &= \frac{1}{2}e^{ax}\sin(b+c)x+\frac{1}{2}e^{ax}\sin(b-c)x. \end{aligned}$$

Therefore, using Example 1.1 (vi), we have

$$\begin{aligned} y_n &= \frac{1}{2}\left(\sqrt{a^2+(b+c)^2}\right)^n e^{ax} \\ &\quad\times \sin\left\{(b+c)x+n\tan^{-1}\frac{b+c}{a}\right\} \\ &\quad+\frac{1}{2}\left(\sqrt{a^2+(b-c)^2}\right)^n e^{ax} \\ &\quad\times \sin\left\{(b-c)x+n\tan^{-1}\frac{b-c}{a}\right\}. \end{aligned}$$

EXAMPLE 1.3

Find second derivatives of the following functions with respect to x:

(i) $x = a(t+\sin t), \qquad y = a(1+\cos t)$

(ii) $x = a\cos^3\theta, \qquad y = b\sin^3\theta.$

Solution: (i) For the given equations, we have

$$\begin{aligned} \frac{dx}{dt} &= a(1+\cos t) = 2a\cos^2\frac{t}{2}, \\ \frac{dy}{dt} &= -a\sin t = -2a\sin\frac{t}{2}\cos\frac{t}{2}. \end{aligned}$$

Therefore

$$\frac{dy}{dx} = \frac{-2a\sin\frac{t}{2}\cos\frac{t}{2}}{2a\cos^2\frac{t}{2}} = -\tan\frac{t}{2}$$

and so

$$\begin{aligned} \frac{d^2y}{dx^2} &= -\frac{1}{2}\sec^2\frac{t}{2}\frac{dt}{dx} \\ &= -\frac{1}{2}\sec^2\frac{t}{2}\left[\frac{1}{2a\cos^2\frac{t}{2}}\right] = -\frac{1}{4a}\sec^4\frac{t}{2}. \end{aligned}$$

(iii) We are given that

$$x = a\cos^3\theta, \qquad y = b\sin^3\theta.$$

Therefore

$$\begin{aligned} \frac{dx}{d\theta} &= -3a\sin\theta\cos^2\theta, \\ \frac{dy}{d\theta} &= 3b\sin^2\theta\cos\theta, \end{aligned}$$

and so

$$\frac{dy}{dx} = \frac{\frac{dy}{d\theta}}{\frac{dx}{d\theta}} = -\frac{b}{a}\tan\theta.$$

Differentiating once more with respect to x, we get

$$\begin{aligned} \frac{d^2y}{dx^2} &= -\frac{b}{a}\sec^2\theta\cdot\frac{d\theta}{dx} \\ &= -\frac{b}{a}\sec^2\theta\left[\frac{1}{-3a\sin\theta\cos^2\theta}\right] \\ &= \frac{b}{3a^2}\operatorname{cosec}\theta\sec^4\theta. \end{aligned}$$

EXAMPLE 1.4

Find the nth derivative of

(i) $\dfrac{x^2-4x+1}{x^3+2x^2-x-2}$ (ii) $\dfrac{1}{x^2-6x+8}$.

Solution: We have

$$\begin{aligned} y &= \frac{x^2-4x+1}{x^3+2x^2-x-2} = \frac{x^2-4x+1}{(x-1)(x+1)(x+2)} \\ &= \frac{13}{3(x+2)}-\frac{3}{x+1}-\frac{1}{3(x-1)} \quad \text{(partial fractions).} \end{aligned}$$

Since $D^n\{(ax+b)^{-1}\} = (-1)^n n!\, a^n(ax+b)^{-n-1}$, we have

$$y_n = (-1)^n n!\left[\frac{13}{3(x+2)^{n+1}}-\frac{3}{(x+1)^{n+1}}-\frac{1}{3(x-1)^{n+1}}\right].$$

(ii) We have

$$y = \frac{1}{x^2 - 6x + 8} = \frac{1}{(x-4)(x-2)}$$
$$= \frac{1}{2(x-4)} - \frac{1}{2(x-2)} \quad \text{(partial fractions).}$$

Therefore, using $D^n\{(ax+b)^{-1}\} = (-1)^n n!\, a^n (ax+b)^{-(n+1)}$, we have

$$y_n = \frac{(-1)^n n!}{2}\left[\frac{1}{(x-4)^{n+1}} - \frac{1}{(x-2)^{n+1}}\right].$$

EXAMPLE 1.5

If $x = \sin t,\ \ y = \sin pt$, show that

$$(1-x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + p^2 y = 0.$$

Solution: We have

$$x = \sin t \quad \text{and} = \sin pt$$

Therefore

$$\frac{dx}{dt} = \cos t, \qquad \frac{dy}{dt} = p\cos pt,$$

and so

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = p\frac{\cos pt}{\cos t}.$$

Differentiating once more with respect to x, we get

$$\frac{d^2y}{dx^2} = p\left[\frac{\cos t(-p\sin pt) - \cos pt(-\sin t)}{\cos^2 t}\right]\cdot\frac{dt}{dx}$$
$$= \frac{p}{\cos^2 t}[\sin t\cos pt - p\sin pt\ \cos t]\cdot\frac{1}{\cos t}.$$

Therefore

$$(1-x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + p^2 y$$
$$= (1-\sin^2 t)\frac{d^2y}{dx^2} - x\frac{p\cos pt}{\cos t} + p^2\sin pt$$
$$= p[\sin t\cos pt - p\sin pt\cos t]\cdot\frac{1}{\cos t}$$
$$-\frac{p\sin t\cos pt}{\cos t} + p^2\sin pt$$
$$= \frac{1}{\cos t}[p\sin t\cos pt - p^2\sin pt\ \cos t$$
$$-p\sin t\ \cos pt - p^2\sin pt\ \cos t] = 0.$$

EXAMPLE 1.6

If $y = a\cos(\log x) + b\sin(\log x)$, show that

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + y = 0.$$

Solution: We have

$$y = a\cos(\log x) + b\sin(\log x).$$

Therefore

$$y_1 = \frac{dy}{dx} = -a\sin(\log x)\frac{1}{x} + b\cos(\log x).\frac{1}{x}$$

or

$$xy_1 = -a\sin(\log x) + b\cos(\log x).$$

Differentiating again with respect to x, we get

$$xy_2 + y_1 = -a\cos(\log x)\frac{1}{x} - b\sin(\log x).\frac{1}{x}$$
$$= -\frac{1}{x}[a\cos(\log x) + b\sin(\log x)] = -\frac{y}{x}.$$

Hence

$$x^2 y_2 + xy_1 + y = 0.$$

EXAMPLE 1.7

If $ax^2 + 2hxy + by^2 = 1$, show that

$$\frac{d^2y}{dx^2} = \frac{h^2 - ab}{(hx+by)^3}.$$

Solution: We have

$$ax^2 + 2hxy + by^2 = 1. \quad (1)$$

Differentiating with respect to x, we get

$$2ax + 2h\left[x\frac{dy}{dx} + y\right] + 2by\frac{dy}{dx} = 0$$

or

$$\frac{dy}{dx} = -\frac{ax+by}{hx+by}.$$

Differentiating again with respect to x, we have

$$\frac{d^2y}{dx^2} = \frac{(h^2-ab)(ax^2+2hxy+by^2)}{(hx+by)^2}$$

$$= \frac{h^2-ab}{(hx+by)^2}, \quad \text{using (1)}.$$

EXAMPLE 1.8

Find the nth derivative of $\tan^{-1}\left(\frac{x}{a}\right)$.

Solution: We have

$$y = \tan^{-1}\left(\frac{x}{a}\right).$$

Therefore

$$y_1 = \frac{a}{x^2+a^2} = \frac{a}{(x+ia)(x-ia)}$$

$$= \frac{1}{2i}\left[\frac{1}{x-ia} - \frac{1}{x+ia}\right]. \quad (1)$$

Since $D^{n-1}\{(x-ia)^{-1}\} = (-1)^{n-1}(n-1)!(x-ia)^{-n}$, differentiating (n−1) times the expression (1) with respect to x, we get

$$y_n = \frac{(-1)^{n-1}(n-1)!}{2i}[(x-ia)^n - (x+ia)^{-n}].$$

Putting $x = r\cos\phi$ and $a = r\sin\phi$, we get

$$y_n = (-1)^{n-1}(n-1)!\, a^{-n}\sin^n\phi\ \sin n\phi,$$

where $\phi = \tan^{-1}\left(\frac{a}{x}\right)$.

1.2 LEIBNITZ'S THEOREM AND ITS APPLICATIONS

This theorem determines the nth differential coefficient of the product of two differentiable functions. The statement of the theorem is as follows:

Theorem 1.1. (Leibnitz's). Let u and v be two functions of x such that nth derivatives $D^n u$ and $D^n v$ exist. Then the nth differential coefficient of their product is given by

$$D^n(uv) = (D^n u)v +^n C_1 D^{n-1}u.Dv +^n C_2 D^{n-2}uD^2v + \ldots +^n C_r D^{n-r}u.D^r v + \ldots +^n C_n uD^n v.$$

Proof: We shall prove the theorem using mathematical induction. For $n = 1$, we have

$$D(uv) = (Du).v + u.Dv.$$

Therefore, the theorem holds for $n = 1$. Let us assume that the theorem holds for $n = m$. Thus,

$$D^m(uv) = (D^m u).v +^m C_1 D^{m-1}u.Dv +^m C_2 D^{m-2}uD^2v + \ldots +^m C_r D^{m-r}u.D^r v + \ldots +^m C_m uD^m v.$$

Differentiating with respect to x, we get

$$\begin{aligned} D^{m+1}(uv) &= [(D^{m+1}u)v + D^m u.Dv] \\ &+^m C_1(D^m u.Dv + D^{m-1}uD^2v) \\ &+^m C_2(D^{m-1}u.D^2v + D^{m-2}uD^3v) + \ldots \\ &+^m C_r(D^{m-r+1}u.D^r v + D^{m-r}u.D^{r+1}v) \\ &+ \ldots + (Du.D^m v + uD^{m+1}v) \\ &= (D^{m+1}u)v + (^mC_0 +^m C_1)D^m u.Dv \\ &+ (^mC_1 +^m C_2)D^{m-1}u.D^2v + \ldots \\ &+ (^mC_r +^m C_{r+1})D^{m-r}u.D^{r+1}v \\ &+ \ldots + uD^{m+1}v \end{aligned}$$

$$= (D^{m+1}u)v + {}^{m+1}C_1 D^m u.Dv$$
$$+ {}^{m+1}C_2 D^{m-1}u.D^2v$$
$$+ \ldots + {}^{m+1}C_{r+1}D^{m-r}u.D^{r+1}v$$
$$+ \ldots + u.D^{m+1}v.$$

Thus, the theorem holds for $n = m + 1$ also. Hence, by mathematical induction, the theorem holds for all positive integral values of n.

EXAMPLE 1.9

Find the nth derivative of $x^{n-1}\log x$.

Solution. We have

$$y = x^{n-1}\log x.$$

Therefore,

$$y_1 = x^{n-1}\left(\frac{1}{x}\right) - (n-1)x^{n-2}\log x$$

or

$$xy_1 = x^{n-1} + (n-1)y.$$

Differentiating both the sides by $(n-1)$ times, we get

$$D^{n-1}(xy_1) = D^{n-1}(x^{n-1}) + D^{n-1}[(n-1)y]$$

or

$$(D^{n-1}y_1)x + {}^{n-1}C_1 D^{n-2}y_1 Dx = (n-1)! + (n-1)y_{n-1}$$

or

$$xy_n + (n-1)y_{n-1} = (n-1)! + (n-1)y_{n-1}$$

or

$$xy_n = (n-1)!,$$

which yields

$$y_n = \frac{(n-1)!}{x}.$$

EXAMPLE 1.10

If $y = a\cos(\log x) + b\sin(\log x)$, show that

$$x^2y_2 + x\,y_1 + y = 0$$

and

$$x_2y_{n+2} + (2n+1)x\,y_{n+1} + (n^2+1)y_n = 0.$$

Solution. We are given that

$$y = a\cos(\log x) + b\sin(\log x).$$

Therefore,

$$y_1 = -\frac{a}{x}\sin(\log x) + \frac{b}{x}\cos(\log x)$$

or

$$xy_1 = -a\sin(\log x) + b\cos(\log x) \qquad (1)$$

Differentiating (1) with respect to x, we get

$$xy_2 + y_1 = -\frac{a}{x}\cos(\log x) - \frac{b}{x}\sin(\log x)$$

or

$$x^2y_2 + xy_1 = -y \qquad (2)$$

Differentiating (2), with respect to x, n times by using Leibnitz's Theorem, we obtain

$$D^n(x^2y_2) + D^n(xy_1) = -D^n y$$

or

$$(D^ny_2)x^2 + {}^nC_1D^{n-1}y_2Dx^2 + {}^nC_2D^{n-2}y_2.D^2x^2$$
$$+ (D^ny_1)x + {}^nC_1D^{n-1}y_1Dx + D^ny = 0$$

or

$$x^2y_{n+2} + 2nxy_{n+1} + \frac{2n(n-1)}{2}y_n + xy_{n+1} + ny_n + y_n = 0$$

or

$$x^2y_{n+2} + (2n+1)xy_{n+1} + (n^2+1)y_n = 0.$$

EXAMPLE 1.11

If $y = e^{a\sin^{-1}x}$, show that $(1-x^2)y_{n+2} - (2n+1)xy_{n+1} - (n^2 + a^2)y_n = 0$.

Solution. We have

$$y = e^{a\sin^{-1}x}. \qquad (1)$$

Therefore,

$$y_1 = e^{a\sin^{-1}x}\left[\frac{a}{\sqrt{1-x^2}}\right]$$

or

$$y_1\sqrt{1-x^2} = ae^{a\sin^{-1}x}$$

or

$$y_1^2(1-x^2) = a^2y^2, \text{ using (1)}.$$

Differentiating with respect to x once more, we get

$$2\,y_1\,y_2(1-x^2) + y_1^2(-2x) = 2\,a^2y\,y_1$$

or

$$y_2(1-x^2) - y_1\,x - a^2\,y = 0. \qquad (2)$$

Using Leibnitz's Theorem, we differentiate (2), n times with respect to x, and get

$$D^n[y_2(1-x^2)] - D^n(y_1x) - a^2D^ny = 0$$

or

$$y_{n+2}(1-x^2) + {}^nC_1y_{n+1}(-2x) + {}^nC_2y_n(-2)$$
$$- y_{n+1}(x) + {}^nC_1y_n - a^2y_n = 0$$

or

$$(1-x^2)y_{n+2}-(2n+1)xy_{n+1}-(n^2+a^2)y_n=0.$$

EXAMPLE 1.12

If $y=e^{\tan^{-1}x}$, show that

$$(1+x^2)\,y_{n+2}+[2(n+1)x-1]\,y_{n+1}+n(n+1)y_n=0.$$

Solution. It is given that

$$y=e^{\tan^{-1}x}.$$

Differentiating with respect to x, we get

$$y_1=e^{\tan^{-1}x}.\ \frac{1}{1+x^2}=\frac{y}{1+x^2}$$

or

$$(1+x^2)y_1=y.$$

Differentiating $(n+1)$ times by using Leibnitz's Theorem, we get

$$D^{n+1}[(1+x^2)y_1]=D^{n+1}y$$

or

$$(1+x^2)y_{n+2}+^{n+1}C_1\,y_{n+1}(2x)+^{n+1}C_2\,y_n.2=y_{n+1}$$

or

$$(1+x^2)y_{n+2}+[2(n+1)x-1]\,y_{n+1}+n(n+1)y_n=0.$$

EXAMPLE 1.13

If $\cos^{-1}\left(\frac{y}{b}\right)=\log\left(\frac{x}{n}\right)^n$, show that

$$x^2\,y_{n+2}+(2n+1)x\,y_{n+1}+2n^2\,y_n=0.$$

Solution. We have

$$\cos^{-1}\left(\frac{y}{b}\right)=\log\left(\frac{x}{n}\right)^n=n\,\log\left(\frac{x}{n}\right)=n[\log x-\log n].$$

Differentiating with respect to x, we get

$$-\frac{1}{\sqrt{1-\frac{y^2}{b^2}}}.\frac{y_1}{b}=\frac{n}{x}$$

or

$$-\frac{y_1}{\sqrt{b^2-y^2}}=\frac{n}{x}$$

or

$$y_1^2\,x^2=n^2(b^2-y^2).$$

Differentiating once more, we get

$$2\,y_1\,y_2\,x^2+2x\,y_1^2=2\,n^2\,y\,y_1$$

or

$$y_2x^2+y_1\,x+n^2\,y=0.$$

Differentiating n times by Leibnitz's Theorem, we get

$$y_{n+2}\,x^2+^nC_1\,y_{n+1}(2x)+^nC_2\,y_n\,2+y_{n+1}x$$
$$+^nC_1\,y_n1+n^2y_n=0$$

or

$$x^2y_{n+2}+(2n+1)x\,y_{n+1}+2n^2\,y_n=0.$$

EXAMPLE 1.14

If $y=\tan^{-1}x$, find $(y_n)_0$.

Solution. Since $y=\tan^{-1}x$, we get $y_1=\frac{1}{1+x^2}$ and so, $(1+x^2)y_1-1=0$

Differentiating once more, we get

$$(1+x^2)y_2+2x\,y_1=0.$$

Now differentiating n times using Leibnitz's Theorem, we obtain

$$y_{n+2}(1+x^2)+ny_{n+1}(2x)+\frac{n(n+1)}{2}y_n.2$$
$$+2x\,y_{n+1}+2ny_n=0$$

or

$$(1+x^2)y_{n+2}+2(n+1)xy_{n+1}+n(n+1)y_n=0.\quad(1)$$

Substituting $x=0$ in the expression for y, y_1, and y_2, we get

$$(y)_0=0,(y_1)_0=1,(y_2)_0=0.$$

Also substituting $x=0$ in (1), we get

$$(y_{n+2})_0=-[n(n+1)](y_n)_0.\qquad(2)$$

Putting $n-4$ and $n-2$ in place of n, respectively, in (2), we get

$$(y_{n-2})_0=-[(n-4)(n-3)](y_{n-4})_0\text{ and}$$
$$(y_n)_0=-[(n-2)(n-1)](y_{n-2})_0$$
$$=-[(n-1)(n-2)][-(n-3)(n-4)](y_{n-4})_0.$$

If n is even, we have

$$(y_n)_0=[-(n-1)(n-2)]$$
$$\times[-(n-3)(n-4)]\cdots[-(3)(2)](y_2)_0$$
$$=0\text{ since }(y_2)_0=0.$$

If n is odd, then

$$(y_n)_0 = [-(n-1)(n-2)] \times [-(n-3)(n-4)]\cdots[-(4)(3)] \times [-(2)(1)](y_1)_0 = (-1)^{\frac{n-1}{2}}(n-1)! \text{ since } (y_1)_0 = 1.$$

EXAMPLE 1.15

If $y = \left[x+\sqrt{1+x^2}\right]^m$, find the value of the nth differential coefficient of y for $x = 0$.

Solution. We have

$$y = \left[x+\sqrt{1+x^2}\right]^m. \quad (1)$$

Therefore,

$$y_1 = m\left[\sqrt{1+x^2}\right]^{m-1}.\left[1+\frac{1}{2}\frac{2x}{\sqrt{1+x^2}}\right]$$
$$= \frac{m}{\sqrt{1+x^2}}\left[x+\sqrt{1+x^2}\right]^m = \frac{my}{\sqrt{1+x^2}}. \quad (2)$$

Squaring, we get

$$y_1^2(1+x^2) - m^2y^2 = 0.$$

Differentiating once again, we get

$$2y_1y_2(1+x^2) + 2xy_1^2 - 2m^2y\, y_1 = 0$$

or

$$y_2(1+x^2) + xy_1 - m^2y = 0. \quad (3)$$

Differentiating (3) n times by Leibnitz's Theorem, we get

$$y_{n+2}(1+x^2) + {}^nC_1y_{n+1}(2x) + {}^nC_2y_n(2) + xy_{n+1} + {}^nC_1y_n - m^2y_n = 0$$

or

$$y_{n+2}(1+x^2)(2n+1)xy_{n+1} + (n^2-m^2)y_n = 0. \quad (4)$$

Substituting $x = 0$ in (1), (2), (3), and (4), we get

$$(y)_0 = 1,$$
$$(y_1)_0 = m,$$
$$(y_2)_0 = m^2, \text{ and}$$
$$(y_{n+2})_0 = (m^2-n^2)(y_n)_0. \quad (5)$$

If n is odd, that is, if $n = 1,3,5,\ldots$, then(5) yields

$$(y_3)_0 = (m^2-1^2)(y_1)_0 = (m^2-1)m,$$
$$(y_5)_0 = (m^2-3^2)(y_3)_0 = (m^2-3^2)(m^2-1^2)m,$$
$$(y_7)_0 = (m^2-5^2)(y_1)_0 = (m^2-5^2)(m^2-3^2) \times (m^2-1^2)m,$$

and so on. Hence, when n is odd, we have

$$(y_n)_0 = [m^2-(n-2)^2] \times [m^2-(n-4)^2]\ldots(m^2-3^2)(m^2-1^2)m$$

If n is even, putting $n = 2,4,6,\ldots$in (5), we get

$$(y_4)_0 = (m^2-2^2)(y_2)_0 = (m^2-2^2)m^2,$$
$$(y_6)_0 = (m^2-4^2)(y_n)_0 = (m^2-4^2)(m^2-2^2)m^2,$$
$$(y_8)_0 = (m^2-6^2)(y_6)_0 = (m^2-6^2)(m^2-4^2) \times (m^2-2^2)m^2,$$

and so on. Hence, when n is even, we get

$$(y_n)_0 = [m^2-(n-2)^2] \times [m^2-(n-4)^2]\ldots(m^2-4^2)(m^2-2^2)m^2.$$

1.3 MISCELLANEOUS EXAMPLES

EXAMPLE 1.16

Find the nth derivatives of (i) $e^x\sin^2 x$. and (ii) $\dfrac{x}{(x-1)(2x+3)}$

Solution. Let $y = e^{-x}\sin^2 x$. Then

$$y = \frac{1}{2}e^{-x}(2\sin^2 x) = \frac{1}{2}e^{-x}(1-\cos 2x) = \frac{1}{2}[e^{-x} - e^{-x}\cos 2x].$$

But $D^ne^{ax} = a^ne^{ax}$ and

$$D^n(e^{ax}\cos(bx+c)) = (a^2+b^2)^{\frac{n}{2}}e^{ax}\cos\left(bx+c-n\tan^{-1}\frac{b}{a}\right).$$

Therefore

$$D^ny = \frac{1}{2}[(-1)^ne^{-x} - (5)^{\frac{1}{2}}e^{-x}\cos(2x+n \times \tan^{-1}(-2))].$$

(ii) Let $y = (ax+b)^{-1}$. Then $y_1 = -a(ax+b)^{-2}$,

$$y_2 = (-1)(-2)a^2(ax+b)^{-3}, \quad \text{and in general}$$
$$y_n = (-1)(-2)\ldots(-n)a^n(ax+b)^{-(n+1)}.$$

Thus

$$D^ny = (-1)^nn!a^n(ax+b)^{-n-1} \quad (1)$$

Now

$$f(x)=\frac{x}{(x-2)(2x+3)}$$
$$=\frac{2}{5(x-1)}+\frac{3}{5(2x+3)}\quad \text{(partial fractions)}.$$

Using (1), we have

$$D^n(x-1)^{-1}=(-1)^n n!(x-1)^{-n-1},$$
$$D^n(2x+3)^{-1}=(-1)^n n!(2x+3)^{-n-1}.$$

Therefore

$$D^n(f)=\frac{1}{5}(-1)^n n![2(x-1)^{-n-1}$$
$$+3\{2^n(2x+3)^{-n-1}\}]$$

EXAMPLE 1.17

Prove that

$$D^n\left[\frac{\log x}{x}\right]=\frac{(-1)^n n!}{x^{n+1}}\left[\log x-1-\frac{1}{2}-\frac{1}{3}-\cdots-\frac{1}{n}\right].$$

Solution. By Leibnitz's Theorem, we have,

$$D^n\left[\frac{\log x}{x}\right]=D^n[x^{-1}\log x]$$
$$=\frac{(-1)^n n!}{x^{n+1}}\log x+{}^nC_1\frac{(-1)^{n-1}(n-1)!}{x^n}\cdot\frac{1}{x}$$
$$+{}^nC_2\frac{(-1)^{n-2}(n-2)!}{x^{n-1}}(-1).\frac{1}{x^2}$$
$$+{}^nC_3\frac{(-1)^{n-3}(n-3)!}{x^{n-2}}(-1)(-2)\cdot\frac{1}{x^3}+\cdots$$
$$+\frac{(-1)^{n-1}(n-1)!}{x^n}\cdot\frac{1}{x}$$
$$=\frac{(-1)^n n!}{x^{n+1}}\left[\log x-1-\frac{1}{2}-\frac{1}{3}-\ldots-\frac{1}{n}\right]$$

EXERCISES

Successive Derivatives

1. If $x^3+y^3=3axy$, show that
$$\frac{d^2y}{dx^2}=-\frac{2a^2xy}{(y^2-ax)^3}.$$

2. If $y=e^{ax}\sin bx$, show that
$$\frac{d^2y}{dx^2}-2a\frac{dy}{dx}+(a^2+b^2)y=0.$$

3. If $y=\tan^{-1}(\sinh x)$, show that
$$y_2-y_1^2\tan y=0.$$

4. Find the nth differential coefficient $\frac{1}{x^2+a^2}$

Ans. $y_n=(-1)^n n! a^{-(n+2)}\sin(n+1)\phi\sin^{n+1}\phi$, where $\phi=\tan^{-1}\left(\frac{a}{x}\right)$.

5. Find the nth derivative of $e^{2x}\cos^2 x\sin x$.

Ans. $$y_n=\frac{1}{4}\left[(13)^{\frac{n}{2}}\sin\left(3x+n\tan^{-1}\left(\frac{3}{2}\right)\right)+(5)^{\frac{n}{2}}\sin\left(x+n\tan^{-1}\left(\frac{1}{2}\right)\right)\right].$$

6. Find the nth derivative of $\frac{x^2}{(x+2)(2-3x)}$

Ans. $$\frac{(-1)^n}{2}\left[\frac{1}{(x+2)^{n+1}}-\frac{(-3)^{n-1}}{(2-3x)^{n+1}}\right].$$

7. Find the nth derivative of $\frac{x}{2x^2+3x+1}$

Ans. $$(-1)^n n!\left[\frac{1}{(x+1)^{n+1}}-\frac{2^n}{(2x+1)^{n+1}}\right].$$

8. Find the nth derivative of $\sin^2 x \cos^3 x$

Ans. $\frac{1}{16}\left[2\cos\left(x+\frac{n\pi}{2}\right) - 3^n\cos\left(3x+\frac{n\pi}{2}\right) - 5^n\cos\left(5x+\frac{n\pi}{2}\right)\right].$

9. Find the nth derivative of $\tan^{-1}\frac{1+x}{1-x}$

Ans. $(-1)^{n-1}(n-1)!\sin n\phi \sin^n\phi$, $\phi = \cot^{-1} x$.

10. If $y = A\sin mx + B\cos mx$, show that

$$y_2 + m^2 y = 0.$$

11. If $p^2 = a^2\cos^2\theta + b^2\sin^2\theta$, show that

$$\frac{d^2p}{d\theta^2} + p = \frac{a^2b^2}{p^3}.$$

Leibnitz`s Theorem

12. Find the nth differential coefficient of $e^x \log x$.

Ans. $e^x[\log x + n_{c_1}x^{-1} - n_{c_2}x^{-2} + n_{c_3}.2\,!x^{-3} - \cdots + (-1)^{n-1}.(n-1)!x^{-n}]$.

13. If $y = (x^2-1)^n$, show that

$$(x^2-1)y_{n+2} + 2xy_{n+1} - n(n+1)y_n = 0.$$

14. If $y = x^2e^x$, show that

$$y_n = \frac{1}{2}n(n-1)y_2 - n(n-2)y_1 + \frac{1}{2}(n-1)(n-2)y.$$

Hint: By Leibnitz's Theorem $y_n = x^2e^x + 2nxe^x + n(n-1)e^x$. Find y_1 and y_2. Substitute the values of x^2e^x , $2xe^x$ and e^x from y, y_1 and y_2 respectively in y_n and get the required result.

15. If $y = (\sin^{-1}x)^2$, show that

(i) $(1-x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} - 2 = 0,$

(ii) $(1-x^2)y_{n+2} - (2n+1)xy_{n+1} - n^2y_n = 0.$

16. If $y = \sin^{-1}x$, show that

$$(y_n)_0 = (n-2)^2(n-4)^2(y_{n-4})_0.$$

Deduce that

(i) $(y_n)_0 = (n-2)^2(n-4)^2\ldots 5^2.3^2.1$ for odd n,

(ii) $(y_n)_0 = 0$ for even n.

17. If $y = e^{a\sin^{-1}x}$ find $(y_n)_0$.

Hint: From Example 1.11,we have $(y_{n+2})_0 = (n^2+a^2)(y_n)_0$.

Ans. $(y_n)_0 = [(n-2)^2+a^2]\ldots(3^2+a^2) \times (1^2+a^2).a$ if n is odd

$(y_n)_0 = [(n-2)^2+a^2]\ldots(4^2+a^2) \times (2^2+a^2) - a^2$ if n in even

18. If $y^{\frac{1}{m}} + y^{-\frac{1}{m}} = 2x$, show that

$$(x^2-1)y_{n+2} + (2n+1)xy_{n+1} + (n^2-m^2)y_n = 0.$$

19. If $y = \log[x + \sqrt{1+x^2}]$, find $(y_n)_0$.

Ans. $(y_n)_0 = (-1)^{\frac{n-1}{2}}(n-2)^2(n-4)^2\ldots \times 3^2.1^2$ if n is odd

$(y_n)_0 = 0$ if n is even

2 Asymptotes and Curve Tracing

The aim of this chapter is to study the shape of a *plane curve* $y = f(x)$. For this purpose, we must investigate the variation of the function f, in the case of unlimited increase and absolute value and of x or y, or both, of a variable point (x, y) on the curve. The study of such variation of the function requires the concept of an asymptote. Before defining an asymptote to a curve, let us define finite- and infinite branches of a plane curve as follows:
Consider the equation of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Solving this equation, we get

$$y = b\sqrt{1 - \frac{x^2}{a^2}} \quad \text{or} \quad y = -b\sqrt{1 - \frac{x^2}{a^2}}.$$

The first equation represents the upper half of the ellipse while the second equation represents the lower half of the ellipse. Thus, the earlier equation represents two branches of the ellipse. Further, both these branches lie within the finite part of the xy-plane bounded by $x = \pm\, a$ and $y = \pm\, b$. Hence, both these branches of the ellipse are *finite*.

Consider now the equation of the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$. Its solution is

$$y = \frac{b}{a}\sqrt{x^2 - a^2} \quad \text{or} \quad y = -\frac{b}{a}\sqrt{x^2 - a^2}.$$

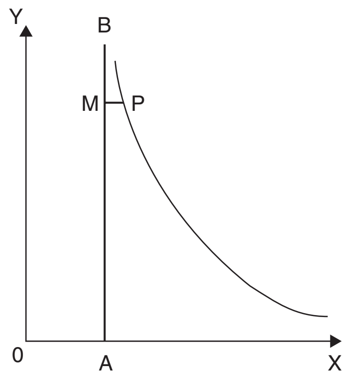
Therefore, y tends to $\pm\infty$ as $x \to \pm\infty$. Hence, both branches of this hyperbola extend to infinity and are therefore called the *infinite branches* of the rectangular hyperbola.

A variable point $P(x, y)$ moves along a curve to infinity if the distance of the point from the origin increases without bound. In other words, a point $P(x, y)$ on an infinite branch of a curve is said to tend to infinity along the curve if either x or y, or both, tend to infinity as $P(x, y)$ moves along the branch of the curve.

Now we are in a position to define an asymptote to a curve.

A straight line, at a finite distance from the origin, is said to be a *rectilinear asymptote* (or simply *asymptote*) of an infinite branch of a curve if the perpendicular distance of a point P on that branch from this straight line tends to zero as P tends to infinity along the branch of the curve.

For example, the line AB will be asymptote of the curve in the following figure if the perpendicular distance PM from the point P to the line AB tends to zero as P tends to infinity along the curve.



2.1 DETERMINATION OF ASYMPTOTES WHEN THE EQUATION OF THE CURVE IN CARTESIAN FORM IS GIVEN

Let

$$y = mx + c \tag{1}$$

be the equation of a straight line. Let $P\,(x, y)$ be an arbitrary point on the infinite branch of the curve $f(x, y) = 0$. We wish to find the values of m and c so that (1) is an asymptote to the curve. Let $PM = p$ be the perpendicular distance of the point $P\,(x, y)$ from (1). Then

$$p = \frac{y - mx - c}{\sqrt{1 + m^2}}.$$

The abscissa x must tend to infinity as the point $P(x, y)$ recedes to infinity along this line. Thus,

$p \to 0$ as $x \to \infty$. Therefore,

$$\lim_{x\to\infty}(y - mx - c) = 0$$

or

$$\lim_{x\to\infty}(y - mx) = c.$$

On the other hand,

$$\frac{y}{x} - m = (y - mx)\frac{1}{x}.$$

Therefore,

$$\lim_{x\to\infty}\left(\frac{y}{x} - m\right) = \lim_{x\to\infty}(y - mx)\lim_{x\to\infty}\frac{1}{x} = c(0) = 0$$

or

$$\lim_{x\to\infty}\frac{y}{x} = m.$$

Hence,

$$m = \lim_{x\to\infty}\frac{y}{x} \quad \text{and} \quad c = \lim_{x\to\infty}(y - mx).$$

Thus, to find asymptotes which are not parallel to the y-axis, we find $\lim_{x\to\infty}\frac{y}{x}$ and $\lim_{x\to\infty}(y - mx)$. If these limits are, respectively, m and c, then $y = mx + c$ is an asymptote.

2.2 THE ASYMPTOTES OF THE GENERAL RATIONAL ALGEBRAIC CURVE

Let $f(x, y) = 0$ be the equation of any rational algebraic curve of the nth degree. Arranging this equation in groups of homogeneous terms in x and y, we get

$$(a_0x^n + a_1x^{n-1}y + a_2x^{n-2}y^2 + \ldots + a_ny^n)$$
$$+ \ldots + (b_1x^{n-1} + b_2x^{n-2}y + \ldots + b_nx^{n-1})$$
$$+ (c_2x^{n-2} + c_3x^{n-3}y + \ldots + c_ny^{n-2})$$
$$+ \ldots + (k_0x + k_1y) + K = 0.$$

This equation can be written as

$$x^n\phi_n\left(\frac{y}{x}\right) + x^{n-1}\phi_{n-1}\left(\frac{y}{x}\right) + x^{n-2}\phi_{n-2}\left(\frac{y}{x}\right)$$
$$+ \ldots + x\phi_1\left(\frac{y}{x}\right) + \phi_0\left(\frac{y}{x}\right) = 0, \qquad (1)$$

where $\phi_r\left(\frac{y}{x}\right)$ is a polynomial in $\frac{y}{x}$ of degree r. Suppose $y = mx + c$ as an asymptote of the curve, where m and c are finite. We have to find m and c. Dividing both sides of equation (1) *by* x^n, we get

$$\phi_n\left(\frac{y}{x}\right) + \frac{1}{x}\phi_{n-1}\left(\frac{y}{x}\right) + \frac{1}{x^2}\phi_{n-2}\left(\frac{y}{x}\right) + \ldots = 0.$$

Proceeding to limits as $x \to \infty$ so that $\lim_{x\to\infty}\frac{y}{x} = m$, we have

$$\lim_{x\to\infty}\left[\phi_n\left(\frac{y}{x}\right) + \frac{1}{x}\phi_{n-1}\left(\frac{y}{x}\right) + \frac{1}{x^2}\phi_{n-2}\left(\frac{y}{x}\right) + \ldots\right] = 0,$$

which yields

$$\phi_n(m) = 0. \qquad (2)$$

Solving the equation (2), we get the slope m of the asymptote $y = mx + c$. But $\lim_{x\to\infty}(y - mx) = c$. Let $y - mx = p$ so that $x \to \infty$, $p \to c$. But $y - mx = p$ implies $\frac{y}{x} = m + \frac{p}{x}$. Substituting this value of $\frac{y}{x}$ in equation (1), we have

$$x^n\phi_n\left(m + \frac{p}{x}\right) + x^{n-1}\phi_{n-1}\left(m + \frac{p}{x}\right)$$
$$+ x^{n-2}\phi_{n-2}\left(m + \frac{p}{x}\right) + \ldots = 0. \qquad (3)$$

Taylor's Theorem expansion of equation (3) yields

$$x^n\left[\phi_n(m) + \frac{p}{x}\phi_n'(m) + \frac{p^2}{2x^2}\phi_n''(m) + \ldots\right]$$
$$+ x^{n-1}\left[\phi_{n-1}(m) + \frac{p}{x}\phi_{n-1}'(m) + \ldots\right]$$
$$+ x^{n-2}\left[\phi_{n-2}(m) + \frac{p}{x}\phi_{n-2}'(m) + \ldots\right] + \ldots = 0.$$

Using equation (2), the said equation reduces to

$$x^{n-1}[p\phi_n'(m) + \phi_{n-1}(m)] + x^{n-2}\left[\frac{p^2}{2!}\phi_n''(m)\right.$$
$$\left. + p\phi_{n-1}'(m) + \phi_{n-2}(m)\right] + \ldots = 0$$

or

$$p\phi_n'(m) + \phi_{n-1}(m) + \frac{1}{x}\left[\frac{p^2}{2!}\phi_n''(m) + p\phi_{n-1}'(m)\right.$$
$$\left. + \phi_{n-2}(m)\right] + \ldots = 0. \qquad (4)$$

Since $x \to \infty$, $p \to c$, we have

$$c\phi_n'(m) + \phi_{n-1}(m) = 0. \qquad (5)$$

Case (i): If $\phi_n(m)$ has no repeated root, then $\phi_n'(m) \neq 0$. Hence, in that case, equation (5) implies

$$c = -\frac{\phi_{n-1}(m)}{\phi_n'(m)}. \qquad (6)$$

If $m_1, m_2, m_3, \ldots$ are the distinct roots of $\phi_n(m) = 0$ and $c_1, c_2, c_3, \ldots$ are the corresponding values of c determined by equation (6), then the asymptotes are

$$y = m_1x + c_1, \quad y = m_2x + c_2, \quad y = m_3x + c_3, \ldots$$

Case (ii): If $\phi_n'(m) = 0$, that is, $\phi_n(m)$ has a repeated root and if $\phi_{n-1}(m) \neq 0$, then equation (6) implies that c is undefined. Hence, there exists no asymptote to the curve in this case.

Case (iii): If $\phi_n'(m) = \phi_{n-1}(m) = 0$. Then equation (5) reduces to an identity and equation (4) reduces to

$$\frac{p^2}{2!}\phi_n''(m)+p\phi_{n-1}'(m)+\phi_{n-2}(m)+\frac{1}{x}[\ldots]+\ldots=0.$$

As $x \to \infty$, $p \to c$ we have

$$\frac{c^2}{2}\phi_n''(m)+c\phi_{n-1}'(m)+\phi_{n-2}(m)=0.$$

If $\phi_n''(m) \neq 0$, then this last quadratic in c gives two values of c. Therefore, there are two asymptotes

$$y = mx + c_1 \text{ and } y = mx + c_2,$$

corresponding to the slope m. Thus, in this case, we have two parallel asymptotes.

Remark 2.1

(i) Since the degree of $\phi_n(m) = 0$ is n at the most, the number of asymptotes, real or imaginary, which are not parallel to y-axis, cannot exceed n. In case the curve has asymptotes parallel to y-axis, then the degree of $\phi_n(m)$ is smaller than n by at least the number of asymptotes parallel to y-axis. Thus, the total number of asymptotes cannot exceed the degree n of the curve.

(ii) Asymptotes parallel to y-axis cannot be found by the said method as the equation of a straight line parallel to y-axis cannot be put in the form $y = mx + c$.

2.3 ASYMPTOTES PARALLEL TO COORDINATE AXES

(i) Asymptotes parallel to y-axis of a rational algebraic curve

Let $f(x, y) = 0$ be the equation of any algebraic curve of the mth degree. Arranging the equation in descending powers of y, we get

$$y^m\phi_0(x)+y^{m-1}\phi_1(x)+y^{m-2}\phi_2(x)+\ldots+\phi_m(x)=0, \quad (1)$$

where $\phi_0(x)$, $\phi_1(x)$, $\phi_2(x)$,… are polynomials in x. Dividing the equation (1) by y^m, we get

$$\phi_0(x)+\frac{1}{y}\phi_1(x)+\frac{1}{y^2}\phi_2(x)+\ldots+\frac{1}{y^m}\phi_m(x)=0. \quad (2)$$

If $x = c$ be an asymptote of the curve parallel to y-axis then $\lim_{y\to\infty} x = c$, where (x, y) lies on the curve (1). Therefore,

$$\lim_{y\to\infty}\left[\phi_0(x)+\frac{1}{y}\phi_1(x)+\frac{1}{y^2}\phi_2(x)+\ldots\right]=0$$

$$\text{or } \phi_0(c)=0$$

so that c is a root of the equation $\phi_0(x) = 0$. If c_1, c_2,… are the roots of $\phi_0(x) = 0$, then $(x - c_1)$, $(x - c_2)$,…are the factors of $\phi_0(x)$. Also $\phi_0(x)$ is the coefficient of the highest power of y, that is, of y^m in equation (1). Thus, we have the following simple rule to determine the asymptotes parallel to y-axis.

The asymptotes parallel to the y-axis are obtained by equating to zero the coefficient of the highest power of y in the given equation of the curve. In case the coefficient of the highest power of y is a constant or if its linear factors are imaginary, then there will be no asymptotes parallel to the y-axis.

(ii) Asymptotes parallel to the *x*-axis of a rational algebraic curve

Proceeding exactly as in case (i) mentioned earlier, we arrive at the following rule to determine the asymptotes parallel to the x-axis:

The asymptotes parallel to the x-axis are obtained by equating to zero the coefficient of the highest power of x in the given equation of the curve. In case the coefficient of the highest power of x is a constant or if its linear factors are imaginary, then there will be no asymptotes parallel to the x-axis.

2.4 WORKING RULE FOR FINDING ASYMPTOTES OF RATIONAL ALGEBRAIC CURVE

In view of the mentioned discussion, we arrive at the following working rule for finding the asymptotes of rational algebraic curves:

1. A curve of degree n may have utmost n asymptotes.
2. The asymptotes parallel to the y-axis are obtained by equating to zero the coefficient of the highest power of y in the given equation of the curve. In case the coefficient of the highest power of y is a constant or if its linear factors are imaginary, then there will be no asymptotes parallel to the y-axis.

   The asymptotes parallel to the x-axis are obtained by equating to zero the coefficient of the highest power of x in the given equation of the curve. In case the coefficient of the highest power of x is a constant or if its linear factors are imaginary, then there will be no asymptotes parallel to the x-axis.

If $y = mx + c$ is an asymptote not parallel to the y-axis, then the values of m and c are found as follows:

(i) Find $\phi_n(m)$ by putting $x = 1, y = m$ in the highest-degree terms of the given equation of the curve. Solve the equation $\phi_n(m) = 0$ for slope (m). If some values are imaginary, reject them.

(ii) Find $\phi_{n-1}(m)$ by putting $x = 1, y = m$ in the next lower-degree terms of the equation of the curve. Similarly $\phi_{n-2}(m)$ may be found taking $x = 1, y = m$ in the next lower-degree terms in the curve and so on.

(iii) If $m_1, m_2, \ldots$ are the real roots of $\phi_n(m)$, then the corresponding values of c, that is, $c_1, c_2, \ldots$ are given by

$$c = -\frac{\phi_{n-1}(m)}{\phi'_n(m)}, \quad m = m_1,\ m_2, \ldots.$$

Then the required asymptotes are

$$y = m_1x + c_1,\ y = m_2x + c_2, \ldots$$

(iv) If $\phi'_n(m) = 0$ for some m but $\phi_{n-1}(m) \neq 0$, then there will be no asymptote corresponding to that value of m.

(v) If $\phi'_n(m) = 0$ and $\phi_{n-1}(m) = 0$ for some value of m, then the value of c is determined from

$$\frac{c^2}{2!}\phi''_n(m) + \frac{c}{1!}\phi'_{n-1}(m) + \phi_{n-2}(m) = 0.$$

This equation will yield two values of c and thus, we will get atmost two parallel asymptotes corresponding to this value of m, provided $\phi''_n(m) \neq 0$.

(vi) Similarly, if $\phi''_n(m) = \phi'_{n-1}(m) = \phi_{n-2}(m) = 0$, then the value of c is determined from

$$\frac{c^3}{3!}\phi'''_n(m) + \frac{c^2}{2!}\phi''_{n-1}(m) + \frac{c}{1!}\phi'_{n-2}(m) + \phi_{n-3}(m) = 0.$$

In this case, we get atmost three parallel asymptotes corresponding to this value of m.

EXAMPLE 2.1

Find the asymptotes of the curve

$$y^2\left(x^2 - a^2\right) = x^2\left(x^2 - 4a^2\right).$$

Solution. The equation of the curve is

$$y^2\left(x^2 - a^2\right) = x^2\left(x^2 - 4a^2\right)$$

or

$$y^2x^2 - x^4 - a^2y^2 + 4a^2x^2 = 0.$$

Since the degree of the curve is 4, it cannot have more than four asymptotes. Equating to zero, the coefficient of the highest power of y, the asymptote parallel to the y-axis is given by $x^2 - a^2 = 0$. Thus, the asymptotes parallel to the y-axis are $x = \pm a$.

Since the coefficient of the highest power of x in the given equation is constant, there is no asymptote parallel to the x-axis.

To find the oblique asymptotes, we put $x = 1$ and $y = m$ in the highest-degree term, that is fourth-degree term $y^2x^2 - x^4$ in the given equation and get $\phi_4(m) = m^2 - 1$. Therefore, slopes of the asymptotes are given by

$$\phi_4(m) = m^2 - 1 = 0.$$

Hence, $m = \pm 1$. Again putting $y = m$ and $x = 1$ in the next highest-degree term, that is, third-degree term, we have $\phi_3(m) = 0$ (since there is no term of degree 3).

Now c is given by

$$c = -\frac{\phi_3(m)}{\phi'_4(m)} = \frac{0}{2m} = 0.$$

Therefore, the oblique asymptotes are $y = x + 0$ and $y = -x + 0$.

Hence, all the four asymptotes of the given curve are $x = \pm a$ and $y = \pm x$.

EXAMPLE 2.2

Find all the asymptotes of the curve

$$f(x,y) = y^3 - xy^2 - x^2y + x^3 + x^2 - y^2 - 1 = 0.$$

Solution. The given curve is of degree 3 and so, it may have atmost three asymptotes. Since the coefficients of the highest power of x and y are constants, the curve has no asymptote parallel to the coordinate axes.

To find the oblique asymptotes, we put $x = 1$ and $y = m$ in the expression containing third-degree terms of $f(x, y)$. Thereby we get

$$\phi_3(m) = m^3 - m^2 - m + 1 = 0.$$

This equation yields $m = 1, 1, -1$. Further, putting $x = 1, y = m$ in the next highest-degree term, we get

$$\phi_2(m) = 1 - m^2.$$

Therefore for $m=-1$, the expression

$$c=-\frac{1-m^2}{3m^2-2m-1}$$

yields $c=0$ and the corresponding asymptote is

$$y=-x+0 \text{ or } y+x=0.$$

For $m=1$, the denominator is zero and so, c cannot be determined by the preceding formula. Putting $x=1$, $y=m$ in the first-degree terms, we have $\phi_1(m)=0$ (since there is no first-degree term). Now for $m=1$, the constant c is given by

$$\frac{c^2}{2}\phi_2''(m)+c\phi_2'(m)+\phi_1(m)=0$$

$$\text{or } (3m-1)c^2-2mc=0$$

$$\text{or } 2c^2-2c=0 \quad \text{for } m=1$$

$$\text{or } c(c-1)=0.$$

Hence, $c=0$ and $c=1$. So the two parallel asymptotes corresponding to $m=1$ are $y=x$ and $y=x+1$. Therefore, the asymptotes to the curve are

$$y+x=0,\ y=x \text{ and } y=x+1.$$

EXAMPLE 2.3

Find the asymptotes of the curve

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

Solution. The degree of the curve is 3. So, there cannot be more than three asymptotes. There is no asymptote parallel to the x-axis. The asymptote parallel to the y-axis is given by $x-2\text{a}=0$, that is, $x=2a$.

To find the oblique asymptotes, we put $x=1$, $y=m$ in the third-degree term and get $\phi_3(m)=m^2-1$ and so, the slope m is given by

$$\phi_3(m)=m^2-1=0.$$

Thus, $m=\pm 1$. Further, putting $x=1$, $y=m$ in the second-degree terms, we get $\phi_2(m)=-2am^2$. Therefore for $m=1$ and $m=-1$, the expression

$$c=-\frac{\phi_2(m)}{\phi_3'(m)}=\frac{2am^2}{2m}=am$$

yields $c=a$, and $-a$ respectively. Hence, the oblique asymptotes are

$$y=x+a \text{ and } y=-x-a.$$

Hence, the three asymptotes of the curve are

$$x=2a,\ x-y+a=0,\ \text{and } x+y+a=0.$$

EXAMPLE 2.4

Find the asymptotes of the curve

$$x^3+3x^2y-4y^3-x+y+3=0.$$

Solution. There is no asymptote parallel to the coordinate axes. To find the oblique asymptotes, we have

$$\phi_3(m)=1+3m-4m^3$$

and so, the slope m is given by

$$\phi_3(m)=1+3m-4m^3=0.$$

Therefore, $m=1,-\frac{1}{2}$, and $-\frac{1}{2}$. For $m=1$, the value of c is given by

$$c=-\frac{\phi_2(m)}{\phi_3''(m)}=-\frac{0}{-12m^2+3}=0.$$

Thus, the asymptote corresponding to $m=1$ is $y=x$ or $x-y=0$.

For $m=-\frac{1}{2}, \phi_3'(m)=0$. So we find $\phi_1(m)$, which is equal to $\phi_1(m)=-1+m$.

Hence, c is given by,

$$\frac{c^2\phi_3''(m)}{2}+c\phi_2'(m)+\phi_1(m)=0$$

or

$$6c^2-\frac{3}{2}=0 \quad \text{or} \quad c^2=\frac{1}{4} \quad \text{or} \quad c=\pm\frac{1}{2}.$$

Thus, the asymptotes corresponding to $m=-\frac{1}{2}$ are

$$y=-\frac{1}{2}x+\frac{1}{2} \text{ and } y=-\frac{1}{2}x-\frac{1}{2}$$

or

$$x+2y-1=0 \text{ and } x+2y+1=0.$$

Hence, the three asymptotes of the curve are $x-y=0$, $x+2y-1=0$, and $x+2y+1=0$.

EXAMPLE 2.5

Find the asymptotes of the curve

$$(x-y)^2(x^2+y^2)-10(x-y)x^2+12y^2+2x+y=0.$$

Solution. The equation of the given curve is

$$(x-y)^2(x^2+y^2)-10(x-y)x^2+12y^2+2x+y=0.$$

The coefficients of x^4 and y^4 are constant. Therefore, the curve has no asymptotes parallel to the axes. Putting $x=1$ and $y=m$ in the fourth-, third- and second-degree terms, we have

$$\phi_4(m)=(1-m)^2(1+m^2)=m^4-2m^3+2m^2-2m+1$$

$$\phi_3(m)=10(m-1),\ \text{and } \phi_2(m)=12m^2.$$

The slopes of the asymptotes are given by

$$\phi_4(m)=(1-m)^2(1+m^2)=0.$$

Therefore, $m=1,1$ are the real roots. Further we have

$$\phi_4'(m)=4m^3-6m^2+4m-2,$$

so that $\phi_4'(m) = 0$ for $m = 1$. Therefore, values of c are given by

$$\frac{c^2}{2}\phi_4''(m) + c\phi_3'(m) + \phi_2(m) = 0,$$

that is,

$$\frac{c^2}{2}\left(12m^2 - 12m + 4\right) + 10c + 12m^2 = 0.$$

For $m = 1$, this equation yields

$$2c^2 + 10c + 12 = 0 \text{ or } c^2 + 5c + 6 = 0.$$

This equation gives $c = -2, -3$. Putting the values of m and c in $y = mx + c$, the asymptotes are given by

$$y = x - 2 \text{ and } y = x - 3.$$

EXAMPLE 2.6

Find the asymptotes of the curve

$$(x + y)^2(x + 2y) + 2(x + y)^2 - (x + 9y) - 2 = 0.$$

Solution. Since the coefficients of the highest-degree term of x and y are constant, the given curve does not have asymptotes parallel to the axes.

To find the oblique asymptotes, we put $x = 1$ and $y = m$ in third-, second- and first-degree terms and get

$$\phi_3(m) = (1 + m)^2(1 + 2m) = 2m^3 + 5m^2 + 4m + 1$$

$$\phi_2(m) = 2(1 + m)^2 = 2m^2 + 4m + 1$$

$$\phi_1(m) = -(1 + 9m) = -9m - 1.$$

Thus,

$$\phi_3'(m) = 6m^2 + 10m + 4,$$

$$\phi_3''(3) = 12m + 10, \text{ and}$$

$$\phi_2'(m) = 4m + 4.$$

The slopes of the asymptotes are given by

$$\phi_3(m) = (1 + m)^2(1 + 2m) = 0,$$

which yields $m = -1, -1$, and $-\frac{1}{2}$. The value of c is given by

$$c = -\frac{\phi_2(m)}{\phi_3'(m)} = -\frac{2m^2 + 4m + 1}{6m^2 + 10m + 4}.$$

For $m = -1$, $\phi_3'(m) = 0$ and so, c cannot be found from this equation. For $m = -\frac{1}{2}$ we have $c = -1$. Thus, the asymptotes corresponding to $m = -\frac{1}{2}$ is

$$y = -\frac{1}{2}x - 1 \text{ or } x + 2y + 2 = 0.$$

For $m = -1$, the value of c is calculated from the relation

$$\frac{c^2}{2}\phi_3''(m) + c\phi_2'(m) + \phi_1(m) = 0$$

or

$$\frac{c^2}{2}(12m + 10) + c(4m + 4) - 9m - 1 = 0$$

or

$$c^2(6m + 5) + c(4m + 4) - 9m - 1 = 0$$

or

$$c^2(-1) + 9 - 1 = 0$$

or

$$c^2 = 8, \text{ which yields } c = \pm 2\sqrt{2}.$$

Thus, the two parallel asymptotes corresponding to the slope $m = -1$ are

$$y = -x + 2\sqrt{2} \text{ and } y = -x - 2\sqrt{2}.$$

Hence, the asymptotes of the curve are

$$x + 2y + 2 = 0,\ y + x = 2\sqrt{2}, \text{ and } y + x = -2\sqrt{2}.$$

EXAMPLE 2.7

Find the asymptotes of the curve

$$6x^2 + xy - 2y^2 + x + 2y + 1 = 0$$

Solution. Since the coefficients of the highest powers of x and y are constants, there is no asymptotes parallel to the axes. To find the oblique asymptotes, we put $x = 1$ and $y = m$ in second- and first-degree terms and get

$$\phi_2(m) = 6 - 2m^2 + m,\ \phi_1(m) = 2m + 1$$

and

$$\phi_2'(m) = -4m + 1.$$

The slopes of the asymptotes are given by

$$\phi_2(m) = 6 - 2m^2 + m = 0$$

and so, $m = 2, -\frac{3}{2}$. The value of c is given by

$$c = -\frac{\phi_1(m)}{\phi_2'(m)} = -\frac{2m + 1}{-4m + 1}.$$

For $m = 2$ and $m = -\frac{3}{2}$, the value of c are $\frac{5}{7}$ and $\frac{2}{7}$ respectively.

Therefore, the asymptotes are

$$y = 2x + \frac{5}{7} \text{ and } y = -\frac{3}{2}x + \frac{2}{7}$$

or

$$14x - 7y + 5 = 0 \text{ and } 21x + 14y - 4 = 0.$$

EXAMPLE 2.8

Find the asymptotes of the curve $\frac{a^2}{x^2} - \frac{b^2}{y^2} = 1$.

Solution. The equation of the given curve is

$$\frac{a^2}{x^2} - \frac{b^2}{y^2} = 1$$

or

$$x^2y^2 - a^2y^2 + b^2x^2 = 0.$$

Since the curve is of degree 4, it cannot have more than four asymptotes.

Equating the coefficient of the highest power of x to zero, we get $y^2 + b^2 = 0$, which yields imaginary asymptotes.

Equating the coefficient of the highest power of y to zero, we get

$$x^2 - a^2 = 0 \text{ or } (x-a)\ (x+a) = 0.$$

Hence, the asymptotes parallel to the y-axis are $x = a$ and $x = -a$. Thus, the only real asymptotes are $x - a = 0$ and $x + a = 0$.

2.5 INTERSECTION OF A CURVE AND ITS ASYMPTOTES

We have seen that the equation of a curve of degree n can be expressed in the form

$$x^n\phi_n\left(\frac{y}{x}\right) + x^{n-1}\phi_{n-1}\left(\frac{y}{x}\right) + x^{n-2}\phi_{n-2}\left(\frac{y}{x}\right) + \ldots = 0. \quad (1)$$

Let

$$y = mx + c \quad (2)$$

be an asymptote to the curve (1). Eliminating y from (1) and (2), we get

$$x^n\phi_n\left(m+\frac{c}{x}\right) + x^{n-1}\phi_{n-1}\left(m+\frac{c}{x}\right) + x^{n-2}\phi_{n-2}\left(m+\frac{c}{x}\right) + \ldots = 0.$$

Expanding by Taylor's Theorem, we get

$$x^n\left[\phi_n(m) + \frac{c}{x}\phi_n'(m) + \frac{c^2}{2!x^2}\phi_n''(m) + \ldots\right] + x^{n-1}\left[\phi_{n-1}(m) + \frac{c}{x}\phi_{n-1}'(m) + \frac{c^2}{2!x^2}\phi_{n-1}''(m) + \ldots\right] + x^{n-2}\left[\phi_{n-2}(m) + \frac{c}{x}\phi_{n-2}'(m) + \frac{c^2}{2!x^2}\phi_{n-2}''(m) + \ldots\right] = 0, \quad (3)$$

that is,

$$x^n\phi_n(m) + x^{n-1}[\phi_n'(m) + \phi_{n-1}(m)] + x^{n-2}\left[\frac{c^2}{2!}\phi_n''(m) + c\phi_{n-1}'(m) + \phi_{n-2}(m)\right] + \ldots = 0.$$

But equation (2) being an asymptote of equation (1), the values of m and c are given by

$$\phi_n(m) = 0 \text{ and } c\phi_n'(m) + \phi_{n-1}(m) = 0.$$

Hence, equation (3) reduces to

$$x^{n-2}\left[\frac{c^2}{2!}\phi_n''(m) + c\phi_{n-1}'(m) + \phi_{n-2}(m)\right] + \ldots = 0,$$

which is of degree $n-2$ and so, yields $(n-2)$ values of x. Hence, the asymptote (2) cuts the curve (1) in $(n-2)$ points.

If the curve has n asymptotes, then they all will intersect the curve in $n(n-2)$ points.

Further, if the equation of the curve of the nth degree can be put in the form $F_n + F_{n-2} = 0$, where F_{n-2} is of degree $n-2$ at the most and F_n consists of n distinct linear factors, then the $n(n-2)$ *points of intersection of the curve* $F_n + F_{n-2} = 0$ *and its n asymptotes (given by* $F_n = 0$*) lie on the curve* $F_{n-2} = 0$.

EXAMPLE 2.9

Find the asymptotes of the curve

$$x^2y - xy^2 + xy + y^2 + x - y = 0$$

and show that they cut the curve in three points that lie on the straight line $x + y = 0$.

Solution. Equating to zero the coefficient of highest power of x, we get $y = 0$. Thus, x-axis is an asymptote to the given curve. Similarly, equating to zero the coefficient of the highest power of y, we get $-x + 1 = 0$ or $x = 1$. Thus, $x = 1$ is the asymptote parallel to y-axis. To find the oblique asymptotes, we put $x = 1$ and $y = m$ in the third- and second-degree terms and get

$$\phi_3(m) = m - m^2,\ \phi_2(m) = m + m^2, \text{ and}$$
$$\phi_3'(m) = 1 - 2m.$$

Then the slopes of the asymptotes are given by

$$\phi_3(m) = m - m^2 = 0,$$

which implies $m = 0$ and $m = 1$. The values of c are given by

$$c = -\frac{\phi_2(m)}{\phi_3'(m)} = -\frac{m + m^2}{1 - 2m}.$$

Thus, the values of c corresponding to $m = 0$ and $m = 1$ are $c = 0$ and $c = 2$, respectively. Therefore, the oblique asymptotes are $y = 0$ and $y = x + 2$. Hence, the asymptotes of the curve are

$$y = 0,\ x = 1, \text{ and } x - y + 2 = 0.$$

The joint equation of the asymptotes is

$$(x-1)y(x-y+2) = 0$$

or

$$x^2y - xy^2 + xy + y^2 - 2y = 0.$$

On the other hand, the equation of the curve can be written as

$$\left(x^2y - xy^2 + xy + y^2 - 2y\right) + y + x = 0,$$

which is of the form $F_n + F_{n-2} = 0$. Hence, the points of intersection which are $n\,(n-2) = 3(1) = 3$ in number lie on the curve $F_{n-2} = x + y = 0$, which is a straight line.

EXAMPLE 2.10

Show that the four asymptotes of the curve

$$\left(x^2 - y^2\right)\left(y^2 - 4x^2\right) + 6x^3 - 5x^2y - 3xy^2 + 2y^3 - x^2 + 3xy - 1 = 0$$

cut the curve in eight points which lie on the circle $x^2 + y^2 = 1$.

Solution. Substituting $x = 1$ and $y = m$ in the fourth- and third-degree terms, respectively, we get

$$\phi_4(m) = \left(1 - m^2\right)\left(m^2 - 4\right) \text{ and}$$

$$\phi_3(m) = 6 - 5m - 3m^2 + 2m^3.$$

Thus,

$$\phi_4'(m) = 10m - 4m^3.$$

The slopes of the asymptotes are given by

$$\phi_4(m) = \left(1 - m^2\right)\left(m^2 - 4\right) = 0$$

and so, $m = \pm 1$, and ± 2. The value of c is given by the expression

$$c = -\frac{\phi_3(m)}{\phi_4'(m)} = \frac{6 - 5m - 3m^2 + 2m^3}{4m^3 - 10m}.$$

The value of c corresponding to $m = 1, -1, 2,$ and -2 are respectively 0, 1, 0,and 1. Hence, the asymptotes are

$$y = x,\ y = -x + 1,\ y = 2x, \text{ and } y = -2x + 1.$$

Since the degree of the given curve is 4, the number of point of intersection is equal to $n(n-2) = 4(4-2) = 8$.

The joint equation of the asymptotes is

$$(y - x)(y + x - 1)(y - 2x)(y + 2x - 1) = 0$$

or

$$\left(y^2 - x^2\right)\left(y^2 - 4x^2\right) - 6x^3 + 5x^2y + 3xy^2 - 2y^3 + y^2 - 3xy + 2x^2 = 0$$

or

$$\left(x^2 - y^2\right)\left(y^2 - 4x^2\right) + 6x^3 - 5x^2y - 3xy^2 + 2y^3 - y^2 + 3xy - 2x^2 = 0.$$

The given equation of the curve can be written as

$$\left(x^2 - y^2\right)\left(y^2 - 4x^2\right) + 6x^3 - 5x^2y - 3xy^2 + 2y^3 - y^2 + 3xy - 2x^2 + \left(x^2 + y^2 - 1\right) = 0,$$

which is of the form $F_n + F_{n-2} = 0$. Hence, the points of intersection lie on $F_{n-2} = 0$, that is, on the circle $x^2 + y^2 - 1 = 0$.

EXAMPLE 2.11

Find the equation of the cubic which has the same asymptotes as the curve $x^3 - 6x^2y + 11xy^2 - 6y^2 + x + y + 1 = 0$, and which touches the axis of y at the origin and passes through the point (3, 2).

Solution. The equation of the curve is

$$x^3 - 6x^2y + 11xy^2 - 6y^2 + x + y + 1 = 0.$$

The curve has no asymptote parallel to the axes. To find the oblique asymptotes, we have

$$\phi_3(m) = 1 - 6m + 11m^2 - 6m^2 = (1 - m)(1 - 2m)(1 - 3m),$$

$$\phi_2(m) = 0,\ \phi_3'(m) = 10m - 6.$$

The slopes of the asymptotes are given by

$$\phi_3(m) = (1 - m)(1 - 2m)(1 - 3m) = 0$$

and so, $m = 1, \frac{1}{2}$, and $\frac{1}{3}$. Further,

$$c = -\frac{\phi_2(m)}{\phi_3'(m)} = 0.$$

Therefore, the asymptotes are

$$y = x,\ y = \frac{x}{2}, \text{ and } y = \frac{x}{3}.$$

The joint equation of the asymptotes is

$$(x - y)(x - 2y)(x - 3y) = 0.$$

The most general equation of any curve having these asymptotes is

$$F_n + F_{n-2} = 0, \quad \text{that is, } F_3 + F_1 = 0$$

or

$$(x - y)(x - 2y)(x - 3y) + ax + by + k = 0,$$

since F_1 is of degree 1.

Since the curve passes through the origin, putting $x = 0$, $y = 0$, in the preceding equation, we get $k = 0$. Thus, the equation of the curve becomes

$$(x - y)(x - 2y)(x - 3y) + ax + by = 0. \quad (1)$$

Equating to zero, the lowest-degree term in (1), we get $ax + by = 0$ as the equation of the tangent at the origin. But y-axis, that is, $x = 0$ is tangent at the origin. Therefore, $b = 0$ and the equation of the curve reduces to

$$(x - y)(x - 2y)(x - 3y) + ax = 0.$$

Since the curve passes through (3, 2), we have

$$(3 - 2)(3 - 4)(3 - 6) + 3a = 0$$

and so, $a = -1$. Hence, the required curve is

$$(x - y)(x - 2y)(x - 3y) - x = 0$$

or

$$x^3 - 6x^2y + 11xy^2 - 6y^3 - x = 0.$$

EXAMPLE 2.12

Show that the eight points of the curve

$$x^4 + 5x^2y^2 + 4y^4 + x^2 - y^2 + x + y + 1 = 0$$

and its asymptotes lie on a rectangular hyperbola.

Solution. The equation of the curve is of degree 4. Therefore, the number of points of intersection with the asymptotes is $n(n - 2) = 4(4 - 2) = 8$. Further, the equation of the given curve can be written as

$$\left(x^2 - y^2\right)\left(x^2 - 4y^2\right) + x^2 - y^2 + x + y + 1 = 0$$

or

$$F_n + F_{n-2} = 0,$$

where

$F_n = \left(x^2 - y^2\right)\left(x^2 - 4y^2\right)$ is of degree 4 and

$F_{n-2} = x^2 - y^2 + x + y + 1 = 0$ is of degree 2.

The asymptotes are given by $F_n = 0$, that is, by $(x^2 - y^2)(x^2 - 4y^2) = 0$. Thus, the asymptotes are $x = \pm y$ and $x = \pm 2y$. The equation $F_{n-2} = 0$, that is, $x^2 - y^2 + x + y + 1 = 0$ is the equation of the curve on which the points of intersection of the asymptotes and the given curve lie. The conic $x^2 - y^2 + x + y + 1 = 0$ is a hyperbola since the sum of the coefficients of x^2 and y^2 is zero. Hence, the eight points of intersection of the given curve with its asymptotes lie on a rectangular hyperbola.

EXAMPLE 2.13

Find the asymptotes of the curve

$$x = \frac{a(1 - t^2)}{1 + t^2},\ y = \frac{at(1 - t^2)}{1 + t^2}.$$

Solution. The equation of the curve is given in parametric form. We eliminate t by dividing and get

$$\frac{x}{y} = \frac{1}{t} \text{ so that } t = \frac{y}{x}.$$

Substituting this value of t in $x = \frac{a(1-t^2)}{1+t^2}$, we obtain

$$x = \frac{a(x^2 - y^2)}{x^2 + y^2}$$

or

$$y^2(a + x) = x^2(a - x). \tag{1}$$

Equating to zero the highest power of y in the equation (1) of the curve, we have $x + a = 0$. Hence, $x + a = 0$ is the asymptote parallel to the y-axis. To find the oblique asymptotes, we put $x = 1$ and $y = m$ in the highest-degree term of $f(x, y)$ to get

$$\phi_3(m) = m^2 + 1 = 0.$$

But the roots of the equation $m^2 + 1 = 0$ are imaginary. Therefore, there is no oblique asymptote. Hence, the only asymptote is $x + a = 0$.

2.6 ASYMPTOTES BY EXPANSION

Let the equation of the given curve be of the form

$$y = mx + c + \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x^3} + \dots . \tag{1}$$

Dividing both sides by x, we get

$$\frac{y}{x} = m + \frac{c}{x} + \frac{A}{x^2} + \frac{B}{x^3} + \frac{C}{x^4} + \dots .$$

Taking limit as $x \to \infty$, we have

$$\lim_{x\to\infty}\left(\frac{y}{x}\right) = m. \tag{2}$$

The equation (1) can also be written as

$$y - mx = c + \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x^3} + \dots .$$

Taking limit as $x \to \infty$, we get

$$\lim_{x\to\infty}(y - mx) = c. \tag{3}$$

It follows (see Article 4.1) from (2) and (3) that $y = mx + c$ is an asymptote of the curve (1). Hence, $y = mx + c$ is an asymptote of a curve, whose equation can be expressed in the form (1) given earlier.

For example, consider the curve

$$f(x, y) = 2x^3 + x^2(2 - y) + x + 1 = 0.$$

The given equation can be written as

$$x^2y = 2x^3 + 2x^2 + x + 1$$

or

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

Hence, $y = 2x + 2$ is an asymptote of the given curve.

2.7 ASYMPTOTES OF THE POLAR CURVES

If α is a root of the equation $f(\theta) = 0$, then $r\sin(\theta - \alpha) = \frac{1}{f'(\alpha)}$ is an asymptote of the polar curve $\frac{1}{r} = f(\theta)$.

Thus, to find the asymptotes of a polar curve, first write down the equation of the curve in the form $\frac{1}{r} = f(\theta)$. Then find the roots of the equation $f(\theta) = 0$. If the roots are $\theta_1, \theta_2, \theta_3, \dots$, find $f'(\theta)$ at $\theta = \theta_1, \theta_2, \theta_3, \dots$. Then the asymptotes of the curve shall be

$$r\sin(\theta-\theta_1)=\frac{1}{f'(\theta_1)},$$

$$r\sin(\theta-\theta_2)=\frac{1}{f'(\theta_2)},\ \text{and so on.}$$

EXAMPLE 2.14

Find the asymptotes of the curve

$$r\sin\theta=2\cos 2\theta.$$

Solution. The equation of the given curve can be written as

$$\frac{1}{r}=\frac{\sin\theta}{2\cos\theta}=f(\theta).$$

Therefore, $f(\theta)=0$ yields $\sin\theta=0$ and so, $\theta=n\pi$, where n is an integer. Since

$$f'(\theta)=\frac{1}{2}\left[\frac{\cos 2\theta\cos\theta-\sin\theta(-2\sin 2\theta)}{\cos^2 2\theta}\right],$$

we have

$$\frac{1}{f'(n\pi)}=\frac{2\cos^2(2n\pi)}{\cos(2n\pi)\cos n\pi+2\sin n\pi\sin 2n\pi}$$

$$=\frac{2}{\cos n\pi}=\frac{2}{(-1)^n}.$$

Hence, the required asymptotes are

$$r\sin(\theta-n\pi)=\frac{2}{(-1)^n}$$

or

$$-r\sin(n\pi-\theta)=\frac{2}{(-1)^n}$$

or

$$-r[(-1)^{n-1}\sin\theta]=\frac{2}{(-1)^n}$$

or

$$r\sin\theta=2.$$

EXAMPLE 2.15

Show that the curve $r=\frac{a}{1-\cos\theta}$ has no asymptotes.

Solution. The equation of the given curve can be written in the form

$$\frac{1}{r}=\frac{1-\cos\theta}{a}=f(\theta).$$

Then $f(\theta)=0$ implies $\cos\theta=1$ and so, $\theta=2n\pi$, where n is an integer. Further,

$$f'(\theta)=\frac{1}{a}\sin\theta$$

and so,

$$f'(2n\pi)=\frac{1}{a}\sin(2n\pi)=0.$$

We know that if α is a root of the equation $f(\theta)=0$, then asymptote corresponding to this asymptotic direction α is given by

$$f'(\alpha).r\sin(\theta-\alpha)=1.$$

So for $\alpha=2n\pi$, the equation of the asymptote is

$$f'(2n\pi).r\sin(\theta-2n\pi)=1.$$

But, we have shown that $f'(2n\pi)=0$. Thus, $0=1$, which is impossible. Hence, there is no asymptote to the given curve.

EXAMPLE 2.16

Find the asymptotes of the curve

$$r=a(\sec\theta+\tan\theta).$$

Solution. We are given that

$$r=a\left(\frac{1}{\cos\theta}+\frac{\sin\theta}{\cos\theta}\right)=\frac{a(1+\sin\theta)}{\cos\theta}.$$

Thus,

$$\frac{1}{r}=\frac{\cos\theta}{a(1+\sin\theta)}=f(\theta).$$

But $f(\theta)=0$ yields $\frac{\cos\theta}{a(1+\sin\theta)}=0$ or $\cos\theta=0$ or $\theta=(2n+1)\frac{\pi}{2}$.

Also,

$$f'(\theta)=\frac{1}{a}\left[\frac{(1+\sin\theta)\ (-\sin\theta)-\cos\theta\cos\theta}{(1+\sin\theta)^2}\right]$$

$$=-\frac{(\sin\theta+1)}{a(1+\sin\theta)^2}.$$

Therefore,

$$f'\left[(2n+1)\frac{\pi}{2}\right]=-\frac{1}{a}\frac{\sin(2n+1)\frac{\pi}{2}+1}{\left[1+\sin(2n+1)\frac{\pi}{2}\right]^2}$$

$$=-\frac{1}{a}\frac{(-1)^n+1}{[1+(-1)^n]^2}$$

and so, the asymptotes are

$$r\sin\left[\theta-(2n+1)\frac{\pi}{2}\right]=\frac{1}{f'\left[(2n+1)\frac{\pi}{2}\right]}$$

or

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

or

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

or

$$r\cos\theta = \frac{a[1+(-1)^n]^2}{1+(-1)^n} = a[1+(-1)^n].$$

Putting $n = 0, 1, 2,\ldots$, the asymptotes of the curve are given by

$$r\cos\theta = 2a \text{ and } r\cos\theta = 0.$$

Thus, we note that there are only two asymptotes of the given curve.

EXAMPLE 2.17

Find the asymptotes of the curve $r = a\tan\theta$.

Solution. The equation of the given curve may be written as

$$\frac{1}{r} = \frac{1}{a}\frac{\cos\theta}{\sin\theta} = f(\theta).$$

Therefore, $f(\theta) = 0$ implies $\cos\theta = 0$ and so, $\theta = (2n+1)\frac{\pi}{2}$. Also

$$f'(\theta) = -\frac{1}{a}\operatorname{cosec}^2\theta.$$

Therefore,

$$f'\left[(2n+1)\frac{\pi}{2}\right] = -\frac{1}{a\left[\sin(2n+1)\frac{\pi}{2}\right]^2} = \frac{-1}{a(-1)^{2n}}.$$

Thus,

$$\frac{1}{f'\left[(2n+1)\frac{\pi}{2}\right]} = a(-1)^{2n-1} = \pm a.$$

The asymptotes are now given by

$$r\sin\left(\theta - (2n+1)\frac{\pi}{2}\right) = \pm a.$$

Proceeding as in the earlier example, we get the asymptotes as

$$r\cos\theta = a \text{ and } r\cos\theta = -a.$$

EXAMPLE 2.18

Find the asymptotes of the following curves:

(i) $r\theta = a$

(ii) $r = \frac{2a}{1-2\cos\theta}$

(iii) $r\sin n\theta = a$.

Solution. (i) From the given equation, we get

$$\frac{1}{r} = \frac{\theta}{a} = f(\theta).$$

Therefore, $f(\theta) = 0$ yields $\frac{\theta}{a} = 0$ or $\theta = 0$.
Also

$$f'(\theta) = \frac{1}{a} \text{ and so, } \frac{1}{f'(\theta)} = a.$$

Thus, the asymptotes are given by

$$r\sin(\theta - 0) = \frac{1}{f'(0)} = a \quad\text{or}\quad r\sin\theta = a.$$

(ii) From the given equation, we get

$$\frac{1}{r} = \frac{1-2\cos\theta}{2a} = f(\theta).$$

Therefore, $f(\theta) = 0$ gives $1 - 2\cos\theta = 0$ or $\cos\theta = \frac{1}{2}$ and so, $\theta = 2n\pi \pm \frac{\pi}{3}$, where n is an integer. Further,

$$f'(\theta) = \frac{1}{2a}(2\sin\theta) = \frac{\sin\theta}{a}.$$

This gives

$$f'\left(2n\pi \pm \frac{\pi}{3}\right) = \frac{1}{a}\sin\left(2n\pi \pm \frac{\pi}{3}\right) = \pm\frac{1}{a}\sin\frac{\pi}{3}$$

$$= \pm\frac{\sqrt{3}}{2a}.$$

Hence, the asymptotes are given by

$$r\sin\left[\theta - \left(2n\pi \pm \frac{\pi}{3}\right)\right] = \frac{1}{f'\left(2n\pi \pm \frac{\pi}{3}\right)} = \pm\frac{2a}{\sqrt{3}}$$

or on simplification,

$$r\sin\left(\theta - \frac{\pi}{3}\right) = \frac{2a}{\sqrt{3}} \text{ and } r\sin\left(\theta + \frac{\pi}{3}\right) = -\frac{2a}{\sqrt{3}}.$$

(iii) The equation of the curve may be written as

$$\frac{1}{r} = \frac{\sin n\theta}{a} = f(\theta).$$

Therefore, $f(\theta) = 0$ implies that $\sin n\theta = 0$ and so, $n\theta = m\pi$, where m is an integer. Thus, $\theta = \frac{m\pi}{n}$. Also,

$$f'(\theta) = \frac{n\cos n\theta}{a}$$

and so,

$$f'\left(\frac{m\pi}{n}\right) = \frac{n\cos m\pi}{a}.$$

Hence, the asymptotes are given by

$$r\sin\left(\theta - \frac{m\pi}{n}\right) = \frac{1}{f'\left(\frac{m\pi}{n}\right)} = \frac{a}{n\cos m\pi},$$

where m is an integer.

2.8 CIRCULAR ASYMPTOTES

Let the equation of a curve be $r = f(\theta)$. If $\lim_{\theta\to\infty} f(\theta) = a$, then the circle $r = a$ is called the *circular asymptote* of the curve $r = f(\theta)$

EXAMPLE 2.19

Find the circular asymptotes of the curves

(i) $r(e^\theta - 1) = a(e^\theta + 1)$.

(ii) $r(\theta + \sin\theta) = 2\theta + \cos\theta$.

(iii) $r = \frac{a\theta}{\theta - 1}$.

Solution. (i) The given equation is

$$r(e^{\theta} - 1) = a(e^{\theta} + 1)$$

or

$$r = \frac{a(e^{\theta} - 1)}{e^{\theta} - 1} = f(\theta).$$

Now

$$\lim_{\theta \to \infty} \frac{a(e^{\theta} - 1)}{e^{\theta} - 1} = a \lim_{\theta \to \infty} \frac{1 + e^{-\theta}}{1 - e^{-\theta}} = a.$$

Hence, $r = a$ is the circular asymptote.

(ii) The equation of the given curve is

$$r = \frac{2\theta + \cos\theta}{\theta + \sin\theta} = f(\theta).$$

Further,

$$\lim_{\theta \to \infty} f(\theta) = \lim_{\theta \to \infty} \frac{2\theta + \cos\theta}{\theta + \sin\theta} = \lim_{\theta \to \infty} \frac{2 + \frac{1}{\theta}\cos\theta}{1 + \frac{\sin\theta}{\theta}}$$

$$= \frac{2}{1 + 0} = 2.$$

Hence, $r = 2$ is the required circular asymptote.

(iii) The given equation is

$$r = \frac{a\theta}{\theta - 1}$$

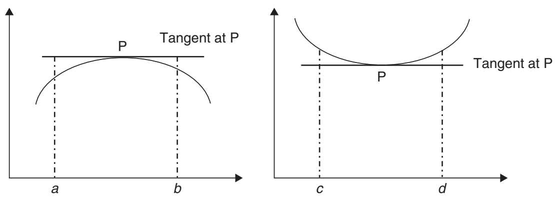
and

$$\lim_{\theta \to \infty} \frac{a\theta}{\theta - 1} = a \lim_{\theta \to \infty} \frac{1}{1 - \frac{1}{\theta}} = a.$$

Hence, $r = a$ is the circular asymptote of the given curve.

2.9 CONCAVITY, CONVEXITY AND SINGULAR POINTS

Consider the curve $y = f(x)$, which is the graph of a single-valued differentiable function in a plane. The curve is said to be *convex upward* or *concave downward* on the interval (a, b) if all points of the curve lie below any tangent to it on this interval. We say that the curve is *convex downward* or *concave upward* on the interval (c, d) if all points of the curve lie above any tangent to it on this interval. Generally, a convex upward curve is called a *convex curve* and a curve convex down is alled a *concave curve*. For example, the curves in figures (a) and (b) are respectively convex and concave curves.



(a) Convex Curve (b) Concave Curve

The following theorems tell us whether the given curve is convex or concave in some given interval.

Theorem 2.1. If at all points of an interval (a, b) the second derivative of the function $f(x)$ is negative, that is, $f''(x) < 0$, then the curve $y = f(x)$ is convex on that interval.

Theorem 2.2. If at all points of an interval (c, d) the second derivative of the function $f(x)$ is positive, that is, $f''(x) > 0$, then the curve $y = f(x)$ is concave on that interval.

A point P on a continuous curve $= f(x)$ is said to be a *point of inflexion* if the curve is convex on one side and concave on the other side of P with respect to any line, not passing through the point P.

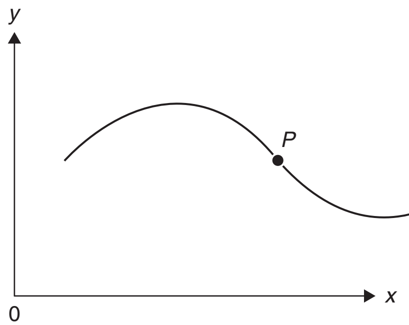
In other words, the point that separates the convex part of a continuous curve from the concave part is called the point of inflexion.

The following theorem gives the sufficient conditions for a given point of a curve to be a point of inflexion.

Theorem 2.3. Let $y = f(x)$ be a continuous curve. If $f''(p) = 0$ or $f''(p)$ does not exist and if the derivative $f''(x)$ changes sign when passing through $x = p$, then the point of the curve with abscissa $x = p$ is the point of inflexion.

Thus at a point of inflexion P, $f''(x)$ is positive on one side of P and negative on the other side. The above theorem implies that at a point of inflexion $f''(x) = 0$ and $f'''(x) \neq 0$.

For example, the point P, in the figure shown below is a point of inflexion.

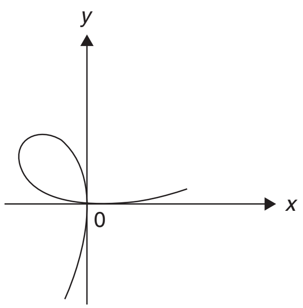


A point through which more than one branches of a curve pass is called a *multiple point* on the curve.

If two branches of curve pass through a point, then that point is called a *double point.* If r branches of a curve pass through a point, then that point is called a multiple *point of order r*.

If two branches of a curve through a double point are real and have different tangents, then the double point is called a *node.*

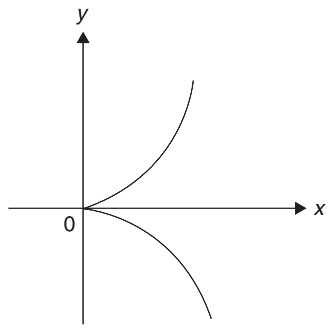
For example, the curve in the figure below has a node at the origin.



(Origin as a Node)

If two branches through a double point P are real and have coincident tangents, then P is called a *cusp*.

For example, the curve in the figure below has a cusp at the origin.



(Origin as a Cusp)

Let $P(x, y)$ be any point on the curve $f(x, y) = 0$. The slope $\frac{dy}{dx}$ of the tangent at P is given by

$$\frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} \quad \text{or} \quad \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dx} = 0,$$

which is a first degree equation in $\frac{dy}{dx}$. Since at a multiple point, the curve must have at least two tangents, therefore $\frac{dy}{dx}$ must have at least two values at a double point. It is possible if and only if

$$\frac{\partial f}{\partial x} = 0 \quad \text{and} \quad \frac{\partial f}{\partial y} = 0.$$

Hence the necessary and sufficient conditions for the existence of multiple points are

$$\frac{\partial f}{\partial x} = 0 \quad \text{and} \quad \frac{\partial f}{\partial y} = 0.$$

EXAMPLE 2.20

Find the points of inflexion of the curve

$$y(a^2 + x^2) = x^3.$$

Solution. The equation of the given curve is

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

Therefore

$$\frac{dy}{dx} = \frac{(a^2 + x^2)3x^2 - 2x^4}{(a^2 + x^2)^2} = \frac{x^4 + 3a^2x^2}{(a^2 + x^2)^2}.$$

Differentiating once more with respect to x, we get

$$\frac{d^2y}{dx^2} = \frac{x(6a^4 + 10a^2x^2 + 4x^4 - 12a^2x^2 - 4x^4)}{(a^2 + x^2)^3}$$

$$= \frac{2xa^2(3a^2 - x^2)}{(a^2 + x^2)^3}.$$

At the point of inflexion, we must have $\frac{d^2y}{dx^2} = 0$ and so

$$\frac{2xa^2(3a^2 - x^2)}{(a^2 + x^2)^3} = 0 \quad \text{and} \quad 2xa^2(3a^2 - x^2) = 0,$$

which yields $x = 0, \ \pm\sqrt{3}a$. Further,

$$\frac{d^3y}{dx^3} = \frac{6a^2(x^4 - 6a^2x^2 + a^4)}{(a^2 + x^2)^4}.$$

If $x = 0$, then $\frac{d^3y}{dx^3} = \frac{6a^6}{a^8} = \frac{6}{a^2} \neq 0$.
If $x = \sqrt{3}a$, then $\frac{d^3y}{dx^3} = -\frac{3}{4}a^2 \neq 0$.
If $x = -\sqrt{3}a$, then $\frac{d^3y}{dx^3} = -\frac{3}{4}a^2 \neq 0$.

Thus all the three values of x corresponds to the points of inflexion.
When $x = 0$, the given equation yields $y = 0$.
When $x = \sqrt{3}a$, the given equation yields $y = \frac{3\sqrt{3}a}{4}$
When $x = -\sqrt{3}a$, the given equation yields $y = -\frac{3\sqrt{3}}{4}a$.
Hence the points of inflexion of the given curve are

$$(0, 0), \left(\sqrt{3}a\ ,\ \frac{3\sqrt{3}}{4}a\right) \text{ and } \left(-\sqrt{3}a\ ,\ \frac{3\sqrt{3}}{4}a\right).$$

EXAMPLE 2.21

Does the curve $y = x^4$ have points of inflexion?

Solution. The equation of the given curve is $y = x^4$. Differentiating with respect to x, we have

$$\frac{dy}{dx} = 4x^3,\ \frac{d^2y}{dx^2} = 12x^2,\ \frac{d^3y}{dx^3} = 24x.$$

Then for the points of inflexion, we must have

$$\frac{d^2y}{dx^2} = 0, \quad \text{that is,} \quad 12x^2 = 0,$$

which yields $x = 0$. But

for $x < 0, \frac{d^2y}{dx^2} > 0$ and therefore the curve is concave,

for $x > 0, \frac{d^2y}{dx^2} > 0$ and therefore the curve is concave.

Since the second derivative does not change sign passing through $x = 0$, the curve has no points of inflexion.

EXAMPLE 2.22

Find the points of inflexion on the curve $y^2 = x(x + 1)^2$.

Solution. The given curve is symmetrical about x axis and gives

$$y = \pm x^{\frac{1}{2}}(x + 1).$$

So, we can proceed with

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

Then

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

$$\frac{d^2y}{dx^2} = \frac{1}{2}\left[\frac{3x^{\frac{1}{2}} - (3x + 1)\frac{1}{2}x^{-\frac{1}{2}}}{x}\right] = \frac{3x - 1}{4x^{\frac{3}{2}}}.$$

To determine the point of inflexion, we put $\frac{d^2y}{dx^2}$ equal to 0. Therefore $3x - 1 = 0$ or $x = \frac{1}{3}$. Further

$$\frac{d^2y}{dx^3} = \frac{3}{8x^{\frac{5}{2}}}(1 - x) \neq 0 \text{ at } x = \frac{1}{3}.$$

Therefore the curve has point of inflexion corresponding to $x = \frac{1}{3}$. Putting $x = \frac{1}{3}$ in the equation of the curve, we have $y = \pm\frac{4}{3\sqrt{3}}$. Hence the points of inflexion on the curve are

$$\left(\frac{1}{3}, \frac{4}{3\sqrt{3}}\right) \text{ and } \left(\frac{1}{3}, -\frac{4}{3\sqrt{3}}\right).$$

EXAMPLE 2.23

Find the points of inflexion and the intervals of convexity and concavity of the Gaussian curve $y = e^{-x^2}$.

Solution. The equation of the Gaussian curve is $y = e^{-x^2}$. Therefore

$$\frac{dy}{dx} = -2xe^{-x^2}, \qquad \frac{d^2y}{dx^2} = 2e^{-x^2}[2x^2 - 1].$$

For the existence of points of inflexion, we must have $\frac{d^2y}{dx^2} = 0$, which yields $x = \pm\frac{1}{\sqrt{2}}$.

Now, since

$$\text{for } x < \frac{1}{\sqrt{2}}, \quad \text{we have } \frac{d^2y}{dx^2} < 0$$

$$\text{for } x > \frac{1}{\sqrt{2}}, \quad \text{we have } \frac{d^2y}{dx^2} > 0,$$

therefore the point of inflexion exists for $x = \frac{1}{\sqrt{2}}$. Putting $x = \frac{1}{\sqrt{2}}$ in the given equation, $y = e^{-\frac{1}{2}}$. Therefore $\left(\frac{1}{\sqrt{2}}, e^{-\frac{1}{2}}\right)$ is a point of inflexion on the curve.

Also

$$\text{for } x < -\frac{1}{\sqrt{2}}, \quad \text{we have } \frac{d^2y}{dx^2} > 0$$

$$\text{for } x > -\frac{1}{\sqrt{2}}, \quad \text{we have } \frac{d^2y}{dx^2} < 0.$$

Thus another point of inflexion exists for the value $x = -\frac{1}{\sqrt{2}}$. Putting $x = -\frac{1}{\sqrt{2}}$ in the equation of the Gaussian curve, we get $y = e^{-\frac{1}{2}}$. Hence the second point of inflexion is $\left(-\frac{1}{\sqrt{2}}, e^{-\frac{1}{2}}\right)$.

EXAMPLE 2.24

Determine whether the curve $y = e^x$ is concave or convex.

Solution. The given exponential curve is $y = e^x$. Then

$$\frac{dy}{dx} = e^x, \quad \frac{d^2y}{dx^2} = e^x > 0 \text{ for the all values of } x.$$

Hence the curve is everywhere concave.

EXAMPLE 2.25

Determine the existence and nature of the double points on the curve

$$f(x,y) = y^2 - (x-2)^2(x-1) = 0.$$

Solution. We have

$$f(x,y) = y^2 - (x-2)^2(x-1) = 0,$$

$$\frac{\partial f}{\partial x} = -(x-2)(3x-4),$$

$$\frac{\partial f}{\partial y} = 2y.$$

Now for the existence of double points, we must have

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0.$$

Hence

$$(x-2)(3x-4) = 0 \quad \text{and } 2y = 0,$$

which yield

$$x = 2, \frac{4}{3} \quad \text{and } y = 0.$$

Thus the possible double points are (2, 0) and $\left(\frac{4}{3}, 0\right)$. But, only (2, 0) satisfies the equation of the curve. To find the nature of the double point (2, 0), we shift the origin to (2, 0). The equation reduces to

$$y^2 = (x+2-2)^2(x+2-1) = x^2(x+1)$$
$$= x^3 + x^2.$$

Equating to zero the lowest degree term, we get $y^2 - x^2 = 0$, which gives $y = \pm x$ as the tangent at (2, 0). Therefore, at the double point (2, 0), there are two real and district tangents. Hence the double point (2, 0) is a node on the given curve.

EXAMPLE 2.26

Does the curve $x^4 - ax^2y + axy^2 + a^2y^2 = 0$ have a node on the origin?

Solution. Equating to zero the lowest degree term in the equation of the given curve, we have

$$a^2y^2 = 0, \text{which yields } \; y = 0, \; 0.$$

Therefore there are two real and coincident tangents at the origin. Hence the given curve has a cusp or conjugate point at the origin and not a node.

2.10 CURVE TRACING (CARTESIAN EQUATIONS)

The aim of this section is to find the appropriate shape of a curve whose equation is given. We shall examine the following properties of the curves to trace it.

1. Symmetry: (i) If the equation of a curve remains unaltered when y is changed to $-y$, then the curve is *symmetrical about the x-axis*. In other words, if the equation of a curve consists of even powers of y, then the curve is symmetrical about the x-axis. For example, the parabola $y^2 = 4ax$ is symmetrical about the x-axis.

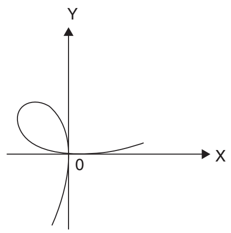
(ii) If the equation of a curve remains unaltered when x is changed to $-x$, then the curve is *symmetrical about the y-axis*. Thus, a curve is symmetrical about the y-axis, if its equation consists of even powers of x. For example, the curve $x^2 + y^2 = a^2$ is symmetrical about the y-axis.

(iii) If the equation of a curve remains unchanged when x is replaced by $-x$ and y is replaced by $-y$, then the curve is *symmetrical in the opposite quadrants*. For example, the curve $xy = c^2$ is symmetrical in the opposite quadrants.

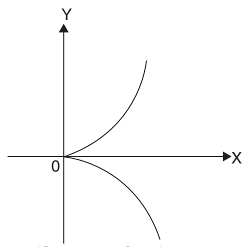
(iv) If the equation of a curve remains unaltered when x and y are interchanged, then the curve is *symmetrical about the line* $y = x$. For example, the folium of Descarte's $x^3 + y^3 = 3axy$ is symmetrical about the line $y = x$.

2. Origin: (i) If the equation of a curve does not contain a constant term, then the curve passes through the origin. In other words, a *curve passes through the origin* if (0, 0) *satisfies the equation of the curve*.

(ii) If the curve passes through the origin, find the *equation of the tangents at the origin by equating to zero the lowest-degree terms* in the equation of the curve. In case there is only one tangent, determine whether the curve lies below or above the tangent in the neighbourhood of the origin. If there are two tangents at the origin, then the *origin is a double point*; if the two tangents are real and distinct, then the origin is a *node*; if the two tangents are real and coincident, then the origin is *cusp*; if the two tangents are imaginary, then the origin is a *conjugate point* or an *isolated point.*



(Origin as a Node)



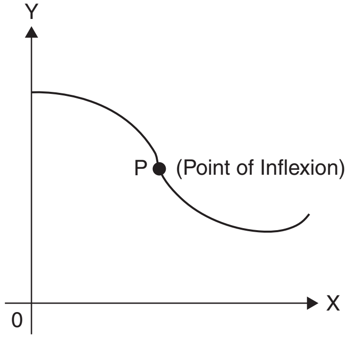
(Origin as a Cusp)

3. Intersection with the Coordinate Axes: To find the points where the curve cuts the coordinate axes, we put $y = 0$ in the equation of the curve to find where the curve cuts the x-axis. Similarly, we put $x = 0$ in the equation to find where the curve cuts the y-axis.

4. Asymptotes: Determine the asymptotes of the curve parallel to the axes and the oblique asymptotes.

5. Sign of the Derivative: Determine the points where the derivative $\frac{dy}{dx}$ vanishes or becomes infinite. This step will yield the points where the tangent is parallel or perpendicular to the x-axis.

6. Points of Inflexion: A point P on a curve is said to be a *point of inflexion* if the curve is concave on one side and convex on the other side of P with respect to any line AB, not passing through the point P.



There will be a point of inflexion at a point P on the curve if $\frac{d^2y}{dx^2} = 0$ but $\frac{d^3y}{dx^3} \neq 0$.

7. **Region, Where the Curve Does Not Exist:** Find out if there is any region of the plane such that no part of the curve lies in it. This is done by solving the given equation for one variable in terms of the other. The curve will not exist for those values of one variable which make the other variable imaginary.

EXAMPLE 2.27

Trace the curve

$$a^2y^2 = x^2(a^2 - x^2).$$

Solution. The equation of the curve is

$$a^2y^2 = x^2(a^2 - x^2).$$

We observe the following:

(i) Since powers of both x and y are even, it follows that the curve is symmetrical about both the axes.

(ii) Since the equation does not contain constant terms, the curve passes through the origin. To find the tangent at the origin, we equate to zero the lowest-degree terms in the given equation. Thus, the tangents at the origin are given by

$$a^2y^2 - a^2x^2 = 0 \text{ or } y = \pm x.$$

Since tangents are distinct, the origin is a node.

(iii) Putting $y = 0$ in the given equation, we get $x = 0$ and $x = \pm a$. Therefore, the curve crosses the x-axis at (0, 0), (a, 0), and ($-a$, 0).

(iv) Shifting the origin to (a, 0), the given equation reduces to

$$a^2y^2 = (x+a)^2[a^2 - (x+a)^2]$$

or

$$a^2y^2 = (x+a)^2(-2ax - x^2).$$

Equating to zero the lowest-degree term, the tangent at the new origin is given by $4a^2x^2 = 0$ or $x = 0$. Thus, the tangent at (a,0) is parallel to the y-axis.

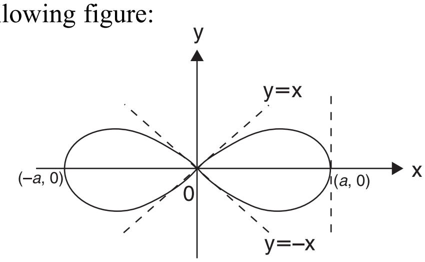
(v) The given equation can be written as

$$y^2 = \frac{x^2(a^2 - x^2)}{a^2}.$$

When $x = 0, y = 0$ and when $x = a, y = a$. When $0 < x < a$, y is real and so, the curve exists in this region. When $x > a$, y^2 is negative and so, y is imaginary. Hence, the curve does not exist in the region $x > a$.

(vi) The given curve has no asymptote.

Hence, the shape of the curve is as shown in the following figure:



EXAMPLE 2.28

Trace the curve

$$xy^2 = 4a^2(2a - x)(\text{Witch of Agnesi}).$$

Solution. We note that

(i) The curve is symmetrical about the x-axis because the equation contains even powers of y.

(ii) Since the equation consists of a constant term, $8a^3$, the curve does not pass through the origin.

(iii) Putting $y = 0$ in the equation, we get $x = 2a$. Therefore, the curve crosses the x-axis at ($2a$, 0). When x = 0, we do not get any value of y. Therefore, the curve does not meet the y-axis.

Shifting *the* origin to ($2a$, 0), the equation of the curve reduces to

$$(x + 2a)y^2 = 4a^2(2a - x - 2a)$$

or

$$y^2x + 2ay^2 + 4a^2x = 0.$$

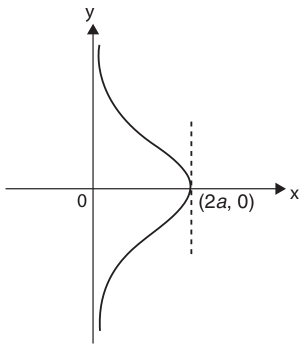
Equating to zero, the lowest-degree terms of this equation, the equation of the tangent at this new origin is given by

$$4a^2x = 0 \text{ or } x = 0.$$

Hence, the tangent at the point ($2a$, 0) to the curve is parallel to y-axis.

(iv) Equating to zero the coefficient of highest power of y, the asymptote parallel to the y-axis is $x = 0$, that is, the y-axis. Further, the curve has no other real asymptote.

(v) The equation of the given curve can be written as
$$y^2 = \frac{4a^2(2a-x)}{x}.$$
Therefore, when $x \to 0$, y approaches ∞ and so, the line $x = 0$ is an asymptote. When $x = 2a$, $y = 0$. When $0 < x < 2a$, the value of y is real and so, the curve exists in the region $0 < x < 2a$. When $x > 2a$, y is imaginary and so, the curve does not exists for $x > 2a$. Similarly, when x is negative, again y is imaginary. Therefore, the curve does not exist for negative x.
In view of the mentioned points, the shape of the curve is as shown in the following figure:



EXAMPLE 2.29

Trace the curve
$$y^2(2a-x) = x^3 \text{ (Cissoid).}$$

Solution. We note that

(i) Since the powers of y in the given equation of the curve are even, the curve is symmetrical about the x-axis.

(ii) Since the equation of the curve does not contain a constant term, the curve passes through the origin. Equating to zero the lowest-degree term in the equation, the tangent at the origin is given by $2ay^2 = 0$. Thus, $y = 0$, $y = 0$ and so, there are two coincident tangents at the origin. Hence, the origin is a *cusp*.

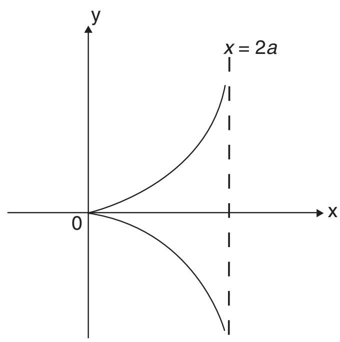
(iii) Putting $x = 0$ in the equation, we get $y = 0$ and similarly, putting $y = 0$, we get $x = 0$. Therefore, the curve meets the coordinate axes only at the origin.

(iv) Equating to zero the highest power of y in the equation of the curve, the asymptote parallel to the y-axis is $x = 2a$. The curve does not have an asymptote parallel to the x-axis or any other oblique asymptote.

(v) The given equation can be written as
$$y^2 = \frac{x^3}{2a-x}.$$

When $x \to 2a$, $y^2 \to \infty$ and so, $x = 2a$ is an asymptote. If $x > 2a$, y is imaginary. Therefore, the curve does not exist beyond $x = 2a$. When $0 < x < 2a$, y^2 is positive and so, y is real. Therefore, the curve exists in the region $0 < x < 2a$. When $x < 0$, again y is imaginary. Therefore, the curve does not exist for a negative x.

In view of the said observations, the shape of the curve is as shown in the following figure:



EXAMPLE 2.30

Trace the curve
$$x^3 + y^3 = 3axy \text{ (Folium of Descartes).}$$
Solution. We observe that

(i) The curve is not symmetrical about the axes. However, the equation of the curve remains unaltered if x and y are interchanged. Hence, the curve is symmetrical about the line $y = x$. It meets this line at $\left(\frac{3a}{2}, \frac{3a}{2}\right)$.

(ii) Since the equation does not contain a constant term, the curve passes through the origin. Equating to zero the lowest-degree term, we get $3axy = 0$. Hence, $x = 0$, $y = 0$ are the tangents at the origin. Thus, both y- and x-axis are tangents to the curve at the origin. Since there are two real and distinct tangents at the origin, the origin is a *node* of the curve.

(iii) The curve intersects the coordinate axes only at the origin.

(iv) If, in the equation of the curve, we take both x and y as negative, then the right-hand side becomes positive while the left-hand side is negative. Therefore, we cannot take both x and y as negative. Thus, the curve does not lie in the third quadrant.

(v) There is no asymptote parallel to the axes. Further, putting $x = 1, y = m$ in the highest-degree term, we have

$$\phi_3(m) = m^3 + 1$$

The slope of the asymptotes are given by $m^3 + 1 = 0$. The real root of this equation is $m = -1$. Also, putting $x = 1, y = m$ in the second-degree term, we have

$$\phi_2(m) = -3am$$

and further,

$$\phi_3'(m) = 3m^2.$$

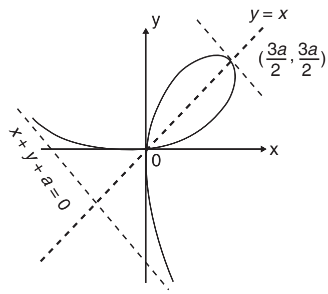
Therefore,

$$c = -\frac{\phi_2(m)}{\phi_3'(m)} = \frac{3am}{3m^2} = \frac{a}{m}.$$

For $m = -1$, we have $c = -a$. Hence, the oblique asymptote is

$$y = -x - a \quad or \quad x + y + a = 0.$$

In view of the earlier facts, the shape of the curve is as shown in the following figure:



EXAMPLE 2.31

Trace the curve

$$y^2(a + x) = x^2(a - x).$$

Solution. We note that

(i) The equation of the curve does not alter if y is changed to $-y$. Therefore, the curve is symmetrical about the x-axis.

(ii) Since the equation does not contain a constant term, the curve passes through the origin. The tangents at the origin are given by
$$ay^2 - ax^2 = 0 \ or \ y = \pm x.$$
Thus, there are two real and distinct tangents at the origin. Therefore, the origin is a *node*.

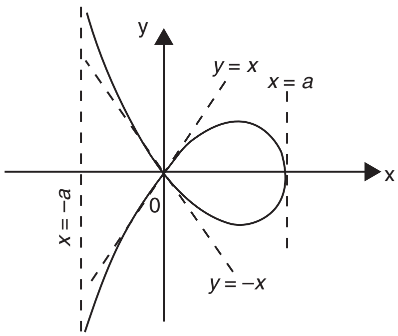
(iii) Putting $y = 0$, we have $x^2(a - x) = 0$ and so, the curve intersects the x-axis at $x = 0$ and $x = a$, that is, at the points (0, 0) and (a, 0). Putting $x = 0$, we get $y = 0$. Thus, the curve intersects the y-axis only at (0, 0). Shifting the origin to (a, 0), the equation of the curve reduces to
$$y^2(2a + x) = -x(x^2 + 2ax + a^2).$$
Equating to zero the lowest-degree term, we get $a^2x = 0$.Hence, at the new origin, $x = 0$ is the tangent. Thus, the tangent at (a, 0) is parallel to the y-axis.

(iv) The equation of the curve can be written as
$$y^2 = \frac{x^2(a - x)}{a + x}.$$
When x lies in $0 < x < a$, y^2 is positive and so, the curve exists in this region. But when $x > a$, y^2 is negative and so, y is imaginary.

Thus, the curve does not exist in the region $x > a$. Further, if $x \to -a$, then $y^2 \to \infty$ and so, $x = -a$ is an asymptote of the curve. If $-a < x < 0$, y^2 is positive and therefore, the curve exists in $-a < x < 0$. When $x < -a$, y^2 is negative and so, the curve does not lie in the region $x < -a$.

(v) Equating to zero the coefficient of the highest power of y in the equation of the curve, we have $x + a = 0$. Thus, $x + a = 0$ is the asymptote parallel to the y-axis. To see whether oblique asymptotes are there or not, we have $\phi_3(m) = m^2 + 1$. But the roots of $m^2 + 1 = 0$ are imaginary. Hence, there is no oblique asymptote.

Thus, the shape of the curve is as shown in the following figure:



EXAMPLE 2.32

Trace the curve

$$x = (y - 1)\ (y - 2)\ (y - 3).$$

Solution. We note that

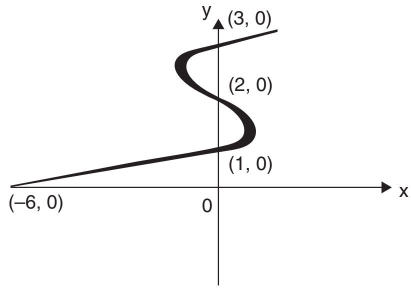
(i) The equation of the curve has odd powers of x and y. Therefore, the curve is not symmetrical about the axes. It is also not symmetrical about $y = x$ or in the opposite quadrants.

(ii) The curve does not pass through the origin.

(iii) Putting $x = 0$ in the given equation, we get $y = 1$, 2, and 3. Thus, the curve cuts the y-axis at $y = 1$, 2, and 3. Similarly, putting $y = 0$, we see that the curve cuts the x-axis at $x = -6$.

(iv) The curve has no linear asymptotes since $y \to \pm\infty$, $x \to \pm\infty$.

(v) When $0 < y < 1$, then all the factors are negative and so, x is negative. When $1 < y < 2$, x is positive. Similarly, when $2 < y < 3$, then x is negative. At $y = 3$, $x = 0$. When $y > 3$, x is positive. When $y < 0$, x is negative. Hence, the shape of the curve is as shown in the following figure:



EXAMPLE 2.33

Trace the curve

$$x^3 + y^3 = a^2x.$$

Solution. We note the following characteristics of the given curve:

(i) Since the equation of the curve contains odd powers of x and y, the curve is not symmetrical about the axes. But if we change the sign of both x and y, then the equation remains unaltered. Therefore, the curve is symmetrical in the opposite quadrants.

(ii) Since the equation of the curve does not have a constant term, the curve passes through the origin. The tangent at the origin is given by $a^2x = 0$. Thus, $x = 0$, that is, y-axis is tangent to the curve at the origin.

(iii) Putting $y = 0$ in the equation, we get $x(x^2 - a^2) = 0$ or $x(x - a)\ (x + a) = 0$. Hence, the curve cuts the x-axis at $x = 0$, $x = a$, and $x = -a$, that is, at the points (0, 0), (a, 0), and ($-a$, 0). On the other hand putting $x = 0$ in the equation, we get $y = 0$. Therefore, the curve cuts the y-axis only at the origin (0, 0).

(iv) The curve does not have any asymptote parallel to the axes. But

$$\phi_3(m) = m^3 + 1, \quad \phi_2(m) = 0.$$

Thus, the slope of the oblique asymptotes is given by $m^3 + 1 = 0$.Thus, the real root is $m = -1$. Also

$$c = -\frac{\phi_2(m)}{\phi_3'(m)} = 0.$$

Therefore, the curve has an oblique asymptote $y = -x$.

(v) From the equation of the curve, we have

$$y^3 = a^2x - x^3.$$

Differentiating with respect to x, we get

$$3y^2\frac{dy}{dx} = a^2 - 3x^2 \text{ or } \frac{dy}{dx} = \frac{a^2 - 3x^2}{3y^2}.$$

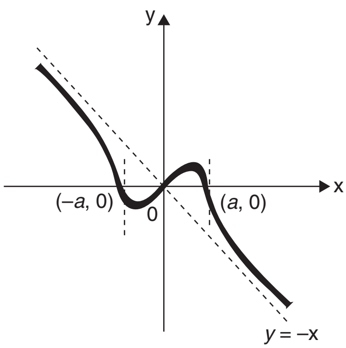
Thus,

$$\left(\frac{dy}{dx}\right)_{(a,0)} = -\infty$$

and so, the tangent at $(a, 0)$ is perpendicular to the x-axis. Similarly, $\left(\frac{dy}{dx}\right)_{(-a,\,0)} = -\infty$ and so, the tangent at $(-a, 0)$ is also perpendicular to the x-axis.

Also we note that $\frac{dy}{dx} = 0$ implies $x = \pm\frac{a}{\sqrt{3}}$. Therefore, the tangents at these points are parallel to the x-axis.

(vi) Also $y^3 = a^2x - x^3 = x(a^2 - x^2)$ implies that y^3 is positive in the region $0 < x < a$. But y^3 is negative in the region $x > a$. The earlier facts imply that the shape of the given curve is as shown in the following figure:



2.11 CURVE TRACING (POLAR EQUATIONS)

To trace a curve with a polar form of equation, we adopt the following procedure:

1. Symmetry: If the equation of the curve does not change when θ is changed into $-\theta$ the curve is *symmetrical about the initial line.*

If the equation of the curve remains unchanged by changing r into $-r$, then *the curve is symmetrical about the pole and the pole is the center of the curve.*

If the equation of the curve remains unchanged when θ is changed to $-\theta$ and r is changed in to $-r$, then the curve is symmetrical about the line $\theta = \frac{\pi}{2}$.

2. Pole: By putting $r = 0$, if we find some real value of θ, then the curve passes through the pole which otherwise not. Further, putting $r = 0$, the real value of θ, if exists, gives the tangent to the curve at the pole.

3. Asymptotes: Find the asymptotes using the method to determine asymptotes of a polar curve.

4. Special Points on the Curve: Solve the equation of the curve for r and find how r varies as θ increases from 0 to ∞ and also as θ decreases from 0 to $-\infty$. Form a table with the corresponding values of r and θ. The points so obtained will help in tracing the curve.

5. Region: Find the region, where the curve does not exist. If r is imaginary in $\alpha < \theta < \beta$, then the curve does not exist in the region bounded by the lines $\theta = \alpha$ and $\theta = \beta$.

6. Value of tanϕ: Find tan ϕ, that is, $r\frac{d\theta}{dr}$, which will indicate the direction of the tangent at any point. If for $\theta = \alpha$, $\phi = 0$ then $\theta = \alpha$ will be tangent to the curve at the point $\theta = \alpha$. On the other hand if for $\theta = \alpha$, $\phi = \frac{\pi}{2}$, then at the point $\theta = \alpha$, the tangent will be perpendicular to the radius vector $\theta = \alpha$.

7. Cartesian Form of the Equation of the Curve: It is useful sometimes to convert the given equation from polar form to cartesian form using the relations $x = r\cos\theta$ and $y = r\sin\theta$.

EXAMPLE 2.34

Trace the curve $r = a \sin 3\theta$.

Solution. We note that

(i) The curve is not symmetrical about the initial line. But if we change θ to $-\theta$ and r to $-r$, then the equation of the curve remains unchanged. Therefore, the curve is symmetrical about the line $\theta = \frac{\pi}{2}$.

(ii) Putting $r = 0$, we get $\sin 3\theta = 0$. Thus, $3\theta = 0, \pi$ or $\theta = 0, \frac{\pi}{3}$. Thus, the curve passes through the pole, and the lines $\theta = 0$ and $\theta = \frac{\pi}{3}$ are tangents to the curve at the pole.

(iii) r is maximum when $\sin 3\theta = 1$ or $3\theta = \frac{\pi}{2}$ or $\theta = \frac{\pi}{6}$. The maximum value of r is a.

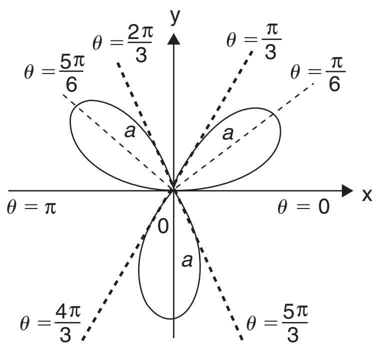
(iv) We have $\frac{dr}{d\theta} = 3a\cos 3\theta$ and so, $\tan\phi = r\frac{d\theta}{dr} = \frac{1}{3}\tan 3\theta$. Thus, $\phi = \frac{\pi}{2}$ when $3\theta = \frac{\pi}{2}$ or $\theta = \frac{\pi}{6}$, and the tangent is perpendicular to the radius vector $\theta = \frac{\pi}{6}$.

(v) Some points on the curve are given below:

$\theta:$	0	$\frac{\pi}{6}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{5\pi}{6}$	π
$r:$	0	a	0	$-a$	0	a	0

One loop of the curve lies in the region $0 < \theta < \frac{\pi}{3}$. The second loop lies in the region $\frac{\pi}{3} < \theta < \frac{2\pi}{3}$ in the opposite direction because r is negative there. The third loop lies in the region $\frac{2\pi}{3} < \theta < \pi$ as r is positive (equal to a) there.

When θ increases from π to 2π, we get again the same branches of the curve. Hence, the shape of the curve is shown in the following figure:



EXAMPLE 2.35

Trace the curve $r = a\,(1-\cos\theta)$ (Cardioid).

Solution. The equation of the given curve is $r = a\,(1-\cos\theta)$. We note the following characteristics of the curve:

(i) The equation of the curve remains unchanged when θ is changed to $-\theta$. Therefore, the curve is symmetrical about the initial line.

(ii) When $r = 0$, we have $1 - \cos\theta = 0$ or $\theta = 0$. Hence, the curve passes through the pole and the line $\theta = 0$ is tangent to the curve at the pole.

(iii) The curve cuts the line $\theta = \pi$ at $(2a, \pi)$.

(iv) $\frac{dr}{d\theta} = a\sin\theta$ and so, $\tan\phi = r\frac{d\theta}{dr} = \frac{r}{a\sin\theta} = \frac{a(1-\cos\theta)}{a\sin\theta} = \tan\frac{\theta}{2}$. If $\frac{\theta}{2} = \frac{\pi}{2}$, then $\phi = 90^{\circ}$. Thus, at the point $\theta = \pi$, the tangent to the curve is perpendicular to the radius vector.

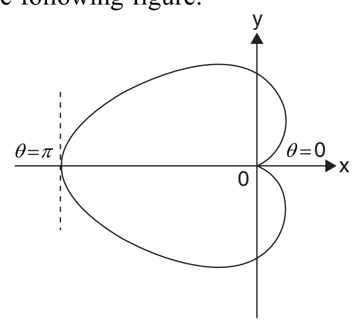
(v) The values of θ and r are:

$\theta:$	0	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	π
$r:$	0	$\frac{a}{2}$	a	$\frac{3a}{2}$	$2a$

We observe that as θ increases from 0 to π, r increases from 0 to $2a$. Further, r is never greater than $2a$. Hence, no portion of the curve lies to the left of the tangent at $(2a, 0)$. Since $|\,r\,| \le 2a$, the curve lies entirely within the circle $r = 2a$.

(vi) There is no asymptote to the curve because for any finite value of θ, r does not tend to infinity.

Hence, the shape of the curve is as shown in the following figure:



EXAMPLE 2.36

Trace the curve $r = a + b\cos\theta$, $a < b$ (Limacon).

Solution. The given curve has the following characteristics:

(i) Since the equation of the curve remains unaltered when θ is changed to $-\theta$, it follows that the curve is symmetrical about the initial line.

(ii) $r = 0$ when $a + b\cos\theta = 0$ or $\theta = \cos^{-1}\left(-\frac{a}{b}\right)$. Since $\frac{a}{b} < 1$, $\cos^{-1}\left(-\frac{a}{b}\right)$ is real. Therefore, the curve passes through the pole and the radius vector $\theta = \cos^{-1}\left(-\frac{a}{b}\right)$ is tangent to the curve at the pole.

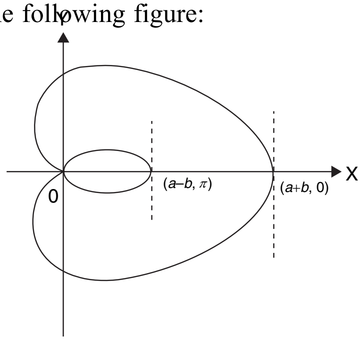
(iii) We note that r is maximum when $\cos\theta = 1$, that is when $\theta = 0$. Thus, the maximum value of r is $a + b$. Thus, the entire curve lies within the circle $r = a + b$. Similarly, r is minimum when $\cos\theta = -1$, that is when $\theta = \pi$. Thus, the minimum value of r is $a - b$, which is negative.

(iv) $\frac{dr}{d\theta} = -b\sin\theta$ and so, $\tan\phi = r\frac{d\theta}{dr} = -\frac{r}{b\sin\theta} = -\frac{a+b\cos\theta}{b\sin\theta}$. Thus, $\phi = 90^\circ$ when $\theta = 0, \pi$. Hence, at the points $\theta = 0$ and $\theta = \pi$, the tangent is perpendicular to the radius vector.

(v) The following table gives the value of r corresponding to the value of θ:

θ: 0	$\frac{\pi}{2}$	$\cos^{-1}\left(-\frac{a}{b}\right)$	$\cos^{-1}\left(-\frac{a}{b}\right) < \theta < \pi$	π
r: $a+b$	a	0	negative	$a-b$

(vi) Since r is not infinite for any value of θ, the given curve has no asymptote:
Hence, the shape of the curve is as shown in the following figure:



EXAMPLE 2.37

Trace the curve $r^2\cos 2\theta = a^2$.

Solution. The equation of the given curve can be written as

$$r^2(\cos^2\theta - \sin^2\theta) = a^2$$

or

$$x^2 - y^2 = a^2 \text{ since } x = r\cos\theta,\ y = r\sin\theta.$$

Therefore, the given curve is a rectangular hyperbola. We note that

(i) The curve is symmetrical about both the axes.

(ii) It does not pass through the origin.

(iii) It cuts the x-axis at $(a, 0)$ and $(-a, 0)$. But it does not meet y-axis.

(iv) Shifting the origin to $(a, 0)$, we get

$$(x + a)^2 - y^2 = a^2 \text{ or } x^2 - y^2 + 2ax = 0.$$

Therefore, the tangent at $(a, 0)$ is given by $2ax = 0$ and so, the tangent at $(a, 0)$ is $x = 0$, the line parallel to the y-axis.

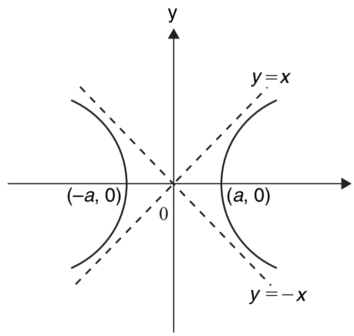
(v) The curve has no asymptote parallel to coordinate axes. The oblique asymptote (verify) are $y = x$ and $y = -x$.

(vi) The equation of the curve can be written as

$$y^2 = x^2 - a^2.$$

When $0 < x < a$, the y^2 is negative and so, y is imaginary. Therefore, the curve does not lie in the region $0 < x < a$. But when $x > a$, y^2 is positive and so, y is real. Thus, the curve exists in the region $x > a$. Further, when $x \to \infty$, $y^2 \to \infty$.

In view of the mentioned facts, the shape of the curve is as shown in the following figure:



2.12 CURVE TRACING (PARAMETRIC EQUATIONS)

If the equation of the curve is given in a parametric form, $x = f(t)$ and $y = \phi(t)$, then eliminate the parameter and obtain a cartesian equation of the curve. Then, trace the curve as dealt with in case of cartesian equations.

In case the parameter is not eliminated easily, a series of values are given to t and the corresponding values of x, y, and $\frac{dy}{dx}$ are found. Then we plot the different points and find the slope of the tangents at these points by the values of $\frac{dy}{dx}$ at the points.

EXAMPLE 2.38

Trace the curve

$$x = a(t + \sin t), \quad y = a(1 + \cos t).$$

Solution. We note that

(i) Since $y = a(1 + \cos t)$ is an even function of t, the curve is symmetrical about the y-axis.

(ii) We have $y = 0$ when $\cos t = -1$, that is when $t = -\pi, \pi$. When $t = \pi$, we have $x = a\pi$. When $t = -\pi$, $x = -a\pi$. Thus, the curve meets the x-axis at $(a\pi, 0)$ and $(-a\pi, 0)$.

(iii) Differentiating the given equation, we get

$$\frac{dx}{dt} = a(1 + \cos t), \quad \frac{dy}{dt} = -a \sin t.$$

Therefore,

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = -\frac{a \sin t}{a(1 + \cos t)}$$

$$= -\frac{2a \sin\frac{t}{2} \cos\frac{t}{2}}{2a \cos^2\frac{t}{2}} = -\tan\frac{t}{2}.$$

Now

$$\left(\frac{dy}{dx}\right)_{t=\pi} = -\tan\frac{\pi}{2} = -\infty.$$

Thus, at the point $(a\pi, 0)$, the tangent to the curve is perpendicular to the x-axis. Similarly, at the point $(-a\pi, 0)$, $\frac{dy}{dx} = \infty$ and hence, at the point $(-a\pi, 0)$, the tangent to the curve is perpendicular to the x-axis.

(iv) y is maximum when $\cos t = 1$, that is, $t = 0$. When $t = 0$ $x = 0$ and $y = 2a$. Thus, the curve cuts the y-axis at $(0, 2a)$. Further,

$$\left(\frac{dy}{dx}\right)_{t=0} = 0$$

and so, at the point (0, 2a), the tangent to the curve is parallel to the x-axis.

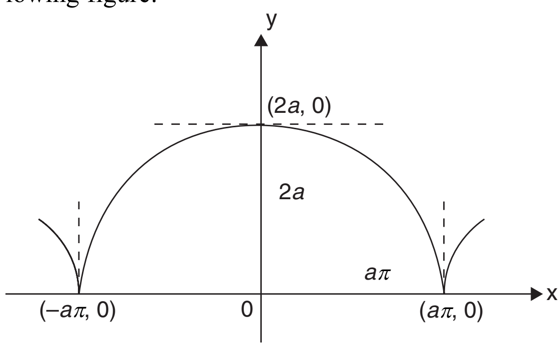
(v) It is clear from the equation that y cannot be negative. Further, no portion of the curve lies in the region $y > 2a$.

(vi) There is no asymptote parallel to the axes.

(vii) The values of x, y corresponding to the values of t are as follows:

t	$-\pi$	$-\frac{\pi}{2}$	0	$\frac{\pi}{2}$	π
x	$-a\pi$	$-a\left(\frac{\pi}{2}+1\right)$	0	$a\left(\frac{\pi}{2}+1\right)$	$a\pi$
y	0	a	$2a$	a	0

Hence, the shape of the curve is as shown in the following figure:



EXAMPLE 2.39

Trace the curve

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

Solution. (i) The parametric equation of the curve are

$$x = a\cos^3 t, \quad y = a\sin^3 t.$$

Therefore,

$$|x| \leq a \text{ and } y \leq a.$$

This implies that the curve lies within the square bounded by the lines $x = \pm a$, $y = \pm a$.

(ii) The equation of the curve can be written as

$$\left(\frac{x^2}{a^2}\right)^{\frac{1}{3}} + \left(\frac{y^2}{a^2}\right)^{\frac{1}{3}} = 1.$$

This equation shows that the curve is symmetrical about both the axes. Also it is symmetrical about the line $y = x$ since interchanging of x and y do not change the equation of the curve.

(iii) The given curve has no asymptote.

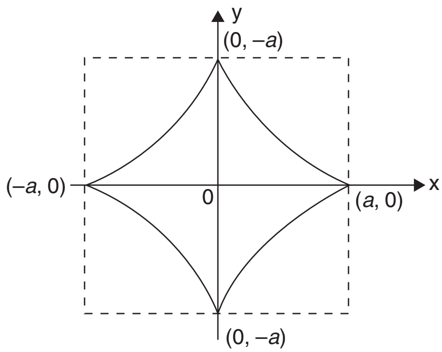
(iv) The curve cuts the x-axis at $(a, 0)$ and $(-a, 0)$. It meets the y-axis at $(0, a)$ and $(0, -a)$. For $x = a$, we have $\cos^3 t = 1$ or $t = 0$. Therefore,

$$\left(\frac{dy}{dx}\right)_{t=0} = \left(\frac{\frac{dy}{dt}}{\frac{dx}{dt}}\right)_{t=0} = (-\tan t)_{t=0} = 0.$$

Hence, at the point $(a, 0)$, the x-axis is the tangent to the curve.

Similarly, at the point $(0, a)$, the y-axis is the tangent to the curve.

Hence, the shape of the curve is as shown in the following figure.



EXERCISES

Find the asymptotes of the following curves:

1. Test the curve $y = x^3$ for concavity/convexity.
   Ans. Concave for $x > 0$ convex for $x < 0$.
2. Find the points of inflexion on the curve $y(a^2 + x^2) = a^2 x$.
   Ans. $(0,\ 0),\ \left(\sqrt{3}a\ ,\ \frac{\sqrt{3}a}{4}\right),\ \left(-\sqrt{3}a\ ,\ -\frac{\sqrt{3}a}{4}\right)$.
3. Show that the points of inflexion on the curve $y^2 = (x-a)^2(x-b)$ lie on the line $3x + a = 4b$.
   Hint: $\frac{d^2y}{dx^2} = 0$ yields $3x + a = 4b$
4. Find the points of inflexion on the curve $x = a(2\theta - \sin\theta),\ y = a(2 - \cos\theta)$.
   Ans. $\left[a\left(4n\pi \pm \frac{2\pi}{3} \mp \frac{\sqrt{3}}{2}\right),\ \frac{3a}{2}\right]$
5. Show that origin is a node on the curve $\frac{a^2}{x^2} - \frac{b^2}{y^2} = 1$.
6. $y^3 + x^2y + 2xy^2 - y + 1 = 0$
   Ans. $y = 0,\ y + x - 1 = 0,\ y + x + 1 = 0$
7. $x^3 + y^3 - 3axy = 0$
   Ans. $x + y + a = 0$
8. $x^3 - 2y^3 + 2x^2y - xy^2 + xy - y^2 + 1 = 0$
   Ans. $x - y = 0,\ x + y + 1 = 0,\ x + 2y - 1 = 0$
9. $3x^3 + 2x^3y - 7xy^2 + 2y^3 - 14xy + 7y^2 + 4x + 5y = 0$
   Ans. $6x - 6y - 7 = 0,\ y = 3x - 1,\ 3x + 6y + 5 = 0$
10. $y^3 = x^3 + ax^2$
    Ans. $y = x + \frac{a}{3}$
11. $y^3 - 3x^2y + xy^2 - 3x^3 + 2y^2 + 2xy + 4x + 5y + 6 = 0$
    Ans. $y = x - 2,\ y = x\sqrt{3} - 1,\ y = -x\sqrt{3} - 1$
12. $x^2y^2 = a^2(x^2 + y^2)$
    Ans. $x = \pm a,\ y = \pm a$
13. $x^2y^3 + x^3y^2 = x^3 + y^3$
    Ans. $x = \pm 1,\ y = \pm 1,\ y = -x$
14. $x^3 + 2x^2y + xy^2 - x^2 - xy + 2 = 0$
    Ans. $x = 0,\ x + y = 0,\ x + y - 1 = 0$
15. $x^3 + x^2y - xy^2 - y^3 - 3x - y - 1 = 0$
    Ans. $y = x,\ y = -x+1,\ x+y+1 = 0$
16. $(x^2 - y^2)(x + 2y + 1) + x + y + 1 = 0$
    Ans. $x + y = 0,\ x - y = 0,\ x + 2y + 1 = 0$
17. (i) $y^2(x-2) = x^2(y-1)$
    Ans. $x = 2,\ y = 1,\ y = x + 1$

    (ii) $x = \frac{a(t+t^3)}{1+t^4},\ y = \frac{a(t-t^3)}{1+t^4}$.
    Hint: Eliminating t, we get $(x^2 + y^2)^2 = a^2(x^2 - y^2)$
    Ans. No asymptote.
18. Show that the asymptotes of the curve
    $$x^2y^2 - a^2(x^2 + y^2) - a^3(x + y) + a^4 = 0$$
    form a square and that the curve passes through two angular points of that square.
    Hint: The four asymptotes are $x = \pm a,\ y = \pm a$. They form a square of length $2a$. The angular points are $(a, a), (a, -a), (-a, a)$, and $(-a, -a)$. The curve passes through two angular points $(a, -a)$ and $(-a, a)$.
19. Show that the points of intersection of the curve
    $$2y^3 - 2xy - 4xy^2 + 4x^3 - 14xy + 6y^2 + 4x^2 + 6y + 1 = 0$$
    and its asymptotes lie on the straight line $8x + 2y + 1 = 0$.
20. Show that the asymptotes of the cubic
    $$x^3 - 2y^3 + xy(2x - y) + y(x - y) + 1 = 0$$
    cut the curve again in three points which lie on the straight line $x - y + 1 = 0$.
    Hint: The asymptotes are $y = x,\ y = -x - 1$, and $y = -\frac{1}{2}x + \frac{1}{2}$. Their joint equation is $x^2 - 2y^3 + 2x^2y - xy^2 + xy - y^2 - x + y = 0$. Subtracting this equation from the equation of the curve, we get $x - y + 1 = 0$.
21. Show that the point of intersection of the curve
    $$4(x^4 + y^4) - 17x^2y^2 - 4x(4y^2 - x^2) + 2(x^2 - 2) = 0$$
    and its asymptotes lie on the ellipse $x^2 + 4y^2 = 4$.

22. Find the equation of the hyperbola passing through the origin and having asymptotes $x + y - 1 = 0$ and $x - y + 2 = 0$.

 Hint: The joint equation of asymptotes is $F_2 = (x - y - 1)(x - y + 2) = 0$. Equation of the curve is $F_n + F_{n-2} = 0$, that is, $F_2 + F_0 = 0$. Thus, F_0 is of a zero degree and so, is a constant. Thus, the equation of the curve is $(x + y - 1)(x - y + 2) + k = 0$. It passes through the origin. So $k = 2$. Hence, the curve is $(x + y - 1)(x - y + 2) + 2 = 0$ or $x^2 - y^2 + x + 3y = 0$.

23. Find the asymptotes of the curve $xy(x^2 - y^2) + x^2 + y^2 = a^2$ and show that the eight points of intersection of the curve and its asymptotes lie on a circle with the origin at the center.

24. Find the equation of the cubic which has the same asymptotes as the curve $x^3 - 6x^2y + 11xy^3 + 4x + 5y + 7 = 0$ and which passes through the points (0, 0), (0,2), and (2, 0).

 Ans. The joint equation of the asymptote is $(x - y)(x - 2y)(x - 3y) = 0$.
 The cubic is $x^3 - 6x^2y + 11xy^2 + - 6y^3 - 4x + 24y = 0$.

25. Find the equation of the straight line on which lie the three points of intersection of the curve $x^3 + 2x^2y - xy^2 - 2y^3 + 4y^2 + 2xy + y - 1 = 0$ and its asymptotes.

 Ans. $x + 3y = 1$.

26. Find the asymptotes of the following polar curves:

 (i) $r\theta \cos\theta = a \cos 2\theta$

 Ans. $r \cos\theta = \frac{2a}{(2k+1)\pi}$

 (ii) $r = a \operatorname{cosec} \theta + b$

 Ans. $r \sin \theta = a$

 (iii) $r = \frac{a}{\log \theta}$

 Ans. $r \sin(\theta - 1) = a$

 (iv) $r = a \sec \theta + b \tan \theta$

 Ans. $r \cos \theta = a + b$, $r \cos \theta = a - b$

 (v) $r(1 - e^{\theta}) = a$

 Ans. $r \sin \theta = -a$

 (vi) $r = \frac{2a\theta}{\theta^2 - \pi^2}$

 Ans. $r \sin \theta = -a$

27. Find the circular asymptote of the curves:

 (i) $r = \frac{3\theta^2 + 2\theta + 1}{2\theta^2 + \theta + 1}$

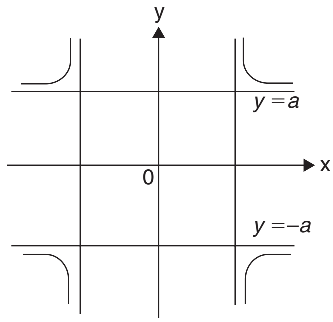
 Ans. $r = \frac{3}{2}$

 (ii) $r = \frac{6\theta^2 + 5\theta - 1}{2\theta^2 - 3\theta + 7}$

 Ans. $r = 3$

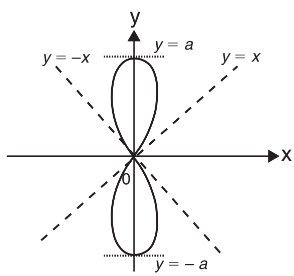
28. Trace the curve $x^2y^2 = a^2(x^2 + y^2)$

Ans.



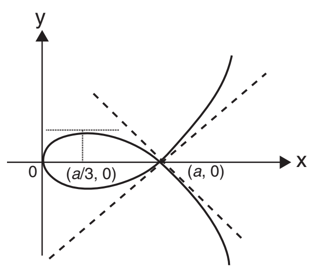
29. Trace the curve $y^2(x^2 + y^2) + a^2(x^2 - y^2) = 0$

Ans.



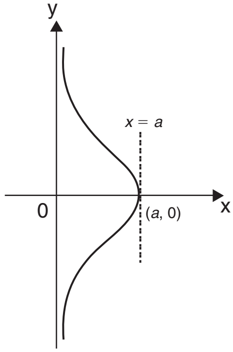
30. Trace the curve $ay^2 = x(x - a)^2$

Ans.



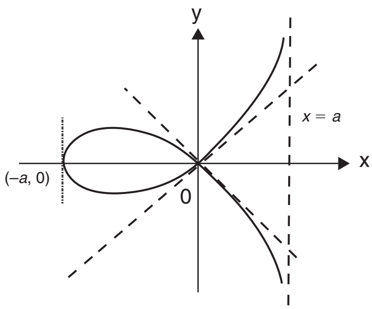
31. Trace the curve $xy = a^2(a - x)$

Ans.



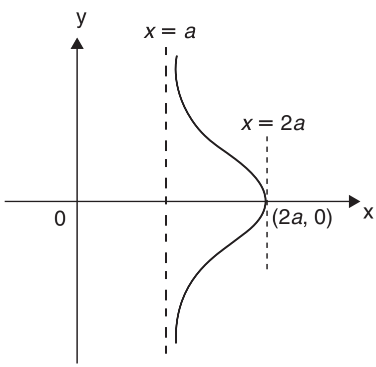
32. Trace the curve $y^2(a - x) = x^2(a + x)$

Ans.



33. Trace the curve $r = a(\cos\theta + \sec\theta)$
Hint: $r^2 = ar(\cos\theta + \sec\theta)$. Therefore, cartesian form is $y^2(x - a) = x^2(2a - x)$

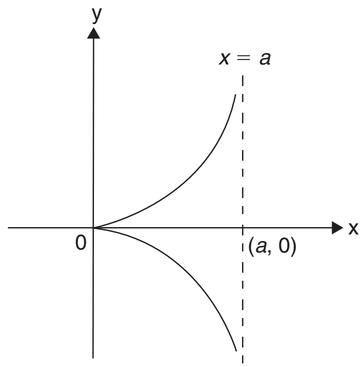
Ans.



34. Trace the curve $r = \frac{a\sin^2\theta}{\cos\theta}$.

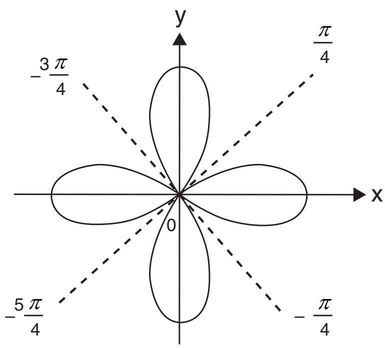
Hint: $r^2 = \frac{ar^2\sin^2\theta}{r\cos\theta} = \frac{ay^2}{x}$ or $x^2 + y^2 = \frac{ay^2}{x}$ or $y^2(a - x) = x^3$

Ans.



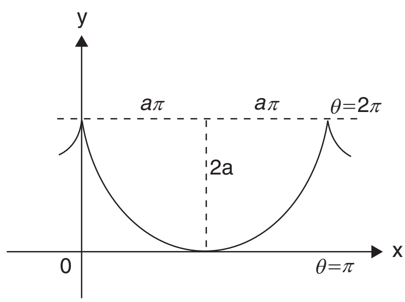
35. Trace the curve $r = a\cos 2\theta$.

Ans.



36. Trace the curve $x = a\,(\theta - \sin\theta)$, $y = a\,(1 + \cos\theta)$.

Ans.



3 Partial Differentiation

Let n be a positive integer and $\Re$ be the set of real numbers. Then, $\Re^n$ is the set of all n-tuples $(x_1, x_2, \ldots, x_n)$, $x_n \in \Re$. Thus,

$\Re$ is the set of all real numbers called, the *real line*,

$\Re^2 = \Re \times \Re = \{(x, y): x, y \in \Re\}$ is a *two-dimensional Cartesian plane*,

$\Re^3 = \Re \times \Re \times \Re = \{(x, y, z): x, y, z \in \Re\}$ is a *three-dimensional Euclidean space*.

Let A be a nonempty subset of $\Re^n$. Then, a function $f: A \to \Re$ is called a *real-valued function of* n *variables* defined on the set A. Thus, f maps $(x_1, x_2, \ldots, x_n)$, $x_i \in \Re$ into a unique real number $f(x_1, x_2, \ldots, x_n)$.

A function f of n variables $x_1, x_2, \ldots, x_n$ is said to tend to a limit λ as $(x_1, x_2, \ldots x_n) \to (a_n, a_2, \ldots a_n)$ if given $\varepsilon > 0$, however small, there exists a real number $\delta > 0$ such that

$$|f(x_1, x_2, \ldots, x_n) - \lambda| < \varepsilon \quad \text{whenever} \quad |(x_1, x_2, \ldots, x_n) - (a_1, a_2, \ldots, a_n)| < \delta$$

or

$$|f(x_1, x_2, \ldots, x_n) - \lambda| < \varepsilon,$$

whenever

$$0 < \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + \ldots + (x_n - a_n)^2} < \delta.$$

In this case, we write

$$\lim_{(x_1, x_2, \ldots, x_n) \to (a_1, a_2, \ldots a_n)} f(x_1, x_2, \ldots, x_n) = \lambda.$$

In what follows, we shall generally deal with functions of two variables.

3.1 CONTINUITY OF A FUNCTION OF TWO VARIABLES

A function $f(x,y)$ is said to be *continuous* at the point (a, b) of its domain if for every $\varepsilon > 0$ there exists a positive number δ such that

$$|f(x, y) - f(a, b)| < \varepsilon \quad \text{whenever} \quad |x - a| < \delta, \quad |y - b| < \delta.$$

Thus, $f(x, y)$ is continuous at a point (a, b)if

$$\lim_{(x,y) \to (a,b)} f(x, y) = f(a, b).$$

3.2 DIFFERENTIABILITY OF A FUNCTION OF TWO VARIABLES

Let $u(x, y)$ be a function of two variables and let Δx and Δy be the increments given to x and y, respectively. A function $u(x, y)$ is said to be *differentiable* at the point (x, y) if it possesses a determinate value in the neighborhood of this point and if

$$\Delta u = A\Delta x + B\Delta y + \varepsilon\rho,$$

where $\rho = |\Delta x| + |\Delta y|$, $\varepsilon \to 0$ as $\rho \to 0$ and A, B are independent of Δx and Δy.

In the preceding definition, ρ may be replaced by η, where $\eta = \sqrt{(\Delta x)^2 + (\Delta y)^2}$.

If the increment ratio

$$\frac{u(x + \Delta x, y) - u(x, y)}{\Delta x}$$

tends to a unique limit as $\Delta x \to 0$, then that limit is called the *partial derivative of u with respect to x* and is written as $\frac{\partial u}{\partial x}$ or u_x.

Similarly $\frac{\partial u}{\partial y}$ can be defined. Thus, if a derivative of a function of several independent variables is found with respect to any one of the independent variables, treating the other as constant, it is said to be a partial derivative of the function with respect to that variable.

3.3 THE DIFFERENTIAL COEFFICIENTS

If in the equation

$$\Delta u = A\ \Delta x + B\ \Delta y + \varepsilon\rho,$$

we suppose that $\Delta y = 0$, then on the assumption that u is differentiable at the point (x, y), we have

$$\Delta u = u(x + \Delta x, y) - u(x, y) = A\ \Delta x + \varepsilon|\Delta x|.$$

Division by Δx yields

$$\frac{u(x+\Delta x,y)-u(x,y)}{\Delta x}=A\pm\varepsilon.$$

Taking the limit as $\Delta x\to 0$, we get

$$\frac{\partial u}{\partial x}=A, \text{since } \varepsilon\to 0 \text{ as } \Delta x\to 0.$$

Similarly, we can show that $\frac{\partial u}{\partial y}=B$. Thus, if $u=u(x, y)$ is differentiable, the partial derivatives $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ are respectively the differential coefficients A and B. Hence,

$$\Delta u=\frac{\partial u}{\partial x}\Delta x+\frac{\partial u}{\partial y}\Delta y+\varepsilon\rho.$$

The differential of the dependant variable du is defined to be the principal part of Δu. Hence,

$$\Delta u=du+\varepsilon\rho.$$

Now as in the case of functions of one variable, the differentials of the independent variables are identical with the arbitrary increments of these variables, that is, $dx=\Delta x$, $dy=\Delta y$. Therefore, $du=\frac{\partial u}{\partial x}dx+\frac{\partial u}{\partial y}dy.$

3.4 DISTINCTION BETWEEN DERIVATIVES AND DIFFERENTIAL COEFFICIENTS

We know that the necessary and sufficient condition for a function $y=f(x)$ to be differentiable at a point x is that it possesses a finite derivative at that point. Thus, for functions of one variable, the existence of derivative $f'(x)$ implies the differentiability of f at that point. But for a function of more than one variable this is not true. We have seen earlier that if $f(x, y)$ is differentiable at (x, y), then the partial derivative of f with respect to x and y exist and are equal to the differential coefficients A and B, respectively. However, the partial derivative may exist at a point when the function is not differentiable at that point. In other words, the *partial derivatives need not always be the differential coefficients.*

EXAMPLE 3.1

Show that the function $f(x,y)=\frac{x^3-y^3}{x^2+y^2}$, where x and y are not simultaneously zero and $f(0, 0)=0$, is not differentiable at (0, 0) but the partial derivatives at (0, 0) exist.

Solution. Suppose that the given function is differentiable at the origin. Then, by definition,

$$f(h,k)-f(0,0)=Ah+Bk+\varepsilon\eta, \qquad (1)$$

where $\eta=\sqrt{h^2+k^2}$ and $\varepsilon\to 0$ as $\eta\to 0$. Putting $h=\eta\cos\theta, k=\eta\sin\theta$ in (1) and dividing throughout by η, we get

$$\cos^3\theta-\sin^3\theta=A\cos\theta+B\sin\theta+\varepsilon.$$

Since $\varepsilon\to 0$ as $\eta\to 0$, we take the limit as $\eta\to 0$ and get

$$\cos^2\theta-\sin^3\theta=A\ \cos\theta+b\ \sin\theta,$$

which is impossible since θ is arbitrary. Hence, the function is *not differentiable* at (0,0). On the other hand,

$$f_x(0,0)=\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{h-0}{h}=1 \text{ and}$$

$$f_y(0,0)=\lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k}=\lim_{k\to 0}\frac{-k-0}{k}=-1.$$

Hence, the partial derivatives at (0,0) exist.

3.5 HIGHER-ORDER PARTIAL DERIVATIVES

Partial derivatives are also, in general, functions of x and y which may possess partial derivatives with respect to both independent variables. Thus,

(i) $\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)=\lim_{\Delta x\to 0}\frac{u_x(x+\Delta x,y)-u_x(x,y)}{\Delta x}$ and

(ii) $\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right)=\lim_{\Delta y\to 0}\frac{u_x(x,y+\Delta y)-u_x(x,y)}{\Delta y}$,

provided that each of these limits exist. The second-order partial derivatives are denoted by $\frac{\partial^2 u}{\partial x^2}$ or u_{xx} and $\frac{\partial^2 u}{\partial y\partial x}$ or u_{yx}. Similarly, we may define higher-order partial derivatives of $\frac{\partial u}{\partial y}$.

EXAMPLE 3.2

If

$$f(x,y)=\frac{xy(x^2-y^2)}{x^2+y^2},\ f(0,0)=0,$$

show that $f_{xy}(0,0)\neq f_{yx}(0,0)$.

Solution. When (x,y) is not the origin, then

$$\frac{\partial f}{\partial x}=y\left[\frac{x^2-y^2}{x^2+y^2}+\frac{4x^2y^2}{(x^2+y^2)^2}\right], \qquad (1)$$

$$\frac{\partial f}{\partial y}=x\left[\frac{x^2-y^2}{x^2+y^2}+\frac{4x^2y^2}{(x^2+y^2)^2}\right]. \qquad (2)$$

On the other hand,

$$f_x(0,0) = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = 0,$$

$$f_y(0,0) = \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = 0.$$

From (1) and (2), we have

$$f_x(0,y) = -y\ (y \neq 0)\ \textit{and}\ f_y(x,0) = x\ (x \neq 0).$$

Therefore,

$$f_{xy}(0,0) = \lim_{h\to 0}\frac{f_y(h,0)-f_y(0,0)}{h} = \lim_{h\to 0}\frac{h}{h} = 1.$$

and

$$f_{yx}(0,0) = \lim_{k\to 0}\frac{f_x(0,k)-f_x(0,0)}{k} = \lim_{k\to 0}\frac{-k}{k} = -1.$$

Hence,

$$f_{xy}\ (0,\ 0) \neq f_{xy}\ (0,\ 0).$$

The question arises: Under what conditions, $f_{xy}(a, b) = f_{yx}(a, b)$?. The following theorems answer this question:

Theorem 3.1. (Young). If (i) f_x and f_y exist in the neighborhood of the point (a, b) and (ii) f_x and f_y are differentiable at (a, b), then $f_{xy}(a, b) = f_{yx}(a, b)$.

Theorem 3.2. (Schwarz). If (i) f_x, f_y, and f_{yx} all exist in the neighborhood of the point (a, b) and (ii) f_{yx} is continuous at (a, b), then f_{xy} also exists at (a, b) and $f_{xy}(a, b) = f_{yx}(a, b)$.

In this chapter, we shall assume in the examples involving f_{xy} and f_{yx} that the partial derivatives of the first two orders of the given function are continuous so that $f_{xy} = f_{yx}$.

EXAMPLE 3.3

If $z = x\log y$, show that

$$\frac{\partial^2 z}{\partial x\partial y} = \frac{\partial^2 z}{\partial y\partial x}.$$

Solution. The given function is $z = x\log y$. Differentiating z partially with respect to x, taking y as constant, we get $\frac{\partial z}{\partial x} = \log y$ and then,

$$\frac{\partial^2 z}{\partial y\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial y}(\log y) = \frac{1}{y}.$$

Differentiating z with respect to y, taking x as constant, we get $\frac{\partial z}{\partial y} = \frac{x}{y}$ and so,

$$\frac{\partial^2 z}{\partial x\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial x}\left(\frac{x}{y}\right) = \frac{1}{y}.$$

Hence,

$$\frac{\partial^2 z}{\partial y\partial x} = \frac{\partial^2 z}{\partial x\partial y} = \frac{1}{y}.$$

EXAMPLE 3.4

If $\frac{1}{u} = \sqrt{x^2+y^2+z^2}$, show that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0.$$

Solution. We are given that

$$u = (x^2+y^2+z^2)^{-\frac{1}{2}}.$$

Therefore,

$$\frac{\partial u}{\partial x} = -\ \frac{1}{2}(x^2+y^2+z^2)^{-\frac{3}{2}}(2x)$$
$$= -x(x^2+y^2+z^2)^{-\frac{3}{2}}.$$

Differentiating partially once again with respect to x, we get

$$\frac{\partial^2 u}{\partial x^2} = \frac{3x}{2}(x^2+y^2+z^2)^{-\frac{5}{2}}(2x) - (x^2+y^2+z^2)^{-\frac{3}{2}}$$
$$= (x^2+y^2+z^2)^{-\frac{5}{2}}[3x^2 - (x^2+y^2+z^2)]$$
$$= u^5(2x^2-y^2-z^2).$$

Similarly, by symmetry in the given function, we have

$$\frac{\partial^2 u}{\partial y^2} = u^5(2y^2-z^2-x^2),$$

$$\frac{\partial^2 u}{\partial z^2} = u^5(2z^2-x^2-y^2).$$

Adding, we get

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = u^5(2x^2-y^2-z^2+2y^2$$
$$-z^2-x^2+2z^2-x^2-y^2)$$
$$= u^5(0) = 0$$

EXAMPLE 3.5

If $x = r\cos\theta$, $y = r\sin\theta$, show that

$$\frac{\partial^2\theta}{\partial x^2} + \frac{\partial^2\theta}{\partial y^2} = 0,\ x, y \neq 0.$$

Solution. We have

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

Therefore,

$$r^2 = x^2+y^2 \text{ and } \theta = \tan^{-1}\frac{y}{x}.$$

Differentiating partially with respect to x, we get

$$\frac{\partial\theta}{\partial x} = \frac{\partial}{\partial x}\left(\tan^{-1}\frac{y}{x}\right) = \frac{1}{1+\left(\frac{y}{x}\right)^2}\left(-\frac{y}{x^2}\right) = -\frac{y}{x^2+y^2}$$

and

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

Similarly differentiating partially with respect to y, we get

$$\frac{\partial\theta}{\partial y} = \frac{\partial}{\partial y}\left(\tan^{-1}\frac{y}{x}\right) = \frac{1}{1+\left(\frac{y}{x}\right)^2}\left(\frac{1}{x}\right)$$

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

and

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

Hence,

$$\frac{\partial^2\theta}{\partial x^2} + \frac{\partial^2 y}{\partial y} = 0.$$

The result holds for $x \neq 0$, $y \neq 0$, otherwise $\frac{\partial^2\theta}{\partial x^2}$ and $\frac{\partial^2\theta}{\partial y^2}$ are of the form $\left(\frac{0}{0}\right)$.

EXAMPLE 3.6

If $u = \sin^{-1}\frac{x}{y} + \tan^{-1}\frac{y}{x}$, show that

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 0.$$

Solution. We have

$$u = \sin^{-1}\frac{x}{y} + \tan^{-1}\frac{y}{x}.$$

Therefore, differentiating partially with respect to x, we get

$$\frac{\partial u}{\partial x} = \frac{1}{\sqrt{1-\left(\frac{x}{y}\right)^2}}\cdot\frac{1}{y} + \frac{1}{1+\left(\frac{y}{x}\right)^2}\left(-\frac{y}{x^2}\right)$$

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

and so,

$$x\frac{\partial y}{\partial x} = \frac{x}{\sqrt{y^2-x^2}} - \frac{xy}{x^2+y^2}. \qquad (1)$$

Now differentiating partially with respect to y, we get

$$\frac{\partial u}{\partial y} = \frac{1}{\sqrt{1-\left(\frac{x}{y}\right)^2}}\left(\frac{-x}{y^2}\right) + \frac{1}{1+\left(\frac{y}{x}\right)^2}\left(\frac{1}{x}\right)$$

$$= \frac{-x}{y\sqrt{y^2-x^2}} + \frac{x}{x^2+y^2}$$

and so,

$$y.\frac{\partial u}{\partial y} = \frac{-x}{\sqrt{y^2-x^2}} + \frac{xy}{x^2+y^2}. \qquad (2)$$

Adding (1) and (2), we have

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 0.$$

EXAMPLE 3.7

If $u = x^2\tan^{-1}\frac{y}{x} - y^2\tan^{-1}\frac{x}{y}$, show that

$$\frac{\partial^2 u}{\partial x\partial y} = \frac{x^2-y^2}{x^2+y^2}.$$

Solution. Differentiating partially the given function with respect to y, we get

$$\frac{\partial u}{\partial y} = x^2\left[\frac{1}{1+\left(\frac{y}{x}\right)^2}\left(\frac{1}{x}\right)\right] - 2y\tan^{-1}\frac{x}{y}\left[\frac{1}{1+\left(\frac{x}{y}\right)^2}\left(\frac{-x}{y^2}\right)\right]$$

$$= \frac{x^3}{x^2+y^2} - 2y\tan^{-1}\frac{x}{y} + \frac{xy^2}{x^2+y^2}$$

$$= x - 2y\tan^{-1}\frac{x}{y}.$$

Now differentiating partially with respect to x, we get

$$\frac{\partial^2 u}{\partial x\partial y} = 1 - 2y\left[\frac{1}{1+\left(\frac{x}{y}\right)^2}\left(\frac{1}{y}\right)\right]$$

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

EXAMPLE 3.8

If $u = \log(x^3+y^3+z^3-3xyz)$, show that

$$\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)^2 u = -\frac{9}{(x+y+z)^2}.$$

Solution. We are given that

$$u = \log(x^3+y^3+z^3-3xyz).$$

Differentiating partially with respect to x, y, and z, we get

$$\frac{\partial u}{\partial x} = \frac{3x^2-3yz}{x^3+y^3+z^3-3xyz},$$

$$\frac{\partial u}{\partial y} = \frac{3y^2-3xz}{x^3+y^3+z^3-3xyz}, \text{ and}$$

$$\frac{\partial u}{\partial z} = \frac{3z^2-3xy}{x^3+y^3+z^3-3xyz}.$$

Therefore,

$$\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)^2 u$$

$$= \left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)u$$

$$=\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)\left(\frac{\partial u}{\partial x}+\frac{\partial u}{\partial y}+\frac{\partial u}{\partial z}\right)$$
$$=\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)\left[\frac{3(x^2+y^2+z^2-yz-zx-xy)}{x^3+y^3+z^3-3xyz}\right]$$
$$=\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)\left(\frac{3}{x+y+z}\right)$$
$$=3\left[\frac{\partial}{\partial x}\left(\frac{1}{x+y+z}\right)+\frac{\partial}{\partial y}\left(\frac{1}{x+y+z}\right)+\frac{\partial}{\partial z}\left(\frac{1}{x+y+z}\right)\right]$$
$$=3\left[\frac{-1}{(x+y+z)^2}-\frac{1}{(x+y+z)^2}-\frac{1}{(x+y+z)^2}\right]$$
$$=\frac{-9}{(x+y+z)^2}.$$

EXAMPLE 3.9

If $u=f(r)$, where $r^2=x^2+y^2$, show that
$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=f''(r)+\frac{1}{r}f'(r).$$

Solution. We have $r^2=x^2+y^2$. Therefore, partial differentiation with respect to x and y yields
$$2r\frac{\partial r}{\partial x}=2x \text{ and } 2r\frac{\partial r}{\partial y}=2y$$
or
$$\frac{\partial r}{\partial x}=\frac{x}{r} \text{ and } \frac{\partial r}{\partial y}=\frac{y}{r}.$$
But $u=f(r)$. Therefore,
$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial r}\cdot\frac{\partial r}{\partial x}=f'(r)\frac{\partial r}{\partial x}=\frac{x}{r}f'(r),$$
$$\frac{\partial^2 u}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{x}{r}f'(r)\right)=\frac{1}{r}f'(r)$$
$$+xf'(r)\left(-\frac{1}{r^2}\frac{\partial r}{\partial x}\right)+\frac{x}{r}f''(r)\frac{\partial r}{\partial x}$$
$$=\frac{1}{r}f'(r)-\frac{x^2}{r^3}f'(r)+\frac{x^2}{r^2}f''(r). \qquad (1)$$
Similarly, due to symmetry, we have
$$\frac{\partial^2 u}{\partial y^2}=\frac{1}{r}f'(r)-\frac{y^2}{r^3}f'(r)+\frac{y^2}{r^2}f''(r). \qquad (2)$$
Adding (1) and (2), we get
$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=\frac{2}{r}f'(r)-\frac{x^2+y^2}{r^3}f'(r)+\frac{x^2+y^2}{r^2}f''(r)$$
$$=\frac{2}{r}f'(r)-\frac{1}{r}f'(r)+f''(r)$$
$$=\frac{1}{r}f'(r)+f''(r).$$

EXAMPLE 3.10

If $z=f(x+ct)+\phi(x–ct)$, show that
$$\frac{\partial^2 z}{\partial t^2}=c^2\frac{\partial^2 z}{\partial x^2}.$$

Solution. We are given that
$$z=f(x+ct)+\phi(x-ct).$$
Therefore,
$$\frac{\partial z}{\partial x}=f'(x+ct).\frac{\partial}{\partial x}(x+xt)+\phi'(x-ct)\frac{\partial}{\partial x}(x-ct)$$
$$=f'(x+ct)+\phi'(x-ct)$$
and
$$\frac{\partial^2 z}{\partial x^2}=f''(x+ct)+\phi''(x-ct).$$
On the other hand,
$$\frac{\partial z}{\partial t}=f'(x+ct).\frac{\partial}{\partial t}(x+ct)+\phi'(x-ct)\frac{\partial}{\partial t}(x-ct)$$
$$=cf'(x+ct)-c\phi'(x-ct)$$
and so,
$$\frac{\partial^2 z}{\partial t^2}=c^2f''(x+ct)+c^2\phi''(x-ct)$$
$$=c^2[f''(x+ct)+\phi''(x-ct)]=c^2\frac{\partial^2 z}{\partial x^2}.$$

EXAMPLE 3.11

If $x=r\cos\theta$, $y=\sin\theta$, show that
$$\frac{\partial^2 r}{\partial x^2}+\frac{\partial^2 r}{\partial y^2}=\frac{1}{r}\left[\left(\frac{\partial r}{\partial x}\right)^2+\left(\frac{\partial r}{\partial y}\right)^2\right].$$

Solution. If $x=r\cos\theta$, $y=\sin\theta$, then $r^2=x^2+y^2$. Therefore, partial differentiation with respect to x yields
$$\frac{\partial r}{\partial x}=\frac{x}{r},\ \frac{\partial^2 r}{\partial x^2}=\frac{r-x\frac{\partial r}{\partial x}}{r^2}=\frac{r^2-x^2}{r^3}$$
$$=\frac{x^2+y^2-x^2}{r^3}=\frac{y^2}{r^3}.$$
Similarly, partial differentiation with respect to y gives
$$\frac{\partial r}{\partial y}=\frac{y}{r},$$
$$\frac{\partial^2 r}{\partial y^2}=\frac{r-y\frac{\partial r}{\partial y}}{r^2}=\frac{r^2-y^2}{r^3}=\frac{x^2+y^2-y^2}{r^3}=\frac{x^2}{r^3}.$$
Therefore,
$$\frac{\partial^2 r}{\partial x^2}+\frac{\partial^2 r}{\partial y^2}=\frac{x^2+y^2}{r^3}=\frac{r^2}{r^3}=\frac{1}{r}.$$

On the other hand,

$$\frac{1}{r}\left[\left(\frac{\partial r}{\partial x}\right)^2+\left(\frac{\partial r}{\partial y}\right)^2\right]=\frac{1}{r}\left[\frac{x^2}{r^2}+\frac{y^2}{r^2}\right]$$

$$=\frac{1}{r}\left(\frac{r^2}{r^2}\right)=\frac{1}{r}.$$

Hence,

$$\frac{\partial^2 r}{\partial x^2}+\frac{\partial^2 r}{\partial y^2}=\frac{1}{r}\left[\left(\frac{\partial r}{\partial x}\right)^2+\left(\frac{\partial r}{\partial y}\right)^2\right].$$

EXAMPLE 3.12

If $x^x\, y^y\, z^z = c$, show that at $x=y=z$,

$$\frac{\partial^2 z}{\partial x \partial y}=-[x\log ex]^{-1}.$$

Solution. We have $x^x\, y^y\, z^z = c$. Here z can be regarded as a function of x and y. Taking logarithm of both sides, we get

$$x\log x+y\log y+z\log z=\log c. \qquad (1)$$

Since z is a function of x and y, differentiating (1) partially with respect to x (taking y as constant), we get

$$x.\frac{1}{x}+\log x+\left[z\frac{1}{z}+\log z\right]\frac{\partial z}{\partial x}=0$$

or

$$1+\log x+(1+\log z)\frac{\partial z}{\partial x}=0.$$

Therefore,

$$\frac{\partial z}{\partial x}=-\frac{1+\log x}{1+\log z} \qquad (2)$$

Similarly, differentiating (1) partially with respect to y we get

$$1+\log y+(1+\log z)\frac{\partial z}{\partial y}=0$$

or

$$\frac{\partial z}{\partial y}=-\frac{1+\log y}{1+\log z}. \qquad (3)$$

Differentiating partially (3) with respect to x, we get

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

$$=-(1+\log y)\frac{\partial}{\partial x}\left(\frac{1}{1+\log z}\right)$$

$$=-(1+\log y)\left[-(1+\log z)^{-2}.\frac{1}{z}.\frac{\partial}{\partial x}\right]$$

$$=\frac{1+\log y}{z(1+\log z)^2}\left[-\frac{1+\log x}{1+\log z}\right], \text{ using (2).}$$

Taking $x=y=z$, we get

$$\frac{\partial^2 z}{\partial x\partial y}=-\frac{(1+\log x)^2}{x(1+\log x)^3}=-\frac{1}{x(1+\log x)}$$

$$=-\frac{1}{x(\log e+\log x)}=-\frac{1}{x\log ex}=-[x\log ex]^{-1}.$$

EXAMPLE 3.13

If $\frac{x^2}{a^2+u}+\frac{y^2}{b^2+u}+\frac{z^2}{c^2+u}=1$, show that

$$\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2=2\left(x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}+z\frac{\partial u}{\partial z}\right).$$

Solution. Here u is a function of three variables x, y, and z. Differentiating the given equation partially with respect to x, we get

$$\frac{(a^2+u)2x-x^2\frac{\partial u}{\partial x}}{(a^2+u)^2}-\frac{y^2}{(b^2+u)}\frac{\partial u}{\partial x}-\frac{z^2}{(c^2+u)^2}.\frac{\partial u}{\partial x}=0$$

or

$$\frac{2x}{a^2+u}-\left[\frac{x^2}{(a^2+u)^2}+\frac{y^2}{(b^2+u)^2}+\frac{z^2}{(c^2+u)^2}\right]\frac{\partial u}{\partial x}=0$$

or

$$\frac{\partial u}{\partial x}=\frac{2x}{(a^2+u)\left[\frac{x^2}{(a^2+u)^2}+\frac{y^2}{(b^2+u)^2}+\frac{z^2}{(c^2+u)^2}\right]}.$$

Thus,

$$\left(\frac{\partial u}{\partial x}\right)^2=\frac{4x^2}{(a^2+u)^2\left[\frac{x^2}{(a^2+u)^2}+\frac{y^2}{(b^2+u)^2}+\frac{z^2}{(c^2+u)^2}\right]^2}$$

$$=\frac{4x^2}{(a^2+u)^2\left[\sum\frac{x^2}{(a^2+u)^2}\right]^2}. \qquad (1)$$

Similarly, due to symmetry,

$$\left(\frac{\partial u}{\partial y}\right)^2=\frac{4y^2}{(b^2+u)^2\left[\sum\frac{y^2}{(b^2+u)^2}\right]^2}, \qquad (2)$$

$$\left(\frac{\partial u}{\partial z}\right)^2=\frac{4z^2}{(c^2+u)^2\left[\sum\frac{z^2}{(c^2+u)^2}\right]^2}. \qquad (3)$$

Adding (1), (2), and (3), we get

$$\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2=\frac{4}{\sum\frac{x^2}{(a^2+u)^2}}. \qquad (4)$$

On the other hand,

$$2x\frac{\partial u}{\partial x}=\frac{4x^2}{(a^2+u)\sum\frac{x^2}{(a^2+u)^2}}.$$

Since $\sum \frac{x^2}{a^2+u} = 1$ (given), we have

$$\sum 2x\frac{\partial u}{\partial x} = \frac{4\sum \frac{x^2}{a^2+u}}{\sum \frac{x^2}{(a^2+u)^2}} = \frac{4}{\sum \frac{x^2}{(a^2+u)^2}}. \qquad (5)$$

From (4) and (5), we get

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 = 2\left(x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z}\right).$$

EXAMPLE 3.14

If $r^2 = x^2 + y^2 + z^2$ and $V = r^m$, show that

$$V_{xx} + V_{yy} + V_{zz} = m(m+1)r^{m-2}.$$

Solution. We have $r^2 = x^2 + y^2 + z^2$. Differentiating partially with respect to x, we get $2r\frac{\partial r}{\partial x} = 2x$ or $\frac{\partial r}{\partial x} = \frac{x}{r}$.

It is given that $V = r^m$. Therefore,

$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial r} \cdot \frac{\partial r}{\partial x} = mr^{m-1}\left(\frac{x}{r}\right) = mxr^{m-2}$$

and

$$\frac{\partial^2 V}{\partial x^2} = m\left[r^{m-2} + x(m-2)r^{m-3}\frac{\partial r}{\partial x}\right]$$
$$= m[r^{m-2} + (m-2)x^2 r^{m-4}].$$

Similarly, due to symmetry, we have

$$\frac{\partial^2 V}{\partial y^2} = m\left[r^{m-2} + (m-2)y^2\, r^{m-4}\right],$$
$$\frac{\partial^2 V}{\partial z^2} = m\left[r^{m-2} + (m-2)z^2 r^{m-4}\right].$$

Hence,

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 3mr^{m-2} + m(m-2)$$
$$\times (x^2 + y^2 + z^2)r^{m-4}$$
$$= 3mr^{m-2} + m(m-2)r^{m-2}$$
$$= r^{m-2}[3m + m^2 - 2m]$$
$$= m(m+1)r^{m-2}.$$

3.6 ENVELOPES AND EVOLUTES

Let α be a parameter which can take all real values and let $f\,(x,\ y,\ \alpha) = 0$ be a family of curves. Suppose that P is a point of intersection of two members $f(x, y, \alpha) = 0$ and $f(x, y, \alpha + \delta\alpha) = 0$ of this family. As $\delta\alpha \to 0$, let P tends to a definite point Q on the member α. The locus of Q (for varying value of α) is called the *envelope of the family*. Thus,

"The envelope of a one parameter family of curves is the locus of the limiting positions of the points of intersection of any two members of the family when one of them tends to coincide with the other which is kept fixed."

The coordinates of the points of intersection of the curves $f\,(x,\ y,\ \alpha) = 0$ and $f\,(x,\ y,\ \alpha + \delta\alpha) = 0$ satisfy the equations

$$f(x,y,\alpha) = 0 \text{ and } f(x,y,\alpha+\delta\alpha) - f(x,y,\alpha) = 0$$

and therefore, they also satisfy

$$f(x,y,\alpha) = 0 \text{ and } \frac{f(x,y,\alpha+\delta\alpha) - f(x,y,\alpha)}{\delta\alpha} = 0.$$

Taking limit as $\delta\alpha \to 0$, it follows that the coordinates of the limiting positions of the point of intersection of the curves $f\,(x, y, \alpha) = 0$ and $f\,(x, y, \alpha + \delta\alpha)$ satisfy the equations

$$f(x,y,\alpha) = 0 \text{ and } \frac{\partial f}{\partial \alpha} = 0.$$

Hence, the equation of the envelope of the family of curves $f(x, y, \alpha) = 0$, where α is a parameter, is determined by eliminating α between the equations $f(x, y, \alpha) = 0$ and $\frac{\partial}{\partial\alpha} f(x,y,\alpha) = 0$.

The *evolute* of a curve is the envelope of the normals to that curve.

EXAMPLE 3.15

Find the envelope of the family of straight lines $y = mx + \frac{a}{m}$, the parameter being m.

Solution. We have

$$y = mx + \frac{a}{m}. \qquad (1)$$

Differentiating with respect to m, we obtain

$$0 = x - \frac{a}{m^2} \text{ or } m = \left(\frac{a}{x}\right)^{\frac{1}{2}}.$$

Putting this value of m in (1), we get

$$y = \left(\frac{a}{x}\right)^{\frac{1}{2}} x + a\left(\frac{a}{x}\right)^{-\frac{1}{2}} = \frac{xa^{\frac{1}{2}}}{x^{\frac{1}{2}}} + \frac{ax^{\frac{1}{2}}}{a^{\frac{1}{2}}} = 2a^{\frac{1}{2}}.x^{\frac{1}{2}},$$

and so, the parabola $y^2 = 4ax$ is the envelope of the family.

EXAMPLE 3.16

Find the envelope of the straight lines $x\cos\alpha + y\sin\alpha = l\sin\alpha\cos\alpha$, where the parameter is the angle α.

Solution. Dividing throughout by sin α cos α we have

$$x \operatorname{cosec} \alpha + y \sec \alpha = l \qquad (1)$$

Differentiating partially with respect to α, we get

$$x(-\cos\alpha\cot\alpha) + y(\sec\alpha\tan\alpha) = 0$$

or

$$\tan\alpha = \frac{x^{\frac{1}{3}}}{y^{\frac{1}{3}}}.$$

This yields

$$\text{coses } \alpha = \sqrt{1+\cot^2\alpha} = \sqrt{1+\frac{y^{\frac{2}{3}}}{x^{\frac{2}{3}}}} = \sqrt{\frac{x^{\frac{2}{3}}+y^{\frac{2}{3}}}{x^{\frac{2}{3}}}}$$

and

$$\sec\alpha = \sqrt{1+\tan^2\alpha} = \sqrt{1+\frac{x^{\frac{2}{3}}}{y^{\frac{2}{3}}}} = \sqrt{\frac{x^{\frac{2}{3}}+y^{\frac{2}{3}}}{y^{\frac{2}{3}}}}.$$

Putting these values in (1), we get

$$\frac{x\sqrt{x^{\frac{2}{3}}+y^{\frac{2}{3}}}}{x^{\frac{1}{3}}} + y\frac{\sqrt{x^{\frac{2}{3}}+y^{\frac{2}{3}}}}{y^{\frac{1}{3}}} = l$$

or

$$x^{\frac{2}{3}} + y^{\frac{2}{3}} = l^{\frac{2}{3}} \text{ (astroid).}$$

EXAMPLE 3.17

Find the envelope of the family of straight lines $y = mx + \sqrt{a^2m^2+b^2}$, the parameter being m.

Solution. The given family of straight lines is

$$y = mx + \sqrt{a^2m^2+b^2}$$

or

$$(y-mx)^2 = a^2m^2 + b^2. \qquad (1)$$

Differentiating partially with respect to m, we get

$$2(y-mx)(-x) = 2a^2m$$

or

$$-2xy + 2mx^2 = 2a^2m$$

or

$$m(a^2-x^2) = -xy$$

or

$$m = -\frac{xy}{a^2-x^2} = \frac{xy}{x^2-a^2}.$$

Putting this value of m in (1), we get

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

or

$$a^4y^2 = a^2x^2y^2 + b^2(x^2-a^2)^2$$

or

$$0 = a^2y^2(x^2-a^2) + b^2(x^2-a^2)^2$$

or

$$0 = a^2y^2 + b^2(x^2-a^2) = a^2y^2 + b^2x^2 - a^2b^2$$

or

$$b^2x^2 + a^2y^2 = a^2b^2$$

or

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \text{ (ellipse)}$$

EXAMPLE 3.18

Find the evolute of the parabola $y = 4ax$

Solution. The equation of parabola is $y^2 = 4ax$. Any normal to the parabola is

$$y = mx - 2am - am^3. \qquad (1)$$

Differentiating partially with respect to m, we get

$$0 = x - 2a - 3am^2,$$

which yields

$$m = \sqrt{\frac{x-2a}{3a}}.$$

Putting this value of m in (1), we have

$$y = \sqrt{\frac{x-2a}{3a}}\left[x - 2a - \frac{a(x-2a)}{3a}\right]$$

or

$$27ay^2 = 4(x-2a)^3.$$

EXAMPLE 3.19

Find the evolute of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

Solution. The parametric equation of the ellipse is

$x = a\cos\theta$, $y = b\sin\theta$, where θ is a parameter.

Therefore,

$$\frac{dx}{d\theta} = -a\sin\theta \text{ and } \frac{dy}{d\theta} = b\cos\theta.$$

Thus,

$$\frac{dy}{dx} = \frac{\frac{dy}{d\theta}}{\frac{dx}{d\theta}} = -\frac{b\cos\theta}{a\sin\theta}.$$

Therefore, slope of the normal to the ellipse at $(a \cos\theta, b \sin\theta)$ is

$$-\frac{dx}{dy} = -\frac{a\sin\theta}{b\cos\theta}.$$

Hence, the equation of the normal to the ellipse at (a cos θ, b sin θ) is

$$y - b\sin\theta = \frac{a\sin\theta}{b\cos\theta}(x - a\cos\theta)$$

or

$$\frac{ax}{\cos\theta} - \frac{by}{\sin\theta} = a^2 - b^2. \qquad (1)$$

Differentiating (1) partially with respect to θ, we get

$$\frac{ax\sin\theta}{\cos^2\theta} + \frac{by\cos\theta}{\sin^2\theta} = 0. \qquad (2)$$

The equation (2) gives

$$\tan^3\theta = -\frac{by}{ax} \quad \text{or} \quad \tan\theta = -\frac{(by)^{\frac{1}{3}}}{(ax)^{\frac{1}{3}}}.$$

Therefore,

$$\sin\theta = \frac{(by)^{\frac{1}{3}}}{\sqrt{(ax)^{\frac{2}{3}}+(by)^{\frac{2}{3}}}} \quad \text{and}$$

$$\cos\theta = -\frac{(ax)^{\frac{1}{3}}}{\sqrt{(ax)^{\frac{2}{3}}+(by)^{\frac{2}{3}}}}$$

or

$$\sin\theta = -\frac{(by)^{\frac{1}{3}}}{\sqrt{(ax)^{\frac{2}{3}}+(by)^{\frac{2}{3}}}} \quad \text{and}$$

$$\cos\theta = \frac{(ax)^{\frac{1}{3}}}{\sqrt{(ax)^{\frac{2}{3}}+(by)^{\frac{2}{3}}}}.$$

Substituting these values of sin θ and cos θ in (1) we get

$$\pm\left[(ax)^{\frac{2}{3}}+(by)^{\frac{2}{3}}\right]^{\frac{3}{2}} = a^2 - b^2$$

or

$$(ax)^{\frac{2}{3}} + (by)^{\frac{2}{3}} = (a^2 - b^2)^{\frac{2}{3}}.$$

3.7 HOMOGENEOUS FUNCTIONS AND EULER'S THEOREM

An expression of the form

$$a_0x^n + a_1x^{n-1}y + a_2x^{n-2}y^2 + \ldots + a_ny^n$$

in which every term is of nth degree is called a *homogeneous function of degree n.* This can also be rewritten as

$$x^n\left[a_0 + a_1\left(\frac{y}{x}\right) + a_2\left(\frac{y}{x}\right)^2 + \ldots + a_{n-1}\left(\frac{y}{x}\right)^{n-1} + a_n\left(\frac{y}{x}\right)^n\right]$$

or $x^n\phi\left(\frac{y}{x}\right)$. Thus, we can define a homogeneous function as follows:

A function *f(x, y)*, which can be expressed as $x^n\phi\left(\frac{y}{x}\right)$, is called a *homogeneous function of degree n in x and* y.

To check whether a function $f(x, y)$ is homogeneous or not, we put tx for x and ty for y in it. If $f(tx, ty) = t^n f(x, y)$, then the function $f(x, y)$ is a homogeneous function of degree n which is otherwise a nonhomogeneous function.

Let $u = x^n f\left(\frac{y}{x}\right)$ be a homogeneous function of x and y of degree n. Then,

$$\frac{\partial u}{\partial x} = nx^{n-1}f\left(\frac{y}{x}\right) + x^n f'\left(\frac{y}{x}\right).\left(\frac{-y}{x^2}\right)$$

$$= x^{n-1}\left[n f\left(\frac{y}{x}\right) - \left(\frac{y}{x}\right)f'\left(\frac{y}{x}\right)\right] = x^{n-1}\phi\left(\frac{y}{x}\right),$$

which is a homogeneous function of degree n – 1. Similarly,

$$\frac{\partial u}{\partial y} = x^n f'\left(\frac{y}{x}\right).\frac{1}{x} = x^{n-1}f'\left(\frac{y}{x}\right) = x^{n-1}\ \psi\left(\frac{y}{x}\right),$$

which is a homogeneous function of degree n – 1. It follows, therefore, that "If u is a homogeneous function of x and y of degree n, then u_x and u_y are also homogeneous functions of x and y of degree n – 1."

Theorem 3.3. (Euler's Theorem). If u is a homogeneous function of x and y of degree n, then

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = nu.$$

Proof: Since u is a homogeneous function of x and y of degree n, it can be expressed as

$$u = x^n f\left(\frac{y}{x}\right).$$

Differentiating partially with respect to x, we have

$$\frac{\partial u}{\partial x} = nx^{n-1}f\left(\frac{y}{x}\right) + x^n f'\left(\frac{y}{x}\right)\left(-\frac{1}{x^2}\right)$$

$$= nx^{n-1}f\left(\frac{y}{x}\right) - yx^{n-2}f'\left(\frac{y}{x}\right).$$

Similarly, the differentiation with respect to y yields

$$\frac{\partial u}{\partial y} = x^n f'\left(\frac{y}{x}\right)\left(\frac{1}{x}\right) = x^{n-1}f'\left(\frac{y}{x}\right).$$

Therefore,

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = nx^n f\left(\frac{y}{x}\right) - yx^{n-1}f'\left(\frac{y}{x}\right) + x^{n-1}yf'\left(\frac{y}{x}\right)$$

$$= nx^n f\left(\frac{y}{x}\right) = nu.$$

Remark 3.1. Euler's theorem on homogeneous functions also holds good for functions of n variables. Thus, "If $u(x_1, x_2, \ldots x_n)$ is a homogeneous function of $x_1, x_2, \ldots x_n$ of degree n, then

$$x_1\frac{\partial u}{\partial x_1}+x_2\frac{\partial u}{\partial x_2}+\ldots+x_n\frac{\partial u}{\partial x_n}=nu."$$

EXAMPLE 3.20

If $u=\tan^{-1}\left(\frac{x^3+y^3}{x+y}\right)$, show that

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=\sin 2u.$$

Solution. The given function is

$$u=\tan^{-1}\left(\frac{x^3+y^3}{x+y}\right).$$

Therefore,

$$\tan u=\left(\frac{x^3+y^3}{x+y}\right)=x^2\frac{1+\left(\frac{y}{x}\right)^3}{1+\frac{y}{x}}=z.$$

Therefore, z is a homogeneous function of degree 2 in x and y. Hence, by Euler's theorem,

$$x\frac{\partial z}{\partial x}+y\frac{\partial z}{\partial y}=2z. \qquad (1)$$

But,

$$\frac{\partial z}{\partial x}=\sec^2 u\frac{\partial u}{\partial x}\text{ and }\frac{\partial z}{\partial y}=\sec^2 u\frac{\partial u}{\partial y}.$$

Therefore, (1) reduces to

$$x\sec^2 u\frac{\partial u}{\partial x}+y\sec^2 u\frac{\partial u}{\partial y}=2\tan u$$

or

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=\frac{2\tan u}{\sec^2 u}=2\sin u\cos u=\sin 2u.$$

EXAMPLE 3.21

If $u=\sin^{-1}\frac{x^2+y^2}{x+y}$, show that

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=\tan u.$$

Solution. We have

$$u=\sin^{-1}\frac{x^2+y^2}{x+y}.$$

Therefore,

$$\sin u=\frac{x^2+y^2}{x+y}=z.$$

Thus,

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

and so, z is a homogeneous function of degree 1 in x and y. Hence by Euler's theorem, we have

$$x\frac{\partial z}{\partial x}+y\frac{\partial z}{\partial y}=z. \qquad (1)$$

But

$$\frac{\partial z}{\partial x}=\cos u\frac{\partial u}{\partial x}\text{ and }\frac{\partial z}{\partial y}=\cos u\frac{\partial u}{\partial y}.$$

Therefore, (1) reduces to

$$x\cos u\frac{\partial u}{\partial x}+y\cos u\frac{\partial u}{\partial y}=\sin u$$

or

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=\tan u.$$

EXAMPLE 3.22

If $u=xf\left(\frac{y}{x}\right)+g\left(\frac{y}{x}\right)$, show that

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}=0.$$

Solution. Let $u_1=x\,f\left(\frac{y}{x}\right)$ and $u_2=g\left(\frac{y}{x}\right)$ so that u u_1+u_2. Then u_1 is a homogeneous function of degree 1 and u_2 is a homogeneous function of degree 0. Therefore, by Euler's theorem, we have

$$x\frac{\partial u_1}{\partial x}+y\frac{\partial u_1}{\partial y}=u_1\text{ and }x\frac{\partial u_2}{\partial x}+y\frac{\partial u_2}{\partial y}=0. \quad (1)$$

But

$$\begin{aligned}x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}&=x\frac{\partial}{\partial x}(u_1+u_2)+y\frac{\partial}{\partial y}(u_1+u_2)\\&=x\frac{\partial u_1}{\partial x}+x\frac{\partial u_2}{\partial x}+y\frac{\partial u_1}{\partial y}+y\frac{\partial u_2}{\partial y}\\&=\left(x\frac{\partial u_1}{\partial x}+y\frac{\partial u_1}{\partial y}\right)+\left(x\frac{\partial u_2}{\partial x}+y\frac{\partial u_2}{\partial y}\right)\\&=u_1+0=u_1\ [\text{using (1)}]. \qquad (2)\end{aligned}$$

Differentiating (2) partially with respect to x and y, we get

$$x\frac{\partial^2 u}{\partial x^2}+\frac{\partial u}{\partial x}+y\frac{\partial^2 u}{\partial x\partial y}=\frac{\partial u_1}{\partial x} \qquad (3)$$

and

$$x\frac{\partial^2 u}{\partial y\partial x}+\frac{\partial u}{\partial x}+y\frac{\partial^2 u}{\partial y^2}=\frac{\partial u_1}{\partial y}. \qquad (4)$$

Multiplying (3) by x and (4) by y and then adding both, we get

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}+x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}$$
$$=x\frac{\partial u_1}{\partial x}+\frac{y\partial u_1}{\partial y}.$$

Using (1) and (2), the last expression reduces to

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}+u_1=u_1$$

or

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}=0.$$

EXAMPLE 3.23

If u is a homogeneous function of x and y of degree n, show that

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}=n(n-1)u.$$

Solution. By Euler's theorem, we have

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=n\,u. \tag{1}$$

Differentiating (1) partially with respect to x, we get

$$x\frac{\partial^2 u}{\partial x^2}+\frac{\partial u}{\partial x}+y\frac{\partial^2 u}{\partial x\partial y}=n\frac{\partial u}{\partial x}$$

or

$$x\frac{\partial^2 u}{\partial x^2}+y\frac{\partial^2 u}{\partial x\partial y}=n\frac{\partial u}{\partial x}-\frac{\partial u}{\partial x}=(n-1)\frac{\partial u}{\partial x}. \tag{2}$$

Again differentiating (1) partially with respect to y, we have

$$x\frac{\partial^2 u}{\partial y\partial x}+y\frac{\partial^2 u}{\partial y^2}+\frac{\partial u}{\partial y}=n\frac{\partial u}{\partial y}$$

or

$$x\frac{\partial^2 u}{\partial y\partial x}+y\frac{\partial^2 u}{\partial y^2}=(n-1)\frac{\partial u}{\partial y}. \tag{3}$$

Multiplying (2) by x and (3) by y and then adding both, we have

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}$$

$$=(n-1)\left(x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}\right)$$

$$=(n-1)nu \text{ [using (1)]}.$$

EXAMPLE 3.24

If $\sin u=\frac{x^2y^2}{x+y}$, show that

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=3\tan u.$$

Solution. We have

$$\sin u=\frac{x^2y^2}{x+y}=x^3\frac{\left(\frac{y}{x}\right)^2}{1+\frac{y}{x}}=v.$$

We observe that v is a homogeneous function of degree 3. Therefore, by Euler's theorem,

$$x\frac{\partial v}{\partial x}+y\frac{\partial v}{\partial y}=3v. \tag{1}$$

Since $v=\sin u$, we have

$$\frac{\partial v}{\partial x}=\cos u\frac{\partial u}{\partial x} \text{ and } \frac{\partial v}{\partial y}=\cos u\frac{\partial u}{\partial y}.$$

Hence, (1) reduces to

$$x\cos u\frac{\partial u}{\partial x}+y\cos u\frac{\partial u}{\partial y}=3\sin u$$

or

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=3\tan u.$$

EXAMPLE 3.25

If $u=\sin^{-1}\frac{x}{y}+\tan^{-1}\frac{y}{x}$, show that

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=0.$$

Solution. We have

$$u=\sin^{-1}\frac{x}{y}+\tan^{-1}\frac{y}{x}=\sin^{-1}\frac{1}{\frac{y}{x}}+\tan^{-1}\frac{y}{x}=x^0f\left(\frac{y}{x}\right).$$

Therefore, u is a homogeneous function of degree 0 in x and y. Hence by Euler's theorem, we have

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=0.u=0.$$

For verification of the result, see Example 3.6.

EXAMPLE 3.26

If $u=\tan^{-1}\frac{x^3+y^3}{x-y}$, show that

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}=\sin 4u-\sin 2u.$$

Solution. We have

$$\tan u=\frac{x^3+y^3}{x-y}=x^2\frac{1+\left(\frac{y}{x}\right)^3}{1-\frac{y}{x}}=v.$$

Then v is a homogeneous function of degree 2 in x and y. Hence, by Euler's theorem,

$$x\frac{\partial v}{\partial x}+y\frac{\partial v}{\partial y}=2v.$$

But

$$\frac{\partial v}{\partial x}=\sec^2 u\frac{\partial u}{\partial x},\ \frac{\partial v}{\partial y}=\sec^2 u\frac{\partial u}{\partial y}.$$

Therefore,

$$x\sec^2 u\frac{\partial u}{\partial x}+y\sec^2 u\frac{\partial u}{\partial y}=2\tan u$$

or

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=2\sin u\cos u=\sin 2u. \qquad (1)$$

Differentiating (1) partially with respect to x and y, respectively, we get

$$x\frac{\partial^2 u}{\partial x^2}+\frac{\partial u}{\partial x}+y\frac{\partial^2 u}{\partial x\partial y}=2\cos 2u\frac{\partial u}{\partial x} \qquad (2)$$

and

$$x\frac{\partial^2 u}{\partial y\partial x}+\frac{\partial u}{\partial y}+y\frac{\partial^2 u}{\partial y^2}=2\cos 2u\frac{\partial u}{\partial y}. \qquad (3)$$

Multiplying (2) by x and (3) by y and then adding both, we get

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}+x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}$$
$$=2\cos 2u\left(x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}\right)$$

or using (1), we have

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}+\sin 2u=2\cos 2u\sin 2u$$

or

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}=2\cos 2u\sin 2u-\sin 2u$$
$$=\sin 4u-\sin 2u.$$

EXAMPLE 3.27

If $u=\frac{x}{y+z}+\frac{y}{z+x}+\frac{z}{x+y}$, show that

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}+z\frac{\partial u}{\partial z}=0.$$

Solution. Replacing x by tx, y by ty, and z by tz, we get

$$u(tx,,ty,tz)=\frac{tx}{ty+tz}+\frac{ty}{tz+tx}+\frac{tz}{tx+ty}$$
$$=1.u(x,y,z)=t^0u(x,y,z).$$

Hence, $u(x, y, z)$ is a homogeneous function of degree zero in x, y, and z. Hence, by Euler's theorem,

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}+z\frac{\partial u}{\partial z}=nu=0.u=0.$$

EXAMPLE 3.28

If $u=e^{x^2+y^2}$, show that

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=2u\log u.$$

Solution. We have $u=e^{x^2+y^2}$ and so,

$$\log u=x^2+y^2=x^2\left[1+\left(\frac{y}{x}\right)^2\right]=v.$$

Thus, v is a homogeneous function of degree 2 in x and y. Hence, by Euler's theorem,

$$x\frac{\partial v}{\partial x}+y\frac{\partial v}{\partial y}=2v.$$

But

$$\frac{\partial v}{\partial x}=\frac{1}{u}\frac{\partial u}{\partial x},\ \frac{\partial v}{\partial y}=\frac{1}{u}\frac{\partial u}{\partial y}.$$

Therefore,

$$\frac{x}{u}\frac{\partial u}{\partial x}+\frac{y}{u}\frac{\partial u}{\partial y}=2\log u$$

or

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=2u\log u.$$

EXAMPLE 3.29

If $u=\sin^{-1}\frac{x+y}{\sqrt{x}+\sqrt{y}}$, show that

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=\frac{1}{2}\tan u.$$

Solution. We have

$$\sin u=\frac{x+y}{\sqrt{x}+\sqrt{y}}=x^{\frac{1}{2}}\frac{1+\frac{y}{x}}{1+\sqrt{\frac{y}{x}}}=v.$$

Therefore, v is a homogeneous function of degree $\frac{1}{2}$ in x and y. Hence, by Euler's theorem,

$$x\frac{\partial v}{\partial x}+y\frac{\partial v}{\partial y}=\frac{1}{2}v.$$

But

$$\frac{\partial v}{\partial x}=\cos u\frac{\partial u}{\partial x}\ \text{and}\ \frac{\partial v}{\partial y}=\cos u\frac{\partial u}{\partial y}.$$

Therefore,

$$x\cos u\frac{\partial u}{\partial x}+y\cos u\frac{\partial u}{\partial y}=\frac{1}{2}\sin u$$

or

$$x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}=\frac{1}{2}\tan u.$$

EXAMPLE 3.30

If $u=\operatorname{cosec}^{-1}\left[\frac{x^{\frac{1}{2}}+y^{\frac{1}{2}}}{x^{\frac{1}{3}}+y^{\frac{1}{3}}}\right]^{\frac{1}{2}}$, show that

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}=\frac{\tan u}{144}(13+\tan^2 u).$$

Solution. We have

$$\operatorname{cosec} u = \left[\frac{x^{\frac{1}{2}} + y^{\frac{1}{2}}}{x^{\frac{1}{3}} + y^{\frac{1}{3}}}\right]^{\frac{1}{2}} = \left[\frac{x^{\frac{1}{2}}\left(1+\left(\frac{y}{x}\right)\right)^{\frac{1}{2}}}{x^{\frac{1}{3}}\left(1+\left(\frac{y}{x}\right)\right)^{\frac{1}{3}}}\right]^{\frac{1}{2}}$$

$$= x^{\frac{1}{12}}\left[\frac{1+\left(\frac{y}{x}\right)^{\frac{1}{2}}}{1+\left(\frac{y}{x}\right)^{\frac{1}{3}}}\right]^{\frac{1}{2}} = v.$$

Therefore, v is a homogeneous function of degree $\frac{1}{12}$ in x and y. Hence, by Euler's theorem,

$$x\frac{\partial v}{\partial x} + y\frac{\partial v}{\partial y} = \frac{1}{12}v.$$

But

$$\frac{\partial v}{\partial x} = -\operatorname{cosec} u \cot u \frac{\partial u}{\partial x},$$

$$\frac{\partial v}{\partial y} = -\operatorname{cosec} u \cot u \frac{\partial u}{\partial y}.$$

Therefore,

$$-x \operatorname{cosec} u \cot u \frac{\partial u}{\partial x} - y \operatorname{cosec} u \cot u \frac{\partial u}{\partial y} = \frac{1}{12}\operatorname{cosec} u$$

or

$$x \cot u \frac{\partial u}{\partial x} + y \cot u \frac{\partial u}{\partial y} = -\frac{1}{12}$$

or

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = \frac{-1}{12}\tan u. \qquad (1)$$

Differentiating (1) partially with respect to x, we get

$$x\frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial x} + y\frac{\partial^2 u}{\partial x \partial y} = -\frac{1}{12}\sec^2 u \frac{\partial u}{\partial x}$$

or

$$x\frac{\partial^2 u}{\partial x^2} + y\frac{\partial^2 u}{\partial x \partial y} = \left(-\frac{1}{12}\sec^2 u - 1\right)\frac{\partial u}{\partial x}. \qquad (2)$$

Similarly, differentiating (1) partially with respect to y, we get

$$y\frac{\partial^2 u}{\partial y^2} + x\frac{\partial^2 u}{\partial y \partial x} = \left(-\frac{1}{12}\sec^2 u - 1\right)\frac{\partial u}{\partial y}. \qquad (3)$$

Multiplying (2) by x and (3) by y and then adding both, we get

$$x^2\frac{\partial^2 u}{\partial x^2} + 2xy\frac{\partial^2 u}{\partial x \partial y} + y^2\frac{\partial^2 u}{\partial y^2}$$

$$= \left(-\frac{1}{12}\sec^2 u - 1\right)\left(x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}\right)$$

$$= \left(-\frac{1}{12}\sec^2 u - 1\right)\left(-\frac{1}{12}\tan u\right)$$

$$= \left(\frac{1}{12}\sec^2 u - 1\right)\left(\frac{1}{12}\tan u\right)$$

$$= \left[\frac{1}{12}\left(1+\tan^2 u\right) - 1\right]\left(\frac{1}{12}\tan u\right)$$

$$= \frac{1}{144}\tan u(13 + \tan^2 u).$$

EXAMPLE 3.31

If $z = x^n f_1\left(\frac{y}{x}\right) + y^{-n} f_2\left(\frac{x}{y}\right)$, show that

$$x^2\frac{\partial^2 z}{\partial x^2} + 2xy\frac{\partial^2 z}{\partial x \partial y} + y^2\frac{\partial^2 z}{\partial y^2} + x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = n^2 z.$$

Solution. Let $u_1 = x^n f_1\left(\frac{y}{x}\right)$ and $u_2 = y^{-n} f_2\left(\frac{x}{y}\right)$ so that $z = u_1 + u_2.$ Then, u_1 is a homogeneous function of degree n and u_2 is a homogeneous function of degree n. Therefore, by Euler's theorem, we have

$$x\frac{\partial u_1}{\partial x} + y\frac{\partial u_1}{\partial y} = nu_1 \text{ and } x\frac{\partial u_2}{\partial x} + y\frac{\partial u_2}{\partial y} = -nu_2. \qquad (1)$$

But

$$x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = x\frac{\partial}{\partial x}(u_1+u_2) + y\frac{\partial}{\partial y}(u_1+u_2)$$

$$= \left(x\frac{\partial u_1}{\partial x} + y\frac{\partial u_1}{\partial y}\right) + \left(x\frac{\partial u_2}{\partial x} + y\frac{\partial u_2}{\partial y}\right)$$

$$= n(u_1 - u_2), \text{ using (1)}. \qquad (2)$$

Differentiating (2) partially with respect to x and y, we get

$$x\frac{\partial^2 z}{\partial x^2} + \frac{\partial z}{\partial x} + y\frac{\partial^2 z}{\partial x \partial y} = n\left(\frac{\partial u_1}{\partial x} - \frac{\partial u_2}{\partial x}\right) \qquad (3)$$

and

$$x\frac{\partial^2 z}{\partial x \partial y} + \frac{\partial z}{\partial y} + y\frac{\partial^2 z}{\partial y^2} = n\left(\frac{\partial u_1}{\partial y} - \frac{\partial u_2}{\partial y}\right). \qquad (4)$$

Multiplying (3) by x and (4) by y and then adding both, we get

$$x^2\frac{\partial^2 z}{\partial x^2} + 2xy\frac{\partial^2 z}{\partial x \partial y} + y^2\frac{\partial^2 z}{\partial y^2} + x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y}$$

$$= nx\left(\frac{\partial u_1}{\partial x} - \frac{\partial u_2}{\partial x}\right) + ny\left(\frac{\partial u_1}{\partial y} - \frac{\partial u_2}{\partial y}\right)$$

$$= n\left[\left(x\frac{\partial u_1}{\partial x} + y\frac{\partial u_1}{\partial y}\right) - \left(x\frac{\partial u_2}{\partial x} + y\frac{\partial u_2}{\partial y}\right)\right]$$

$$= n[nu_1 + nu_2] = n^2(u_1 + u_2) = n^2 z.$$

3.8 DIFFERENTIATION OF COMPOSITE FUNCTIONS

Let $x = \phi(t)$ and $y = \psi(t)$. If $u = f(x, y)$, then u is called a *composite function of the single variable t*.

Similarly, if $x=\phi(u,v)$, $y=\psi(u,v)$, and $F=f(x,y)$, then F is called a *composite function of two variables u and v.*

Theorem 3.4. If $u=f(x,y)$, where $x=\phi(t)$ and $y=(t)$, then

$$\frac{du}{dt}=\frac{\partial u}{\partial x}\cdot\frac{dx}{dt}+\frac{\partial u}{\partial y}\cdot\frac{dy}{dt}.$$

($\frac{du}{dt}$ is called the *total differential coefficient* or *total derivative of u with respect to t*).

Proof: Let δt be the small increment given to t and δx, δy, and δu be the corresponding increments in x, y, and u, respectively. We have

$$u=f(x,y), \qquad (1)$$

$$u+\delta u=f(x+\delta x,y+\delta y). \qquad (2)$$

Subtracting (1) from (2), we get

$$\begin{aligned}\delta u&=f(x+\delta x,y+\delta y)-f(x,y)\\&=[f(x+\delta x,y+\delta y)-f(x,y+\delta y)]\\&\quad+[f(x,y+\delta y)-f(x,y)].\end{aligned}$$

Thus,

$$\begin{aligned}\frac{\delta u}{\delta t}&=\frac{f(x+\delta x,y+\delta y)-f(x,y+\delta y)}{\delta t}\\&\quad+\frac{f(x,y+\delta y)-f(x,y)}{\delta t}\\&=\frac{f(x+\delta x,y+\delta y)-f(x,y+\delta y)}{\delta x}\cdot\frac{\delta x}{\delta t}\\&\quad+\frac{f(x,y+\delta y)-f(x,y)}{\delta y}\cdot\frac{\delta y}{\delta t}.\end{aligned}$$

If $\delta t\to 0$, δx and δy both tend to zero and so,

$$\begin{aligned}\lim_{\delta t\to 0}\frac{\delta u}{dt}&=\lim_{\delta x\to 0}\left[\frac{f(x+\delta x,y+\delta y)-f(x,y+\delta y)}{\delta x}\right]\lim_{\delta t\to 0}\frac{\delta x}{\delta t}\\&\quad+\lim_{\delta y\to 0}\left[\frac{f(x,y+\delta y)-f(x,y)}{\delta y}\right]\lim_{\delta t\to 0}\frac{\delta x}{\delta t}\\&=\frac{\partial f}{\partial x}.\lim_{\delta t\to 0}\frac{\delta x}{\delta t}+\frac{\partial f}{\partial y}.\lim_{\delta t\to 0}\frac{\delta y}{\delta t}.\end{aligned}$$

or

$$\frac{du}{dt}=\frac{\partial f}{\partial x}\cdot\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}=\frac{\partial u}{\partial x}\frac{dx}{dt}+\frac{\partial u}{\partial y}\frac{dy}{dt}. \qquad (3)$$

Corollary 1. If u is a function of x and y and y is a function of x, then

$$\frac{du}{dx}=\frac{\partial u}{\partial x}+\frac{\partial u}{\partial y}\cdot\frac{dy}{dx}.$$

The result follows taking $t=x$ in (3).

Corollary 2. If $z=f(x,y)$, $x=\phi(u,v)$, and $y=\psi(u,v)$, then

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial u}$$

and

$$\frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial v}.$$

Corollary 3. If $u=f(x,y)=c$, then by Corollary 1, we have

$$0=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\cdot\frac{dy}{dx}$$

or

$$\frac{dy}{dx}=-\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}.$$

Differentiating once more with respect to x, we get

$$\frac{d^2y}{dx^2}=-\frac{f_{xx}f_y^2-2f_xf_yf_{xy}+f_{yy}f_x^2}{f_y^3}.$$

EXAMPLE 3.32

If $x^y+y^x=a^b$, find $\frac{dy}{dx}$.

Solution. Let

$$f(x,y)=x^y+y^x-a^b.$$

Since from the given relation $x^y+y^x-a^b=0$, it follows that $f(x,y)=0$. Therefore,

$$\frac{dy}{dx}=-\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}=\frac{yx^{y-1}+y^x\log y}{x^y\log x+xy^{x-1}}.$$

EXAMPLE 3.33

If $f(x,y)=0$ and $\phi(y,z)=0$, show that

$$\frac{\partial f}{\partial y}\cdot\frac{\partial\phi}{\partial z}\cdot\frac{dz}{dx}=\frac{\partial f}{\partial x}\cdot\frac{\partial\phi}{\partial y}.$$

Solution. Since $f(x,y)=0$, we have

$$\frac{dy}{dx}=-\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}. \qquad (1)$$

Again, since $\phi(y,z)=0$, we have

$$\frac{dz}{dy}=-\frac{\frac{\partial\phi}{\partial y}}{\frac{\partial\phi}{\partial z}}. \qquad (2)$$

Multiplication of (1) and (2) yields

$$\frac{dy}{dx}\cdot\frac{dz}{dy}=\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}\cdot\frac{\frac{\partial\phi}{\partial y}}{\frac{\partial\phi}{\partial z}}$$

or

$$\frac{dz}{dx}=\frac{\frac{\partial f}{\partial x}\cdot\frac{\partial\phi}{\partial y}}{\frac{\partial f}{\partial y}\cdot\frac{\partial\phi}{\partial z}}$$

or

$$\frac{dz}{dx}\cdot\frac{\partial f}{\partial y}\cdot\frac{\partial\phi}{\partial z}=\frac{\partial f}{\partial x}\cdot\frac{\partial\phi}{\partial y}.$$

EXAMPLE 3.34

If z is a function of x and y and $x=e^u+e^{-v}$ and $y=e^{-u}-e^{v}$, show that

$$\frac{\partial z}{\partial u}-\frac{\partial z}{\partial v}=x\frac{\partial z}{\partial x}-y\frac{\partial z}{\partial y}.$$

Solution. The function z is a composite function of u and v. Therefore,

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}.\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}.\frac{\partial y}{\partial u}=\frac{\partial z}{\partial x}.e^u-\frac{\partial z}{\partial y}.e^{-u}\qquad(1)$$

and

$$\frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}.\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}.\frac{\partial y}{\partial v}=-\frac{\partial z}{\partial x}.e^{-v}-\frac{\partial z}{\partial y}.e^{v}.\qquad(2)$$

Subtracting (2) from (1), we have

$$\frac{\partial z}{\partial u}-\frac{\partial z}{\partial v}=(e^u-e^{-v})\frac{\partial z}{\partial x}+(e^{-u}+e^{v})\frac{\partial z}{\partial y}$$
$$=x\frac{\partial z}{\partial x}+y\frac{\partial z}{\partial y}.$$

EXAMPLE 3.35

Prove that $\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=\frac{\partial^2 u}{\partial \xi^2}+\frac{\partial^2 u}{\partial \eta^2}$, where

$$x=\xi\cos\alpha-\eta\sin\alpha,\ y=\xi\sin\alpha+\eta\cos\alpha.$$

Solution. We have

$$x=\xi\cos\alpha-\eta\sin\alpha,\ y=\xi\sin\alpha+\eta\cos\alpha.$$

Now

$$\frac{\partial u}{\partial \xi}=\frac{\partial u}{\partial x}\cdot\frac{\partial x}{\partial \xi}+\frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial \xi}=\cos\alpha\frac{\partial u}{\partial x}+\sin\alpha\frac{\partial u}{\partial y},$$

which yields

$$\frac{\partial}{\partial \xi}(u)=\left(\cos\alpha\frac{\partial}{\partial x}+\sin\alpha\frac{\partial}{\partial y}\right)(u)$$

and so,

$$\frac{\partial}{\partial \xi}=\cos\alpha\frac{\partial}{\partial x}+\sin\alpha\frac{\partial}{\partial y}.$$

Thus,

$$\frac{\partial^2 u}{\partial \xi^2}=\frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \xi}\right)=\left(\cos\alpha\frac{\partial}{\partial x}+\sin\alpha\frac{\partial}{\partial y}\right)$$
$$\times\left(\cos\alpha\frac{\partial u}{\partial x}+\sin\alpha\frac{\partial u}{\partial y}\right)$$
$$=\cos\alpha\frac{\partial}{\partial x}\left(\cos\alpha\frac{\partial u}{\partial x}+\sin\alpha\frac{\partial u}{\partial y}\right)$$
$$+\sin\alpha\frac{\partial}{\partial y}\left(\cos\alpha\frac{\partial u}{\partial x}+\sin\alpha\frac{\partial u}{\partial y}\right)$$
$$=\cos^2\alpha\frac{\partial^2 u}{\partial x^2}+2\sin\alpha\cos\alpha\frac{\partial^2 u}{\partial x\partial y}+\sin^2\alpha\frac{\partial^2 u}{\partial y^2}.\ (1)$$

On the other hand,

$$\frac{\partial u}{\partial \eta}=\frac{\partial u}{\partial x}\cdot\frac{\partial x}{\partial \eta}+\frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial \eta}=-\sin\alpha\frac{\partial u}{\partial x}+\cos\alpha\frac{\partial u}{\partial y},$$

which gives

$$\frac{\partial}{\partial \eta}=\left(-\sin\alpha\frac{\partial}{\partial x}+\cos\alpha\frac{\partial}{\partial y}\right).$$

Therefore,

$$\frac{\partial^2 u}{\partial \eta^2}=\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \eta}\right)$$
$$=\left(-\sin\alpha\frac{\partial}{\partial x}+\cos\alpha\frac{\partial}{\partial y}\right)$$
$$\times\left(-\sin\alpha\frac{\partial u}{\partial x}+\cos\alpha\frac{\partial u}{\partial y}\right)$$
$$=-\sin\alpha\frac{\partial}{\partial x}\left(-\sin\alpha\frac{\partial u}{\partial x}+\cos\alpha\frac{\partial u}{\partial y}\right)$$
$$+\cos\alpha\frac{\partial}{\partial y}\left(-\sin\alpha\frac{\partial u}{\partial x}+\cos\alpha\frac{\partial u}{\partial y}\right)$$
$$=\sin^2\alpha\frac{\partial^2 u}{\partial x^2}=-2\sin\alpha\cos\alpha\frac{\partial^2 u}{\partial x\partial y}$$
$$+\cos^2\alpha\frac{\partial^2 u}{\partial y^2}.\qquad(2)$$

Adding (1) and (2), we get

$$\frac{\partial^2 u}{\partial \xi^2}+\frac{\partial^2 u}{\partial \eta^2}=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}.$$

EXAMPLE 3.36

If $y^3-3ax^2+x^3=0$, show that

$$\frac{d^2y}{dx^2}+\frac{2a^2x^2}{y^2}=0.$$

Solution. Let

$$f(x,y)=y^3-3ax^2+x^3=0.$$

Therefore, by Corollary 3 of Theorem 3.4

$$\frac{d^2y}{dx^2} = -\frac{f_{xx}f_y^2 - 2f_xf_yf_{xy} + f_{yy}f_x^2}{f_y^3}. \quad (1)$$

But in the present case,

$$f_x(x,y) = -6ax + 3x^2, f_y(x,y) = 3y^2, f_{xy}(x,y) = 0,$$
$$f_{xx}(x,y) = -6a + 6x, \text{ and } f_{yy}(x,y) = 6y.$$

Therefore, (1) reduces to

$$\frac{d^2y}{dx^2} = -\frac{6(x-a)9y^4 + (3x^2 - 6ax)^2 6y}{27y^6} = -2\frac{a^2x^2}{y^5}.$$

Hence,

$$\frac{d^2y}{dx^2} + 2\frac{a^2x^2}{y^5} = 0.$$

EXAMPLE 3.37

If $u = f(y - z, z - x, z - y)$, show that

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} + \frac{\partial u}{\partial z} = 0.$$

Solution. Substituting $y - z, = A$, $z - x = B$, and $x - y = C$, we have

$$u = f(A, B, C).$$

Thus, u is a composite function of A, B, and C. Therefore,

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial A}\cdot\frac{\partial A}{\partial x} + \frac{\partial u}{\partial B}\cdot\frac{\partial B}{\partial x} + \frac{\partial u}{\partial C}\cdot\frac{\partial C}{\partial x}$$
$$= \frac{\partial u}{\partial A}(0) + \frac{\partial u}{\partial B}(-1) + \frac{\partial u}{\partial C}(1) = -\frac{\partial u}{\partial B} + \frac{\partial u}{\partial C}, (1)$$

$$\frac{\partial u}{\partial y} = \frac{\partial u}{\partial A}\cdot\frac{\partial A}{\partial y} + \frac{\partial u}{\partial B}\cdot\frac{\partial B}{\partial y} + \frac{\partial u}{\partial C}\cdot\frac{\partial C}{\partial y} = \frac{\partial u}{\partial A} - \frac{\partial u}{\partial C}, \quad (2)$$

$$\frac{\partial u}{\partial z} = \frac{\partial u}{\partial A}\cdot\frac{\partial A}{\partial z} + \frac{\partial u}{\partial B}\cdot\frac{\partial B}{\partial z} + \frac{\partial u}{\partial C}\cdot\frac{\partial C}{\partial z} = -\frac{\partial u}{\partial A} + \frac{\partial u}{\partial B} \quad (3)$$

Adding (1), (2), and (3), we have

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} + \frac{\partial u}{\partial z} = 0.$$

EXAMPLE 3.38

If $u = f(r, s, t)$ and $r = \frac{x}{y}$, $s = \frac{y}{z}$, $t = \frac{z}{x}$, show that

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = 0$$

Solution. The function $u = f(r, s, t)$ is a composite function of r, s, and t. We have

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r}\cdot\frac{\partial r}{\partial x} + \frac{\partial u}{\partial s}\cdot\frac{\partial s}{\partial x} + \frac{\partial u}{\partial t}\cdot\frac{\partial t}{\partial x}$$
$$= \frac{\partial u}{\partial r}\left(\frac{1}{y}\right) + \frac{\partial u}{\partial s}(0) + \frac{\partial u}{\partial t}\left(-\frac{z}{x^2}\right)$$

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r}\cdot\frac{\partial r}{\partial x} + \frac{\partial u}{\partial s}\cdot\frac{\partial s}{\partial x} + \frac{\partial u}{\partial t}\cdot\frac{\partial t}{\partial x}$$
$$= \frac{\partial u}{\partial r}\left(\frac{1}{y}\right) + \frac{\partial u}{\partial s}(0) + \frac{\partial u}{\partial t}\left(-\frac{z}{x^2}\right)$$
$$= \frac{\partial u}{y\partial r} - \frac{z}{x^2}\frac{\partial u}{\partial t},$$

$$\frac{\partial u}{\partial y} = \frac{\partial u}{\partial r}\cdot\frac{\partial r}{\partial y} + \frac{\partial u}{\partial s}\cdot\frac{\partial s}{\partial y} + \frac{\partial u}{\partial t}\cdot\frac{\partial t}{\partial y}$$
$$= -\frac{x}{y^2}\frac{\partial u}{\partial r} + \frac{1}{z}\frac{\partial u}{\partial s}, \text{ and}$$

$$\frac{\partial u}{\partial z} = \frac{\partial u}{\partial r}\cdot\frac{\partial r}{\partial z} + \frac{\partial u}{\partial s}\cdot\frac{\partial s}{\partial z} + \frac{\partial u}{\partial t}\cdot\frac{\partial t}{\partial z}$$
$$= -\frac{y}{z^2}\frac{\partial u}{\partial s} + \frac{1}{x}\frac{\partial u}{\partial t}.$$

Therefore,

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = \frac{x}{y}\frac{\partial u}{\partial r} - \frac{z}{x}\frac{\partial u}{\partial t} - \frac{x}{y}\frac{\partial u}{\partial r}$$
$$+ \frac{y}{z}\frac{\partial u}{\partial s} - \frac{y}{z}\frac{\partial u}{\partial s} + \frac{z}{x}\frac{\partial u}{\partial t} = 0.$$

EXAMPLE 3.39

If $f(x, y) = \theta(u, v)$, where $u = x^2 - y^2$ and $v = 2xy$, show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 4(x^2 + y^2)\left(\frac{\partial^2\theta}{\partial u^2} + \frac{\partial^2\theta}{\partial v^2}\right).$$

Solution. Here $f(x, y)$ is composite function of u and v. Therefore,

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\cdot\frac{\partial v}{\partial x} = \frac{\partial\theta}{\partial u}(2x) + \frac{\partial\theta}{\partial v}(2y)$$
$$= 2\left(x\frac{\partial}{\partial u} + y\frac{\partial}{\partial v}\right)\theta,$$

which yields

$$\frac{\partial}{\partial x} = 2\left(x\frac{\partial}{\partial u} + y\frac{\partial}{\partial v}\right).$$

Therefore,

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = 2\frac{\partial}{\partial x}\left[x\frac{\partial\theta}{\partial u} + y\frac{\partial\theta}{\partial v}\right]$$
$$= 4\left(x\frac{\partial}{\partial u} + y\frac{\partial}{\partial v}\right)\left(x\frac{\partial\theta}{\partial u} + y\frac{\partial\theta}{\partial v}\right)$$
$$= 4\left(x^2\frac{\partial^2\theta}{\partial u^2} + xy\frac{\partial^2\theta}{\partial u\partial v} + xy\frac{\partial^2\theta}{\partial v\partial u} + y^2\frac{\partial^2\theta}{\partial u^2}\right)$$
$$= 4\left(x^2\frac{\partial^2\theta}{\partial u^2} + 2xy\frac{\partial^2\theta}{\partial u\partial v} + y^2\frac{\partial^2\theta}{\partial v^2}\right). \quad (1)$$

Further,

$$\frac{\partial f}{\partial y}=\frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial y}+\frac{\partial f}{\partial v}\cdot\frac{\partial v}{\partial y}=\frac{\partial \theta}{\partial u}(-2y)+\frac{\partial \theta}{\partial v}(2x)$$

$$=-2\left(y\frac{\partial \theta}{\partial u}-x\frac{\partial \theta}{\partial v}\right)$$

and so,

$$\frac{\partial}{\partial y}=-2\left(y\frac{\partial}{\partial u}-x\frac{\partial}{\partial v}\right).$$

Therefore,

$$\frac{\partial^2 f}{\partial y^2}=\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)=4\left(y\frac{\partial}{\partial u}-x\frac{\partial}{\partial v}\right)\left(y\frac{\partial \theta}{\partial u}-x\frac{\partial \theta}{\partial v}\right)$$

$$=4\left(y^2\frac{\partial^2\theta}{\partial u^2}-xy\frac{\partial^2\theta}{\partial u\partial v}-xy\frac{\partial^2\theta}{\partial v\partial u}+x^2\frac{\partial^2\theta}{\partial v^2}\right)$$

$$=4\left(y^2\frac{\partial^2\theta}{\partial u^2}-2xy\frac{\partial^2\theta}{\partial x\partial y}+x^2\frac{\partial^2\theta}{\partial v^2}\right). \qquad (2)$$

Adding (1) and (2), we get

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}=4(x^2+y^2)\left[\frac{\partial^2\theta}{\partial u^2}+\frac{\partial^2\theta}{\partial v^2}\right].$$

EXAMPLE 3.40

Let $u(x,\ y)$ be a function possessing continuous partial derivatives of the first two orders and let $x+y=2e^{\theta}\cos\phi$ and $x-y=2ie^{\theta}\sin\phi$. Show that

$$\frac{\partial^2 u}{\partial\theta^2}+\frac{\partial^2 u}{\partial\phi^2}=4xy\frac{\partial^2 u}{\partial x\partial y}.$$

Solution. The function u is a composite function of θ and ϕ. We are given that $x+y=2e^{\theta}\cos\phi$ and $x-y=2ie^{\theta}\sin\phi$. Solving these equations, we get

$$x=e^{\theta}(\cos\phi+i\sin\phi)=e^{\theta}e^{i\phi} \text{ and}$$

$$y=e^{\theta}(\cos\phi-i\sin\phi)=e^{\theta}e^{-\phi}.$$

Therefore,

$$\frac{\partial u}{\partial\theta}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial\theta}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial\theta}=e^{\theta}\left(e^{i\phi}\frac{\partial u}{\partial x}+e^{-i\phi}\frac{\partial u}{\partial y}\right)$$

$$=e^{\theta}\left(e^{i\phi}\frac{\partial}{\partial x}+e^{-i\phi}\frac{\partial}{\partial y}\right)(u),$$

and so,

$$\frac{\partial}{\partial\theta}=e^{\theta}\left(e^{i\phi}\frac{\partial}{\partial x}+e^{-i\phi}\frac{\partial}{\partial y}\right).$$

Therefore,

$$\frac{\partial^2 u}{\partial\theta^2}=\frac{\partial}{\partial\theta}\left(\frac{\partial u}{\partial\theta}\right)=e^{2\phi}\left(e^{i\theta}\frac{\partial}{\partial x}+e^{-i\phi}\frac{\partial}{\partial y}\right)\left(e^{i\phi}\frac{\partial u}{\partial x}+e^{-i\phi}\frac{\partial u}{\partial y}\right)$$

$$=e^{2\theta}\left[e^{2i\phi}\frac{\partial^2 u}{\partial x^2}+2\frac{\partial^2 u}{\partial x\partial y}+e^{-2i\phi}\frac{\partial^2 u}{\partial y^2}\right].$$

Similarly,

$$\frac{\partial^2 u}{\partial\phi^2}=-e^{2\theta}\left[e^{2i\phi}\frac{\partial^2 u}{\partial x^2}-2\frac{\partial^2 u}{\partial x\partial y}+e^{-2i\phi}\frac{\partial^2 u}{\partial y^2}\right].$$

Adding the earlier two expressions, we get

$$\frac{\partial^2 u}{\partial\theta^2}+\frac{\partial^2 u}{\partial\phi^2}=4e^{2\theta}\left(\frac{\partial^2 u}{\partial x\partial y}\right)=4xy\frac{\partial^2 u}{\partial x\partial y}.$$

3.9 TRANSFORMATION FROM CARTESIAN TO POLAR COORDINATES AND VICE VERSA

In this section, we deal with some important transformations of relations from one coordinate system to another coordinate system.

EXAMPLE 3.41

Transform the Laplace equation

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}=0$$

into polar coordinates.

Solution. The function V is a function of x and y, where x and y are expressed in polar coordinates as $x=r\cos\theta$ and $y=r\sin\theta$. Thus, V is a function of r and θ also. We have

$$r=\sqrt{x^2+y^2} \text{ and } \theta=\tan^{-1}\left(\frac{y}{x}\right).$$

Therefore,

$$\frac{\partial r}{\partial x}=\frac{x}{\sqrt{x^2+y^2}}=\frac{r\cos\theta}{r}=\cos\theta,$$

$$\frac{\partial r}{\partial y}=\frac{y}{\sqrt{x^2+y^2}}=\frac{r\sin\theta}{r}=\sin\theta,$$

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

$$=-\frac{r\sin\theta}{r^2}=-\frac{\sin\theta}{r}, \text{ and}$$

$$\frac{\partial\theta}{\partial y}=\frac{1}{1+\frac{y^2}{x^2}}\left(\frac{1}{x}\right)=\frac{x}{x^2+y^2}=\frac{r\cos\theta}{r^2}=\frac{\cos\theta}{r}.$$

Now,

$$\frac{\partial V}{\partial x}=\frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial x}+\frac{\partial V}{\partial \theta}\cdot\frac{\partial \theta}{\partial x}=\cos\theta\frac{\partial V}{\partial r}-\frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\quad\text{and}$$

$$\frac{\partial V}{\partial y}=\frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial y}+\frac{\partial V}{\partial \theta}\cdot\frac{\partial \theta}{\partial y}=\sin\theta\frac{\partial V}{\partial r}+\frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}.$$

Therefore,

$$\frac{\partial}{\partial x}=\left(\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\quad\text{and}$$

$$\frac{\partial}{\partial y}=\left(\sin\theta\frac{\partial}{\partial r}+\frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\right).$$

Hence,

$$\begin{aligned}\frac{\partial^2 V}{\partial x^2}&=\left(\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\cos\theta\frac{\partial V}{\partial r}-\frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\right)\\
&=\cos\theta\frac{\partial}{\partial r}\left(\cos\theta\frac{\partial V}{\partial r}-\frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\right)\\
&\quad-\frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left(\cos\theta\frac{\partial V}{\partial r}-\frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\right)\\
&=\cos\theta\left(\cos\theta\frac{\partial^2 V}{\partial r^2}+\frac{\sin\theta}{r^2}\frac{\partial V}{\partial \theta}-\frac{\sin\theta}{r}\frac{\partial^2 V}{\partial r\partial \theta}\right)\\
&\quad-\frac{\sin\theta}{r}\left(\cos\theta\frac{\partial^2 V}{\partial \theta\partial r}-\sin\theta\frac{\partial V}{\partial r}\right.\\
&\quad\left.-\frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}-\frac{\sin\theta}{r}\frac{\partial^2 V}{\partial \theta^2}\right)\\
&=\cos^2\theta\frac{\partial^2 V}{\partial r^2}-2\frac{\sin\theta\cos\theta}{r^2}\frac{\partial^2 V}{\partial r\partial \theta}\\
&\quad+\frac{\sin\theta}{r^2}\frac{\partial^2 V}{\partial \theta^2}+\frac{\sin^2\theta}{r}\frac{\partial V}{\partial r}+\frac{2\sin\theta}{r^2}\frac{\partial V}{\partial \theta}\qquad(1)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2 V}{\partial y^2}&=\left(\sin\theta\frac{\partial}{\partial r}+\frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\sin\theta\frac{\partial V}{\partial r}+\frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}\right)\\
&=\sin\theta\frac{\partial}{\partial r}\left(\sin\theta\frac{\partial V}{\partial r}+\frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}\right)\\
&\quad+\frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial V}{\partial r}+\frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}\right)\\
&=\sin\theta\left(\sin\theta\frac{\partial^2 V}{\partial r^2}-\frac{\cos\theta}{r^2}\frac{\partial V}{\partial \theta}+\frac{\cos\theta}{r}\frac{\partial^2 V}{\partial r\partial \theta}\right)\\
&\quad+\frac{\cos\theta}{r}\left(\sin\theta\frac{\partial^2 V}{\partial \theta\partial r}+\cos\theta\frac{\partial V}{\partial r}\right.\\
&\quad\left.-\frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}+\frac{\cos\theta}{r}\frac{\partial^2 V}{\partial \theta^2}\right)\\
&=\sin^2\theta\frac{\partial^2 V}{\partial r^2}+\frac{\sin\theta\cos\theta}{r}\frac{\partial^2 V}{\partial r\partial \theta}+\frac{\cos\theta\sin\theta}{r^2}\frac{\partial V}{\partial \theta}\\
&\quad+\frac{\cos\theta\sin\theta}{r}\frac{\partial^2 V}{\partial \theta\partial r}+\frac{\cos^2\theta}{r^2}\cdot\frac{\partial^2 V}{\partial \theta^2}+\frac{\cos^2\theta}{r}\frac{\partial V}{\partial r}\\
&\quad-\frac{\sin\theta\cos\theta}{r^2}\frac{\partial V}{\partial \theta}.\qquad(2)\end{aligned}$$

Adding (1) and (2), we obtain

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}=\frac{\partial^2 V}{\partial r^2}+\frac{1}{r}\cdot\frac{\partial V}{\partial r}+\frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2},$$

which is the required result.

EXAMPLE 3.42

Transform the expression

$$\left(x\frac{\partial V}{\partial x}+y\frac{\partial V}{\partial y}\right)^2+(a^2-x^2-y^2)\left\{\left(\frac{\partial V}{\partial x}\right)^2+\left(\frac{\partial V}{\partial y}\right)^2\right\}$$

by substituting $x=r\cos\theta$ and $y=r\sin\theta$.

Solution. If V is a function of x, y, then

$$\frac{\partial V}{\partial r}=\frac{\partial V}{\partial x}\cdot\frac{\partial x}{\partial r}+\frac{\partial V}{\partial y}\cdot\frac{\partial y}{\partial r}=\frac{x}{r}\frac{\partial V}{\partial x}+\frac{y}{r}\frac{\partial V}{\partial y}$$

or

$$r\frac{\partial V}{\partial r}=x\frac{\partial V}{\partial x}+y\frac{\partial V}{\partial y}=\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}\right)V.$$

Thus,

$$r\frac{\partial}{\partial r}=x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}.$$

Similarly,

$$\frac{\partial}{\partial \theta}=x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}.$$

Now,

$$\frac{\partial V}{\partial x}=\frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial x}+\frac{\partial V}{\partial \theta}\cdot\frac{\partial \theta}{\partial x}=\cos\theta\frac{\partial V}{\partial r}-\frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\quad\text{and}$$

$$\frac{\partial V}{\partial y}=\sin\theta\frac{\partial V}{\partial r}+\frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}.$$

Therefore,

$$\left(\frac{\partial V}{\partial x}\right)^2+\left(\frac{\partial V}{\partial y}\right)^2=\left(\frac{\partial V}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial V}{\partial \theta}\right)^2,$$

and the given expression is equal to

$$\left(r\frac{\partial V}{\partial r}\right)^2+(a^2-r^2)\left[\left(\frac{\partial V}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial V}{\partial \theta}\right)^2\right]$$

$$=a^2\left(\frac{\partial V}{\partial r}\right)^2+\left(\frac{a^2}{r^2}-1\right)\left(\frac{\partial V}{\partial \theta}\right)^2.$$

EXAMPLE 3.43

If $x = r\cos\theta$ and $y = r\sin\theta$, prove that

$$(x^2-y^2)\left(\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}\right)+4xy\frac{\partial^2 u}{\partial x\partial y} = r^2\frac{\partial^2 u}{\partial r^2}-r\frac{\partial u}{\partial r}-\frac{\partial^2 u}{\partial\theta^2},$$

where u is any twice-differentiable function of x and y.

Solution. We have

$$\frac{\partial u}{\partial r}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial r} = \cos\theta\frac{\partial u}{\partial x}+\sin\theta\frac{\partial u}{\partial y}=\frac{x}{r}\frac{\partial u}{\partial x}+\frac{y}{r}\frac{\partial u}{\partial y}$$

and so,

$$r\frac{\partial u}{\partial r}=x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}. \qquad (1)$$

Therefore,

$$r\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) = \left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}\right)\left(x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}\right)$$

$$=x\frac{\partial}{\partial x}\left(x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}\right)+y\frac{\partial}{\partial y}\left(x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}\right)$$

$$=x^2\frac{\partial^2 u}{\partial x^2}+xy\frac{\partial^2 u}{\partial x\partial y}+xy\frac{\partial^2 u}{\partial y\partial x}+y^2\frac{\partial^2 u}{\partial y^2}+x\frac{\partial u}{\partial x}+y\frac{\partial u}{\partial y}$$

$$=x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}+x\frac{\partial y}{\partial x}+y\frac{\partial u}{\partial y}.$$

Therefore,

$$r^2\frac{\partial^2 u}{\partial r^2}-r\frac{\partial u}{\partial r}=x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}, \text{ using (1).} \qquad (2)$$

Again,

$$\frac{\partial u}{\partial\theta}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial\theta}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial\theta}=x\frac{\partial u}{\partial y}-y\frac{\partial u}{\partial x}.$$

Therefore,

$$\frac{\partial^2 u}{\partial\theta^2}=\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)\left(x\frac{\partial u}{\partial y}-y\frac{\partial u}{\partial x}\right)$$

$$=x\frac{\partial}{\partial y}\left(x\frac{\partial u}{\partial y}-y\frac{\partial u}{\partial x}\right)-y\frac{\partial}{\partial x}\left(x\frac{\partial u}{\partial y}-y\frac{\partial u}{\partial x}\right)$$

$$=x^2\frac{\partial^2 u}{\partial y^2}-2xy\frac{\partial^2 u}{\partial y\partial x}+y^2\frac{\partial^2 u}{\partial x^2}-x\frac{\partial u}{\partial x}-y\frac{\partial u}{\partial y}. \qquad (3)$$

Subtracting (3) from (2), we get the required result.

EXAMPLE 3.44

If $x = r\cos\theta$, $y = r\sin\theta$, and $z = f(x, y)$, prove that

$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial r}\cos\theta-\frac{1}{r}\frac{\partial z}{\partial\theta}\sin\theta.$$

Solution. Since $x = r\cos\theta$ and $y = r\sin\theta$, we have $r=\sqrt{x^2+y^2}$ and $\theta=\tan^{-1}\left(\frac{y}{x}\right)$. Therefore,

$$\frac{\partial r}{\partial x}=\frac{x}{r}=\frac{r\cos\theta}{r}=\cos\theta,$$

$$\frac{\partial r}{\partial y}=\frac{y}{r}=\frac{r\sin\theta}{r}=\sin\theta,$$

$$\frac{\partial\theta}{\partial x}=-\frac{y}{x^2+y^2}=\frac{-r\sin\theta}{r^2}=\frac{-\sin\theta}{r}, \text{ and}$$

$$\frac{\partial\theta}{\partial y}=\frac{y}{x^2+y^2}=\frac{r\cos\theta}{r^2}=\frac{\cos\theta}{r}.$$

Now,

$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial r}\cdot\frac{\partial r}{\partial x}+\frac{\partial z}{\partial\theta}\cdot\frac{\partial\theta}{\partial x} = \cos\theta\frac{\partial z}{\partial r}-\frac{\sin\theta}{r}\frac{\partial z}{\partial\theta}.$$

3.10 TAYLOR'S THEOREM FOR FUNCTIONS OF SEVERAL VARIABLES

In view of Taylor's theorem for functions of one variable, it is not unnatural to expect the possibility of expanding a function of more than one variable $f(x+h,\ y+k,\ z+\ l,\ldots)$ in a series of ascending powers of $h, k, l, \ldots$. To fix the ideas, we consider here a function of two variables only, the reasoning in the general case being precisely the same.

Theorem 3.5. (Taylor). If $f(x, y)$ and all its partial derivatives of order n are finite and continuous for all points (x,y) in the domain $a \le x \le a+h$ and $b \le y \le b+k$, then $f(a+h,\ b+k) = f(a,b) + +R_n$, $\frac{1}{2!}d^2f(a,b)+\ldots+\frac{1}{(n-1)}d^{n-1}f(a,b)+R_n$, where

$$R_n=\frac{1}{n!}d^nf(a+\theta h,\ b+\theta k),\ 0<\theta<1$$

and

$$df=h\frac{\partial f}{\partial x}+k\frac{\partial f}{\partial y}.$$

Proof: Consider a circular domain of center (a,b) and radius large enough for the point $(a+h, b+k)$ to be also with in the domain. The partial derivatives of the order n of $f(x, y)$ are continuous in the domain. Write

$$x=a+ht \text{ and } y=b+kt$$

so that as t ranges from 0 to 1, the point (x, y) moves along the line joining the point (a, b) to the point $(a+h,\ b+k)$. Then

$$f(x,y) = f(a+ht, b+kt) = \phi(t). \qquad (1)$$

Now,

$$\phi'(t) = \frac{\partial f}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial f}{\partial y}\cdot\frac{dy}{dt} = h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} = df.$$

Similarly,

$$\phi''(t) = d^2f, \ldots, \phi^{(n)}(t) = d^nf.$$

Thus, $\phi(t)$ and its n derivatives are continuous functions of t in the interval $0 \le t \le 1$ and so, by Maclaurin's theorem for function of one variable, we have

$$\phi(t) = \phi(0) + t\,\phi'(0) + \frac{t^2}{2!}\phi''(0) + \ldots + \frac{t^n}{n!}\phi^{(n)}(\theta\, t), \qquad (2)$$

where $0 < \theta < 1$. Further, the relation (1) yields

$$\phi(1) = f(a+h, b+k),$$
$$\phi(0) = f(a,b),$$
$$\phi'(0) = df(a,b),$$
$$\phi''(0) = d^2f(a,b), \text{ and}$$
$$\phi^{(n)}(\theta\, t) = d^nf(a+\theta\, h, b+\theta\, k).$$

Hence, (2) yields

$$f(a+h, b+k) = f(a,b) + df(a,b) + \frac{1}{2!}d^2f(a,b) + \ldots + \frac{1}{(n-1)!}d^{n-1}f(a,b) + R_n,$$

where

$$R_n = \frac{1}{n!}d^nf(a+\theta\, h,\ b+\theta\, k),\ 0 < \theta < 1.$$

Remark 3.2. Taylor's expansion does not necessarily hold if the partial derivatives of the order n of the function are not continuous in the domain. The theorem can be extended easily to any number of variables.

Theorem 3.6. (Maclaurin). Under the conditions of the Taylor's theorem

$$f(x,y) = f(0,0) + df(0,0) + \frac{1}{2!}d^2f(0,0) + \ldots + \frac{1}{(n-1)!}d^{n-1}f(0,0) + R_n,$$

where

$$R_n = \frac{1}{n!}d^nf(\theta\, x, \theta\, y),\ 0 < \theta < 1.$$

Proof: By Taylor's theorem, we have

$$f(a+h, b+k) = f(a,b) + df(a,b) + \frac{1}{2!}d^2f(a,b) + \ldots + \frac{1}{(n-1)!}d^{n-1}f(a,b) + R_n,$$

where

$$R_n = \frac{1}{n!}d^nf(a+\theta\, h, b+\theta\, k),\ 0 < \theta < 1.$$

Putting

$a = b = 0,\ h = x,$ and $k = y$, we get

$$f(x,y) = f(0,0) + df(0,0) + \frac{1}{2!}d^2f(0,0) + \ldots + \frac{1}{(n-1)!}d^{n-1}f(0,0) + R_n,$$

where

$$R_n = \frac{1}{n!}d^nf(\theta\, x, \theta\, y),\ 0 < \theta < 1.$$

EXAMPLE 3.45

If $f(x,y) = (|xy|)^{\frac{1}{2}}$, prove that the Taylor's expansion about the point (x, x) is not valid in any domain which includes the origin. Give reasons.

Solution. We are given that

$$f(x,y) = (|x\, y|)^{\frac{1}{2}}.$$

Therefore,

$$f_x(0,\ 0) = \lim_{h\to 0}\frac{f(h,0) - f(0,\ 0)}{h} = \lim_{h\to 0}\frac{0}{h} = 0,$$

$$f_y(0,\ 0) = \lim_{k\to 0}\frac{f(0,k) - f(0,\ 0)}{k} = \lim_{k\to 0}\frac{0}{k} = 0.$$

Also,

$$f_x(x,y) = \begin{cases} \frac{1}{2}\left(\left|\frac{y}{x}\right|\right)^{\frac{1}{2}}, & x > 0 \\ -\frac{1}{2}\left(\left|\frac{y}{x}\right|\right)^{\frac{1}{2}}, & x < 0 \end{cases}$$

and

$$f_y(x,y) = \begin{cases} \frac{1}{2}\left(\left|\frac{x}{y}\right|\right)^{\frac{1}{2}}, & y > 0 \\ -\frac{1}{2}\left(\left|\frac{x}{y}\right|\right)^{\frac{1}{2}}, & y < 0 \end{cases}$$

Thus,

$$f_x(x,x) = f_y(x,x) = \begin{cases} \frac{1}{2}, & x > 0 \\ -\frac{1}{2}, & x < 0. \end{cases}$$

If Taylor's expansion about (x,x) for $n = 1$ were possible, then we should have

$$f(x+h, x+h) = f(x,x) + h[f_x(x+\theta h, x+\theta h) + f_y(x+\theta h, x+\theta h)]$$

or

$$|x+h| = \left\{ \begin{array}{ll} |x| + h & \text{if } x+\theta h > 0 \\ |x| - h & \text{if } x-\theta h < 0 \\ |x| & \text{if } x-\theta h = 0 \end{array} \right\}. \quad (1)$$

Now if the domain $(x, x; x+h, x+h)$ includes the origin, then x and $x+h$ are of opposite signs. Thus, either $|x+h| = x+h$, $|x| = -x$ or $|x+h| = -(x\ h)$, $|x| = x$. But under these conditions, none of the inequalities (1) holds. Hence, the expansion is not valid.

EXAMPLE 3.46

Expand $x^2y + 3y - 2$ in powers of (x – 1) and$(y+2)$ using Taylor's theorem for several variables.

Solution. For all points in the domain $a \le x \le a+h$ and $b \le y \le b+k$, the Taylor's theorem asserts that

$$f(x,y) = f(a+h,\ b+k) = f(a,b) + df(a,b) + \frac{1}{2!}d^2f(a,b) + \frac{1}{3!}d^3f(a,b) + \frac{1}{4!}d^4f(a,b) + \ldots$$

In the present example, $a = 1$ and $b = -2$. Thus,

$f_x(x,y) = x^2y + 3y - 2$ which yields $f(1,-2) = -10$,

$f_x(x,y) = 2xy$ which yields $f_x(1,-2) = -4$,

$f_y(x,y) = x^2 + 3$ which yields $f_y(1,-2) = 4$,

$f_{xx}(x,y) = 2y$ which yields $f_{xx}(1,-2) = -4$,

$f_{xy}(x,y) = 2x$ which yields $f_{xy}(1,-2) = 2$,

$f_{yy}(x,y) = 0$ which yields $f_{yy}(1,-2) = 0$,

$f_{xxx}(x,y) = 0$ which yields $f_{xxx}(1,-2) = 0$,

$f_{yyy}(x,y) = 0$ which yields $f_{yyy}(1,-2) = 0$,

and

$$f_{yxx}(1,-2) = f_{xxy(}(1,-2) = 2.$$

All other higher derivatives are zero. Hence,

$$\begin{aligned} f(x,y) &= f(a,b) + \left[(x-1)\frac{\partial f}{\partial x}(1,-2) + (y+2)\frac{\partial f}{\partial y}(1,-2)\right] \\ &\quad + \frac{1}{2\,!}\left[(x-1)\frac{\partial}{\partial x} + (y+2)\frac{\partial}{\partial y}\right]^2 f(1,-2) \\ &\quad + \frac{1}{3\,!}\left[(x-1)\frac{\partial}{\partial x} + (y+2)\frac{\partial}{\partial y}\right]^3 f(1,-2) \\ &= -10 - 4(x-1) + 4(y+2) \\ &\quad + \frac{1}{2\,!}[-4(x-1)^2 + 4(x-1)(y+2)] \\ &\quad + \frac{1}{3\,!}[(x-1)^3(0) + 3(x-1)^2(y+2)(2) \\ &\quad + 3(x-1)(y+2)^2(0) + (y+2)^2(0)] \\ &= -10 - 4(x-1) + 4(y+2) - 2(x-1)^2 \\ &\quad + 2(x-1)(y+2) + (x-1)^2(y+2). \end{aligned}$$

EXAMPLE 3.47

Expand $\sin(xy)$ in power of $(x-1)$ and $\left(y - \frac{\pi}{2}\right)$ up to and including second-degree terms.

Solution. We want to expand $\sin(xy)$ about the point $\left(1, \frac{\pi}{2}\right)$. By Taylor's theorem, we have

$$\begin{aligned} f(x,y) &= f(a+h,\ b+k) \\ &= f(a,b) + \left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right] f(a,b) \\ &\quad + \frac{1}{2}\left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right]^2 f(a,b) + \ldots \end{aligned}$$

But

$f(x,y) = \sin xy$ implies $f\left(1, \frac{\pi}{2}\right) = 1$,

$f_x(x,y) = y\cos xy$ implies $f_x\left(1, \frac{\pi}{2}\right) = 0$,

$f_y(x,y) = x\cos xy$ implies $f_y\left(1, \frac{\pi}{2}\right) = 0$,

$f_{xx}(x,y) = -y^2 \sin xy$ implies $f_{xx}\left(1, \frac{\pi}{2}\right) = -\frac{\pi^2}{4}$,

$f_{xy}(x,y) = -xy \sin xy$ implies $f_{xy}\left(1, \frac{\pi}{2}\right) = \left(y - \frac{\pi}{2}\right)$,

and

$f_{yy}(x,y) = -x^2 \sin xy$ implies $f_{yy}\left(1, \frac{\pi}{2}\right) = -1$.

Hence,

$$f(x,y) = 1 - \frac{\pi^2}{8}(x-1)^2 - \frac{\pi}{2}(x-1)\left(y-\frac{\pi}{2}\right)$$
$$-\frac{1}{2}\left(y-\frac{\pi}{2}\right)^2.$$

EXAMPLE 3.48

Expand e^{xy} at (1,1) in powers of $(x-1)$ and $(y-1)$.

Solution. We have $f(x, y) = e^{xy}$. By Taylor's theorem, we have

$$f(x,y) = f(a+h,\ b+k)$$
$$= f(a,b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(a,b)$$
$$+ \frac{1}{2\,!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a,b) + \dots .$$

But,

$f(x,y) = e^{xy}$ implies $f(1,1) = e$,
$f_x(x,y) = ye^{xy}$ implies $f_x(1,1) = e$,
$f_y(x,y) = xe^{xy}$ implies $f_y(1,1) = e$,
$f_{xx}(x,y) = y^2e^{xy}$ implies $f_{xx}(1,1) = e$,
$f_{xy}(x,y) = xye^{xy} + e^{xy}$ implies $f_{xy}(1,1) = e + e = 2e$, and
$f_{yy}(x,y) = x^2e^{xy}$ implies $f_{yy}(1,1) = e$.

We have $h = x - a = x - 1$ and $k = y - b = y - 1$. Hence,

$$f(x,y) = f(1,\ 1) + (x-a)f_x(1,\ 1) + (y-k)$$
$$\times f_y(1,\ 1) + \frac{1}{2\,!}[(x-1)^2 f_{xx}(1,1)$$
$$+ (y-1)^2 f_{yy}(1,1) + 2(x-1)(y-1)$$
$$\times f_{xy}(1,\ 1)$$
$$= e + (x-1)e + (y-1)e + \frac{1}{2!}[(x-1)^2 e$$
$$+ 4(x-1)(y-1)e + (y-1)^2 e] + \dots$$
$$= e\left\{1 + (x-1) + (y-1) + \frac{1}{2!}[(x-1)^2\right.$$
$$\left. + 4(x-1)(y-1) + (y-1)^2\right\} + \dots$$

EXAMPLE 3.49

Expand $e^{ax}\sin by$ in power of x and y as far as terms of third degree.

Solution. We have $f(x, y) = e^{ax}\sin by$. By Taylor's theorem for function of two variables,

$$f(x,y) = f(a+h, b+k)$$
$$= f(a,b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(a,b)$$
$$+ \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a,b)$$
$$+ \frac{1}{3!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^3 f(a,b) + \dots .$$

We wish to expand the function about (0, 0). So $h = x - 0 = x$ and $k = y - 0 = y$.

$$f(x,y) = f(0+h, 0+k)$$
$$= f(0,0) + \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right) f(0,0)$$
$$+ \frac{1}{2!}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^2 f(0,0)$$
$$+ \frac{1}{3!}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^3 f(0,0) + \dots$$

But,

$f(x,y) = e^{ax}\sin\ by$ implies $f(0,0) = 0$,
$f_x(x,y) = ae^{ax}\sin by$ implies $f_x(0,0) = 0$,
$f_y(x,y) = be^{ax}\cos\ by$ implies $f_y(0,0) = b$,
$f_{xx}(x,y) = a^2e^{ax}\sin\ by$ implies $f_{xx}(0,0) = 0$,
$f_{xy}(x,y) = abe^{ax}\cos\ by$ implies $f_{xy}(0,0) = ab$,
$f_{yy}(x,y) = -b^2e^{ax}\sin\ by$ implies $f_{yy}(x,y) = 0$,
$f_{xxx}(x,y) = a^3e^{ax}\sin\ by$ implies $f_{xxx}(0,0) = 0$,
$f_{xxy}(x,y) = a^2be^{ax}\cos\ by$ implies $f_{xxy}(0,0) = a^2b$,
$f_{xyy}(x,y) = -b^2ae^{ax}\sin\ by$ implies $f_{xyy}(0,0) = 0$,
$f_{yyy}(x,y) = -b^3e^{ax}\cos\ by$ implies $f_{yyy}(0,0) = -b^3$,

and so on. Hence,

$$f(x,y) = by + abxy + \frac{1}{3!}(3a^2bx^2y - b^3y^3) + \dots$$

EXAMPLE 3.50

Expand $f(x,y) = \tan^{-1}\frac{y}{x}$ in the neighborhood of (1,1) up to third-degree terms. Hence, compute $f(1.1,0.9)$ approximately.

Solution. We note that

$$f(x,y) = \tan^{-1}\frac{y}{x} \text{ implies } f(1,1) = \frac{\pi}{4},$$

$$f_x(x,y) = \frac{-y}{x^2+y^2} \text{ implies } f_x(1,1) = -\frac{1}{2},$$

$$f_y(x,y) = \frac{x}{x^2+y^2} \text{ implies } f_y(1,1) = \frac{1}{2},$$

$$f_{xx}(x,y) = \frac{2xy}{(x^2+y^2)^2} \text{ implies } f_x(1,1) = \frac{1}{2},$$

$$f_{xy}(x,y) = \frac{y^2-x^2}{(x^2+y^2)^2} \text{ implies } f_{xy}(1,1) = 0,$$

$$f_{yy}(x,y) = -\frac{2xy}{(x^2+y^2)^2} \text{ implies } f_{yy}(1,1) = -\frac{1}{2},$$

$$f_{xxx}(x,y) = \frac{2y(y^2-3x^2)}{(x^2+y^2)^3} \text{ implies } f_{xxx}(1,1) = -\frac{1}{2},$$

$$f_{xxy}(x,y) = \frac{2x(x^2-3y^2)}{(x^2+y^2)^3} \text{ implies } f_{xxy}(1,1) = -\frac{1}{2},$$

$$f_{xyy}(x,y) = \frac{2y(3x^2-y^2)}{(x^2+y^2)^3} \text{ implies } f_{xyy}(1,1) = \frac{1}{2}, \text{ and}$$

$$f_{yyy}(x,y) = \frac{2x(3y^2-x^2)}{(x^2+y^2)^3} \text{ implies } f_{yyy}(1,1) = \frac{1}{2}.$$

Therefore, by Taylor's theorem, we have

$$\begin{aligned}
f(x,y) &= \tan^{-1}\frac{y}{x}\\
&= f(1,1) + \left[(x-1)\frac{\partial}{\partial x} + (y-1)\frac{\partial}{\partial y}\right] f(1,1)\\
&\quad + \frac{1}{2!}\left[(x-1)\frac{\partial}{\partial x} + (y-1)\frac{\partial}{\partial y}\right]^2 f(1,1)\\
&\quad + \frac{1}{3!}\left[(x-1)\frac{\partial}{\partial x} + (y-1)\frac{\partial}{\partial y}\right]^3 f(1,1) + \ldots\\
&= \frac{\pi}{4} + \left[(x-1)\left(-\frac{1}{2}\right) + (y-1)\left(\frac{1}{2}\right)\right]\\
&\quad + \frac{1}{2}\left[(x-1)^2\left(\frac{1}{2}\right) + 2(x-1)(y-1)(0)\right.\\
&\quad \left. + (y-1)^2\left(-\frac{1}{2}\right)\right]\\
&\quad + \frac{1}{3!}\left[(x-1)^3\left(-\frac{1}{2}\right) + 3(x-1)^2(y-1)\left(-\frac{1}{2}\right)\right.\\
&\quad \left. + 3(x-1)(y-1)^2\left(\frac{1}{2}\right) + (y-1)^3\left(\frac{1}{2}\right)\right] + \ldots\\
&= \frac{\pi}{4} - \frac{1}{2}[(x-1)-(y-1)] + \frac{1}{4}[(x-1)^2 + (y-1)^2]\\
&\quad - \frac{1}{12}\left[(x-1)^3 + 3(x-1)^2(y-1)\right.\\
&\quad \left. - 3(x-1)(y-1)^3 - (y-1)^3\right] + \ldots.
\end{aligned}$$

Putting $x = 1.1$ and $y = 0.9$, we get

$$f(1.1, 0.9) = 0.6857.$$

3.11 EXTREME VALUES

A function $f(x, y)$ of two independent variables x and y is said to have an *extreme value* at the point (a, b) if the increment

$$\Delta f = f(a+h, b+k) - f(a,b)$$

preserves the same sign for all values of h and k whose moduli do not exceed a sufficiently small positive number η.

If Δf is negative, then the extreme value is a *maximum* and if Δf is positive, then the extreme value is a *minimum.*

Necessary and sufficient conditions for extreme values

By Taylor's theorem, we have

$$\begin{aligned}
\Delta f &= f(a+b, b+k) - f(a,b)\\
&= h\frac{\partial f}{\partial x}(a,b) + k\frac{\partial f}{\partial y}(a,b)\\
&\quad + \text{terms of second and higher order.}
\end{aligned}$$

Now by taking h and k sufficiently small, the first-order terms can be made to govern the sign of the right-hand side and therefore, of the left-hand side of the previous expansion. Hence, the change in the sign of h and k would change the sign of the left-hand side, that is, of Δf. But if the sign of Δf changes, $f(x, y)$ cannot have an extreme point at (a, b). Hence, as a first condition for the extreme value, we must have

$$h\frac{\partial f}{\partial x}(a,b) + k\frac{\partial f}{\partial y}(a,b) = 0.$$

Since the arbitrary increments h and k are independent of each other, we must have

$$\frac{\partial f}{\partial x}(a,b) = 0 \text{ and } \frac{\partial f}{\partial y}(a,b) = 0,$$

which are necessary conditions for the existence of extreme points. However, these are not sufficient conditions for the existence of extreme points.

Further, a point (a, b) is called a *stationary point* if $f_x(a, b) = f_y(a, b) = 0$. The value $f(a, b)$ is called a *stationary value*.

To find sufficient conditions, let (a, b) be an interior point of the domain of f such that f admits the second-order continuous partial derivatives in the neighborhood of (a, b). Suppose that $f_x(a, b)$ $f_y(a, b) = 0$. We further, suppose that

$$r = \frac{\partial^2 f}{\partial x^2}, \quad s = \frac{\partial^2 f}{\partial x \partial y}, \text{ and}$$

$$t = \frac{\partial^2 f}{\partial y^2}, \text{ when } x = a \text{ and } y = b.$$

Thus,

$$f_{xx}(a,b) = r,\ f_{xy}(a,b) = s, \text{ and } f_{yy}(a,b) = t.$$

If $(a+b, b+k)$ is any point in the neighborhood of (a, b), then by Taylor's theorem, we have

$$\begin{aligned}\Delta f &= f(a+b, b+k) - f(a,b) \\ &= hf_x(a,b) + kf_y(a,b) + \frac{1}{2}[h^2 f_{xx}(a,b) \\ &\quad + 2hkf_{xy}(a,b) + k^2 f_{yy}(a,b)] \ + R_3 \\ &= \frac{1}{2}[rh^2 + 2hks + tk^2] + R_3,\end{aligned}$$

where R_3 consists of terms of third and higher orders of small quantities. Thus, by taking h and k sufficiently small, now the second order terms can be made to govern the sign of the right-hand side and therefore, of the left-hand side of the previous expansion. But

$$\begin{aligned}\frac{1}{2}[rh^2 + 2hks + tk^2] &= \frac{1}{2r}[r^2h^2 + 2hkrs + rtk^2] \\ &= \frac{1}{2r}[r^2h^2 + 2hkrs \\ &\quad + rtk^2 + k^2s^2 - h^2s^2] \\ &= \frac{1}{2r}[(rh+sk)^2 + k^2(rt - s^2)].\end{aligned}$$

Since $(rh+sk)^2$ is always positive, it follows that Δf is positive if $rt - s^2$ is positive. Now $rt - s^2 > 0$ if both r and t have the same sign. Thus, the sign of Δf shall be that of r. Therefore, if $rt - s^2$ is positive, we have a maximum or a minimum accordingly, as both r and t are either negative or positive. This condition was first pointed out by Lagrange and is known as *Lagrange's condition*. However, if $rt = s^2$, then $rh^2 + 2hks + tk^2$ becomes $\frac{1}{r}(hr + ks)^2$ and is, therefore, of the same sign as r or t unless $\frac{h}{k} = -\frac{s}{r}$, say, for which $(hr + ks)^2$ vanishes.

In such a case, we must consider terms of higher order in the expansion of $f(a+h, b+k)$. Thus, we may state that

1. The value $f(a, b)$ is an extreme value of $f(x, y)$ if $f_x(a, b) = f_y(a, b) = 0$ and if $rt - s^2 > 0$. The value is *maximum* or *minimum* accordingly as $f_{xx}(a, b)$ or $f_{yy}(a, b)$ is negative or positive.
2. If $rt - s^2 < 0$, then $f(x, y)$ has no extreme value at (a, b). The point (a, b) is a *saddle point* in this case.
3. If $rt - s^2 = 0$, the case is doubtful and requires terms of higher order in the expansion of the function.

EXAMPLE 3.51

Show that the function $f(x, y) = y^2 + x^2y + x^4$ has a minimum value at the origin.

Solution. We have

$$f(x,\ y) = y^2 + x^2y + x^4.$$

Therefore,

$$\begin{aligned} f_x &= 2xy + 4x^3 \text{ which yields } f_x(0,0) = 0, \\ f_y &= 2y + x^2 \text{ which yields } f_y(0,0) = 0, \\ f_{xx} &= 2y + 12x^2 \text{ which yields } f_{xx}(0,0) = 0, \\ f_{yy} &= 2 \text{ which yields } f_{yy}(0,0) = 2, \text{ and} \\ f_{xy} &= 2x \text{ which yields } f_{xy}(0,0) = 0. \end{aligned}$$

Hence, at the origin, we have $rt - s^2 = 0$. Thus, further investigation is needed in the case. We write

$$f(x,y) = y^2 + x^2y + x^4 = \left(y + \frac{1}{2}x^2\right)^2 + \frac{3x^4}{4}.$$

Then,

$$\Delta f = f(h,k) - f(0,0) = \left(k + \frac{h^2}{2}\right)^2 + \frac{3h^2}{4},$$

which is always greater than zero for all values of h and k. Hence, $f(x, y)$ has a *minimum value at the origin*.

EXAMPLE 3.52

Show that the function

$$u = xy + \frac{a^3}{x} + \frac{a^3}{y}$$

has a minimum value at (a, a).

Solution. We have

$$u = xy + \frac{a^3}{x} + \frac{a^3}{y}.$$

Therefore,

$$\frac{\partial u}{\partial x} = y - \frac{a^3}{x^2} \text{ yields } f_x(a,a) = 0,$$

$$\frac{\partial u}{\partial y} = x - \frac{a^3}{y^2} \text{ yields } f_y(a,a) = 0,$$

$$\frac{\partial^2 u}{\partial x^2} = \frac{2a^3}{x^3} \text{ yields } f_{xx}(a,a) = 2,$$

$$\frac{\partial^2 u}{\partial x \partial y} = 1 \text{ yields } f_{xy}(a,a) = 1, \text{ and}$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{2a^3}{y^3} \text{ and so, } f_{yy}(a,a) = 2.$$

We observe that $rt - s^2 = 4 - 1 = 3$ (positive) and r and t too positive. Therefore, u has the minimum at (a, a). Thus, the minimum value of u is

$$u\,(a, a) = a^2 + a^2 + a^2 = 3a^2.$$

EXAMPLE 3.53

Show that the function $f\,(x, y) = 2x^4 - 3x^2\,y + y^2$ does not have a maximum or a minimum at (0, 0).

Solution. The given function is

$$f(x, y) = 2x^4 - 3x^2y + y^2.$$

Therefore,

$$\frac{\partial f}{\partial x} = 8x^3 - 6xy, \text{ which implies } f_x(0,0) = 0,$$

$$\frac{\partial f}{\partial y} = -3x^2 + 2y, \text{ which implies } f_y(0,0) = 0,$$

$$\frac{\partial^2 f}{\partial x^2} = 24x^2 - 6y, \text{ which implies } f_{xx}(0,0) = 0,$$

$$\frac{\partial^2 f}{\partial x \partial y} = -6x, \text{ which implies } f_{xy}(0,0) = 0, \text{ and}$$

$$\frac{\partial^2 f}{\partial y^2} = 2, \text{ which implies } f_{yy}(0,0) = 2.$$

Thus, $rt - s^2 = 0$ and so, further investigation is required. We have

$$f(x,y) = (x^2 - y)(2x^2 - y),\ f(0,\ 0) = 0.$$

Therefore,

$$\Delta f = f(x,y) - f(0,0) = (x^2 - y)(2x^2 - y).$$

Thus, Δf is positive, for $y < 0$ or $x^2 > y > 0$ and Δf is negative, for $y > x^2 > \frac{y}{2} > 0$. Thus, Δf does not keep the same sign in the neighborhood of (0, 0). Hence, the function does not have a maximum or a minimum at (0, 0).

EXAMPLE 3.54

Examine the function $\sin x + \sin y + \sin(x + y)$ for extreme points.

Solution. The given function is $f\,(x, y) = \sin x + \sin y + \sin\,(x + y)$. Therefore,

$$f_x = \cos x + \cos(x+y),$$

$$f_y = \cos y + \cos(x+y),$$

$$f_{xx} = -\sin x - \sin(x+y),$$

$$f_{xy} = -\sin(x+y), \text{ and}$$

$$f_{yy} = -\sin y - \sin(x+y).$$

For extreme points, we must have $f_x = f_y = 0$ and so,

$$\cos x + \cos(x+y) = 0 \qquad (1)$$

and

$$\cos y + \cos(x+y) = 0 \qquad (2)$$

Subtracting (2) from (1), we get $\cos x = \cos y$ and so, $x = y$. Also then, $\cos x + \cos 2x = 0$ which yields $\cos 2x = -\cos x = \cos(\pi - x)$ and so, $2x = \pi - x$ or $x = \frac{\pi}{3}$. Thus, $\left(\frac{\pi}{3}, \frac{\pi}{3}\right)$ is a stationary point. Now

$$r = f_{xx}\left(\frac{\pi}{3}, \frac{\pi}{3}\right) = -\frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} = -\sqrt{3} \text{ (negative)},$$

$$s = f_{xy}\left(\frac{\pi}{3}, \frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}$$

$$t = f_{yy}\left(\frac{\pi}{3}, \frac{\pi}{3}\right) = -\frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} = -\sqrt{3} \text{ (negative)}.$$

Thus,

$$rt - s^2 = 3 - \frac{3}{4} = \frac{9}{4} \text{ (positive)}$$

and r is negative. Hence, the given function has a maximum value at $\left(\frac{\pi}{3}, \frac{\pi}{3}\right)$ given by

$$f\left(\frac{\pi}{3}, \frac{\pi}{3}\right) = \sin\frac{\pi}{3} + \sin\frac{\pi}{3} + \sin\frac{2\pi}{3}$$

$$= \frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2} = \frac{3\sqrt{3}}{2}.$$

EXAMPLE 3.55

Examine the following surface for high- and low points:

$$z = x^2 + xy + 3x + 2y + 5.$$

Solution. We have

$$\frac{\partial z}{\partial x} = 2x + y + 3, \quad \frac{\partial z}{\partial y} = x + 2,$$

$$\frac{\partial^2 z}{\partial x^2} = 2, \quad \frac{\partial^2 z}{\partial x \partial y} = 1, \text{ and } \frac{\partial^2 z}{\partial y^2} = 0.$$

For an extreme point, we must have $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0$ and so,

$$2x + y + 3 = 0 \text{ and } x + 2 = 0.$$

Solving these equations, we get $x = -2$, $y = 1$. Thus, z can have a maximum or a minimum only at (–2, 1). Further,

$$r = \frac{\partial^2 z}{\partial x^2}(-2, 1) = 2,$$

$$s = \frac{\partial^2 z}{\partial x \partial y}(-2, 1) = 1, \text{ and}$$

$$t = \frac{\partial^2 z}{\partial y^2}(-2, 1) = 0.$$

Therefore, $rt - s^2 = -1$ (negative) and so, the stationary value of z at (–2, 1) is neither a maximum nor a minimum. Hence, the surface has no high- or low point.

EXAMPLE 3.56

Locate the stationary points of $x^4 + y^4 - 2x^2 + 4xy - 2y^2$ and determine their nature.

Solution. We have

$$f(x,y) = x^4 + y^4 - 2x^2 + 4xy - 2y^2.$$

Therefore,

$$f_x = 4x^3 - 4x + 4y \text{ and } f_y = 4y^3 + 4x - 4y.$$

The stationary points are given by

$$f_x = 4x^3 - 4x + 4y = 0 \quad (1)$$

$$f_y = 4y^3 + 4x - 4y = 0 \quad (2)$$

Adding (1) and (2), we get

$$x^3 + y^3 = 0 \text{ or } (x + y)(x^2 - xy + y^2) = 0.$$

Therefore, either $y = -x$ or $x^2 - xy + y^2 = 0$. Putting $y = -x$ in (1), we get $x(x^2 - 2) = 0$, which yields $x = 0$, $\sqrt{2}$, or $-\sqrt{2}$. The value of y corresponding to these values are 0, $-\sqrt{2}$, and $\sqrt{2}$. Thus, the points $(0,0)$, $(\sqrt{2}, -\sqrt{2})$, and $(-\sqrt{2}, \sqrt{2})$ satisfy (1) and (2). On the other hand, from equation (1) and $x^2 - xy + y^2 = 0$ we get (0, 0) as the only real root. Thus, the stationary points are

$$(0,0), \; (\sqrt{2}, -\sqrt{2}), \text{ and } (-\sqrt{2}, \sqrt{2}).$$

Also,

$$f_{xx} = 12x^2 - 4, \; f_{xy} = 4, \text{ and } f_{yy} = 12y^2 - 4.$$

At (0,0), we have

$$r = f_{xx}(0,0) = -4,$$

$$s = f_{xy}(0,0) = 4, \text{ and}$$

$$t = f_{yy}(0,0) = -4,$$

and so, $rt - s^2 = 0$. Thus, at (0, 0), the case is doubtful. The given equation can be written as

$$f(x,y) = x^4 + y^4 - 2(x - y)^2.$$

So,

$$f(0,0) = 0 \text{ and } f(h,k) = h^4 + k^4 - 2(h - k)^2.$$

We observe that for small quantities of h and k,

$$\Delta f = f(h,k) - f(0,0) = h^4 + k^4 - 2(h - k)^2$$

is greater than 0, if $h = k$ and less than 0, if h $h \neq k$. Since Δf does not preserve the sign, the function has no extreme value at the origin.

At $(\sqrt{2}, \; -\sqrt{2})$, we have $r = 20$, $s = 4$, and $t = 0$ so that $rt - s^2 = 384$ (positive). Since r is positive, $f(x, y)$ has a minimum at this point.

At $(-\sqrt{2}, \; \sqrt{2})$, we have $r = 20$, $s = 4$, and $t = 20$. Thus, $rt - s^2$ is positive. Since r is positive, $f(x, y)$ has a minimum at $(-\sqrt{2}, \; \sqrt{2})$ also.

EXAMPLE 3.57

Find the minimum value of $x^2 + y^2 + z^2$ when $ax + by + cz = p$.

Solution. Let $f(x, y, z) = x^2 + y^2 + z^2$. From the relation $ax + by + cz = p$, we get $z = \frac{p - ax - by}{c}$. Putting this value of z in $f(x, y, z)$, we get

$$f(x,y,z) = x^2 + y^2 + \left(\frac{p - ax - by}{c}\right)^2$$

as a function of two variables x and y. Then,

$$f_x = 2x - \frac{2a}{c^2}(p - ax - by) \text{ and}$$

$$f_y = 2y - \frac{2b}{c^2}(p - ax - by).$$

For extreme points, we must have $f_x = f_y = 0$. Thus,

$$2x - \frac{2a}{c^2}(p - ax - by) = 0 \text{ and}$$

$$2y - \frac{2b}{c^2}(p - ax - by) = 0.$$

Solving these equations, we get

$$x = \frac{ap}{a^2 + b^2 + c^2} \text{ and } y = \frac{bp}{a^2 + b^2 + c^2}.$$

Now,

$$f_{xx} = 2 + \frac{2a^2}{c^2},\ f_{xy} = \frac{2ab}{c^2},\ \text{and } f_{yy} = 2 + \frac{2b^2}{c^2},$$

so that

$$rt - s^2 = 4\left(1 + \frac{a^2}{c^2}\right)\left(1 + \frac{b^2}{c^2}\right) - \frac{4a^2b^2}{c^4}$$

$$= 4\left(1 + \frac{a^2}{c^2} + \frac{b^2}{c^2}\right) \text{ (positive).}$$

Also $r = f_{xx}$ is positive. Therefore, $f(x, y)$ has a minimum at $\left(\frac{ap}{a^2+b^2+c^2}, \frac{bp}{a^2+b^2+c^2}\right)$ and the minimum value is

$$\text{Min. } f(x,y,z) = \frac{p^2}{a^2 + b^2 + c^2}.$$

EXAMPLE 3.58

Show that the function

$$f(x,y) = x^2 - 2xy + y^2 + x^3 - y^3 + x^5$$

has neither a maximum nor a minimum at (0, 0).

Solution. For the given function,

$$f_x = 2x - 2y + 3x^2 + 5x^4 \text{ and } f_y = -2x + 2y - 3y^2.$$

$$f_{xx} = 2 + 6x + 20x^3 \text{ and } f_{xy} = -2,\ f_{yy} = 2 - 6y.$$

For a stationary value of $f(x, y)$, we must have f_x $f_y = 0$. Thus,

$$2x - 2y + 3y^2 + 5x^4 = 0 \text{ and } -2x + 2y - 3y^2 = 0.$$

The origin (0,0) satisfies these equations. Further, $r = f_{xx}(0,0) = 2$, $s = f_{xy}(0,0) = -2$, $t = f_{yy}(0,0) = 2$, and so, $rt - s^2 = 0$. Hence, further investigations are required. We rewrite the equation as

$$f(x,y) = (x-y)^2 + (x-y)(x^2 + xy + y^2) + x^5.$$

We note that $f(0, 0) = 0$. But,

$$\Delta f = f(h,k) - f(0,0) = f(h,k)$$

$$= (h-k)^2 + (h-k)(h^2 + hk + k^2) + k^5.$$

In the neighborhood of (0, 0), if $h = k$, then $\Delta f = k^5$; which is positive, when $k > 0$ and negative, when $k < 0$. Thus, Δf does not keep the same sign in the neighborhood of (0, 0). Hence, $f(x, y)$ cannot have a maximum or a minimum at the point (0, 0).

EXAMPLE 3.59

Find the dimensions of the rectangular box, open at the top, of maximum capacity whose surface is 432 sq. cm.

Solution. Let x, y, and z cm be the dimensions of the box and S be its surface. Then

$$S = xy + 2yz + 2zx = 432 \text{ (given)} \quad (1)$$

and

$$V = xyz. \quad (2)$$

We have to maximize V. From (1), we have

$$z = \frac{432 - xy}{2y + 2x}. \quad (3)$$

Therefore, (2) reduces to

$$V = xy\left(\frac{432 - xy}{2y + 2x}\right) = \frac{432xy - x^2y^2}{2y + 2x}.$$

Now,

$$\frac{\partial V}{\partial x} = \frac{(2y + 2x)(432y - 2xy^2) - 2(432xy - x^2y^2)}{(2y + 2x)^2}$$

$$= \frac{864y^2 - 4xy^3 - 2x^2y^2}{(2x + 2y)^2},$$

$$\frac{\partial V}{\partial y} = \frac{(2x + 2y)(432x - 2x^2y) - 2(432xy - x^2y^2)}{(2y + 2x)^2}$$

$$= \frac{864x^2 - 4x^3y - 2x^2y^2}{(2x + 2y)^2}.$$

For stationary points, we must have $\frac{\partial V}{\partial x} = \frac{\partial V}{\partial y} = 0$. So,

$$864 - 4xy - 2x^2 = 0 \quad (4)$$

$$864 - 4xy - 2y^2 = 0. \quad (5)$$

Subtracting (5) from (4), we get $y = \pm x$. Substituting $x = y$ in (5), we get

$$864 - 4y^2 - 2y^2 = 0 \text{ or } y^2 = \frac{864}{6} = 124.$$

Thus, $x = y = 12$ and (3) implies z = 6. It can be verified that $rt - s^2 > 0$ and that r is positive for these values. Hence, the dimensions of the box are x = y = 12*cm* and z = 6*cm*.

EXAMPLE 3.60

Examine $x^3y^2(1 - x - y)$ for extreme points.

Solution. We have

$$f(x,y) = x^3y^2(1 - x - y).$$

Therefore,

$$\frac{\partial f}{\partial x} = 3x^2y^2(1 - x - y) + x^2y^2(-1)$$

$$= 3x^2y^2 - 4x^3y^2 - 3x^2y^3 \text{ and}$$

$$\frac{\partial f}{\partial y} = 2x^3y(1 - x - y) + x^2y^2(-1)$$

$$= 2x^3y - 2x^4y - 3x^3y^2.$$

For a maximum or a minimum of f, we must have $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$. Therefore,

$x^2y^2(3 - 4x - 3y) = 0$ and $x^3y(2 - 2x - 3y) = 0.$

Solving these equations, we get the stationary points (0, 0) and $\left(\frac{1}{2}, \frac{1}{3}\right)$. Further,

$$r = \frac{\partial^2 f}{\partial x^2} = 6xy^2 - 12x^2y^2 - 6xy^3 = 6xy^2(1 - 2x - y),$$

$$s = \frac{\partial^2 f}{\partial x \partial y} = x^2y(6 - 8x - 9y), \text{ and}$$

$$t = \frac{\partial^2 f}{\partial y^2} = 2x^3(1 - x - 3y).$$

Therefore,

(i) at (0,0), $r = 0$, $t = 0$, and $s = 0$, and so, $rt - s^2 = 0$.

But,

$$\Delta f = f(h,k) - f(0,0) = h^3k^2(1 - h - k).$$

Sign is governed by h^3k^2 which is positive, if $h > 0$ and negative, if $h < 0$. Since Δf does not keep the same sign in the neighborhood of (0, 0), the given function does not have a maximum or a minimum value at (0, 0).

(ii) at $\left(\frac{1}{2}, \frac{1}{3}\right)$, we have

$$r = \frac{\partial^2 f}{\partial x^2}\left(\frac{1}{2}, \frac{1}{3}\right) = -\frac{1}{9},$$

$$s = \frac{\partial^2 f}{\partial x \partial y}\left(\frac{1}{2}, \frac{1}{3}\right) = \frac{-1}{12}, \text{ and}$$

$$t = \frac{\partial^2 f}{\partial y^2}\left(\frac{1}{2}, \frac{1}{3}\right) = -\frac{1}{8}.$$

Therefore,

$$rt - s^2 = \left(-\frac{1}{9}\right)\left(-\frac{1}{8}\right) - \left(\frac{-1}{12}\right)^2 = \frac{1}{72} - \frac{1}{144}$$

$$= \frac{1}{144} \text{ (positive)}.$$

But r is negative. Hence, $f(x, y)$ has a maximum at $\left(\frac{1}{2}, \frac{1}{3}\right)$. The maximum value is $f\left(\frac{1}{2}, \frac{1}{3}\right) = \frac{1}{8} \cdot \frac{1}{9}\left(1 - \frac{1}{2} - \frac{1}{3}\right) = \frac{1}{432}$.

EXAMPLE 3.61

Find the points where the function $x^3 + y^3 - 3axy$ has a maximum or a minimum.

Solution. We have

$$f_x = 3x^2 - 3ay, \; f_y = 3y^2 - 3ax,$$

$$f_{xx} = 6x, \; f_{yy} = 6y, \text{ and } f_{xy} = -3a.$$

For extreme points, we have $f_x = f_y = 0$ and so,

$$3x^2 - 3ay = 0 \text{ and } 3y^2 - 3ax = 0.$$

Solving the earlier equations, we get two stationary points (0, 0) and (a, a). Further,

$$rt - s^2 = 36xy - 9a^2.$$

At (0, 0), $rt - s^2 = -9a^2$ (negative). Therefore, there is no extreme point at the origin.

At (a,a), we have $rt - s^2 = 36a^2 - 9a^2 = 27a^2 > 0$. Also r at (a, a) is equal to $6a$. If a is positive, then r is positive and $f(x, y)$ will have a minimum at (a, a). If a is negative, then r is negative and so, $f(x, y)$ will have a maximum at (a, a) for $a < 0$.

EXAMPLE 3.62

Prove that the rectangular solid of maximum volume which can be inscribed in a sphere is a cube.

Solution. Let x, y, and z be the length, breadth, and height of a rectangular solid. Then, the volume of the solid is

$$V = xyz. \tag{1}$$

Now each diagonal of the rectangular solid passes through the center of the sphere. Therefore, each diagonal is the diameter of the sphere, that is,

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

or

$$x^2 + y^2 + z^2 = d^2$$

or

$$z = \sqrt{d^2 - x^2 - y^2}. \tag{2}$$

Therefore, (1) reduces to

$$V = xy\sqrt{d^2 - x^2 - y^2}$$

or

$$V^2 = x^2y^2\left(d^2 - x^2 - y^2\right)$$

$$= x^2y^2d^2 - x^4y^2 - x^2y^4 = f(x, y).$$

Then,

$$\frac{\partial f}{\partial x} = 2xy^2d^2 - 4x^3y^2 - 2xy^4 = 2xy^2\left(d^2 - 2x^2 - y^2\right),$$

$$\frac{\partial f}{\partial y} = 2x^2yd^2 - 2x^4y - 4x^2y^3 = 2x^2y\left(d^2 - x^2 - 2y^2\right),$$

$$\frac{\partial^2 f}{\partial x^2} = 2d^2y^2 - 12x^2y^2 - 2y^4,$$

$$\frac{\partial^2 f}{\partial y^2} = 2d^2x^2 - 12x^2y^2 - 2x^4, \text{ and}$$

$$\frac{\partial^2 f}{\partial x \partial y} = 4xyd^2 - 8x^3y - 8xy^3.$$

For stationary points, we have $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$. Therefore,

$$d^2 - 2x^2 - y^2 = 0 \text{ and}$$
$$d^2 - x^2 - 2y^2 = 0. \quad (3)$$

Solving the preceding equations, we get $y = x$. Substituting $y = x$ in (3), we get $x = \frac{d}{\sqrt{3}}$. Thus, $x = y = \frac{d}{\sqrt{3}}$. Hence, from (2), we have $z = \frac{d}{\sqrt{3}}$. Thus, the stationary point is $\left(\frac{d}{\sqrt{3}}, \frac{d}{\sqrt{3}}\right)$. At $\left(\frac{d}{\sqrt{3}}, \frac{d}{\sqrt{3}}\right)$,

$$r = -\frac{8d^4}{9} \text{ (negative)}, s = -\frac{4d^4}{9}, \text{ and } t = \frac{-8d^4}{9}.$$

Therefore, $rt - s^2 = \frac{64d^8}{81} - \frac{16d^8}{81} = \frac{48}{81}d^8 = \frac{16d^8}{27} > 0$. Since r is negative, it follows that $f(x, y)$ or V^2 has a maximum value at $\left(\frac{d}{\sqrt{3}}, \frac{d}{\sqrt{3}}\right)$. Hence, V is maximum when $x = y = z$. Consequently, the solid is a cube.

EXAMPLE 3.63

A rectangular box, open at the top, is to have a volume of 32 cubic feet. Determine the dimensions of the box requiring least material for its construction.

Solution. Let S be the surface, and x, y, and z in feet be the edges of the box. Then,

$$S = xy + 2yz + 2zx \quad (1)$$

and

$$V = xyz = 32 \text{ cubic feet (given)}. \quad (2)$$

From (2), we have $z = \frac{32}{xy}$ and so,

$$S = xy + 2(y + x)\frac{32}{xy} = xy + 64\left(\frac{1}{x} + \frac{1}{y}\right).$$

Then,

$$\frac{\partial S}{\partial x} = y - \frac{64}{x^2}, \quad \frac{\partial S}{\partial y} = x - \frac{64}{y^2},$$

$$\frac{\partial^2 S}{\partial x^2} = \frac{128}{x^3}, \quad \frac{\partial^2 S}{\partial x \partial y} = 1, \text{ and } \frac{\partial^2 S}{\partial y^2} = \frac{128}{y^3}.$$

The stationary values are given by

$$\frac{\partial S}{\partial x} = y - \frac{64}{x^2} = 0 \text{ and } \frac{\partial S}{\partial y} = x - \frac{64}{y^2} = 0.$$

Solving these equations, we get $x = y = 4$. Putting these values in (1), we get $z = 2$. Further, at (4, 4), we have $rt - s^2 = 3$ (positive) and r at (4, 4) is 2 (positive). Therefore, S is minimum for (4, 4). The dimensions of the box are $x = 4$, $y = 4$, $z = 2$.

EXAMPLE 3.64

Find the points on the surface $z^2 = xy + 1$ nearest to the origin.

Solution. If r is the distance from (0, 0, 0) of any point (x, y, z) on the given surface, then

$$r^2 = (x - 0)^2 + (y - 0)^2 + (z - 0)^2 = x^2 + y^2 + z^2$$
$$= x^2 + y^2 + xy + 1, \text{ using the equation of the given surface.}$$

Thus, we have a function of two variables given by

$$r^2 = x^2 + y^2 + xy + 1 = f(x, y), \text{ say.}$$

Then,

$$\frac{\partial f}{\partial x} = 2x + y, \quad \frac{\partial f}{\partial y} = 2y + x,$$

$$\frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial y^2} = 2, \text{ and } \frac{\partial^2 f}{\partial x \partial y} = 1.$$

The stationary points are given by $\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$ and therefore,

$$2x + y = 0 \text{ and } 2y + x = 0.$$

Solving the preceding equations, we get $x = y = 0$ and then, $z^2 = xy + 1$ yields $z = \pm 1$. Thus, the stationary points are $(0, 0, \pm 1)$. Further, at these points, $r = 2$, $s = 1$, and $t = 2$ and so, $rt - s^2 = 3$ (positive). Since r is positive, the value is minimum at $(0, 0, \pm 1)$.

3.12 LAGRANGE'S METHOD OF UNDETERMINED MULTIPLIERS

Let $u = \phi(x_1, x_2, \ldots, x_n)$ be a function of n variables $x_1, x_2, \ldots, x_n$, which are connected by m equations

$$f_1(x_1, x_2, \ldots, x_n) = 0, \; f_2(x_1, x_2, \ldots, x_n) = 0, \; \ldots, \; f_m(x_1, x_2, \ldots, x_n) = 0,$$

so that only $n - m$ of the variables are independent. For a maximum or a minimum value of u, we must have

$$du = \frac{\partial u}{\partial x_1}dx_1 + \frac{\partial u}{\partial x_2}dx_2 + \frac{\partial u}{\partial x_3}dx_3 + \ldots + \frac{\partial u}{\partial x_n}dx_n = 0.$$

Also, differentiating the given m equations connecting the variables, we have

$$df_1 = \frac{\partial f_1}{\partial x_1}dx_1 + \frac{\partial f_1}{\partial x_2}dx_2 + \frac{\partial f_1}{\partial x_3}dx_3 + \ldots + \frac{\partial f_1}{\partial x_n}dx_n = 0$$

$$df_2 = \frac{\partial f_2}{\partial x_1}dx_1 + \frac{\partial f_2}{\partial x_2}dx_2 + \frac{\partial f_2}{\partial x_3}dx_3 + \ldots + \frac{\partial f_2}{\partial x_n}dx_n = 0$$

....

....

$$df_m = \frac{\partial f_m}{\partial x_1}dx_1 + \frac{\partial f_m}{\partial x_2}dx_2 + \frac{\partial f_m}{\partial x_3}dx_3 + \ldots + \frac{\partial f_m}{\partial x_n}dx_n = 0.$$

Multiplying the earlier $(m+1)$ equations, obtained on differentiation, by $1, \lambda_1, \lambda_2, \ldots, \lambda_m$, respectively, and then adding all, we get an equation which may be written as

$$P_1dx_1 + P_2dx_2 + P_3dx_3 + \ldots + P_ndx_n = 0, \quad (1)$$

where,

$$P_r = \frac{\partial u}{\partial x_r} + \lambda_1\frac{\partial f_1}{\partial x_r} + \lambda_2\frac{\partial f_2}{\partial x_r} + \lambda_3\frac{\partial f_3}{\partial x_r} + \ldots + \lambda_m\frac{\partial f_m}{\partial x_r}.$$

The m quantities $\lambda_1, \lambda_2, \ldots, \lambda_m$ are at our choice. Let us choose them so as to satisfy the m linear equations.

$$P_1 = P_2 = P_3 = \ldots = P_m = 0.$$

Then, the equation (1) reduces to

$$P_{m+1}dx_{m+1} + P_{m+2}dx_{m+2} + \ldots + P_ndx_n = 0.$$

It is indifferent which of the $n-$m of the n variables are regarded as independent. So, suppose that the variables $x_{m+1}, x_{m+2}, \ldots, x_n$ are independent, then as the $n-m$ quantities $dx_{m+1}, dx_{m+2}, \ldots, dx_n$ are all independent, their coefficients must be separately zero. Thus, we obtain the additional $n-$m equations as follows:

$$P_{m+1} = 0,\ P_{m+2} = 0, \ldots, P_n = 0.$$

In this way, we get $(m+n)$ equations

$$f_1 = 0,\ f_2 = 0, \ldots, f_m = 0 \text{ and}$$

$$P_1 = 0,\ P_2 = 0, \ldots, P_n = 0,$$

which together with relation $u = \phi(x_1, x_2, \ldots, x_n)$ determine the m multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the values of n variables $x_1, x_2, \ldots, x_n$ for which the maximum and minimum values of u are possible.

The drawback of the Lagrange's method of undetermined multipliers is that it does not determine the nature of the stationary point.

EXAMPLE 3.65

Find the point of the circle $x^2+y^2+z^2=k^2$ and $lx+my+nz=0$ at which the function $u=ax^2+by^2+cz^2+2fyz+2gzx+2hxy$ attains its greatest and the least value.

Solution. We have

$$u = ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy, \quad (1)$$

$$f_1 = lx + my + nz = 0, \text{ and} \quad (2)$$

$$f_2 = x^2 + y^2 + z^2 = k^2. \quad (3)$$

For extreme points, we must have $du = 0$. So,

$$(ax + gz + hy)dx + (hx + by + fz)dy + (gx + fy + cz)dz = 0. \quad (4)$$

Also differentiating (2) and (3), we get

$$ldx + mdy + ndz = 0, \text{ and} \quad (5)$$

$$xdx + ydy + zdz = 0. \quad (6)$$

Multiplying (4), (5), and (6) by 1, λ_1, and λ_2, respectively, and then by adding all and equating to zero the coefficients of dx, dy, and dz, we get

$$ax + hy + gz + \lambda_1 l + \lambda_2 x = 0, \quad (7)$$

$$hx + by + fz + \lambda_1 m + \lambda_2 y = 0, \text{ and} \quad (8)$$

$$gx + fy + cz + \lambda_1 n + \lambda_2 z = 0. \quad (9)$$

Multiplying (7), (8), and (9) by x, y, and z, respectively, and then adding all, we get

$$u + \lambda_2 = 0 \text{ or } \lambda_2 = -u.$$

Putting $\lambda_2 = -u$ in (7), (8), and (9), we obtain

$$(a-u)x + hy + gz + \lambda_1 l = 0, \quad (10)$$

$$hx + (b-u)y + fz + \lambda_1 m = 0, \text{ and} \quad (11)$$

$$gx + fy + (c-u)z + \lambda_1 n = 0. \quad (12)$$

Also,

$$lx + my + nz + \lambda_1.0 = 0. \quad (13)$$

Eliminating x, y, z, and λ_1 from (10), (11), (12), and (13), we get

$$\begin{vmatrix} a-u & h & g & l \\ h & b-u & f & m \\ g & f & c-u & n \\ l & m & n & 0 \end{vmatrix} = 0,$$

which gives the maximum or minimum value of u.

EXAMPLE 3.66

Prove that the volume of the greatest rectangular parallelopiped that can be inscribed in the ellipsoid $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ is $\frac{8abc}{3\sqrt{3}}$.

Solution. Let (x, y, z) denote the coordinates of the vertex of the rectangular parallelopiped which lies in the positive octant and let V denote its volume. Volume V is given by $V = 8xyz$. Its maximum value is to be determined under the condition that it is inscribed in the ellipsoid $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$. Thus, we have

$$V = 8xyz, \text{ and} \quad (1)$$

$$f_1 = \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} = 1. \quad (2)$$

For an extreme value, we must have

$$dV = yzdx + zxdy + xydz = 0. \quad (3)$$

Also differentiating (2), we get

$$df_1 = \frac{x}{a^2}dx + \frac{y}{b^2}dy + \frac{z}{c^2}dz = 0. \quad (4)$$

Multiplying (3) and (4) by 1 and λ, respectively, and then adding both and equating the coefficients of dx, dy, and dz to zero, we get

$$yz + \frac{\lambda x}{a^2} = 0, \quad (5)$$

$$zx + \frac{\lambda y}{b^2} = 0, \quad (6)$$

and

$$xy + \frac{\lambda z}{c^2} = 0. \quad (7)$$

From (5), (6), and (7), we get

$$\lambda = -\frac{a^2yz}{x} = -\frac{b^2zx}{y} = -\frac{c^2xy}{z}$$

and so,

$$\frac{a^2yz}{x} = \frac{b^2zx}{y} = \frac{c^2xy}{z}.$$

Dividing throughout by xyz, we get

$$\frac{a^2}{x^2} = \frac{b^2}{y^2} = \frac{c^2}{z^2}.$$

Then, equation (2) yields

$$3\frac{x^2}{a^2} = 1 \quad \text{or} \quad x = \tfrac{a}{\sqrt{3}},$$

$$3\frac{y^2}{b^2} = 1 \quad \text{or} \quad y = \tfrac{b}{\sqrt{3}}, \text{ and}$$

$$3\frac{z^2}{c^2} = 1 \quad \text{or} \quad z = \tfrac{c}{\sqrt{3}}.$$

Thus, the stationary value is at the point $\left(\frac{a}{\sqrt{3}}, \frac{b}{\sqrt{3}}, \frac{c}{\sqrt{3}}\right)$.

Differentiating partially the equation (2) with respect to x, taking y as constant, we get

$$\frac{2x}{a^2} + \frac{2z}{c^2}\frac{\partial z}{\partial x} = 0 \text{ and so, } \frac{\partial z}{\partial x} = -\frac{c^2x}{a^2z}.$$

Now,

$$\frac{\partial V}{\partial x} = 8yz + 8xy\frac{\partial z}{\partial x} = 8yz + 8xy\left(-\frac{c^2x}{a^2z}\right) = 8yz - \frac{8c^2x^2y}{a^2z}$$

and so,

$$\frac{\partial^2 V}{\partial x^2} = 8y\left(-\frac{c^2x}{a^2z}\right) - \frac{16c^2xy}{a^2z} - \frac{8c^2x^2y}{a^2z}.\frac{c^2x}{a^2z},$$

which is negative. Hence, V is maximum at $\left(\frac{a}{\sqrt{3}}, \frac{b}{\sqrt{3}}, \frac{c}{\sqrt{3}}\right)$ and

$$\text{Max } V = \frac{8abc}{3\sqrt{3}}.$$

EXAMPLE 3.67

Solve Example 3.63 using Lagrange's method of undetermined multipliers.

Solution. We have

$$S = xy + 2yz + 2zx \quad \text{and} \quad (1)$$

$$V = xyz = 32. \quad (2)$$

For S to be minimum, we must have

$$dS = (y+2z)dx + (x+2z)dy + 2(x+y)dz = 0. \quad (3)$$

Also, from (2), since V is constant, we have

$$yzdx + zxdy + xydz = 0. \quad (4)$$

Multiplying (3) by 1 and (4) by λ and then adding both and equating to zero the coefficients of dx, dy, and dz, we get

$$(y + 2z) + \lambda yz = 0 \quad (5)$$

$$(x + 2z) + \lambda xz = 0 \quad (6)$$

$$2x + 2y + \lambda xy = 0. \quad (7)$$

Multiplying (5) by x and (6) by y and subtracting, we get

$$2zx - 2zy = 0 \text{ or } x = y,$$

since $z = 0$ is not admissible due to the fact that depth cannot be zero.

Similarly, from the equations (6) and (7), we get $y = 2z$. Thus, for a stationary value, the dimensions of the box are

$$x = y = 2z = 4, \ [\text{using } (2)].$$

Proceeding, as in Example 3.65, we note that $\frac{\partial^2 f}{\partial x^2} = 2$ (positive) and $rt - s^2 > 0$. Thus, at (4, 4, 2), S has a minimum. Hence, the required dimensions are $x = 4$, $y = 4$, and $z = 2$.

EXAMPLE 3.68

Investigate the maximum- and minimum radii vector of the sector of "surface of elasticity" $(x^2 + y^2 + z^2) = a^2x^2 + b^2y^2 + z^2c^2$, made by the plane $lx + my + nz = 0$.

Solution. On differentiating, we get

$$xdx + ydy + ndz = 0 \quad (1)$$

$$a^2xdx + b^2ydy + c^2zdz = 0 \quad (2)$$

$$ldx + mdy + ndz = 0. \quad (3)$$

Multiplying (1), (2), and (3) by 1, λ_1, and λ_2, respectively, and adding and equating to zero the coefficients of dx, dy, and dz, we get

$$x + a^2x\lambda_1 + l\lambda_2 = 0 \quad (4)$$

$$y + b^2y\lambda_1 + m\lambda_2 = 0 \quad (5)$$

$$z + c^2z\lambda_1 + n\lambda_2 = 0. \quad (6)$$

Multiplying (4), (5), and (6) by x, y, and z, respectively, and adding we get

$$(x^2 + y^2 + z^2) + (a^2x^2 + b^2y^2 + c^2z^2)\lambda_1 + (lx + my + nz)\lambda_2 = 0$$

or

$$r^2 + \lambda_1 r^4 = 0 \text{ or } \lambda_1 = -\frac{1}{r^2}.$$

Putting this value of λ_1 in (4), (5), and (6), we get

$$x = \frac{\lambda_2 l r^2}{a^2 - r^2},\ y = \frac{\lambda_2 m r^2}{b^2 - r^2},\ \text{and } z = \frac{\lambda_2 n r^2}{c^2 - r^2}.$$

Substituting these values of x, y, and z in $lx + my$ $nz = 0$, we get

$$\frac{\lambda_2 l^2 r^2}{a^2 - r^2} + \frac{\lambda_2 m^2 r^2}{b^2 - r^2} + \frac{\lambda_2 n^2 r^2}{c^2 - r^2} = 0,$$

or

$$\frac{l^2}{a^2 - r^2} + \frac{m^2}{b^2 - r^2} + \frac{n^2}{c^2 - r^2} = 0,$$

which is an equation in r giving the required values.

EXAMPLE 3.69

Find the length of the axes of the section of the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ by the plane $lx + my + nz = 0$.

Solution. We have to find the extreme values of the function $r^2 = x^2 + y^2 + z^2$ subject to the conditions

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \text{ and } lx + my + nz = 0.$$

Differentiation yields

$$x\,dx + y\,dy + z\,dz = 0 \quad (1)$$

$$\frac{x}{a^2}dx + \frac{y}{b^2}dy + \frac{z}{c^2}dz = 0 \quad (2)$$

$$ldx + mdy + ndz = 0. \quad (3)$$

Multiplying (1), (2), and (3) by 1, λ_1, and λ_2, respectively, adding and then equating to zero the coefficients of dx, dy, and dz, we get

$$x + \lambda_1\frac{x}{a^2} + \lambda_2 l = 0, \quad (4)$$

$$y + \lambda_1\frac{y}{a^2} + \lambda_2 m = 0, \text{ and} \quad (5)$$

$$z + \lambda_1\frac{z}{a^2} + \lambda_2 n = 0. \quad (6)$$

Multiplying (4), (5), and (6) by x, y, and z and adding, we obtain

$$(x^2 + y^2 + z^2) + \lambda_1\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) + \lambda_2(lx + my + nz) = 0$$

or

$$r^2 + \lambda_1 = 0, \text{ which gives } \lambda_1 = -r^2.$$

Hence, from (4), (5), and (6), we have

$$x = \frac{\lambda_2 l}{\frac{r^2}{a^2} - 1},\ y = \frac{\lambda_2 m}{\frac{r^2}{b^2} - 1},\ \text{and } z = \frac{\lambda_2 n}{\frac{r^2}{c^2} - 1}.$$

Putting these values of x, y, and z in $lx + my + nz$ 0, we get

$$\lambda_2\left(\frac{l^2a^2}{r^2 - a^2} + \frac{m^2b^2}{r^2 - b^2} + \frac{n^2c^2}{r^2 - c^2}\right) = 0.$$

Since $\lambda_2 \neq 0$, the equation giving the values of r^2, the squares of the length of the semi-axes, is

$$\frac{l^2a^2}{r^2 - a^2} + \frac{m^2b^2}{r^2 - b^2} + \frac{n^2c^2}{r^2 - c^2} = 0.$$

EXAMPLE 3.70

If a, b, and, c are positive and

$$u = \frac{a^2x^2 + b^2y^2 + c^2z^2}{x^2y^2z^2},$$

$$ax^2 + by^2 + cz^2 = 1,$$

show that a stationary value of u is given by

$$x^2 = \frac{\mu}{2a(\mu + a)},\ y^2 = \frac{\mu}{2b(\mu + b)},\ \text{and } z^2 = \frac{\mu}{2c(\mu + c)},$$

where μ is the positive root of the cubic

$$\mu^3 - (bc + ca + ab)\mu - 2abc = 0.$$

Solution. We have

$$u = \frac{a^2x^2 + b^2y^2 + c^2z^2}{x^2y^2z^2}, \quad (1)$$

$$ax^2 + by^2 + cz^2 = 1. \quad (2)$$

Differentiating (1), we get

$$\Sigma\frac{1}{x^3}\left(\frac{b^2}{z^2} + \frac{c^2}{y^2}\right)dx = 0,$$

which on multiplication by $x^2y^2z^2$ yields

$$\Sigma\frac{1}{x}(b^2y^2 + c^2z^2)dx = 0. \quad (3)$$

Differentiating (2), we get

$$\Sigma\, axdx = 0. \quad (4)$$

Using Lagrange's multipliers 1 and μ, we get

$$\frac{1}{x}(b^2y^2 + c^2z^2) = \mu ax \text{ or } b^2y^2 + c^2z^2 = \mu ax^2, \quad (5)$$

$$\frac{1}{y}(c^2z^2 + a^2x^2) = \mu by \text{ or } c^2z^2 + a^2x^2 = \mu by^2, \quad (6)$$

and

$$\frac{1}{z}(a^2x^2+b^2y^2)=\mu cz \quad \text{or} \quad a^2x^2+b^2y^2=\mu cz^2. \quad (7)$$

Then, (6) + (7) – (5) yields

$$2a^2x^2 = \mu(by^2+cz^2-ax^2)$$
$$= \mu(1-2ax^2), \text{ using}(2).$$

Thus,

$$2a(a+\mu)x^2 = \mu \quad \text{or} \quad x^2 = \frac{\mu}{2a(\mu+a)}.$$

Similarly, we obtain

$$y^2 = \frac{\mu}{2b(\mu+b)} \quad \text{and} \quad z^2 = \frac{\mu}{2c(\mu+c)}.$$

Substituting these values of x^2, y^2, and z^2 in (2), we have

$$\frac{\mu}{2(a+\mu)}+\frac{\mu}{2(b+\mu)}+\frac{\mu}{2(c+\mu)}=1$$

or

$$\mu^3-(bc+ca+ab)\mu-2abc=0. \qquad (8)$$

Since a, b, and c are positive, any one of (5), (6), or (7) shows that μ must be positive. Hence, μ is a positive root of (8).

3.13 JACOBIANS

If $u_1, u_2, \ldots, u_n$ are n functions of n variables $x_1, x_2, \ldots, x_n$, then the determinant

$$\begin{vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \cdots & \cdots & \frac{\partial u_1}{\partial x_n} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \cdots & \cdots & \frac{\partial u_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial u_n}{\partial x_1} & \frac{\partial u_n}{\partial x_2} & \cdots & \cdots & \frac{\partial u_n}{\partial x_n} \end{vmatrix}$$

is called the *Jacobian* of $u_1, u_2, \ldots, u_n$ with regard to $x_1, x_2, \ldots, x_n$. This determinant is often denoted by $\frac{\partial(u_1,u_2,\ldots,u_n)}{\partial(x_1,x_2,\ldots,x_n)}$ or $J\ (u_1, u_2, \ldots, u_n)$.

3.14 PROPERTIES OF JACOBIAN

Theorem 3.7. If U, V are functions of u and v, where u and v are themselves functions of x and y, then

$$\frac{\partial(U,V)}{\partial(x,y)} = \frac{\partial(U,V)}{\partial(u,v)} \cdot \frac{\partial(u,v)}{\partial(x,y)}.$$

Proof: Let

$U=f(u,v)$, $V=F(u,v)$, $u=\phi(x,y)$, and $v=\psi(x,y)$.

Then,

$$\frac{\partial U}{\partial x} = \frac{\partial U}{\partial u}\cdot\frac{\partial u}{\partial x}+\frac{\partial U}{\partial v}\cdot\frac{\partial v}{\partial x},$$
$$\frac{\partial U}{\partial y} = \frac{\partial U}{\partial u}\cdot\frac{\partial u}{\partial y}+\frac{\partial U}{\partial v}\cdot\frac{\partial v}{\partial y},$$
$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial u}\cdot\frac{\partial u}{\partial x}+\frac{\partial V}{\partial v}\cdot\frac{\partial v}{\partial x}, \text{ and}$$
$$\frac{\partial V}{\partial y} = \frac{\partial V}{\partial u}\cdot\frac{\partial u}{\partial y}+\frac{\partial V}{\partial v}\cdot\frac{\partial v}{\partial y}.$$

and so,

$$\frac{\partial(U,V)}{\partial(u,v)}\cdot\frac{\partial(u,v)}{\partial(x,y)}$$
$$= \begin{vmatrix} \frac{\partial U}{\partial u} & \frac{\partial U}{\partial v} \\ \frac{\partial V}{\partial u} & \frac{\partial V}{\partial v} \end{vmatrix} \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$
$$= \begin{vmatrix} \frac{\partial U}{\partial u}\cdot\frac{\partial u}{\partial x}+\frac{\partial U}{\partial v}\cdot\frac{\partial v}{\partial x} & \frac{\partial U}{\partial u}\cdot\frac{\partial u}{\partial y}+\frac{\partial U}{\partial v}\cdot\frac{\partial v}{\partial y} \\ \frac{\partial V}{\partial u}\cdot\frac{\partial u}{\partial x}+\frac{\partial V}{\partial v}\cdot\frac{\partial v}{\partial x} & \frac{\partial V}{\partial u}\cdot\frac{\partial u}{\partial y}+\frac{\partial V}{\partial v}\cdot\frac{\partial v}{\partial y} \end{vmatrix}$$
$$= \begin{vmatrix} \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} \end{vmatrix}$$
$$= \frac{\partial(U,V)}{\partial(x,y)}.$$

Theorem 3.8. If J is the Jacobian of the system u, v with regard to x, y, and J' is the Jacobian of x, y with regard to u, v, then $JJ'=1$.

Proof: Let $u=f(x,y)$ and $v=F(x,y)$. Suppose that these are solved for x and y giving $x=\phi(u,v)$ and $y=\psi(u,v)$. Differentiating $u=f(x,y)$ with respect to u and v, we have

$$1 = \frac{\partial u}{\partial x}\cdot\frac{\partial x}{\partial u}+\frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial u} \text{ and } 0 = \frac{\partial u}{\partial x}\cdot\frac{\partial x}{\partial v}+\frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial v}.$$

Similarly, differentiating $v=F(x,y)$ with respect to u and v, we get

$$0 = \frac{\partial v}{\partial x}\cdot\frac{\partial x}{\partial u}+\frac{\partial v}{\partial y}\cdot\frac{\partial y}{\partial u} \text{ and } 1 = \frac{\partial v}{\partial x}\cdot\frac{\partial x}{\partial v}+\frac{\partial v}{\partial y}\cdot\frac{\partial y}{\partial v}.$$

Therefore,

$$JJ' = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$
$$= \begin{vmatrix} \frac{\partial u}{\partial x}\cdot\frac{\partial x}{\partial u}+\frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial u} & \frac{\partial u}{\partial x}\cdot\frac{\partial x}{\partial v}+\frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial v} \\ \frac{\partial v}{\partial x}\cdot\frac{\partial x}{\partial u}+\frac{\partial v}{\partial y}\cdot\frac{\partial y}{\partial u} & \frac{\partial v}{\partial x}\cdot\frac{\partial x}{\partial v}+\frac{\partial v}{\partial y}\cdot\frac{\partial y}{\partial v} \end{vmatrix}$$
$$= \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

EXAMPLE 3.71

If $x = r\cos\theta$ and $y = r\sin\theta$, show that

$$\text{(i)}\quad \frac{\partial(x,y)}{\partial(r,\theta)} = r \text{ and (ii) } \frac{\partial(r,\theta)}{\partial(x,y)} = \frac{1}{r}.$$

Solution. (i) We have

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix}$$

$$= r\cos^2\theta + r\sin^2\theta = r.$$

(ii) We have

$$r^2 = x^2 + y^2 \text{ and } \tan\theta = \frac{y}{x}.$$

Differentiating partially with respect to x and y, we get

$$2r\frac{\partial r}{\partial x} = 2x \text{ and so, } \frac{\partial r}{\partial x} = \frac{x}{r},$$

$$2r\frac{\partial r}{\partial y} = 2y \text{ and so, } \frac{\partial r}{\partial y} = \frac{y}{r},$$

$$\sec^2\theta\frac{\partial \theta}{\partial x} = -\frac{y}{x^2} \text{ and so, } \frac{\partial \theta}{\partial x} = -\frac{y}{x^2\sec^2\theta} = -\frac{y}{r^2\cos^2\theta\sec^2\theta} = -\frac{y}{r^2},$$

$$\sec^2\theta\frac{\partial \theta}{\partial y} = \frac{1}{x}, \text{ and so, } \frac{\partial \theta}{\partial y} = \frac{1}{x\sec^2\theta} = \frac{\cos^2\theta}{x} = \frac{x^2}{r^2}\frac{1}{x} = \frac{x}{r^2}.$$

Therefore,

$$\frac{\partial(r,\theta)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{x}{r} & \frac{y}{r} \\ -\frac{y}{r^2} & \frac{x}{r^2} \end{vmatrix}$$

$$= \frac{x^2}{r^3} + \frac{y^2}{r^3} = \frac{x^2+y^2}{r^3} = \frac{r^2}{r^3} = \frac{1}{r}.$$

EXAMPLE 3.72

If $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, and $z = r\cos\theta$, show that

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = r^2\sin\theta.$$

Solution. We have

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix}$$

$$= \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}$$

$$= \cos\theta(r^2\sin\theta\cos\theta\cos^2\phi + r^2\sin\theta\cos\theta\sin^2\phi)$$
$$+ r\sin\theta(r\sin^2\theta\cos^2\phi + r\sin^2\theta\sin^2\phi)$$
$$= r^2\sin\theta\cos^2\theta + r^2\sin^3\theta$$
$$= r^2\sin\theta(\cos^2\theta + \sin^2\theta) = r^2\sin\theta.$$

EXAMPLE 3.73

If $u = x + y + z$, $uv = y + z$, and $uvw = z$, show that

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = u^2v.$$

Solution. We have

$$z = uvw,$$
$$y = uv - z = uv - uvw \text{ and}$$
$$x = u - y - z = u - uv + uvw - uvw = u - uv.$$

Therefore,

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix} = \begin{vmatrix} 1-v & -u & 0 \\ v-vw & u-uw & -uv \\ vw & uw & uv \end{vmatrix}$$

$$= \begin{vmatrix} 1-v & -u & 0 \\ v & u & 0 \\ vw & uw & uv \end{vmatrix} = uv\begin{vmatrix} 1-v & -u \\ v & u \end{vmatrix}$$

$$= uv(u - uv + uv) = u^2v.$$

EXAMPLE 3.74

If $u_1 = \frac{x_2x_3}{x_1}$, $u_2 = \frac{x_3x_1}{x_2}$, and $u_3 = \frac{x_1x_2}{x_3}$, show that

$$\frac{\partial(u_1,u_2,u_3)}{\partial(x_1,x_2,x_3)} = 4.$$

Solution. We have

$$\frac{\partial(u_1,u_2,u_3)}{\partial(x_1,x_2,x_3)} = \begin{vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \frac{\partial u_1}{\partial x_3} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \frac{\partial u_2}{\partial x_3} \\ \frac{\partial u_3}{\partial x_1} & \frac{\partial u_3}{\partial x_2} & \frac{\partial u_3}{\partial x_3} \end{vmatrix}$$

$$= \begin{vmatrix} -\frac{x_2x_3}{x_1^2} & \frac{x_3}{x_1} & \frac{x_2}{x_1} \\ \frac{x_3}{x_2} & -\frac{x_3x_1}{x_2^2} & \frac{x_1}{x_2} \\ \frac{x_2}{x_3} & \frac{x_1}{x_3} & -\frac{x_1x_2}{x_3^2} \end{vmatrix}$$

$$= \frac{1}{x_1^2x_2^2x_3^2}\begin{vmatrix} -x_2x_3 & x_3x_1 & x_1x_2 \\ x_2x_3 & -x_3x_1 & x_1x_2 \\ x_2x_3 & x_3x_1 & -x_1x_2 \end{vmatrix}$$

$$= \frac{1}{x_1^2x_2^2x_3^2}\begin{vmatrix} 0 & 0 & 2x_1x_2 \\ x_2x_3 & -x_3x_1 & x_1x_2 \\ x_2x_3 & x_3x_1 & -x_1x_2 \end{vmatrix}$$

using $R_1 \to R_1 + R_2$

$$= \frac{2x_1x_2}{x_1^2x_2^2x_3^2}\begin{vmatrix} x_2x_3 & -x_3x_1 \\ x_2x_3 & x_3x_1 \end{vmatrix}$$

$$= \frac{2x_1x_2}{x_1^2x_2^2x_3^2}(2x_1x_2x_3^2) = 4.$$

EXAMPLE 3.75

If $u = \frac{y^2}{2x}$ and $v = \frac{x^2+y^2}{2x}$, find $\frac{\partial(u,v)}{\partial(x,y)}$.

Solution. We have

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} -\frac{y^2}{2x^2} & \frac{y}{x} \\ \frac{1}{2} - \frac{y^2}{2x^2} & \frac{y}{x} \end{vmatrix}$$

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

EXAMPLE 3.76

If $u = \frac{x+y}{1-xy}$ and $v = \tan^{-1}x + \tan^{-1}y$, find $\frac{\partial(u,v)}{\partial(x,y)}$.

Solution. We have

$$\frac{\partial u}{\partial x} = \frac{(1-xy)-(x+y)(-y)}{(1-xy)^2} = \frac{1-xy+xy+y^2}{(1-xy)^2}$$

$$= \frac{1+y^2}{(1-xy)^2},$$

$$\frac{\partial u}{\partial y} = \frac{(1-xy)-(x+y)(-x)}{(1-xy)^2} = \frac{1-xy+x^2+xy}{(1-xy)^2}$$

$$= \frac{1+x^2}{(1-xy)^2},$$

$$\frac{\partial v}{\partial x} = \frac{1}{1+x^2}, \text{ and } \frac{\partial v}{\partial y} = \frac{1}{1+y^2}.$$

Therefore,

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{1+y^2}{(1-xy)^2} & \frac{1+x^2}{(1-xy)^2} \\ \frac{1}{1+x^2} & \frac{1}{1+y^2} \end{vmatrix} = \frac{1}{(1-xy)^2} - \frac{1}{(1-xy)^2} = 0.$$

EXAMPLE 3.77

If $u = 2xy$, $v = x^2 - y^2$, $x = r\cos\theta$, and $y = r\sin\theta$, find $\frac{\partial(u,v)}{\partial(r,\theta)}$.

Solution. We have

$$\frac{\partial(u,v)}{\partial(r,\theta)} = \frac{\partial(u,v)}{\partial(x,y)} \cdot \frac{\partial(x,y)}{\partial(r,\theta)}$$

$$= \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} \cdot \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix}$$

$$= \begin{vmatrix} 2y & 2x \\ 2x & -2y \end{vmatrix} \cdot \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix}$$

$$= -4(y^2 + x^2).r = -4r^3.$$

3.15 NECESSARY AND SUFFICIENT CONDITIONS FOR JACOBIAN TO VANISH

The following two theorems, stated without proof, provide necessary and sufficient condition for the Jacobian to vanish.

Theorem 3.9. If u_1, $u_2, \ldots$, u_n are n-differentiable functions of the n-independent variables $x_1, x_2, \ldots, x_n$ and there exists an identical, differentiable functional relation $\phi(u_1, u_2, \ldots, u_n) = 0$, which does not involve x_i explicitly, then the Jacobian $\frac{\partial(u_1, u_2, \ldots, u_n)}{\partial(x_1, x_2, \ldots, x_n)}$ vanishes identically provided that ϕ as a function of the u_i has no stationary values in the domain considered.

Theorem 3.10. If $u_1, u_2, \ldots, u_n$ are n functions of the n variables $x_1, x_2, \ldots, x_n$, say, $u_m = f_m(x_1, x_2, \ldots, x_n)$, $m = 1, 2, \ldots, n$, and if $\frac{\partial(u_1, u_2, \ldots, u_n)}{\partial(x_1, x_2, \ldots, x_n)} = 0$, then if all the differential coefficients are continuous, there exists a functional relation connecting some or all of the variables and which is independent of $x_1, x_2, \ldots, x_n$.

EXAMPLE 3.78

If $u = x + 2y + z$, $v = x - 2y + 3z$, and $w = 2xy - xz + 4yz - 2z^2$, show that $\frac{\partial(u,v,w)}{\partial(x,y,z)} = 0$ and find a relation between u, v, and w.

Solution. We have

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix}$$

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

$$= \begin{vmatrix} 1 & 0 & 0 \\ 1 & -4 & 2 \\ 2y-z & 2x+6z-4y & -x-2y-3z \end{vmatrix} = 0.$$

Hence, a relation between u, v, and w exists. Now,

$$u+v=2x+4z=2(x+2z)$$
$$u-v=4y-2z=2(2y-z)$$
$$w=x(2y-z)+2z(2y-z)$$
$$=(x+2z)(2y-z)=\frac{1}{4}(u+v)(u-v).$$

Therefore,

$$4w=(u+v)(u-v)$$

is the required relation connecting u, v, and w.

EXAMPLE 3.79

If $f(0)=0$ and $f'(x)=\frac{1}{1+x^2}$, show that

$$f(x)+f(y)=f\left(\frac{x+y}{1-xy}\right).$$

Solution. Suppose that

$$u=f(x)+f(y) \text{ and } v=\frac{x+y}{1-xy}.$$

Then,

$$\frac{\partial(u,v)}{\partial(x,y)}=\begin{vmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{vmatrix}=\begin{vmatrix}\frac{1}{1+x^2} & \frac{1}{1+y^2}\\ \frac{1+y^2}{(1-xy)^2} & \frac{1+x^2}{(1-xy)^2}\end{vmatrix}=0.$$

Therefore, u and v are connected by a functional relation. Let $u=\phi(v)$, that is,

$$f(x)+f(y)=\phi\left(\frac{x+y}{1-xy}\right).$$

Putting $y=0$, we get

$f(x)+f(0)=\phi(x)$ or $f(x)=\phi(x)$, since $f(0)=0$.

Hence,

$$f(x)+f(y)=f\left(\frac{x+y}{1-xy}\right).$$

3.16 DIFFERENTIATION UNDER THE INTEGRAL SIGN

In the following theorem of Leibnitz, we shall show that under suitable conditions, the derivative of the integral and the integral of the derivative are equal. The result is useful to determine the value of a definite integral by differentiating the integrand with respect to a quantity of which the limits of integration are independent.

Theorem 3.11. (Leibnitz's Rule): Let $f(x,\alpha)$ and $f_x(x,\alpha)$ be continuous functions of x and α. Then,

$$\frac{d}{d\alpha}\left[\int_a^b f(x,\alpha)dx\right]=\int_a^b\frac{\partial}{\partial\alpha}[f(x,\alpha)]dx,$$

where the limits a and b are independent of α.

Proof: Let $F(\alpha)=\int_a^b f(x,\alpha)dx$. Then

$$F(\alpha+\delta\alpha)=\int_a^b f(x,\alpha+\delta\,\alpha)dx$$

and so,

$$F(\alpha+\delta\alpha)-F(\alpha)=\int_a^b [f(x,\alpha+\delta\alpha) - f(x,\alpha)]dx. \quad (1)$$

But by Lagrange's mean value theorem, we have

$$f(\alpha+\delta\alpha)-f(\alpha)=\delta\alpha\frac{\partial}{\partial\alpha}f(x,\alpha+\theta\delta\alpha), 0<\theta<1.$$

Hence, (1) reduces to

$$\frac{F(\alpha+\delta\alpha)-F(\alpha)}{\delta\alpha}=\int_a^b\frac{\partial}{\partial\alpha}f(x,\alpha+\theta\delta\alpha)dx.$$

Therefore,

$$\lim_{\delta\alpha\to 0}\frac{F(\alpha+\delta\alpha)-F(\alpha)}{\delta\alpha}=\int_a^b\frac{\partial}{\partial\alpha}f(x,\alpha)dx$$

or

$$\frac{d}{d\alpha}[F(\alpha)]=\frac{d}{d\alpha}\left[\int_a^b f(x,\alpha)dx\right]=\int_a^b\frac{\partial}{\partial\alpha}f(x,\alpha)dx.$$

Remark 3.3. If the limits of integration a and b are not independent of α, then

$$\frac{d}{d\alpha}\left[\int_a^b f(x,\alpha)dx\right]=\int\frac{\partial}{\partial\alpha}[f(x,\alpha)]dx + \frac{db}{d\alpha}f(b,\alpha)-\frac{da}{d\alpha}f(a,\alpha).$$

EXAMPLE 3.80

Show that

$$\int_0^\infty\frac{\tan^{-1}ax}{x(1+x^2)}dx=\frac{\pi}{2}\log(1+a),\ a\geq 0.$$

Solution. Let

$$F(a)=\int_0^\infty\frac{\tan^{-1}ax}{x(1+x^2)}dx.$$

Then by Leibnitz's Rule,

$$\frac{d}{da}\left[\int_0^\infty \frac{\tan^{-1} ax}{x(1+x^2)}dx\right]$$

$$= \int_0^\infty \frac{\partial}{\partial a}\left[\frac{\tan^{-1} ax}{x(1+x^2)}\right]dx$$

$$= \int_0^\infty \frac{1}{x(1+x^2)}\cdot\frac{1}{1+a^2x^2}.xdx$$

$$= \int_0^\infty \frac{dx}{(1+x^2)(1+a^2x^2)}$$

$$= \frac{1}{1-a^2}\int_0^\infty \left[\frac{1}{1+x^2} - \frac{a^2}{1+a^2x^2}\right]dx$$

(by partial fractions)

$$= \frac{1}{1-a^2}[\tan^{-1} x]_0^\infty - \frac{a^2}{1-a^2}\int_0^\infty \frac{dx}{1+a^2x^2}$$

$$= \frac{\pi}{2(1-a^2)} - \frac{1}{1-a^2}\int_0^\infty \frac{dx}{x^2+\frac{1}{a^2}}$$

$$= \frac{\pi}{2(1-a^2)} - \frac{1}{1-a^2}\cdot\frac{1}{\frac{1}{a}}\left[\tan^{-1}\frac{x}{\frac{1}{a}}\right]_0^\infty$$

$$= \frac{\pi}{2(1-a^2)} - \frac{a}{1-a^2}\frac{\pi}{2}$$

$$= \frac{\pi}{2(1-a^2)}[1-a] = \frac{\pi}{2(1+a)}.$$

Integrating both sides with respect to a, we get

$$F(a) = \int_0^\infty \frac{\tan^{-1} ax}{x(1+x^2)}dx = \frac{\pi}{2}\log(1+a) + c. \quad (1)$$

Also $F(0) = 0$. Therefore, (1) yields $0 = \frac{\pi}{2}\log 1 + c$ and so, $c = 0$. Hence,

$$\int_0^\infty \frac{\tan^{-1} ax}{x(1+x^2)}dx = \frac{\pi}{2}\log(1+a).$$

EXAMPLE 3.81

Evaluate $\int_0^1 \frac{x^\alpha - 1}{\log x}dx$, $\alpha \geq 0$ using differentiation under the integral sign.

Solution. Let $F(\alpha) = \int_0^1 \frac{x^\alpha-1}{\log x}dx$. Then by Leibnitz's Rule,

$$\frac{d}{d\alpha}\left[\int_0^1 \frac{x^\alpha - 1}{\log x}dx\right] = \int_0^1 \frac{\partial}{\partial\alpha}\left[\frac{x^\alpha - 1}{\log x}\right]dx$$

$$= \int_0^1 \frac{1}{\log x}.x^\alpha \log x dx$$

$$= \int_0^1 x^\alpha dx = \left[\frac{x^{\alpha+1}}{\alpha+1}\right]_0^1 = \frac{1}{1+\alpha}.$$

Integrating with respect to α, we get

$$F(\alpha) = \int_0^1 \frac{x^\alpha - 1}{\log x}dx$$

$$= \log(1+\alpha) + c \text{ (a constant of integration)}$$

But when $\alpha = 0$, $F(0) = \int_0^1 0\ dx = 0$. Therefore, $0 = \log 1 + c = c$. Hence,

$$F(\alpha) = \int_0^1 \frac{x^\alpha - 1}{\log x}dx = \log(1+\alpha).$$

EXAMPLE 3.82

Evaluate the integral

$$\int_0^\infty e^{-x}\frac{\sin \alpha x}{x}dx.$$

Solution. We cannot compute this integral directly because the anti-derivative of the function $e^{-x}\frac{\sin\alpha x}{x}$ is not expressible in terms of elementary functions. So we use Leibnitz's Rule to evaluate it. Let

$$F(\alpha) = \int_0^\infty e^{-x}.\frac{\sin\alpha x}{x}dx.$$

Then by Leibnitz's Rule, we have

$$\frac{d}{d\alpha}F(\alpha) = \frac{d}{d\alpha}\left[\int_0^\infty e^{-x}\frac{\sin\alpha}{x}dx\right]$$

$$= \int_0^\infty \frac{\partial}{\partial\alpha}\left[e^{-x}\frac{\sin\alpha x}{x}\right]dx$$

$$= \int_0^\infty e^{-x}\frac{1}{x}\cos\alpha x.x\ dx$$

$$= \int_0^\infty e^{-x}\cos\alpha x\ dx = \frac{1}{1+\alpha^2}.$$

Integrating, we get

$$F(\alpha) = \tan^{-1}\alpha + c. \qquad (1)$$

But,

$$F(0) = \int_0^\infty e^{-x}\frac{\sin 0x}{x}dx = \int_0^\infty 0\, dx = 0.$$

Therefore, (1) yields

$$0 = \tan^{-1} 0 + c \text{ and so, } c = 0.$$

Hence,

$$F(\alpha) = \int_0^\infty e^{-x}\frac{\sin \alpha x}{x}dx = \tan^{-1}\alpha.$$

EXAMPLE 3.83

Evaluate $\int_0^a \frac{\log(1+ax)}{1+x^2}dx$ and hence, show that $\int_0^1 \frac{\log(1+x)}{1+x^2}dx = \frac{\pi}{8}\log 2.$

Solution. We note that the limits of integration are not independent of the parameter a. Therefore, the formula mentioned in Remark after Theorem 3.11 is applicable. Let

$$F(a) = \int_0^a \frac{\log(1+ax)}{1+x^2}dx.$$

Then,

$$\frac{d}{da}\left[\int_0^a \frac{\log(1+ax)}{1+x^2}dx\right]$$

$$= \int_0^a \frac{\partial}{\partial a}\left[\frac{\log(1+ax)}{1+x^2}\right]dx + \frac{\log(1+a^2)}{1+a^2}\frac{d}{da}(a)$$

$$- \frac{\log(1+a.0)}{1+0}\cdot\frac{d}{da}(0)$$

$$= \int_0^a \frac{x}{(1+x^2)(1+ax)}dx + \frac{\log(1+a^2)}{1+a^2}$$

$$= \frac{1}{1+a^2}\int_0^a \left[\frac{-a}{1+ax} + \frac{x+a}{1+x^2}\right]dx$$

$$+ \frac{\log(1+a^2)}{1+a^2} \text{(by partial fractions)}$$

$$= \frac{1}{1+a^2}\left[-\log(1+ax) + \frac{1}{2}\log(1+x^2)\right.$$

$$\left. + a\tan^{-1}x\right]_0^a + \frac{\log(1+a^2)}{1+a^2}$$

$$= \frac{1}{1+a^2}\left[-\frac{1}{2}\log(1+a^2) + a\tan^{-1}a\right]$$

$$+ \frac{\log(1+a^2)}{1+a^2}$$

$$= \frac{1}{1+a^2}\left[\frac{1}{2}\log(1+a^2) + a\tan^{-1}a\right].$$

Integrating with respect to a, we get

$$F(a) = \int_0^a \frac{\log(1+ax)}{1+x^2}dx$$

$$= \frac{1}{2}\int \log(1+a^2)\frac{1}{1+a^2}da$$

$$+ \int \frac{a\tan^{-1}a}{1+a^2}da + c$$

$$= \frac{1}{2}\left[\log(1+a^2)\tan^{-1}a\right.$$

$$\left. - \int \frac{2a}{1+a^2}\tan^{-1}a\, da\right]$$

$$+ \int \frac{a\tan^{-1}a}{1+a^2}da + c$$

$$= \frac{1}{2}\log(1+a^2)\tan^{-1}a + c.$$

Substituting $a = 0$, we have $F(0) = 0$. Therefore,

$$0 = \frac{1}{2}\log 1 \ \tan^{-1} 0 + c \text{ and so, c = 0.}$$

Hence,

$$F(a) = \int_0^a \frac{\log(1+ax)}{1+x^2}dx = \frac{1}{2}\log(1+a^2)\tan^{-1}a.$$

Substituting $a = 1$, we get

$$\int_0^1 \frac{\log(1+x)}{1+x^2}dx = \frac{\pi}{8}\log 2.$$

EXAMPLE 3.84

Prove that

$$\int_{\frac{\pi}{2}-\alpha}^{\frac{\pi}{2}} \sin\theta\cos^{-1}(\cos\alpha \text{ cosec } \theta)d\theta = \frac{\pi}{2}(1-\cos\alpha).$$

Solution. Here the limits involve the parameter α. Let

$$F(\alpha) = \int_{\frac{\pi}{2}-\alpha}^{\frac{\pi}{2}} \sin\theta \cos^{-1}(\cos\alpha \operatorname{cosec}\theta)d\theta.$$

Then,

$$\begin{aligned}
\frac{d}{d\alpha}[F(\alpha)] &= \int_{\frac{\pi}{2}-\alpha}^{\frac{\pi}{2}} \frac{\partial}{\partial\alpha}[\sin\theta\cos^{-1}(\cos\alpha \operatorname{cosec}\theta)]d\theta \\
&\quad + \frac{d}{d\alpha}\left(\frac{\pi}{2}\right)\left[\sin\frac{\pi}{2}\cos^{-1}\left(\cos\alpha \operatorname{cosec}\frac{\pi}{2}\right)\right] \\
&\quad - \frac{d}{d\alpha}\left(\frac{\pi}{2}-\alpha\right)\left[\sin\left(\frac{\pi}{2}-\alpha\right)\cos^{-1}\left(\cos\alpha\operatorname{cosec}\left(\frac{\pi}{2}-\alpha\right)\right)\right] \\
&= \int_{\frac{\pi}{2}-\alpha}^{\frac{\pi}{2}} \frac{\sin\alpha d\theta}{\sqrt{1-\cos^2\alpha\operatorname{cosec}^2\theta}} + \sin\left(\frac{\pi}{2}-\alpha\right) \\
&\quad \times \cos^{-1}\left[\cos\alpha\operatorname{cosec}\left(\frac{\pi}{2}-\alpha\right)\right] \\
&= \int_{\frac{\pi}{2}-\alpha}^{\frac{\pi}{2}} \frac{\sin\alpha\sin\theta}{\sqrt{\sin^2\theta-\cos^2\alpha}}d\theta + \cos\alpha\cos^{-1}(1) \\
&= \int_0^{\sin\alpha} \frac{\sin\alpha}{\sqrt{\sin^2\alpha - t^2}}dt \text{ taking } \cos\theta = t \\
&= \left[\sin\alpha\sin^{-1}\left(\frac{t}{\sin\alpha}\right)\right]_0^{\sin\alpha} = \frac{\pi}{2}\sin\alpha.
\end{aligned}$$

Integrating with respect to α, we get

$$F(\alpha) = -\frac{\pi}{2}\cos\alpha + c.$$

But $F(0) = 0$, therefore,

$$0 = -\frac{\pi}{2} + c \text{ or } c = \frac{\pi}{2}.$$

Hence,

$$F(\alpha) = -\frac{\pi}{2}\cos\alpha + \frac{\pi}{2} = \frac{\pi}{2}[1-\cos\alpha].$$

EXAMPLE 3.85

If $y = \int_0^x f(t)\sin[k(x-t)]dt$, show that it satisfies the differential equation $\frac{d^2y}{dx^2} + k^2y = kf(x)$.

Solution. We have

$$y = \int_0^x f(t)\sin[k(x-t)]dt.$$

The upper limit in this integral involves the parameter x. So, using Leibnitz's Rule, we have

$$\begin{aligned}
\frac{dy}{dx} &= \int_0^x \frac{\partial}{\partial x}[f(t)\sin[k(x-t)]dt \\
&\quad + f(x)\sin[k(x-x)]\frac{d}{dx}(x) \\
&\quad - f(0)\sin[k(x-0)]\frac{d}{dx}(0) \\
&= \int_0^x kf(t)\cos[k(x-t)]dt.
\end{aligned}$$

Using once more the Leibnitz's Rule, we get

$$\begin{aligned}
\frac{d^2y}{dx^2} &= \int_0^x \frac{\partial}{\partial x}[kf(t)\cos[k(x-t)]dt \\
&\quad + kf(x)\cos[k(x-x)]\frac{d}{dx}(x) \\
&\quad - kf(0)\cos[k(x-0)]\frac{d}{dx}(0). \\
&= -k^2\int_0^x f(t)\sin[k(x-t)]dt + kf(x) \\
&= -k^2y + kf(x).
\end{aligned}$$

Hence,

$$\frac{d^2y}{dx^2} + k^2y = kf(x).$$

EXAMPLE 3.86

By successive use of Leibnitz's Rule to $\int_0^1 x^m dx$, evaluate $\int_0^1 x^m(\log x)^n dx$.

Solution. We have

$$I = \int_0^1 x^m dx = \left|\frac{x^{m-1}}{m+1}\right|_0^1 = \frac{1}{m+1}.$$

Therefore, using Leibnitz's Rule, we get

$$\frac{d}{dm}\left(\frac{1}{m+1}\right) = \int_0^1 \frac{\partial}{\partial m}(x^m)dx$$

or

$$-\left(\frac{1}{m+1}\right)^2 = \int_0^1 x^m \log x dx.$$

Applying again the Leibnitz's Rule, we get

$$\frac{d}{dm}\left(\frac{-1}{(m+1)^2}\right) = \int_0^1 \frac{\partial}{\partial m}(x^m \log x) dx$$

or

$$\frac{(-1)(-2)}{(m+1)^3} = \int_0^1 x^m (\log x)^2 dx.$$

Repeated use of Leibnitz's Rule yields

$$\frac{(-1)(-2)(-3)}{(m+1)^4} = \int_0^1 x^m (\log x)^3 dx$$

....

....

$$\frac{(-1)(-2)(-3)\ldots(-n)}{(m+1)^{n+1}} = \int_0^1 x^m (\log x)^n dx$$

or

$$\frac{(-1)^n n!}{(m+1)^{n+1}} = \int_0^1 x^m (\log x)^n dx.$$

3.17 APPROXIMATION OF ERRORS

In numerical computation, the quantity [True value – Approximate Value] is called the *error*.

We come across the following types of errors in numerical computation.

1. *Inherent Error (initial error).* Inherent error is the quantity which is already present in the statement (data) of the problem before its solution. This type of error arises due to the use of approximate value in the given data because there are limitations of the mathematical tables and calculators. This type of error can also be there due to mistakes by human. For example, some one can write, by mistake, 67 instead of 76. The error in this case is called *transposing* error.
2. *Round – off Error.* This error arises due to rounding off the numbers during computation and occur due to the limitation of computing aids. However, this type of error can be minimized by
   (i) Avoiding the subtraction of nearly equal numbers or division by a small number.
   (ii) Retaining at least one more significant figure at each step of calculation.
3. *Truncation Error.* If is the error caused by using approximate formulas during computation such as the one that arise when a function $f(x)$ is evaluated from an infinite series for x after truncating it at certain stage.

   For example, we will see that in Newton – Raphson Method for finding the roots of an equation, if x is the true value of the root of $f(x) = 0$ and x_0 and h are approximate value and correction respectively, then by Taylor's Theorem,

   $$f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2!} f''(x_0) + \cdots + = 0.$$

   To find the correction h, we truncate the series just after first derivative. Therefore some error occurs due to this truncation.
4. *Absolute Error.* If x is the true value of a quantity and x_0 is the approximate value, then $|x - x_0|$ is called the *absolute error.*
5. *Relative Error.* If x is the true value of a quantity and x_0 is the approximate value, then $\left(\frac{x-x_0}{x}\right)$ is called the *relative error.*
6. *Percentage Error.* If x is the true value of quantity and x_0 is the approximate value, then $\left(\frac{x-x_0}{x}\right) \times 100$ is called the *percentage error.* Thus, percentage error is 100 times the relative error.

3.18 GENERAL FORMULA FOR ERRORS

Let

$$u = f(u_1,\ u_2, \ldots, u_n) \tag{1}$$

be a function of $u_1,\ u_2, \ldots, u_n$ which are subject to the errors $\Delta u_1,\ \Delta u_2,\ \ldots, \Delta u_n$ respectively.

Let Δu be the error in u caused by the errors $\Delta u_1, \Delta u_2, \dots, \Delta u_n$ in $u_1,\ u_2, \dots, u_n$ respectively. Then

$$u + \Delta u = f(u_1 + \Delta u_1,\ u_2 + \Delta u_2, \dots, u_n + \Delta u_n) \qquad (2)$$

Expanding the right hand side of (2) by Taylor's Theorem for a function of several variables, we have

$$u + \Delta u = f(u_1, u_2, \dots, u_n) + \left(\Delta u_1 \frac{\partial}{\partial u_1} + \cdots + \Delta u_n \frac{\partial}{\partial u_n}\right) f + \frac{1}{2}\left(\Delta u_1 \frac{\partial}{\partial u_1} + \cdots + \Delta u_n \frac{\partial}{\partial u_n}\right)^2 f + \cdots$$

Since the errors are relatively small, we neglect the squares, product and higher powers and have

$$u + \Delta u = f(u_1, u_2, \dots, u_n) + \left(\Delta u_1 \frac{\partial}{\partial u_1} + \cdots + \Delta u_n \frac{\partial}{\partial u_n}\right) f \qquad (3)$$

Subtracting (1) from (3), we have

$$\Delta u = \frac{\partial f}{\partial u_1}\Delta u_1 + \frac{\partial f}{\partial u_2}\Delta u_2 + \cdots + \frac{\partial f}{\partial u_n}\Delta u_n$$

or

$$\Delta u = \frac{\partial u}{\partial u_1}\Delta u_1 + \frac{\partial u}{\partial u_2}\Delta u_2 + \cdots + \frac{\partial u}{\partial u_n}\Delta u_n,$$

which is known as ***general formula for error***. We note that the right hand side is simply the total derivative of the function u.

For a relative error E_r of the function u, we have

$$E_r = \frac{\Delta u}{u} = \frac{\partial u}{\partial u_1}\frac{\Delta u_1}{u} + \frac{\partial u}{\partial u_2}\frac{\Delta u_2}{u} + \cdots + \frac{\partial u}{\partial u_n}\frac{\Delta u_n}{u}.$$

EXAMPLE 3.87

If $u = \frac{5xy^2}{z^3}$ and error in x, y, z are 0.001, compute the relative maximum error $(E_r)_{\max}$ in u when $x = y = z = 1$.

Solution. We have $u = \frac{5xy^2}{z^3}$. Therefore

$$\frac{\partial u}{\partial x} = \frac{5y^2}{z^3}, \quad \frac{\partial u}{\partial y} = \frac{10xy}{z^3}, \quad \frac{\partial u}{\partial z} = -\frac{15xy^2}{z^4}$$

and so

$$\Delta u = \frac{5y^2}{z^3}\Delta x + \frac{10xy}{z^3}\Delta y - \frac{15xy^2}{z^4}\Delta z$$

But it is given that $\Delta x = \Delta y = \Delta z = 0.001$ and $x = y = z = 1$. Therefore

$$(\Delta u)_{\max} \approx \left|\frac{5y^2}{z^3}\Delta x\right| + \left|\frac{10xy}{z^3}\Delta y\right| + \left|\frac{15xy^2}{z^4}\Delta z\right| = 5(0.001) + 10(0.001) + 15(0.001) = 0.03.$$

Thus the relative maximum error $(E_r)_{\max}$ is given by

$$(E_r)_{\max} = \frac{(\Delta u)_{\max}}{u} = \frac{0.03}{u} = \frac{0.03}{5} = 0.006.$$

EXAMPLE 3.88

The diameter and altitude of a can in the shape of a right circular cylinder are measured as 4 cm and 6 cm respectively. The possible error in each measurement is 0.1 cm. Find approximately the maximum possible error in the values computed for the volume and the lateral surface.

Solution. If x and y denote the diameter and the height of the can, then volume of the can is given by $V = \frac{\pi}{4}x^2 y$ and so

$$\frac{\partial V}{\partial x} = \frac{\pi}{2}xy \quad \text{and} \quad \frac{\partial V}{\partial y} = \frac{\pi}{4}x^2.$$

Therefore the error formula yields

$$\Delta V = \frac{\partial V}{\partial x}\Delta x + \frac{\partial V}{\partial y}\Delta y = \frac{\pi}{2}(xy\Delta x) + \frac{\pi}{4}(x^2\Delta y).$$

Putting $x = 4$, $y = 6$, $\Delta x = \Delta y = 0.1$, we get

$$\Delta V = (1.2)\pi + (0.4)\pi = 1.6\,\pi \text{ cm}^3.$$

Further, the lateral surface is given by $S = \pi xy$ and so

$$\frac{\partial S}{\partial x} = \pi y \quad \text{and} \quad \frac{\partial S}{\partial y} = \pi x.$$

Therefore

$$\Delta S = \pi(y\Delta x + x\Delta y).$$

Putting the values of x, y, Δx and Δy, we get

$$\Delta S = \pi(0.6 + 0.4) = \pi \text{ cm}^2.$$

EXAMPLE 3.89

The height h and the semi-vertical angle α of a cone are measured and from them the total area A of the cone (including the base) is calculated. If h and α are in error by small quantities δh *and* $\delta\alpha$ respectively, find corresponding error in the area. Show further that $\alpha = \frac{\pi}{6}$, an error of 1 percent in h will be approximately compensated by an error of -0.33 degree in α.

Solution. Radius of the base $= r = h\tan\alpha$. Further, slant height $= l = h\sec\alpha$. Therefore

$$\begin{aligned}\text{Total area } &= \pi r^2 + \pi r\,l = \pi r\,(r + l)\\ &= \pi h \tan\alpha(h \tan\alpha + h \sec\alpha)\\ &= \pi h^2(\tan^2\alpha + \sec\alpha\tan\alpha).\end{aligned}$$

Then the error in A is given by

$$\begin{aligned}\delta A &= \frac{\partial A}{\partial h}\delta h + \frac{\partial A}{\partial \alpha}\delta\alpha\\ &= 2\pi h(\tan^2\alpha + \sec\alpha\tan\alpha)\delta h\\ &\quad + \pi h^2(2\tan\alpha\sec^2\alpha + \sec^3\alpha + \sec\alpha\tan^2\alpha)\delta\alpha\end{aligned}$$

For the second part of the question,

$$\alpha = \frac{\pi}{6}, \quad \delta h = \frac{h}{100}.$$

Therefore

$$\begin{aligned}\delta A &= 2\pi h\left[\frac{1}{3} + \frac{2}{3}\right]\frac{h}{100}\\ &\quad + \pi h^2\left(\frac{2}{\sqrt{3}}\left(\frac{4}{3}\right) + \frac{8}{3\sqrt{3}} + \frac{2}{3\sqrt{3}}\right)\delta\alpha\\ &= \frac{\pi h^2}{50} + 2\sqrt{3}\pi h^2\delta\alpha \qquad (1)\end{aligned}$$

But after compensation $\delta A = 0$. Therefore (1) implies

$$\delta\alpha = -\frac{1}{100\sqrt{3}} \text{ radians } = -\frac{57.3^\circ}{173.2} = -0.33^\circ$$

EXAMPLE 3.90

The time T of a complete oscillation of a simple pendulum of length L is governed by the equation $T = 2\pi\sqrt{L/g}$, g is constant, find the approximate error in the calculated value of T corresponding to the error of 2% in the value of L.

Solution. We have

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

Taking logarithm, we get

$$\log T = \log 2\pi + \frac{1}{2}\log\, l - \frac{1}{2}\log g \qquad (1)$$

Differentiating (1), we get

$$\frac{1}{T}\delta T = \frac{1}{2}\frac{\delta l}{l} - \frac{1}{2}\frac{\delta g}{g}$$

or

$$\begin{aligned}\frac{\delta T}{T}\times 100 &= \frac{1}{2}\left[\frac{\delta l}{l}\times 100 - \frac{1}{2}\frac{\delta g}{g}\times 100\right]\\ &= \frac{1}{2}[2 - 0] = 1.\end{aligned}$$

Hence the approximate error is 1%.

3.19 MISCELLANEOUS EXAMPLES

EXAMPLE 3.91

If $z = f(x, y)$ and u, v are two variables such that $u = lx + my, v = ly - mx$. Prove that

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = \left(l^2 + m^2\right)\left[\frac{\partial^2 z}{\partial u^2} + \frac{\partial^2 z}{\partial v^2}\right]$$

Solution. We have

$$u = lx + my, \quad v = ly - mx,$$
$$\frac{\partial u}{\partial x} = l, \ \frac{\partial v}{\partial x} = -m$$
$$\frac{\partial u}{\partial y} = m, \ \frac{\partial v}{\partial y} = l.$$

Therefore

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial z}{\partial v} \cdot \frac{\partial v}{\partial x} = l\frac{\partial z}{\partial u} - m\frac{\partial z}{\partial v} \quad (1)$$

$$\frac{\partial z}{\partial y} = \frac{\partial z}{\partial u} \cdot \frac{\partial u}{\partial y} + \frac{\partial z}{\partial v} \cdot \frac{\partial v}{\partial y} = m\frac{\partial z}{\partial u} + l\frac{\partial z}{\partial v} \quad (2)$$

From (1) and (2), we have

$$\frac{\partial}{\partial x} = l\frac{\partial}{\partial u} - m\frac{\partial}{\partial v}$$

and

$$\frac{\partial}{\partial y} = m\frac{\partial}{\partial u} + l\frac{\partial}{\partial v}.$$

Therefore

$$\begin{aligned}\frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \left(l\frac{\partial}{\partial u} - m\frac{\partial}{\partial v}\right)\left(l\frac{\partial z}{\partial u} - m\frac{\partial z}{\partial v}\right)\\ &= l\frac{\partial}{\partial u}\left(l\frac{\partial z}{\partial u} - m\frac{\partial z}{\partial v}\right) - m\frac{\partial}{\partial v}\left(l\frac{\partial z}{\partial u} - m\frac{\partial z}{\partial v}\right)\\ &= l^2\frac{\partial^2 z}{\partial u^2} - lm\frac{\partial^2 z}{\partial u \partial v} - lm\frac{\partial^2 z}{\partial v \partial u} + m^2\frac{\partial^2 z}{\partial v^2}\\ &= l^2\frac{\partial^2 z}{\partial u^2} + m^2\frac{\partial^2 z}{\partial v^2} - 2lm\frac{\partial^2 z}{\partial u \partial v} \quad (3)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2 z}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) = \left(m\frac{\partial}{\partial u} + l\frac{\partial}{\partial v}\right)\left(m\frac{\partial z}{\partial u} + l\frac{\partial z}{\partial v}\right)\\ &= m\frac{\partial}{\partial u}\left(m\frac{\partial z}{\partial u} + l\frac{\partial z}{\partial v}\right) + l\frac{\partial}{\partial v}\left(m\frac{\partial z}{\partial u} + l\frac{\partial z}{\partial v}\right)\\ &= m^2\frac{\partial^2 z}{\partial u^2} + lm\frac{\partial^2 z}{\partial u \partial v} + lm\frac{\partial^2 z}{\partial v \partial u} + l^2\frac{\partial^2 z}{\partial v^2}\\ &= l^2\frac{\partial^2 z}{\partial v^2} + m^2\frac{\partial^2 z}{\partial u^2} + 2lm\frac{\partial^2 z}{\partial u \partial v}. \quad (4)\end{aligned}$$

Adding (3) and (4), we get

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = (l^2 + m^2)\left(\frac{\partial^2 z}{\partial u^2} + \frac{\partial^2 z}{\partial v^2}\right).$$

EXAMPLE 3.92

(a) If $f(x, y) = (1 - 2\,xy + y^2)^{-1/2}$, show that

$$\frac{\partial}{\partial x}\left[(1 - x^2)\frac{\partial f}{\partial x}\right] + \frac{\partial}{\partial y}\left[y^2\frac{\partial f}{\partial y}\right] = 0.$$

(b) If $V = f(2x - 3y, 3y - 4z, 4z - 2x)$, compute the value of $6V_x + 4\,V_y + 3V_z$.

Solution. (a) We have, $f = (1 - 2xy + y^2)^{-\frac{1}{2}}$. Then

$$\begin{aligned}\frac{\partial f}{\partial x} &= -\frac{1}{2}(1 - 2xy + y^2)^{-\frac{3}{2}}(-2y)\\ &= y(1 - 2xy + y^2)^{-\frac{3}{2}},\\ \frac{\partial^2 f}{\partial x^2} &= -\frac{3}{2}y(1 - 2xy + y^2)^{-\frac{5}{2}}(-2y)\\ &= 3y^2(1 - 2xy + y^2)^{-\frac{5}{2}}.\end{aligned}$$

Therefore

$$\begin{aligned}&\frac{\partial}{\partial x}\left[(1 - x^2)\frac{\partial f}{\partial x}\right]\\ &= \frac{\partial}{\partial x}(1 - x^2)\frac{\partial f}{\partial x} + (1 - x^2)\frac{\partial^2 f}{\partial x^2}\\ &= (-2x)\frac{\partial f}{\partial x} + (1 - x^2)\frac{\partial^2 f}{\partial x^2}\\ &= -2xy(1 - 2xy + y^2)^{-\frac{3}{2}}\\ &\quad + 3(1 - x^2)y^2(1 - 2xy + y^2)^{-\frac{5}{2}}. \quad (1)\end{aligned}$$

Similarly differentiating partially with respect to y, we get

$$\frac{\partial f}{\partial y} = (x-y)(1-2xy+y^2)^{-\frac{3}{2}},$$

$$\frac{\partial^2 f}{\partial y^2} = -(1-2xy+y^2)^{-\frac{3}{2}} + 3(x-y)^2(1-2xy+y^2)^{-\frac{5}{2}}.$$

Therefore

$$\begin{aligned}\frac{\partial}{\partial y}\left[y^2\frac{\partial f}{\partial y}\right] &= \frac{\partial}{\partial y}(y^2)\frac{\partial f}{\partial y} + y^2\frac{\partial^2 f}{\partial y^2}\\ &= 2y(x-y)(1-2xy+y^2)^{-\frac{3}{2}}\\ &\quad + y^2[-(1-2xy+y^2)^{-\frac{3}{2}}\\ &\quad + 3(x-y)^2(1-3xy+y^2)^{-\frac{5}{2}}]\\ &= y(1-2xy+y^2)^{-\frac{3}{2}}[3y(x-y)^2\\ &\quad \times (1-2xy+y^2)^{-1} + (2x-3y)] \qquad (2)\end{aligned}$$

Adding (1) and (2), we get the required result.

(b) We have

$$V = f(2x-3y,\ 3y-4z,\ 4z-2x).$$

Let

$$r = 2x-3y,\ \ s = 3y-4z \text{ and } t = 4z-2x.$$

Then

$$V = f(r,s,t).$$

Further,

$$\begin{aligned}\frac{\partial V}{\partial x} &= \frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial x} + \frac{\partial V}{\partial s}\frac{\partial s}{\partial x} + \frac{\partial V}{\partial t}\cdot\frac{\partial t}{\partial x}\\ &= 2\frac{\partial V}{\partial r} + 0 - 2\frac{\partial V}{\partial t} = 2\frac{\partial V}{\partial r} - 2\frac{\partial V}{\partial t} \qquad (1)\end{aligned}$$

$$\begin{aligned}\frac{\partial V}{\partial y} &= \frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial y} + \frac{\partial V}{\partial s}\frac{\partial s}{\partial y} + \frac{\partial V}{\partial t}\cdot\frac{\partial t}{\partial y}\\ &= -3\frac{\partial V}{\partial r} + 3\frac{\partial V}{\partial s} + 0 = 3\frac{\partial V}{\partial r} + 3\frac{\partial V}{\partial s} \qquad (2)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial V}{\partial z} &= \frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial z} + \frac{\partial V}{\partial s}\frac{\partial s}{\partial z} + \frac{\partial V}{\partial t}\cdot\frac{\partial t}{\partial z}\\ &= 0 - 4\frac{\partial V}{\partial s} + 4\frac{\partial V}{\partial t} = -4\frac{\partial V}{\partial s} + 4\frac{\partial V}{\partial t}.\end{aligned}$$

The relations (1), (2) and (3) yields

$$\begin{aligned}&6V_x + 4V_y + 3V_z\\ &= 6\left(2\frac{\partial V}{\partial r} - 2\frac{\partial V}{\partial t}\right) + 4\left(-3\frac{\partial V}{\partial r} + 3\frac{\partial V}{\partial s}\right)\\ &\quad + 3\left(-4\frac{\partial V}{\partial s} + 4\frac{\partial V}{\partial t}\right) = 0.\end{aligned}$$

EXAMPLE 3.93

(a) If $u = \sin^{-1}\left(\frac{3x^2+4y^2}{3x+4y}\right)$, prove that $x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = \tan u$.

(b) If $u = x^3 + y^3 + z^3 + 3xyz$, show that

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = 3u.$$

(c)If $u = \log\left(\frac{x^2+y^2}{x+y}\right)$, prove that $x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 1$.

Solution. (a) We have

$$\sin u = \frac{3x^2+4y^2}{3x+4y} = z, \ \ \text{say}.$$

Thus

$$z = \frac{3x^2}{3x}\left(\frac{1+\frac{4}{3}\left(\frac{y}{x}\right)^2}{1+\frac{4}{3}\left(\frac{y}{x}\right)}\right) = x\left(\frac{1+\frac{4}{3}\left(\frac{y}{x}\right)^2}{1+\frac{4}{3}\left(\frac{y}{x}\right)}\right),$$

and so z is a homogeneous function of degree 1 in x and y. Hence, by Euler's Theorem, we have

$$x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = z. \qquad (1)$$

But

$$\frac{\partial z}{\partial x} = \cos u\frac{\partial u}{\partial x} \quad \text{and} \quad \frac{\partial z}{\partial y} = \cos u\frac{\partial u}{\partial y}.$$

Hence (1) reduces to

$$x\cos u\frac{\partial u}{\partial x} + y\cos u\frac{\partial u}{\partial y} = \sin u$$

or

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = \tan u.$$

(b) We have

$$u = x^3 + y^3 + z^3 + 3xyz.$$

Replacing x by tx, y by ty and z by tz, we get

$$u(tx, ty, tz) = t^3x^3 + t^3y^3 + t^3z^3 + 3txtytz$$
$$= t^3(x^3 + y^3 + z^3 + 3xyz) = t^3u(x, y, z).$$

Hence $u(x, y, z)$ is a homogeneous function of degree 3 in $u(x, y, z)$. Therefore, by Euler's Theorem,

$$x\frac{\partial u}{\partial z} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = 3u.$$

(c) We have

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

Therefore

$$e^u = \frac{x^2 + y^2}{x + y},$$

which is homogeneous function of degree 1 in x and y. Therefore, by Euler's Theorem, we have

$$x\frac{\partial}{\partial x}(e^u) + y\frac{\partial}{\partial y}(e^u) = e^u$$

or

$$x\, e^u\frac{\partial u}{\partial x} + y\, e^u\frac{\partial u}{\partial y} = e^u$$

or

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 1.$$

EXAMPLE 3.94

Find the stationary points of $x^2 - xy + y^2 - 2x + y$.

Solution. We have $f(x, y) = x^2 - xy + y^2 - 2x + y$,

$$f_x = 2x - y - 2,\ f_y = -x + 2y + 1.$$

Therefore the stationary points are given by

$$f_x = f_y = 0$$

and so

$$2x - y - 2 = 0 \quad \text{and} \quad -x + y + 1 = 0.$$

Solving these equations, we get the stationary point as $\left(\frac{3}{5}, \frac{-4}{5}\right)$.

EXAMPLE 3.95

The temperature $u(x, y, z)$ at any point in space is $u = 400xyz^2$. Find the highest temperature on surface of the sphere $x^2 + y^2 + z^2 = 1$.

Solution. We have

$$u\ (x, y, z) = 400\ xyz^2 = 400\ xy\ (1 - y^2 - x^2)$$
$$= 400\ xy - 400\ xy^3 - 400\ x^3y,$$

which is a function of two variables x and y. Then

$$u_x = 400y - 400y^3 - 1200x^2y,$$
$$u_y = 400x - 1200xy^2 - 400x^3.$$

For extreme points, we must have $u_x = u_y = 0$. Thus $1 - y^2 - 3x^2 = 0$ and $1 - 3y^2 - x^2 = 0$. Solving these equations, we get $x = \pm\frac{1}{2},\ y = \pm\frac{1}{2}$. Thus we have four stationary points $\left(\frac{1}{2}, \frac{1}{2}\right)$, $\left(-\frac{1}{2}, -\frac{1}{2}\right)$, $\left(\frac{1}{2}, -\frac{1}{2}\right)$ and $\left(-\frac{1}{2}, \frac{1}{2}\right)$. Also

$$u_{xx} = -2400xy, u_{xy} = 400 - 1200y^2 - 1200x^2$$
$$u_{yy} = -2400xy.$$

At $\left(\frac{1}{2}, \frac{1}{2}\right)$ and $\left(-\frac{1}{2}, -\frac{1}{2}\right)$, $rt - s^2$ is positive and r is negative. Therefore maximum exist at these points. Further, $x = \frac{1}{2}, y = \frac{1}{2}$, give $z = \sqrt{1 - x^2 - y^2} = \frac{1}{\sqrt{2}}$. Therefore

$$\max u\ (x, y, z) = 400\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 50.$$

EXAMPLE 3.96

A flat circular plate is heated so that the temperature at any point (x, y) is $u(x, y) = x^2 + 2y^2 - x$. Find the coldest point on the plate.

Solution. We have $u\ (x, y) = x^2 + 2y^2 - x$, so that

$$u_x = 2x - 1, \quad u_y = 4y.$$

Then $u_x = u_y = 0$ imply $x = \frac{1}{2}, y = 0$. Also $u_{xx} = 2, u_{yy} = 4$ and $u_{xy} = 0$. Then $rt - s^2 = 8(+\text{ve})$ and $r = 2$. Therefore u is minimum at $\left(\frac{1}{2}, 0\right)$. Therefore the coldest point is $\left(\frac{1}{2}, 0\right)$.

EXAMPLE 3.97

Find the minimum value of $x^2 + y^2 + z^2$ when $x + y + z = 3a$.

Solution. Special case of Example 3.57 (putting $a = b = c = 1$ and $p = 3a$ in that example).

EXAMPLE 3.98

(a) If $u = x + y$ and $y = uv$, find the Jacobian $\frac{\partial(x,y)}{\partial(u,v)}$.

(b) Show that the functions $u = x^2 + y^2 + z^2$, $v = x + y + z$, $w = yz + zx + xy$ are not independent of one another.

Solution. (a) We are given that $u = x + y, \;\; y = uv$. Therefore

$$x = u - y = u - uv \quad \text{and} \quad y = uv.$$

Therefore

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 - v & -u \\ v & u \end{vmatrix}$$
$$= u(1 - v) + uv = u.$$

(b) We have

$$u = x^2 + y^2 + z^2, \quad v = x + y + z,$$
$$w = yz + zx + xy.$$

Then

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix} = \begin{vmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \\ z + y & z + x & x + y \end{vmatrix}$$
$$= 2x[(x + y) - (z + x) - 2y[(x + y)$$
$$- (z + y)] + 2z[(z + x) - (z + y)] = 0.$$

Since Jacobian $J(u, v, w) = 0$, there exists a functional relation connecting some or all of the variables x, y and z. Hence u, v, w are not independent.

EXAMPLE 3.99

(a) If $x = e^u \cos v$ and $y = e^u \sin v$, show that $J \cdot J' = 1$.

(b) Verify the chain rule for Jacobians if $x = u$, $y = u \tan v, \; z = w$.

Solution. (a) We have

$$x = e^u \cos v, \quad y = e^u \sin v.$$

Then

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} e^u \cos v & -e^u \sin v \\ e^u \sin v & e^u \cos u \end{vmatrix}$$
$$= e^{2u} \cos^2 u + e^{2u} \sin^2 v = e^{2u}.$$

One the other hand the given equations yield

$$x^2 + y^2 = e^{2u} \quad \text{and} \quad v = \tan^{-1} \frac{y}{x}.$$

Therefore

$$2e^{2u} \frac{\partial u}{\partial x} = 2x \text{ which yields } \frac{\partial u}{\partial x} = \frac{x}{e^{2u}},$$
$$2e^{2u} \frac{\partial u}{\partial y} = 2y \text{ which yields } \frac{\partial u}{\partial y} = \frac{y}{e^{2u}},$$
$$\frac{\partial v}{\partial x} = \frac{-y}{x^2 + y^2} = \frac{-y}{e^{2u}}, \quad \frac{\partial v}{\partial y} = \frac{x}{x^2 + y^2} = \frac{x}{e^{2u}}.$$

Therefore

$$J' = \frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{x}{e^{2u}} & \frac{y}{e^{2u}} \\ -\frac{y}{e^{2u}} & \frac{x}{e^{2u}} \end{vmatrix} = \frac{2x^2}{e^{4u}} + \frac{2y^2}{e^{4u}} = \frac{e^{2u}}{e^{4u}} = \frac{1}{e^{2u}}.$$

Hence $JJ' = 1$.

(b) We are given thet

$$x = u, \; y = u \tan v, \;\; z = w \tag{1}$$

Then

$$J = \frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}$$
$$= \begin{vmatrix} 1 & 0 & 0 \\ \tan v & u \sec^2 v & 0 \\ 0 & 0 & 1 \end{vmatrix} = u \sec^2 v.$$

Also, from (1), we have

$$u = x\ ,\ v = \tan^{-1}\frac{y}{x} \quad \text{and} \quad w = z.$$

Therefore

$$J' = \frac{\partial(u,v,w)}{\partial(x,y,z)}$$
$$= \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix} - \begin{vmatrix} 1 & 0 & 0 \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} & 0 \\ 0 & 0 & 1 \end{vmatrix}$$
$$= \frac{x}{x^2+y^2} = \frac{1}{x\left[1+\left(\frac{y}{x}\right)^2\right]}$$
$$= \frac{1}{u(1+\tan^2 v)}, \quad \text{since } \frac{y}{u} = \tan v$$
$$= \frac{1}{u\sec^2 v}.$$

Hence

$$J\,J' = 1, \text{ which proves the chain rule.}$$

EXAMPLE 3.100

Assuming $\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$, prove that $\int_0^\infty e^{-x^2}\cos\ 2ax\ dx = \frac{\sqrt{\pi}}{2}e^{-a^2}$.

Solution. Let

$$F(a) = \int_0^\infty e^{-x^2}\cos 2ax\ dx.$$

Then

$$F'(a) = \int_0^\infty e^{-x^2}\frac{\partial}{\partial a}(\cos 2ax)dx$$
$$= \int_0^\infty -2xe^{-x^2}\sin 2ax\ dx$$
$$= [e^{-x^2}\sin 2ax]_0^\infty - 2a\int e^{-x^2}\cos 2ax\ dx$$
$$= -2aF(a).$$

Therefore

$$\frac{F'(a)}{F(a)} = -2a.$$

Integrating, we get

$$\log F(a) = -\frac{2a^2}{2} = -a^2 + \log c$$

or

$$\log\frac{F(a)}{c} = -a^2$$

or

$$F(a) = c\ e^{-a^2}.$$

Putting $a = 0$, we get $c = F(0) = \int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$.
Hence

$$F(a) = \int_0^\infty e^{-a^2}\cos 2ax\ dx = \frac{\sqrt{\pi}}{2}e^{-a^2}.$$

EXERCISES

1. Let $f\colon \Re^2 \to \Re$ be a continuous function. Define $\phi\colon \Re^2 \to \Re$ by
$$\phi(x,\ y) = \begin{cases} f(x,\ y) \text{ if } (x,\ y) \neq (0,0) \\ f(x,\ y)+1 \text{ if } (x,\ y) = (0,0). \end{cases}$$
Show that ϕ is not continuous at (0, 0).
Hint: $\phi(x,\ y) = f(x,\ y)+1$ for (0, 0) and so, $\phi(0,\ 0) = f(0,\ 0)+1$. Since f is continuous, it is continuous at $(0,0)$ also. So $\lim_{(x,y)\to(0,0)} f(x,y) = f(0,0)$. Then, $\lim_{(x,y)\to(0,0)}\phi(x,y) = \lim_{(x,y)\to(0,0)} f(x,y) = f(0,\ 0)$. Thus, $\lim_{(x,y)\to(0,0)}\phi(x,y) \neq \phi(0,0)$. Hence, ϕ is not continuous at $(0,0)$.
2. Show that $f(x,\ y) = x^2 + y - 1$ is continuous at $(1,\ -2)$
3. Prove that the function $f(x,y) = (|xy|)^{\frac{1}{2}}$ is not differentiable at the point (0, 0) but that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist at the origin and have value zero. Deduce that these two partial derivatives are continuous except at the origin.
Hint: If $f(x,\ y)$ is differentiable at (0, 0), then
$$\Delta f = f(h,k) = h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} + \varepsilon\sqrt{h^2+k^2},$$

where $\varepsilon \to 0$ as $\sqrt{h^2+k^2} \to 0$. Since $f_x = f_y = 0$ at (0, 0), $f(h,k) = 0 + 0 + \varepsilon\sqrt{h^2+k^2}$ and so, $\varepsilon = \frac{[|hk|]^{\frac{1}{2}}}{\sqrt{h^2+k^2}}$. Put $h = \rho\cos\theta$ and $k = \rho\sin\theta$, so that $\sqrt{h^2+k^2} = \rho$. Then, $\varepsilon = \sqrt{|\sin\theta\cos\theta|}$ and so, $\lim_{\rho\to 0}\varepsilon = \sqrt{|\sin\theta\cos\theta|} = 0$, which is absurd. Hence, f is not differentiable at (0, 0).

4. If $u = x\phi\left(\frac{y}{x}\right) + \psi\left(\frac{y}{x}\right)$, prove that
$$x^2\frac{\partial^2 u}{\partial x^2} + 2xy\frac{\partial^2 u}{\partial x\partial y} + y^2\frac{\partial^2 u}{\partial y^2} = 0.$$
5. If $u = x^y$, show that $\frac{\partial^3 u}{\partial x^2\partial y} = \frac{\partial^3 u}{\partial x\partial y\partial x}$.
6. If $z = f\left(\frac{y}{x}\right)$, show that $x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = 0$.
7. If $z = \frac{x^2+y^2}{x+y}$, show that
$$\left(\frac{\partial z}{\partial x} - \frac{\partial z}{\partial y}\right)^2 = 4\left(1 - \frac{\partial z}{\partial x} - \frac{\partial z}{\partial y}\right).$$
8. If $u = (1 - 2xy + y^2)^{-\frac{1}{2}}$, show that
$$\frac{\partial}{\partial x}\left\{(1-x^2)\frac{\partial u}{\partial x}\right\} + \frac{\partial}{\partial y}\left\{y^2\frac{\partial u}{\partial y}\right\} = 0.$$
9. Find the value of $\frac{1}{a^2}\frac{\partial^2 z}{\partial x^2} + \frac{1}{b^2}\frac{\partial^2 z}{\partial y^2}$, where $a^2x^2 + b^2y^2 - c^2z^2 = 0$.

   Ans. $\frac{1}{c^2 z}$.
10. If $u = \frac{y}{z} + \frac{z}{x} + \frac{x}{y}$, show that $x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = 0$.
11. If $u = e^x(x\cos y - y\sin y)$, show that $u_{xx} + u_{yy} = 0$.
12. Find the envelope of a system of concentric- and coaxial ellipses of constant area.

    Hint: Area of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ is πab, given to be constant say equal to πc^2 and so, $ab = c^2$ or $b = \frac{c^2}{a}$. Putting $b = \frac{c^2}{a}$ in $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, we get $\frac{x^2}{a^2} + \frac{y^2}{c^4a^2} = 0$. Differentiating partially with respect to a gives $a = \frac{c^2 x}{y}$. Putting this value of a in $\frac{x^2}{a^2} + \frac{y^2}{c^4a^2} = 0$, we get $2xy = c^2$, which is the required equation of the envelope.
13. Find the envelope of the straight line $\frac{x}{a} + \frac{y}{b} = 1$, where $a^2 + b^2 = c^2$.

    Ans. $\sqrt{x} + \sqrt{y} = \sqrt{c}$.
14. Show that the evolute of the rectangular hyperbola $xy = c^2$ is
$$(x+y)^{\frac{2}{3}} - (x-y)^{\frac{2}{3}} = (4c)^{\frac{2}{3}}.$$
15. If $u = \log\frac{x^4+y^4}{x+y}$, show that $x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 3$.

    Hint: $e^u = \frac{x^4+y^4}{x+y}$ (A homogeneous function of degree 3 in x and y).
16. If $u = \sin^{-1}\left(\frac{x+y}{\sqrt{x}+\sqrt{y}}\right)$, show that
$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = \frac{1}{2}\tan u.$$
    Hint: $\sin u = \frac{x+y}{\sqrt{x}+\sqrt{y}}$ is a homogeneous function of degree $\frac{1}{2}$.
17. If $u = \sin^{-1}\left(\frac{x+y}{\sqrt{x}+\sqrt{y}}\right)$, show that
$$x^2\frac{\partial^2 u}{\partial x^2} + 2xy\frac{\partial^2 u}{\partial x\partial y} + y^2\frac{\partial^2 u}{\partial y^2} = \frac{-\sin u\cos 2u}{4\cos^3 u}.$$
18. If $u = \tan^{-1}\left(\frac{y^2}{x}\right)$, show that
$$x^2\frac{\partial^2 u}{\partial x^2} + 2xy\frac{\partial^2 u}{\partial x\partial y} + y^2\frac{\partial^2 u}{\partial y^2} = -\sin 2u\sin^2 u.$$
19. If $u = \sin^{-1}\left(\frac{x^3+y^3+z^3}{ax+by+cz}\right)$, show that
$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = 2\tan u.$$
20. If $u = \cos^{-1}\frac{x+y}{\sqrt{x}+\sqrt{y}}$, show that
$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = -\frac{1}{2}\cot u.$$
21. If $u = \tan^{-1}\left(\frac{y}{x}\right)$, where $x = e^t - e^{-t}$ and $y = e^t + e^{-t}$, find $\frac{du}{dt}$.

    Ans. $\frac{-2}{e^{2x}+e^{-2x}}$.
22. If $z = e^{ax+by}f(ax - by)$, show that
$$b\frac{\partial z}{\partial x} + a\frac{\partial z}{\partial y} = 2ab\,z.$$
23. If $\sqrt{1-x^2} + \sqrt{1-y^2} = a(x-y)$, find $\frac{dy}{dx}$.

    Ans. $\sqrt{\frac{1-y^2}{1-x^2}}$.
24. If $y^3 - 3ax^2 + x^3 = 0$, show that
$$\frac{d^2y}{dx^2} + \frac{2a^2x^2}{y^4} = 0.$$
25. If $x = r\cos\theta$ and $y = r\sin\theta$, show that
$$\frac{\partial^2 r}{\partial x^2} + \frac{\partial^2 r}{\partial y^2} = \frac{1}{r}\left[\left(\frac{\partial r}{\partial x}\right)^2 + \left(\frac{\partial r}{\partial y}\right)^2\right].$$

26. If $x+y=2e^{\theta}\cos\phi$ and $x-y=2\sqrt{-1}\,e^{\theta}\sin\phi$, show that
$$\frac{\partial^2 V}{\partial\theta^2}+\frac{\partial^2 V}{\partial\phi^2}=4xy\frac{\partial^2 V}{\partial x\partial y}.$$
27. Expand $e^x\log(1+y)$ in a Taylor's series in the neighborhood of the point (0, 0).
Ans. $y+xy-\frac{y^2}{2}+\frac{x^2y}{2}-\frac{xy^2}{2}+\frac{1}{3}y^2-\ldots$.
28. Expand $e^x\cos y$ in powers of x and y up to third-degree terms.
Ans. $1+x+\frac{1}{2}(x^2-y^2)+\frac{1}{6}(x^3-3xy^2)+\ldots$.
29. Show that for $0<\theta<1$, $\sin x\sin y=xy-\frac{1}{6}[(x^3+3xy^2)\cos\theta x\sin\theta y+(y^3+3x^2y)\sin\theta x\cos\theta y]$.
Hint: Use Maclaurin's theorem.
30. If the perimeter of a triangle is constant, show that its area is maximum when the triangle is equilateral.
Hint: $2s=a+b+c$ $c=2s-a-b$ and
$$\Delta=\sqrt{s(s-a)(s-b)(s-c)}$$
$$=\sqrt{s(s-a)(s-b)(a+b-s)}$$
Take $f(a,b)=\Delta^2=s(s-a)(s-b)(a+b-s)$ and find f_a and f_b etc, and proceed.
31. Find the points (x, y), where the function $xy(1-x-y)$ is either maximum or minimum.
Ans. $\left(\frac{1}{3},\frac{1}{3}\right)$.
32. Find a point within a triangle such that the sum of the squares of its distances from the three vertices is the minimum.
Ans. $\left(\frac{x_1+x_2+x_3}{3},\frac{y_1+y_2+y_3}{3}\right)$ (centroid of the triangle).
33. Find the point on the plane $2x+3y-z=12$ that is nearest to the origin.
Hint: Distance form the origin is
$$l=\sqrt{x^2+y^2+(2x+3y-12)^2}.$$
Put $f(x,y)=x^2+y^2+(2x+3y-12)^2$ and proceed.
Ans. $\left(\frac{12}{7},\frac{18}{7},\frac{-6}{7}\right)$.
34. Discuss the maxima and minima of
$$u=2\sin\frac{1}{2}(x+y)\cos\frac{1}{2}(x-y)+\cos(x+y).$$
35. Find the extreme values of $2(x^2-y^2)-x^4+y^4$.
Ans. Max at $(\pm1,0)$ and Min at $(0,\pm1)$.
36. Find the maximum value of $x^m y^n z^p$ subject to the condition $x+y+z=a$.
Hint: Use Lagrange's method of undetermined multipliers.
Ans. $x=\frac{ma}{m+n+p}$, $y=\frac{na}{m+n+p}$, and $z=\frac{pa}{m+n+p}$.
$$\text{Max. value}=\frac{m^m n^n p^{m+n+p}}{(m+n+p)^{m+n+p}}.$$
37. Divide 24 into three parts such that the continued product of the first, square of the second, and the cube of the third part may be maximum.
Hint: Find the Max of xy^2z^3 subject to the condition $x+y+z=24$. Also can be obtained from Exercise 36 by putting $a=24$, $m=1$, $n=2$, and $p=3$.
38. The temperature T at any point (x, y, z) in space is $T=400xyz^2$. Find the highest temperature on the surface of the unit sphere $x^2+y^2+z^2=1$.
Ans. 50.
39. If $x=c\cos u\cosh v$ and $y=c\sin u\sinh v$, show that
$$\frac{\partial(x,y)}{\partial(u,v)}=\frac{1}{2}c^2(\cos 2u-\cosh 2v).$$
40. If $u=\frac{x}{y-z}$, $v=\frac{y}{z-x}$, and $w=\frac{z}{z-y}$, show that
$$\frac{\partial(u,v,w)}{\partial(x,y,z)}=0.$$
41. If $x=r\cos\theta$, $y=r\sin\theta$, and $z=z$, evaluate $\frac{\partial(x,y,z)}{\partial(r,\theta,z)}$.
Ans. r.
42. If $x=uv$ and $y=\frac{u+v}{u-v}$, determine $\frac{\partial(u,v)}{\partial(x,y)}$.
Ans. $\frac{(u-v)^2}{4uv}$.
43. The roots of the equation in λ
$$(\lambda-x)^3+(\lambda-y)^3+(\lambda-z)^3=0$$
are u, v, and w. Show that
$$\frac{\partial(u,v,w)}{\partial(x,y,z)}=-2\frac{(y-z)(z-x)(x-y)}{(v-w)(w-u)(u-v)}.$$
Hint: The equation simplifies to
$$\lambda^3-(x+y+z)\lambda^2+(x^2+y^2+z^2)\lambda-\frac{1}{3}(x^3+y^3+z^3)=0.$$
Let $x+y+z=\xi$, $x^2+y^2+z^2=\eta$, and $x^3+y^3+z^3=\xi$. Then $u+v+w=\xi$, $uv+vw+uw=\eta$, and $uvw=\xi$. Find $\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)}$ and $\frac{\partial(\xi,\eta,\zeta)}{\partial(u,v,w)}$. Then $\frac{\partial(u,v,w)}{\partial(x,y,z)}=\frac{\partial(u,v,w)}{\partial(\xi,\eta,\zeta)}\cdot\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)}$.

44. If $U=x+y-z$, $V=x-y+z$, and $W=x^2+y^2+z^2-2yz$, show that U, V and W are connected by a functional relation, and find that functional relation.
Hint: Show that $\frac{\partial(U,V,W)}{\partial(x,y,z)}=0$. Further, $U+V=2x$, $U-V=2(y-z)$. Then $(U+V)^2+(U-V)^2=4(x^2+y^2+z^2-2yz)=4W$.

45. If $u=x^2-2y$, $v=x+y+z$, and $w=x-2y=x-2y+3z$, find $\frac{\partial(u,v,w)}{\partial(x,y,z)}$.
Ans. $10x-2$.

46. Evaluate $\int_0^{\pi}\log(1+a\cos x)dx$ using Leibnitz's Rule.
Ans. $\pi\log\left(\frac{1}{2}+\frac{1}{2}\sqrt{1-a^2}\right)$.

47. Show that $\int_0^{\pi}\frac{\log(1+\sin\alpha\cos x)}{\cos x}dx=\pi\alpha$.

48. Show that $\int_0^{\infty}\frac{e^{-x}}{x}(1-e^{-ax})dx=\log(1+a),(a>-1)$.

49. Differentiating $\int_0^{x}\frac{dx}{x^2+a^2}=\frac{1}{a}\tan^{-1}\frac{x}{a}$ under the integral sign, find the value of $\int_0^{x}\frac{dx}{(x^2+a^2)^2}$.

50. Evaluate $\int_0^{\frac{\pi}{2}}\frac{dx}{(a^2\cos^2 x+b^2\sin^2 x)^2}$.
Hint: Putting $\tan x=t$, we get

$$\int_0^{\frac{\pi}{2}}\frac{dx}{a^2\cos^2 x+b^2\sin^2 x}=\frac{\pi}{2ab}.$$

Use Leibnitz's Rule, first differentiating with respect to a and then with respect to b. We shall get

$$\int_0^{\frac{\pi}{2}}\frac{\cos^2 x\,dx}{(a^2\cos^2 x+b^2\sin^2 x)^2}=\frac{\pi}{4a^3b}\text{ and}$$

$$\int_0^{\frac{\pi}{2}}\frac{\sin^2 x\,dx}{(a^2\cos^2 x+b^2\sin^2 x)^2}=\frac{\pi}{4ab^3}.$$

Adding these two results, we get the value of the given integral as $\frac{\pi(a^2+b^2)}{4a^3b^3}$.

UNIT II

Matrices

4 Matrices

There are many situations in pure and applied mathematics, theory of electrical circuits, aerodynamics, nuclear physics, and astronomy in which we have to deal with algebraic structures and rectangular array of numbers or functions. These arrays will be called matrices. The aim of this chapter is to study algebra of matrices along with its application to solve system of linear equations.

4.1 CONCEPTS OF GROUP, RING, FIELD AND VECTOR SPACE

Definition 4.1 Let S be a non-empty set. Then a mapping $f: S \times S \to S$ is called a *binary operation* in S.

A non-empty set along with one or more binary operations defined on it is called an *algebraic structure*.

Definition 4.2 A non-empty set G together with a binary operation f: $G \times G \to G$ defined on it and denoted by $*$ is called a *group* if the following axioms are satisfied:

(G_1) **Associativity**: For $a, b, c, \in G$,

$$(a * b) * c = a * (b * c)$$

(G_2) **Existence of Identity**: There exists an element e in G such that for all $a \in G$,

$$a * e = e * a = a$$

(G_3) **Existence of Inverse Element:** For each element $a \in G,$ there exists an element $b \in G$, such that

$$a * b = b * a = e.$$

Definition 4.3 Let G be a group. If for every pair a, $b \in G$,

$$a * b = b * a,$$

then G is called a *commutative* (or *abelian*) group.

If $a * b \neq b * a$, then G will be called *non-abelian* or *non-commutative* group.

Definition 4.4 The number of elements in a group G is called the *order of the group* G and is denoted by O (G). A group having a finite number of elements is called a *finite group*.

EXAMPLE 4.1

Let $\mathbb{Z}$ be the set of all integers and let f: $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ defined by $f\ (a,\ b) = a\ *\ b = a + b$ be binary operation in $\mathbb{Z}$. Then

(i) $a + (b + c) = (a + b) + c$ for all $a, b, c \in \mathbb{Z}$

(ii) $a + 0 = 0 + a = a$ for all $a \in \mathbb{Z}$ and so 0 acts as an additive identity.

(iii) $a + (-\ a) = (-\ a) + a = 0$ for $a \in \mathbb{Z}$ and so $(-\ a)$ is the inverse of a.

(iv) $a + b = b + a$, $a, b \in \mathbb{Z}$ (Commutativity).

Hence, $(\mathbb{Z}, +)$ is an infinite additive abelian group.

EXAMPLE 4.2

The set of all integers $\mathbb{Z}$ cannot be a group under multiplication operation $f(a, b) = ab$. In fact, ±1 are the only two elements in $\mathbb{Z}$ which have inverses.

EXAMPLE 4.3

The set of even integers [0, ±2, ±4, …] is an additive abelian group under addition.

EXAMPLE 4.4

The set of vectors V form an additive abelian group under addition.

EXAMPLE 4.5

We shall note in article 13.10 on matrices that the set of all $m \times n$ matrices form an additive abelian group.

EXAMPLE 4.6

The set $\{-1, 1\}$ is a multiplicative abelian group of order 2.

Definition 4.5 Let S be a set with binary operation $f(m, n) = mn$, then an element $a \in S$ is called

(i) **Left cancellative** if

$$ax = ay \Rightarrow x = y \text{ for all } x, y \in S,$$

(ii) **Right cancellative** if

$$xa = ya \Rightarrow x = y \text{ for all } x, y \in S.$$

If any element a is both left- and right cancellative, then it is called *cancellative* (or *regular*). If every element of a set S is regular, then we say that *cancellation law* holds in S.

Theorem 4.1 If G is a group under the binary operation $f(ab) = a * b = ab$ then *for* $a, b, c \in G$,

$$ab = ac \Rightarrow b = c \quad \text{(left cancellation law)}$$
$$ba = ca \Rightarrow b = c \quad \text{(right cancellation law)}$$

(Thus *cancellation law holds in a group*).

Proof: Since G is a group and $a \in G$, there exists an element $c \in G$ such that $ac = ca = e$. Therefore,

$$\begin{aligned} ab = ac &\Rightarrow c(ab) = c(ac) \\ &\Rightarrow (ca)b = c(ac) \\ &\Rightarrow eb = ce \\ &\Rightarrow b = c. \end{aligned}$$

Similarly, we can show that

$$ba = ca \Rightarrow b = c.$$

Theorem 4.2 Let G be a group. Then,

(a) The idetity element of G is unique.
(b) Every $a \in G$ has a unique inverse.
(c) For every $a \in G$, $(a^{-1})^{-1} = a$
(d) For all $a, b \in G$

$$(ab)^{-1} = b^{-1}a^{-1}.$$

Proof: (a) Suppose that there are two identity elements e and e' in G. Then,

$$ee' = e \text{ since } e' \text{ is an identity element,}$$

and

$$ee' = e' \text{ since } e \text{ is an identity element.}$$

Hence $e = e'$.

(b) Suppose that an arbitrary element a in G has two inverses b and c. Then, $ab = ba = e$ and $ac = ca = e$. Therefore,

$$(ba)c = ec = c$$

and

$$b(ac) = be = b.$$

But, by associativity in G,

$$(ba)c = b(ac).$$

Hence $b = c$.

(c) Since G is a group, every element $a \in G$ has its inverse a^{-1}. Then, $a^{-1}a = e$. Now

$$a^{-1}(a^{-1})^{-1} = e = a^{-1}a$$

By left cancellation law, it follows that $(a^{-1})^{-1} = a$.

(d) We have

$$\begin{aligned} (ab)(b^{-1}a^{-1}) &= a(bb^{-1})a^{-1} = aea^{-1} \\ &= aa^{-1} = e. \end{aligned}$$

Similarly

$$\begin{aligned} (b^{-1}a^{-1})(ab) &= b^{-1}(a^{-1}a)b \\ &= b^{-1}eb = b^{-1}b = e. \end{aligned}$$

Thus

$$(ab)(b^{-1}a^{-1}) = (b^{-1}a^{-1})(ab) = e.$$

Hence, by the definition of inverse,

$$(ab)^{-1} = b^{-1}a^{-1}.$$

Definition 4.6 A subset H of a group G is said to a *subgroup* of G, if under the binary operation in G, H itself forms a group.

Every group G has two trivial subgroups, G itself and the identity group $\{e\}$.

The non-trivial subgroups of G are called *proper subgroups* of G.

EXAMPLE 4.7

The additive group $\mathbb{R}$ of real numbers is a subgroup of the additive group $\mathbf{C}$ of complex numbers.

Regarding subgroups, we have

Theorem 4.3 A non-empty subset H of a group G is a subgroup of G if and only if

(i) $a, b \in H \Rightarrow ab \in H$,

(ii) $a \in H \Rightarrow a^{-1} \in H$.

Conditions (i) and (ii) may be combined into a single one and we have "A non-empty subset H of a group G is a subgroup of G if and only if $a, b \in H \Rightarrow ab^{-1} \in H$."

Theorem 4.4 The intersection of two subgroups of a group is again a subgroup of that group.

Definition 4.7 Let G and H be two groups with binary operations ϕ: $G \times G \to G$ and ψ: $H \times H \to H$, respectively, then a mapping f: $G \to H$ is said to be a *group homomorphism* if for all $a, b \in G$,

$$f(\phi(a, b)) = \psi(f(a), f(b)). \qquad (1)$$

Thus if G is additive group and H is multiplicative group, then (1) becomes

$$f(a + b) = f(a).f(b).$$

If, in addition f is bijective, then f is called an *isomorphism*.

EXAMPLE 4.8

Let Z be additive group of integers. Then the mapping f: $Z \to H$, where H is the additive group of even integers defined by $f(a) = 2a$ for all $a \in Z$ is a group homomorphism. In fact, for $a, b \in Z$

$$f(a + b) = 2(a + b) = 2a + 2b = f(a) + f(b).$$

Also

$$f(a) = f(b) \Rightarrow 2a = 2b \Rightarrow a = b,$$

and so f is one-one homomorphism (called *monomorphism*).

Definition 4.8 Let G and H be two groups. If f: $G \to H$ is a homomorphism and e_H denotes the identity element of H, then the subset

$$K = \{x: x \in G, f(x) = e_H\}$$

of G is called the *kernel* of the homomorphism f.

Definition 4.9 A non-empty set R with two binary operation '+' and '.' is called a *ring* if the following conditions are satisfied.

(i) **Associativity of '+'**: if $a, b, c \in R$, then

$$a + (b + c) = (a + b) + c$$

(ii) **Existence of Identity for '+'**: There exists an element 0 in R such that

$$a + 0 = 0 + a = a \quad \text{for all } a \in R$$

(iii) **Existence of inverse with respect to '+'**: To each element $a \in R$, there exists an element $b \in R$ such that

$$a + b = b + a = 0$$

(iv) **Commutativity of '+'**: If $a, b \in R$, then

$$a + b = b + a$$

(v) **Associativity of '.'**: If $a, b, c \in R$, then

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c$$

(vi) **Distributivity of '+' over '.'**: If $a, b, c \in R$, then

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

(Left distributive law)

and

$$(a + b) \cdot c = a \cdot c + b \cdot c$$

(Right distributive law)

Let R be a ring, if there is an element 1 in R such that $a.1 = 1.a = a$ for every $a \in R$, then R is called a *ring with unit element.*

If R is a ring such that $a.b = b.a$ for every $a, b \in R$, then R is called *commutative ring*.

A ring R is said to be a *ring without zero divisors* if $a.b = 0 \Rightarrow a = 0$ or $b = 0$.

EXAMPLE 4.9

We have seen that $(\mathbb{Z}, +)$ is an abelian group. Further, if $a, b, c \in \mathbb{Z}$ then

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c$$
$$a \cdot (b + c) = a \cdot b + a \cdot c$$
$$(a + b) \cdot c = a \cdot c + b \cdot c$$
$$a \cdot 1 = 1 \cdot a = a$$
$$a \cdot b = b \cdot a$$
$$a \cdot b = 0 \Rightarrow a = 0 \text{ or } b = 0.$$

Hence $\mathbb{Z}$ is commutative ring with unity which is *without zero divisor*.

EXAMPLE 4.10

The set of even integers is a commutative ring but there does not exist any element b satisfying $b \cdot a = a \cdot b = a$ for $a \in R$, Hence, it is a ring without unity.

EXAMPLE 4.11

We shall see later on that the set of $n \times n$ matrices form a non-commutation ring with unity. This ring is a *ring with zero divisors*. For example, if $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, then their product is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. But none of the given matrix is zero.

Definition 4.10 A commutative ring with unity is called an *integral domain* if it has no zero divisor.

For example, *ring of integers is an integral domain*.

Definition 4.11 A ring R with unity is said to be a *division ring* (or *skew field*) if every non-zero element of R has a multiplicative inverse.

Definition 4.12 A commutative division ring is called a *field*.

For example, the set of rational number $\mathbb{Q}$ under addition and multiplication forms a field. Similarly, $\mathbb{R}$ and $\mathbb{C}$ are also fields. Every field is an integral domain but the converse is not true. For example, the set of integers form an integral domain but is not a field. An important result is that

"*Every finite integral domain is a field.*"

Definition 4.13 A subset S of ring R is called a *subring* of R if S is a ring under the binary operations in R. Thus, S will be a *subring* of R if

(i) $a, b \in S \Rightarrow a - b \in S$,

(ii) $a, b \in S \Rightarrow ab \in S$.

For example, the set of real numbers is a subring of the ring of complex numbers.

Definition 4.14 A mapping $f : R \to R'$ from the ring R into the ring R' is said to be a *ring homomorphism* if

(i) $f(a+b) = f(a) + f(b)$,

(ii) $f(ab) = f(a).\, f(b)$,

for all $a, b \in R$.

If, in addition, f is one-to-one and onto then f is called *ring isomorphism*.

Definition 4.15 A non-empty set V is said to be a *Vector Space* over the field F if

(i) V is an additive abelian group.

(ii) If for every $\alpha \in F$, $v \in V$, there is defined an element αv, called scalar multiple of α and v, in V subject to

$$\alpha(v + \omega) = \alpha v + \alpha\omega,$$
$$(\alpha + \beta)v = \alpha v + \beta v,$$
$$\alpha(\beta v) = (\alpha\beta)\, v,$$
$$1v = v,$$

for all $\alpha, \beta \in F$, $v, w \in$ V, where 1 represents the unit elements of F under multiplication.

In the above definition, the elements of V are called *vectors* whereas the elements of F are called *scalars*.

EXAMPLE 4.12

Let $V_2 = \{(x, y): x, y \in \mathbb{R}\}$ be a set of ordered pairs. Define addition and scalar multiplication on V_2 by

$$(x, y) + (x', y') = (x + x', y + y'),$$

and

$$\alpha(x, y) = (\alpha x, \alpha y).$$

Then V_2 is an abelian group under addition operation defined earlier and

$$\begin{aligned} \alpha[(x,\ y) + (x',\ y')] &= \alpha(x + x',\ y + y') \\ &= (\alpha x + \alpha x',\ \alpha y + \alpha y') \\ &= (\alpha x + \alpha y) + (\alpha x' + \alpha y') \\ &= \alpha(x,\ y) + \alpha(x',\ y'), \\ (\alpha + \beta)(x,\ y) &= ((\alpha + \beta)x, (\alpha + \beta)y) \\ &= (\alpha x + \beta x, \alpha y + \beta y) \\ &= (\alpha x, \alpha y) + (\beta x, \beta y) \\ &= \alpha(x, y) + \beta(x, y), \\ \alpha(\beta(x, y)) &= (\alpha\beta)\ (x, y), \\ 1.(x, y) &= (x, y). \end{aligned}$$

Hence, V_2 is a vector space over $\mathbb{R}$. It is generally denoted by $\mathbb{R}^2$.

Similarly, the set of n-tuples $(x_1, x_2,\ldots, x_n)$ form a vector space over $\mathbb{R}$ and is denoted by V_n or $\mathbb{R}^n$.

Definition 4.16 Let V be a vector space over the field K and W be a subset of V. If W is a vector space under the operations of V, then it is called a *vector subspace* of V.

Thus, W will be a subspace of V if

(i) W is a subgroup of V,

(ii) $\lambda \in$ F, $\omega \in W$ imply $\lambda\, \omega \in W$.

The conditions (i) and (ii) can be combined into a single condition, namely, $\lambda_1, \lambda_2 \in F$ and $w_1, w_2 \in W$ imply $\lambda_1\, w_1 + \lambda_2\, w_2 \in W$.

Definition 4.17 Let V be a vector space over a field F and let $v_1, v_2, \ldots, v_n \in V$. Then any element of the form $\alpha_1 v_1 + \alpha_2\, v_2 + \ldots + \alpha_n\, v_n$, $\alpha_i \in F$ is called a *linear combination* over F of $v_1, v_2, \ldots v_n$.

Definition 4.18 Let S be a non-empty subset of the vector space V. Then the *linear span* of S, denoted by L(S), is the set of all linear combinations of finite sets of the elements of S.

Definition 4.19 Let V be a vector space over a field F. Then $v_1, v_2, \ldots, v_n \in V$ are said to be *linearly independent* over F if for scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in F$, $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$ implies $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$.

Definition 4.20 Let V be a vector space. Then $v_1, v_2, \ldots, v_n \in V$ are called *linearly dependent* if there exist $\lambda_1, \lambda_2, \ldots, \lambda_n \in F$, not all of them zero, such that $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$.

Thus, $v_1, v_2, \ldots, v_n$ are linearly dependent if they are not linearly independent.

Definition 4.21 An infinite subset S of a vector space V over a field F is said to be *linearly independent* if every finite subset of S is linearly independent.

Theorem 4.5 $L(S)$ is a subspace of V.

Proof: Let $v, w \in L(S)$. Then

$$v = \lambda_1 s_1 + \lambda_2 s_2 + \ldots + \lambda_n s_n,$$
$$w = \mu_1 t_1 + \mu_2 t_2 + \ldots + \mu_m t_m,$$

where $\lambda's$ and $\mu's$ are scalars and s_i and t_i are in S.

Therefore, for $\alpha, \beta \in$ F, we have

$$\begin{aligned}\alpha v + \beta w &= \alpha(\lambda_1 s_1 + \lambda_2 s_2 + \ldots + \lambda_n s_n) \\ &\quad + \beta(\mu_1 t_1 + \ldots + \mu_m t_m) \\ &= (\alpha\lambda_1)s_1 + (\alpha\lambda_2)s_2 + \ldots + (\alpha\lambda_n)s_n \\ &\quad + (\beta\mu_1)t_1 + \ldots + (\beta\mu_m)t_m \in \mathrm{L(S)}.\end{aligned}$$

Hence $L(S)$ is subspace of V.

Further, if S and T are subsets of a vector space V, then

(i) $S \subset T \Rightarrow L(S) \subset L(T)$

(ii) $L(S \cup T) = L(S) \cup L(T)$

(iii) $L(L(S)) = L(S)$

EXAMPLE 4.13

Let $v_1 = (1, 0)$ and $v_2 = (1, 0)$ be vectors in the vector space $R^2 = \{(x, y) : x, y \in R\}$. If $\lambda_1, \lambda_2 \in R$, then

$$\begin{aligned}\lambda_1 v_1 + \lambda_2 v_2 = 0 &\Rightarrow \lambda_1(1,0) + \lambda_2(0,1) = 0 \\ &\Rightarrow (\lambda_1, 0) + (0, \lambda_2) = 0 \\ &\Rightarrow (\lambda_1, \lambda_2) = 0 \\ &\Rightarrow \lambda_1 = \lambda_2 = 0.\end{aligned}$$

Hence, v_1, v_2 are linearly independent.

EXAMPLE 4.14

Let $v_1 = (1, 0, 1)$, v_2 $(0, 1, 0)$ and $v_3 = (1, 1, 1)$. Then we note that

$$\begin{aligned}v_1 + v_2 - v_3 &= (1,0,1) + (0,1,0) - (1,1,1) \\ &= (1+0-1, 0+1-1, 1+0-1) \\ &= (0,0,0).\end{aligned}$$

Hence, v_1, v_2 and v_3 are linearly dependent.

Theorem 4.6 Let V be a vector space over a field F. If $v_1, v_2, v_3, \ldots, v_n$ are linearly independent elements of V, then every element in their span has a unique representation in the form $\lambda_1 v_1 + \lambda_2 v_2 + \ldots \lambda_n v_n$ with $\lambda_i \in F$.

Proof: Every element in the linear span is of the form $\lambda_1\, v_1 + \lambda_2\, v_2 + \ldots\, \lambda_n\, v_n$. Suppose that there are following two representations for an element:

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = \mu_1 v_1 + \mu_2 v_2 + \ldots + \mu_n v_n$$

and so

$$(\lambda_1 - \mu_1)v_1 + (\lambda_2 - \mu_2)v_2 + \ldots + (\lambda_n - \mu_n)v_n = 0.$$

Since v_1, v_2, $v_3 \ldots v_n$ are linearly independent, we have

$$\lambda_1 - \mu_1 = 0, \lambda_2 - \mu_2 = 0, \ldots, \lambda_n - \mu_n = 0,$$

which yield

$$\lambda_1 = \mu_1, \quad \lambda_2 = \mu_2, \ldots, \lambda_n = \mu_n.$$

Hence, representation of every element in the span is unique.

Theorem 4.7 If $v_1, v_2, v_3, \ldots, v_n$ are in V, then either they are linearly independent or some v_k is linear combination of the preceding ones $v_1, v_2, v_3, \ldots, v_{k-1}$.

Proof: If $v_1, v_2, v_3, \ldots, v_n$ are linearly independent, we are done. So suppose that $v_1, v_2, v_3, \ldots, v_n$ are linearly dependent. Thus, $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$, where not all of $\alpha_1, \alpha_2, \ldots, \alpha_n$, are zero. Let k be the largest integer for which $\alpha_k \neq 0$. Since $\alpha_i = 0$ for $i > k$, $\alpha_1 v_1 + \ldots + \alpha_k v_k = 0$. Since $\alpha_k \neq 0$, we have

$$\begin{aligned} v_k &= \alpha_k^{-1}(-\alpha_1 v_1 - \alpha_2 v_2 - \ldots - \alpha_{k-1} v_{k-1}) \\ &= \left(-\alpha_k^{-1}\alpha_1\right) v_1 + \ldots + \left(-\alpha_k^{-1}\alpha_{k-1}\right) v_{k-1}. \end{aligned}$$

Hence, v_k is a linear combination of its predecessors.

Theorem 4.8 A system of vectors in a vector space is linearly dependent if and only if any one of the vectors in that system can be represented as a linear combination of the other vectors in the system.

Proof: Suppose that V is a vector space over the field F and let $v_1, v_2, v_3, \ldots, v_n \in$ V be linearly dependent. Then, by definition,

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0,$$

where not all of the λ_i are zero. Suppose that $\lambda_1 \neq 0$, then

$$v_1 = -\frac{\lambda_2}{\lambda_1} v_2 - \frac{\lambda_3}{\lambda_1} v_3 - \ldots - \frac{\lambda_n}{\lambda_1} v_n.$$

Hence v_1 is linear combination of other vector in the system.

Conversely, suppose that v_1 is linear combination of $v_2, v_3, \ldots, v_n$, that is

$$v_1 = \lambda_2 v_2 + \lambda_3 v_3 + \ldots + \lambda_n v_n, \quad \lambda_i \in F$$

and so

$$(-1)v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \ldots + \lambda_n v_n = 0.$$

Since the first coefficient is non-zero, it follows that $v_1, v_2, v_3, \ldots, v_n$is a linearly dependent system.

Theorem 4.9 If a subsystem of a finite system of vectors in a vector space is linearly dependent, then the whole system is linearly dependent.

Proof: Let $v_1, v_2, v_3, \ldots, v_n \in V$ be a finite system of vectors in V. Suppose that $v_1, v_2, v_3, \ldots, v_k$, $k < n$ is linearly dependent. Therefore,

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k + 0 v_{k+1} + 0 v_{k+2} + \ldots + 0 v_n = 0,$$

where not all of $\lambda_1, \lambda_2, \ldots, \lambda_k$ are zero. Hence, $v_1, v_2, v_3, \ldots, v_n$ is linearly dependent.
It follows from Theorem 4.9 that any superset of a linearly dependent set is also linearly dependent.

EXAMPLE 4.15

Show that the set

$$\{(1,1,0), (0,1,1), (1,0,-1), (1,1,1)\}$$

is linearly dependent.

Solution. We note that

$$1(1,1,0) - 1(0,1,1) - 1(1,0,-1) = (0,0,0)$$

Hence the set

$$\{(1,1,0), (0,1,1), (1,0,-1)\}$$

is linearly dependent. Being superset of this linearly dependent set, the given set is also linearly dependent.

Definition 4.22 Let S be subset of a vector space V. If every element of V can be written as the linear combination of the elements of S, then S is called *generator* of V.

For example, let $V_2 = \{(x, y): x, y \in \mathbb{R}\}$ be the vector space and let

$$v_1 = (1,0), v_2 = (0,1)$$

be vectors in V_2. If $(x, y) \in V_2$ be arbitrary, then

$$\begin{aligned} (x,y) &= x\ (1,0) + y\ (0,1) \\ &= xv_1 + yv_2 \end{aligned}$$

Hence $S = \{v_1, v_2\}$ generates V_2.

Definition 4.23 Let S be subset of a vector space V. If

(i) S generates V, that is, $L(S) = V$, and

(ii) the elements of S are linearly independent, then S is called *basis* of the vector space V.

For example {(1,0), (0,1)} is a basis of the vector space

$$V_2 = \{(x, y) : x, y \in \mathbb{R}\}.$$

Definition 4.24 The number of elements in the basis of a vector space is called the *dimension* of that vector space.

For example, the dimension of V_2 is 2.

If the number of elements in the basis of a vector space is finite, then the vector space is called *finite dimensional vector space*.

EXAMPLE 4.16

Let F be a field and

$$F^{(n)} = \{(x_1, x_2, \ldots, x_n) : \ x_i \in F\}$$

be a set of *n-tuples*. If we define addition and scalar multiplication in $F^{(n)}$ by

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)$$
$$= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

and

$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

Then $F^{(n)}$ becomes a vector space (linear space).

Let

$$e_1 = (1, 0, \ldots, 0),$$
$$e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1).$$

Then,

$$\begin{aligned}(x_1, x_2, \ldots, x_n) &= x_1(1, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) \\ &\quad + \ldots + x_n(0, 0, \ldots, 1) \\ &= x_1 e_1 + x_2 e_2 + \ldots + x_n e_n\end{aligned}$$

and so $\{e_1, e_2, \ldots, e_n\}$ generates $F^{(n)}$. Further

$$\begin{aligned}&\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n = 0 \\ &\quad \Rightarrow \lambda(1, 0, \ldots, 0) + \lambda_2(0, 1, \ldots, 0) \\ &\quad\quad + \ldots + \lambda_n(0, 0, \ldots, 1) = 0 \\ &\quad \Rightarrow (\lambda_1, \lambda_2, \ldots, \lambda_n) = 0 \\ &\quad \Rightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0\end{aligned}$$

and so $\{e_1, e_2, \ldots, e_n\}$ is linearly independent. Hence, $\{e_1, e_2, \ldots, e_n\}$ is the basis of $F^{(n)}$ and so $F^{(n)}$ is n-dimensional.

4.2 MATRICES

Definition 4.25 A rectangular array of mn real or complex numbers, arranged in m rows and n columns, is called an $m \times n$ matrix.

An $m \times n$ matrix A is represented by the symbol

$$A = \begin{bmatrix} a_{11} & a_{12} & .. & .. & a_{1n} \\ a_{21} & a_{22} & .. & .. & a_{2n} \\ .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. \\ a_{m1} & a_{m2} & .. & .. & a_{mn} \end{bmatrix} \text{ or }$$

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

where a_{ij} denotes the element in the *ith* row and *jth* column of the matrix. Each of the mn number constituting a $m \times n$ matrix is called *entry* (or *element*) of the matrix A. We generally abbreviate the symbol of the matrix A by $A = [a_{ij}]_{m \times n}$ or simply by $[a_{ij}]$. Further, if a matrix $A = [a_{ij}]$ has m rows and n columns, then it is said to be of order $m \times n$.

Definition 4.26 Two matrices $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{m \times n}$ over a field F($\mathbb{R}$ or $\mathbb{C}$) are said to be *equal* if

(i) they are both of the same type, that is, have the same number of rows and columns.

(ii) the elements in the corresponding places of the two matrices are equal, that is $a_{ij} = b_{ij}$ for all pairs of *i, j*.

We observe that the relation of equality of two matrices is an equivalence relation. In fact,

(i) If A is any matrix, A = A (Reflexivity)

(ii) If A = B, then B = A (Symmetry)

(iii) If A = B and B = C, then A = C (Transitivity).

Definition 4.27 The elements a_{ii} of a matrix $A = [a_{ij}]$ are called the *diagonal elements* of A.

Definition 4.28 A matrix in which the number of rows is equal to the number of columns is called a *square matrix*.

If A is a square matrix having n rows and n columns then it is also called a *matrix of order n*.

Definition 4.29 If the matrix A is of order n, the elements $a_{11}, a_{22}, \ldots, a_{nn}$ are said to constitute the *main diagonal* of A and the elements $a_{n1}, a_{n-12}, \ldots, a_{1n}$ constitute its *secondary diagonal*.

Definition 4.30 A square matrix $A = [a_{ij}]$ is said to be a *diagonal matrix* if each of its non-diagonal element is zero, that is, if $a_{ij} = 0$ whenever $i \neq j$.

A diagonal matrix whose diagonal elements, in order, are $d_1, d_2, \ldots, d_n$ is denoted by Diag $[d_1, d_2, \ldots, d_n]$ or Diag $[a_{11}, a_{22}, \ldots, a_{nn}]$ if $A = [a_{ij}]$.

Definition 4.31 A diagonal matrix, whose diagonal elements are all equal is called a *scalar matrix*. For example, the matrix

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

is a scalar matrix of order 3.

Definition 4.32 A scalar matrix of order n, each of whose diagonal element is equal to 1 is called a *unit matrix* or *identity matrix* of order n and is denoted by I_n.

For example, the matrix

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

is a unit matrix of order 4.

Definition 4.33 A matrix, rectangular or square, each of whose entry is zero is called a *zero matrix* or a *null matrix* and is denoted by **0**.

Definition 4.34 A matrix having 1 row and n column is called a *row matrix* (or a *row vector*). For example, the matrix

$$\begin{bmatrix} 2 & 3 & 5 & 6 & 2 \end{bmatrix}$$

is a row matrix.

Definition 4.35 If a matrix has m rows and 1 column, it is called a *column matrix* (or a *column vector*).

For example, the matrix

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

is a column matrix.

Definition 4.36 A *submatrix* of a given matrix A is defined to be either A or any array obtained on deleting some rows or columns or both of the matrix A.

Definition 4.37 A square submatrix of a square matrix is called a *principal submatrix* if its diagonal elements are also the diagonal elements of the matrix A.

Thus to obtain principal submatrix, it is necessary to delete corresponding rows and columns. For example, the matrix

$$\begin{bmatrix} 3 & 1 \\ 4 & 3 \end{bmatrix}$$

is a principal submatrix of the matrix

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 0 \\ 6 & 4 & 3 & 2 \\ 1 & 2 & 4 & 1 \end{bmatrix}.$$

Definition 4.38 A principal square submatrix is called *leading submatrix* if it is obtained by deleting only some of the last rows and the corresponding columns. For example,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

is the *leading principal submatrix* of the matrix

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 0 \\ 6 & 4 & 3 & 2 \\ 1 & 2 & 4 & 1 \end{bmatrix}.$$

4.3 ALGEBRA OF MATRICES

Matrices allow the following basic operations:

(a) Multiplication of a matrix by a scalar.

(b) Addition and subtraction of two matrices.

(c) Product of two matrices.

However, the concept of dividing a matrix by another matrix is undefined.

Definition 4.39 Let α be a scalar (real or complex) and A be a given matrix. Then the *multiplication* of $A=[a_{ij}]$ by the scalar α is defined by

$$\alpha A = \alpha[a_{ij}] = [\alpha a_{ij}],$$

that is, each element of A is multiplied by the scalar α. The order of the matrix so obtained will be the same as that of the given matrix A.

For example

$$4\begin{bmatrix} 3 & 1 & 2 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 12 & 4 & 8 \\ 8 & 4 & 0 \end{bmatrix}.$$

Definition 4.40 Two matrices $A=[a_{ij}]$ and $B=[b_{ij}]$ are said to be *comparable* (*conformable*) for addition/subtraction if they are of the same order.

Definition 4.41 Let A and B be two matrices of the same order, say m x n. Then, the *sum of the matrices* A and B is defined by

$$C=[c_{ij}] = A + B = [a_{ij}] + [b_{ij}] = [a_{ij}+b_{ij}].$$

Thus,

$$c_{ij} = a_{ij} + b_{ij}, 1 \le i \le m, 1 \le j \le n,$$

The order of the new matrix C is same as that of A and B. Similarly,

$$C = A - B = [a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}]$$

Thus,

$$c_{ij} = a_{ij} - b_{ij} \quad for \quad 1 \le i \le m, \quad 1 \le j \le n$$

Definition 4.42 If $A_1, A_2, \ldots, A_n$ are n matrices which are conformable for addition and $\lambda_1, \lambda_2, \ldots \lambda_n$ are scalars, then $\lambda_1 A_1 + \lambda_2 A_2 + \ldots + \lambda_n A_n$ is called a *linear combination* of the matrices $A_1, A_2, \ldots, A_n$.

Let $A=[a_{ij}], B=[b_{ij}], C=[c_{ij}]$,be $m \times n$ matrices with entries from the complex numbers. Then the following properties hold:

(a) $A+B=B+A$ (Commutative law for addition)

(b) $(A+B)+C=A+(B+C)$ (Associative law for addition)

(c) $A+\mathbf{0}=\mathbf{0}+A=A$ (Existence of additive identity)

(d) $A+(-A)=(-A)+A=\mathbf{0}$ (Existence of inverse)

Thus the *set of matrices form an additive commutative group*.

4.4 MULTIPLICATION OF MATRICES

Definition 4.43 Two matrices $A=[a_{ij}]_{m \times n}$ and $B=[b_{ij}]_{p \times q}$ are said to *comparable* or *conformable* for the product AB if $n=p$, that is, if the number of columns in A is equal to the number of rows in B.

Definition 4.44 Let $A=[a_{ij}]_{m \times n}$ and $B=[b_{ij}]_{p \times q}$ be two matrices. Then, the product AB is the matrix $C=[c_{ij}]_{m \times q}$ such that

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{1n}b_{nj}$$
$$= \sum_{k=1}^{n} a_{ik}\, b_{kj} \text{ for } 1 \le i \le m, 1 \le j \le n.$$

Note that the c_{ij} [the (i, j)th element of AB] has been obtained by multiplying the ith row of A, namely $(a_{i1}, a_{i2}, \ldots, a_{in})$ with the jth column of B, namely

$$\begin{bmatrix} b_{1j} \\ b_{2j} \\ \cdots \\ \cdots \\ b_{nj} \end{bmatrix}$$

Remark 4.1 In the product AB, the matrix A is called *prefactor* and B is called *postfactor*.

EXAMPLE 4.17

Construct an example to show that product of two non-zero matrices may be a zero matrix.

Solution. Let

$$A = \begin{bmatrix} x & 0 \\ y & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix}.$$

Then A and B are both 2×2 matrices. Hence, they are conformable for product. Now,

$$AB = \begin{bmatrix} x & 0 \\ y & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix}$$
$$= \begin{bmatrix} 0+0 & 0+0 \\ 0+0 & 0+0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Definition 4.45 When a product $AB=\mathbf{0}$ such that neither A nor B is $\mathbf{0}$ then the factors A and B are called *divisors of zero*.

The above example shows that in the algebra of matrices, there exist divisors of zero, whereas in the algebra of complex numbers, there is no zero divisor.

EXAMPLE 4.18

Taking

$$A = \begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \; B = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix},$$

show that matrix multiplication is not, in general, commutative.

Solution. Both A and B are 3×3 matrices. Therefore, both AB and BA are defined. We have,

$$AB = \begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 9 & 13 \\ -1 & 2 & 4 \\ -2 & 2 & 4 \end{bmatrix}$$

and

$$BA = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 12 & 11 \\ -1 & 7 & 8 \\ -2 & -1 & 5 \end{bmatrix}.$$

Hence $AB \neq BA$.

EXAMPLE 4.19

Give an example to show that cancellation law does not hold, in general, in matrix multiplication.

Solution. Let

$$A = \begin{bmatrix} 0 & 4 \\ 0 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 4 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

Then A and B are conformable for multiplication. Similarly, A and C are also conformable for multiplication. Thus,

$$AB = \begin{bmatrix} 0 & 4 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$AC = \begin{bmatrix} 0 & 4 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence $AB = AC$, $A \neq \mathbf{0}$ without having $B = C$ and so we cannot ordinarily cancel A from $AB = AC$ even if $A \neq \mathbf{0}$.

Remark 4.2 The above examples show that in matrix algebra

(a) The commutative law $AB = BA$ does not hold true.

(b) There exist divisors of zero, that is, there exists matrices A and B such that $AB = \mathbf{0}$ but neither A nor B is zero.

(c) The cancellation law does not hold in general, that is, $AB = AC$, $A \neq 0$ does not imply in general that $B = C$.

4.5 ASSOCIATIVE LAW FOR MATRIX MULTIPLICATION

If $A = [a_{ij}]_{m\times n}$ $B = [b_{jk}]_{n\times p}$, and $C = [c_{kl}]_{n\times p}$ are three matrices with entries from the set of complex numbers, then

$$(AB)C = A(BC).$$

(Associative Law for Matrix Multiplication).

4.6 DISTRIBUTIVE LAW FOR MATRIX MULTIPLICATION

If $A = [a_{ij}]_{m\times n}, B = [b_{jk}]_{n\times p}$ and $C = [c_{kl}]_{p\times q}$ are three matrices with elements from the set of complex numbers, then

$$A(B + C) = AB + AC$$

(Distributive Law for Matrix Multiplication).

Definition 4.46 The matrices A and B are said to be *anticommutative* or *anticommute* if $AB = -BA$.

For example, each of the Pauli Spin matrices (used in the study of electron spin in quantum mechanics)

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

where $i^2 = -1$ anticommute with the others. In fact,

$$\sigma_x\sigma_y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix},$$

$$\sigma_y\sigma_x = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix},$$

and so $\sigma_x\sigma_y = -\sigma_y\sigma_x$.

Definition 4.47 If A and B are matrices of order n, then the matrix $AB - BA$ is called the *commutator* of A and B.

Definition 4.48 The sum of the main diagonal elements $a_{ii},\ i = 1, 2, \ldots, n$ of a square matrix A is called the *trace* or *spur* of A.
Thus,

$$tr\ A = a_{11} + a_{22} + \ldots + a_{nn}.$$

Theorem 4.10 Let A and B be square matrices of order n and λ be a scalar. Then

(a) $tr\ (\lambda A) = \lambda\ tr\ A$,

(b) $tr\ (A + B) = tr\ A + tr\ B$,

(c) $tr\ (AB) = tr\ (BA)$.

Proof: Let

$$A = [a_{ij}]_{n\times n} \text{ and } B = [\,b_{ij}]_{n\times n}.$$

(a) We have

$$\lambda A = [\lambda\ a_{ij}]_{n\times n}$$

and so

$$tr\ (\lambda A) = \sum_{I=1}^{n} \lambda a_{ii} = \lambda \sum_{i=1}^{n} a_{ii} = \lambda\ tr\ A.$$

(b) We have

$$A + B = [a_{ij} + b_{ij}]_{n\times n}$$

and so

$$tr(\mathrm{A} + \mathrm{B}) = \sum_{i=1}^{n} [a_{ii} + b_{ii}] = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = tr\ A + tr\ B.$$

(c) We have

$$AB = [c_{ij}]_{n\times n}$$

where

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

and

$$BA = [d_{ij}]_{n\times n},$$

where

$$d_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj}.$$

Then

$$tr\ (AB) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik}\ b_{ki}\right) = \sum_{k=1}^{n}\left(\sum_{i=1}^{n} a_{ik}\ b_{ki}\right) = \sum_{k=1}^{n}\left(\sum_{i=1}^{n} b_{ki}\ a_{ik}\right) = \sum_{k=1}^{n} d_{kk} = tr\ (BA).$$

EXAMPLE 4.20

If A and B are matrices of the same order say n, show that the relation $AB - BA = I_n$ does not hold good.

Solution. Suppose on the contrary that the relation $AB - BA = I_n$ holds true. Since A and B are of same order, AB and BA are also of order n. Therefore,

$$tr(AB - BA) = tr\ I_n$$
$$\Rightarrow tr\ AB - tr\ BA = tr\ I_n.$$

Since $tr\ AB = tr\ BA$, we have

$$0 = tr\ I_n = 1 + 1 + 1 \ldots + 1 = n,$$

which is absurd. Hence, the given relation does not hold good.

Definition 4.49 An $n \times n$ matrix A is said to be *nilpotent* if $A^n = 0$ for some positive integer n.

The smallest positive integer n, for which $A^n = 0$, is called the *degree of nilpotence* of A.

For example, the matrix

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

is nilpotent and the degree of nilpotence is 4. Similarly, the matrix

$$A = \begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix}$$

is nilpotent with degree of nilpotence 2. In fact,

$$A^2 = \begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix}\begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix}$$

$$= \begin{bmatrix} 36-36 & 54-54 \\ -24+24 & -36+36 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

It can be shown that every 2×2 nilpotent matrix A such that $A^2 = 0$ may be written in the form

$$\begin{bmatrix} \lambda\mu & \mu^2 \\ -\lambda^2 & -\lambda\mu \end{bmatrix},$$

where λ, μ are scalars. If A is real then λ, μ are also real.

Definition 4.50 A square matrix A is said to be *involutory* if $A^2 = I$

For example, the matrix

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

is involutory.

Theorem 4.11 A matrix A is involutory if and only if $(I+A)(I-A)=0$.

Proof: Suppose first that A is involutory, then

$$A^2 = I$$

or

$$I - A^2 = 0$$

or

$$I^2 - A^2 = 0 \text{ since } I^2 = I$$

or

$$(I+A)(I-A) = 0 \text{ since } AI = IA.$$

Conversely, let

$$(I+A)(I-A) = 0$$

Then,

$$I^2 - IA + AI - A^2 = 0$$

or

$$I^2 - A^2 + 0 = 0$$

or

$$I^2 - A^2 = 0$$

or

$$A^2 = I^2 = I.$$

Definition 4.51 A square matrix A is said to be *idempotent* if $A^2 = A$.

For example, I_n is idempotent.

4.7 TRANSPOSE OF A MATRIX

Definition 4.52 A matrix obtained by interchanging the corresponding rows and columns of a matrix A is called the *transpose matrix* of A.

The transpose of a matrix A is denoted by A^T (or by A'). Thus, if $A = [a_{ij}]_{m\times n}$, then $A^T = [a_{ji}]_{n\times m}$ is an $n \times m$ matrix. For example, the transpose of the matrix

$$\begin{bmatrix} 1 & 0 & 2 \\ 3 & 7 & 4 \\ 1 & 2 & 8 \end{bmatrix}$$

is

$$\begin{bmatrix} 1 & 3 & 1 \\ 0 & 7 & 2 \\ 2 & 4 & 8 \end{bmatrix}.$$

Further,

(i) The transpose of a row matrix is a column matrix. For example, if $A = [1\ 2\ 4\ 3\ ]$, then

$$A^T = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 3 \end{bmatrix}.$$

(ii) The transpose of a column matrix is a row matrix. For example, if

$$A = \begin{bmatrix} 3 \\ 8 \\ 3 \end{bmatrix},$$

then

$$A^T = [\ 3 \quad 8 \quad 3\ ]$$

(iii) If A is $m \times n$ matrix, then A^T is an $n \times m$ matrix. Therefore, the product AA^T, $A^T A$ are both defined and are of order $m \times m$ and $n \times n$, respectively.

If $A = [a_{ij}]_{m\times n}$ and $B = [b_{ij}]_{m\times n}$ are matrices of the same order and if λ is a scalar, then the transpose of matrix has the following properties:

(a) $(A^T)^T = A$

(b) $(\lambda A)^T = \lambda A^T$

(c) $(A+B)^T = A^T + B^T$

(d) $(AB)^T = B^T A^T$ (Reversal law).

4.8 SYMMETRIC, SKEW-SYMMETRIC, AND HERMITIAN MATRICES

Definition 4.53 A square matrix A is said to be *symmetric* if $A = A^T$.

Thus, $A=[a_{ij}]_{n\times n}$ is symmetric if $a_{ij}=a_{ji}$ for $i\le i\le n$, $1\le j\le n$.

Definition 4.54 A square matrix $A=[a_{ij}]_{n\times n}$ is said to be *skew symmetric* if $a_{ij}=-a_{ji}$ for all i and j.

Thus square matrix is skew-symmetrical if $A=-A^T$. For example,

$$\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$$

is symmetric matrix whereas the matrix

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

is a skew-symmetric matrix.

Properties of Symmetric and Skew-Symmetric Matrices

(a) In a skew-symmetric matrix A, all diagonal elements are zero. In fact, if A is skew-symmetric, then

$$\begin{aligned} a_{ij} &= -a_{ji} \text{ for all } i \text{ and } j. \\ \Rightarrow a_{ii} &= -a_{ii} \\ \Rightarrow a_{ii} &= 0. \end{aligned}$$

(b) The matrix which is both symmetric and skew-symmetric must be a null matrix. In fact, if $A=[a_{ij}]$ is symmetric, then

$$a_{ij}=a_{ji} \text{ for all } i \text{ and } j.$$

Further, if $A=[a_{ij}]$ is skew-symmetric, then $a_{ij}=-a_{ji}$ for all i and j. Adding, we get $2a_{ij}=0$ for all i and j and so $a_{ij}=0$ for all i and j. Hence, A is a *null matrix*. Thus, “Null matrix is the only matrix which is both symmetric and skew-symmetric.”

(c) For any square matrix A, $A+A^T$ is a symmetric matrix and $A-A^T$ is a skew-symmetric matrix. In fact, we note that

(i) $\left(A+A^T\right)^T=A^T+\left(A^T\right)^T=A^T+A=A+A^T$ and so $A+A^T$ is symmetric.

(ii) $\left(A-A^T\right)^T=A^T-\left(A^T\right)^T=A^T-A$
$= -(A-A^T)$

and so $A-A^T$ is skew-symmetric.

(d) Every square matrix A can be expressed uniquely as the sum of a symmetric and a skew-symmetric matrix. To show it, set

$$P=\frac{1}{2}\left(A+A^T\right) \text{ and } Q=\frac{1}{2}\left(A-A^T\right).$$

Then

$$\begin{aligned} P^T &= \left[\frac{1}{2}\left(A+A^T\right)\right]^T=\frac{1}{2}\left(A+A^T\right)^T \\ &=\frac{1}{2}\left(A^T+\left(A^T\right)^T\right)=\frac{1}{2}\left(A^T+A\right) \\ &=\frac{1}{2}\left(A+A^T\right)=P. \end{aligned}$$

and so P is symmetric. Further,

$$\begin{aligned} Q^T &= \left[\frac{1}{2}\left(A-A^T\right)\right]^T = \frac{1}{2}\left(A-A^T\right)^T \\ &=\frac{1}{2}\left(A^T-\left(A^T\right)^T\right)=\frac{1}{2}\left(A^T-A\right) \\ &=-\frac{1}{2}\left(A-A^T\right) = -Q, \end{aligned}$$

and so Q is skew-symmetric. Also $P+Q=A$. Thus A can be expressed as the sum of a symmetric and a skew-symmetric matrix.

To establish the uniqueness of the expression, let $A=P_1+Q_1$, where P_1 is symmetric and Q_1 is skew-symmetric. It is sufficient to show that $P_1=P$ and $Q_1=Q$. We have,

$$A^T=(P_1+Q_1)^T=P_1^T+Q_1^T=P_1-Q_1$$

Thus

$$A+A^T=2P_1 \text{ or } P_1=\frac{1}{2}\left(A+A^T\right) = P.$$

Also,

$$\begin{aligned} Q_1 &= A-P_1=A-\frac{1}{2}\left(A+A^T\right) \\ &=\frac{1}{2}\left(A-A^T\right) = Q. \end{aligned}$$

Hence, the expression is unique.

(e) If A is a square matrix, then $A+A^T$ and AA^T are symmetric matrices. These facts follow from

(i) $(A+A^T)^T= A^T+(A^T)^T=A^T+A$
$=A+A^T$

and

(ii) $(AA^T)^T=(A^T)^TA^T=AA^T$.

(f) If A and B are two symmetric matrices, then AB-BA is a skew-symmetric matrix.

In fact,

$$(AB - BA)^T = (AB)^T - (BA)^T$$
$$= B^T A^T - A^T B^T \text{(Reversal Law)}$$
$$= BA - AB$$
$$= -(AB - BA).$$

(g) If A is a symmetric (skew-symmetric), then $B^T\ AB$is a symmetric (skew-symmetric) matrix. In fact, if A is symmetric, $A^T = A$ and so

$$\left(B^T AB\right)^T = B^T A^T \left(B^T\right)^T = B^T A^T B$$
$$= B^T AB$$

and if A is skew-symmetric, then

$$\left(B^T AB\right)^T = B^T A^T \left(B^T\right)^T = B^T(-A)B = -B^T AB.$$

EXAMPLE 4.21

Express the matrix

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 0 & 2 \\ 7 & 2 & 5 \end{bmatrix}$$

as the sum of a symmetric matrix and a skew-symmetric matrix.

Solution. We know that every square matrix A can be expressed as the sum of symmetric matrix $\frac{1}{2}(A + A^T)$ and a skew-symmetric matrix $\frac{1}{2}(A - A^T)$. In the present case

$$\frac{1}{2}\left(A + A^T\right) = \frac{1}{2}\left[\begin{pmatrix} 1 & 2 & 4 \\ 3 & 0 & 2 \\ 7 & 2 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 3 & 7 \\ 2 & 0 & 2 \\ 4 & 2 & 5 \end{pmatrix}\right]$$
$$= \frac{1}{2}\begin{pmatrix} 2 & 5 & 11 \\ 5 & 0 & 4 \\ 11 & 4 & 10 \end{pmatrix} = \begin{pmatrix} 1 & \frac{5}{2} & \frac{11}{2} \\ \frac{5}{2} & 0 & 2 \\ \frac{11}{2} & 2 & 5 \end{pmatrix}$$

and

$$\frac{1}{2}\left(A - A^T\right) = \frac{1}{2}\left[\begin{pmatrix} 1 & 2 & 4 \\ 3 & 0 & 2 \\ 7 & 2 & 5 \end{pmatrix} - \begin{pmatrix} 1 & 3 & 7 \\ 2 & 0 & 2 \\ 4 & 2 & 5 \end{pmatrix}\right]$$
$$= \frac{1}{2}\begin{pmatrix} 0 & -1 & -3 \\ 1 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{2} & -\frac{3}{2} \\ \frac{1}{2} & 0 & 0 \\ \frac{3}{2} & 0 & 0 \end{pmatrix}.$$

Hence

$$\begin{pmatrix} 1 & 2 & 4 \\ 3 & 0 & 2 \\ 7 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & \frac{5}{2} & \frac{11}{2} \\ \frac{5}{2} & 0 & 2 \\ \frac{11}{2} & 2 & 5 \end{pmatrix} + \begin{pmatrix} 0 & -\frac{1}{2} & -\frac{3}{2} \\ \frac{1}{2} & 0 & 0 \\ \frac{3}{2} & 0 & 0 \end{pmatrix}.$$

Definition 4.55 A matrix obtained from a given matrix A by replacing its elements by the corresponding conjugate complex numbers is called the *conjugate* of A and is denoted by $\bar{A}$.

Thus if $A = [a_{ij}]_{m \times n}$, then $\bar{A} = [\overline{a_{ij}}]_{m \times n}$, where $\bar{a}_{ij}$ denotes the complex conjugate of a_{ij}.

Definition 4.56 A matrix whose all elements are real is called a *real* matrix.

If A is a real matrix, then obviously $A = \bar{A}$ Further if A and B are two matrices, then

(a) $(\bar{\bar{A}}) = A$
(b) $\overline{(A + B)} = \bar{A} + \bar{B}$
(c) $\overline{(\lambda A)} = \bar{\lambda}\,\bar{A}$
(d) $\overline{(AB)} = \bar{A}\,\bar{B}$, where λ a complex number.

Definition 4.57 The transpose of the conjugate of a matrix A is called *transposed conjugate* or *tranjugate* of A and is denoted by A^{θ} or sometimes by A^*.

We observe that

$$(\bar{A})^T = \overline{(A^T)}.$$

For example, let

$$A = \begin{bmatrix} 2 & 1 + 2i & 3 + 4i \\ 1 + i & 7 & 2 + i \\ 3 + 2i & 4 + i & 3 + 3i \end{bmatrix},$$

then

$$\bar{A} = \begin{bmatrix} 2 & 1-2i & 3-4i \\ 1-i & 7 & 2-i \\ 3-2i & 4-i & 3-3i \end{bmatrix}$$

and

$$A^{\theta} = \begin{bmatrix} 2 & 1-i & 3-2i \\ 1-2i & 7 & 4-i \\ 3-4i & 2-i & 3-3i \end{bmatrix}.$$

Let A and B the matrices, then the tranjugate of the matrix possesses the following properties:

(a) $(A^{\theta})^{\theta} = A$
(b) $(A+B)^{\theta} = A^{\theta} + B^{\theta}$, A and B being of the same order.
(c) $(\lambda A)^{\theta} = \bar{\lambda} A^{\theta}$, λ being a complex number.
(d) $(AB)^{\theta} = B^{\theta} A^{\theta}$, A and B being conformable to multiplication.

Definition 4.58 A square matrix $A = [a_{ij}]$ is said to be *Hermitian* if $a_{ij} = \overline{a_{ji}}$ for all i and j.
Thus, a matrix is Hermitian if and only if $A = A^{\theta}$. We note that

(a) A real Hermitian matrix is a real symmetric matrix.
(b) If A is Hermitian, then

$$a_{ii} = \overline{a_{ii}} \text{ for all } i,$$

and so a_{ii} is real for all i. Thus, every diagonal element of a Hermitian matrix must be real.

Definition 4.59 A square matrix $A = [a_{ij}]$ is said to be Skew-Hermitian if $a_{ij} = \overline{a_{ji}}$ for all i and j. Thus, a matrix is Skew-Hermitian if $A = -A^{\theta}$. We observe that

(a) A real Skew-Hermitian matrix is nothing but a real Skew-symmetric matrix.
(b) If A is Skew-Hermitian matrix, then $a_{ii} = -\bar{a}_{ii}$ or $a_{ii} + \bar{a}_{ii} = 0$ and so a_{ii} is either a pure imaginary number or must be zero. Thus the diagonal element of a Skew-Hermitian matrix must be a pure imaginary number or zero.

For example,

$$\begin{bmatrix} 2 & 3-4i & 2+3i \\ 3+4i & 0 & 7-5i \\ 2-3i & 7+5i & 4 \end{bmatrix}$$

is an Hermitian matrix, whereas, the matrix

$$\begin{bmatrix} 0 & 3+4i \\ -3+4i & i \end{bmatrix}$$

is Skew-Hermitian.
It can be shown easily that if A is any square matrix, then $A + A^{\theta}$, AA^{θ}, $A^{\theta}A$ are Hermitian and $A - A^{\theta}$ is Skew-Hermitian.

EXAMPLE 4.22

Show that every square matrix can be uniquely expressible as the sum of a Hermitian matrix and Skew-Hermitian matrix.

Solution. As mentioned above, if A is any square matrix, then $A + A^{\theta}$ is Hermitian and $A - A^{\theta}$ is Skew-Hermitian. Therefore, $\frac{1}{2}(A + A^{\theta})$ and $\frac{1}{2}(A - A^{\theta})$ are Hermitian and Skew- – Hermitian, respectively, so that

$$A = \frac{1}{2}(A + A^{\theta}) + \frac{1}{2}(A - A^{\theta}),$$

which proves first part of our result. The uniqueness can be proved easily and is left to the reader.

EXAMPLE 4.23

Show that every square matrix A can be uniquely expressed as $P + i\,Q$ where P and Q, are Hermitian matrices.

Solution. We take

$$P = \frac{1}{2}(A + A^{\theta}) \text{ and } Q = \frac{1}{2i}(A - A^{\theta}).$$

Then $A = P + iQ$. Further,

$$\begin{aligned} P^{\theta} &= \left(\frac{1}{2}(A + A^{\theta})\right)^{\theta} = \frac{1}{2}(A + A^{\theta})^{\theta} \\ &= \frac{1}{2}A^{\theta} + \frac{1}{2}(A^{\theta})^{\theta} = \frac{1}{2}(A^{\theta} + A) \\ &= \frac{1}{2}(A + A^{\theta}) = P, \end{aligned}$$

showing that P is Hermitian. Similarly,

$$\begin{aligned} Q^{\theta} &= \left[\frac{1}{2i}\left(A - A^{\theta}\right)\right]^{\theta} \\ &= \left(-\frac{1}{2i}\right)\left(A - A^{0}\right)^{\theta} \\ &= -\frac{1}{2i}\left\{A^{\theta} - (A^{\theta})^{\theta}\right\} \\ &= -\frac{1}{2i}\left(A^{\theta} - A\right) \\ &= \frac{1}{2i}\left(A - A^{\theta}\right) = Q, \end{aligned}$$

showing that Q is also Hermitian. Thus $A = P + iQ$, where P and Q are Hermitian.

4.9 LOWER AND UPPER TRIANGULAR MATRICES

Definition 4.60 A square matrix $A = [a_{ij}]$, in which all elements above the main diagonal are zero, is called a *lower triangular matrix*.

Thus a matrix A is lower triangular if $a_{ij} = 0$ for $i < j$.

Definition 4.61 A square matrix $A = [a_{ij}]$, in which all elements below the main diagonal are zero, is called an *upper triangular matrix*.

Thus a matrix A is upper triangular matrix if $a_{ij} = 0$ for $i > j$.

For example,

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \textit{and} \begin{bmatrix} 1 & 0 & 0 \\ 6 & -5 & 0 \\ 0 & 4 & 1 \end{bmatrix}$$

are, respectively, upper triangular and lower triangular matrices.

4.10 ADJOINT OF A MATRIX

Definition 4.62 Let $A = [a_{ij}]$ be a square matrix of order n. Then the *cofactor* of a_{ij} is defined as

$$A_{ij} = cof\ (a_{ij}) = (-1)^{i+j}\ |\ M_{ij}\ |,$$

where M_{ij} is the matrix obtained by deleting ith row and jth column of the matrix A.

For example, if

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 0 & 4 \\ 3 & 6 & 5 \end{bmatrix},$$

then

$$\begin{aligned} \text{cof}(a_{23}) &= cof(4) = (-1)^5 \begin{vmatrix} 1 & 3 \\ 3 & 6 \end{vmatrix} \\ &= -(6-9) = 3 \\ \text{cof}(a_{32}) &= \text{cof}(6) = (-1)^5 \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0. \end{aligned}$$

Definition 4.63 Let $A=[a_{ij}]$ be a square matrix of order n. Then the *cofactor matrix* of A is defined to be the matrix $[A_{ij}]$, where A_{ij} denotes the cofactor of the entry a_{ij} in $|A|$.

For example, if

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & -3 & 1 \\ -1 & -1 & 3 \end{bmatrix},$$

then

$$\begin{aligned} A_{11} &= (-1)^2\ (-9+1) = -8, \\ A_{12} &= (-1)^3\ (0+1) = -1, \\ A_{13} &= (-1)^4\ (0-3) = -3, \\ A_{21} &= (-1)^3\ (3+0) = -3, \\ A_{22} &= (-1)^4\ (6+0) = 6, \\ A_{23} &= (-1)^5\ (-2+1) = 1, \\ A_{31} &= (-1)^4\ (1+0) = 1, \\ A_{32} &= (-1)^5\ (2+0) = -2, \\ A_{33} &= (-1)^6\ (-6+0) = -6. \end{aligned}$$

Hence, the confactor matrix of A is given by

$$[A_{ij}] = \begin{bmatrix} -8 & -1 & -3 \\ -3 & 6 & 1 \\ 1 & -2 & -6 \end{bmatrix}.$$

Definition 4.64 The *adjoint* of a square matrix $A = [a_{ij}]$ of order n is defined to be the transpose of the cofactor matrix of A. Thus

$$\text{adj}\ A = \begin{bmatrix} A_{11} & A_{21} & \dots & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & \dots & A_{n2} \\ A_{13} & A_{23} & \dots & \dots & A_{n3} \\ \dots & \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & \dots & A_{nn} \end{bmatrix}.$$

EXAMPLE 4.24

Find the adj A if

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & -3 & 1 \\ -1 & -1 & 3 \end{bmatrix}.$$

Solution. We have seen earlierthat the cofactor matrix of A is

$$[A_{ij}] = \begin{bmatrix} -8 & -1 & -3 \\ -3 & 6 & 1 \\ 1 & -2 & -6 \end{bmatrix}.$$

Therefore,

$$\text{adj}A = [A_{ij}]^T = \begin{bmatrix} -8 & -1 & -3 \\ -3 & 6 & 1 \\ 1 & -2 & -6 \end{bmatrix}^T = \begin{bmatrix} -8 & -3 & 1 \\ -1 & 6 & -2 \\ -3 & 1 & -6 \end{bmatrix}.$$

Theorem 4.12 Let A be an $n \times n$ matrix. Then

$$A(adj\ A) = A|A|I_n = (adj\ A)\ A.$$

Proof: Since both A and *adj* A are square matrices of order n, the products A(adj A) and (adj A) A are defined. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{bmatrix}.$$

Then the cofactor matrix of A is

$$[A_{ij}] = \begin{bmatrix} A_{11} & A_{12} & \dots & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & \dots & A_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & \dots & A_{nn} \end{bmatrix}$$

and therefore,

$$adj\ A = \begin{bmatrix} A_{11} & A_{21} & \dots & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & \dots & A_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & \dots & A_{nn} \end{bmatrix}.$$

Thus

$$A\ (adj\ A) = \begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{bmatrix} \times \begin{bmatrix} A_{11} & A_{21} & \dots & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & \dots & A_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & \dots & A_{nn} \end{bmatrix}$$

But, we know that

$$a_{i1}\ A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = \begin{cases} |A| \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}.$$

Therefore,

$$A(\text{adj } A) = \begin{bmatrix} |A| & 0 & \dots & \dots & 0 \\ 0 & |A| & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & |A| \end{bmatrix} = |\ A\ | \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 1 \end{bmatrix} = |A|\ I_n.$$

Similarly,

$$(\text{adj } A)A = |A|I_n.$$

Hence

$$A\ (\text{adj } A) = |A|\ I_n = (\text{adj } A)\ A.$$

Corollary 4.1 If $|A| \neq \mathbf{0}$, then

$$A\left(\frac{1}{|A|}\text{adj } A\right) = I_n = \left(\frac{1}{|A|}\text{adj } A\right) A.$$

4.11 THE INVERSE OF A MATRIX

Definition 4.65 A square matrix A of order n is said to be *invertible* if there exists another square matrix B of order n such that

$$AB = BA = I_n.$$

The matrix B is then called the *inverse of* A. If there exists no such matrix B, then A is called *non-invertible* (*singular*). The inverse of A is denoted by A^{-1}. For example, if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$, then

$$AB = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$BA = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus $AB = BA = I_2$. Hence A is invertible and its inverse is B.

Theorem 4.13 The inverse of a square matrix is unique.

Proof: Suppose on the contrary that B and C are two inverse of a matrix A. Then

$$AB = BA = I_n \tag{18}$$

and

$$AC = CA = I_n. \tag{19}$$

Thus, we have

$$\begin{aligned} B &= B\,I_n && \text{(property of identity matrix)} \\ &= BAC && [\text{using}(19)] \\ &= (BA)\,C && \text{(Associative Law)} \\ &= I_n C && [\text{using}(18)] \\ &= C. \end{aligned}$$

Hence, inverse of A is unique.

Definition 4.66 A square matrix A is called *non-singular* if $|A| \neq 0$.

The square matrix A will be called *singular* if $|A| = 0$.

Theorem 4.14 A square matrix A is invertible if and only if it is non-singular.

Proof: The condition is necessary. Let A be invertible and let B be the inverse of A so that

Therefore, $\quad AB = I = BA.$

$$|A|\,|B| = |I| = 1.$$

Hence $|A| \neq 0$.

The condition is sufficient. Let A be non singular. Therefore, $|A| \neq 0$. Let

$$B = \frac{1}{|A|}(\text{adj } A). \text{ Then}$$

$$AB = A\left(\frac{1}{|A|}\text{adj } A\right) = \frac{1}{|A|}\,(A \text{ adj } A)$$

$$= \frac{1}{|A|}\,[\,|A|\,I] = I.$$

Similarly, $BA = I$. Hence, $AB = BA = I$ and so $B = \frac{1}{|A|}(\text{adj } A)$ is the inverse of A.

Theorem 4.15 Let A and B be two non-singular matrices of the same order. Then AB is non-singular and

$$(AB)^{-1} = B^{-1}\,A^{-1}.$$

Proof: Since

$$|AB| = |A|\,|B| \neq 0,$$

it follows that AB is non-singular and so invertible. Moreover,

$$\begin{aligned} (AB)\,\left(B^{-1}\,A^{-1}\right) &= A\left(B\,B^{-1}\right)A^{-1} \\ &= AIA^{-1} = AA^{-1} = I \end{aligned}$$

and

$$\begin{aligned} \left(B^{-1}A^{-1}\right)(AB) &= B^{-1}\left(A^{-1}A\right)B = B^{-1}I\,B \\ &= B^{-1}B = I. \end{aligned}$$

Hence

$$(AB)\left(B^{-1}A^{-1}\right) = I = \left(B^{-1}A^{-1}\right)(AB),$$

which proves that $B^{-1}\,A^{-1}$ is the inverse of AB, that is,

$$(AB)^{-1} = B^{-1}A^{-1}$$

Theorem 4.16 If A is a non-singular matrix, then

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T.$$

(Thus operations of transposing and inversion commute).

Proof: We note that

$$A^T\left(A^{-1}\right)^T = \left(A^{-1}A\right)^T = I^T = I$$

and

$$\left(A^{-1}\right)^T A^T = \left(AA^{-1}\right)^T = I^T = I.$$

Hence

$$A^T\left(A^{-1}\right)^T = I = \left(A^{-1}\right)^T A^T$$

and so

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T.$$

Theorem 4.17 If a matrix A is invertible, then A^θ is invertible and

$$\left(A^\theta\right)^{-1} = \left(A^{-1}\right)^\theta.$$

Proof: We have

$$A^\theta\left(A^{-1}\right)^\theta = \left(A^{-1}A\right)^\theta = I^\theta = I$$

and

$$\left(A^{-1}\right)^\theta A^\theta = \left(A\,A^{-1}\right)^\theta = I^\theta = I.$$

Thus, $\left(A^{-1}\right)^\theta$ is the inverse of the A^θ.

4.12 METHODS OF COMPUTING INVERSE OF A MATRIX

1. Method of an Adjoint Matrix

If A is non-singular square matrix, then we have

$$A\left(\frac{1}{|A|}\text{adj } A\right) = I = \left(\frac{1}{|A|}\text{adj } A\right)A.$$

This relation yields

$$A^{-1} = \frac{1}{|A|}\text{adj } A.$$

EXAMPLE 4.25

Find the inverse of the matrix

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

Solution. We have

$$|A| = \begin{vmatrix} 3 & -3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{vmatrix}$$

$$= 3\,(-3+4) + 3\,(\ 2\ ) + 4\,(-2) = 1.$$

Cofactor of the entries are

$$A_{11} = 1, \quad A_{12} = -2 \quad A_{13} = -2$$
$$A_{21} = -1, \quad A_{22} = 3 \quad A_{23} = 3$$
$$A_{31} = 0 \quad A_{32} = -4 \quad A_{33} = -3.$$

Therefore, the cofactor matrix is

$$\left[A_{ij}\right] = \begin{bmatrix} 1 & -2 & -2 \\ -1 & 3 & 3 \\ 0 & -4 & -3 \end{bmatrix}$$

and so

$$\text{adj } A = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 3 & -4 \\ -2 & 3 & -3 \end{bmatrix}.$$

Hence

$$A^{-1} = \frac{1}{|A|}\text{adj } A = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 3 & -4 \\ -2 & 3 & -3 \end{bmatrix}.$$

EXAMPLE 4.26

Find A^{-1} if

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 0 & 2 \\ 4 & -2 & 5 \end{bmatrix}.$$

Solution. We have $|A| = -4$. The cofactor matrix is

$$\left[A_{ij}\right] = \begin{bmatrix} 4 & -7 & -6 \\ -8 & 9 & 10 \\ 4 & -5 & -6 \end{bmatrix}$$

and so

$$\text{adj } A = \begin{bmatrix} 4 & -8 & 4 \\ -7 & 9 & -5 \\ -6 & 10 & -6 \end{bmatrix}.$$

Hence

$$A^{-1} = -\frac{1}{4}\begin{bmatrix} 4 & -8 & 4 \\ -7 & 9 & -5 \\ -6 & 10 & -6 \end{bmatrix} = \begin{bmatrix} -1 & 2 & -1 \\ \frac{7}{4} & \frac{-9}{4} & \frac{5}{4} \\ \frac{3}{2} & \frac{-5}{2} & \frac{3}{2} \end{bmatrix}$$

2. Method Using Definition of Inverse

Let B be the inverse of matrix A, which is non-singular. Then, $AB = I$, that is,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & \dots & b_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 1 \end{bmatrix}$$

Multiplying the matrices on the left and then comparing the corresponding entries we can find b_{11}, $b_{12}, \dots, b_{nn}$. Then, B will be the inverse of A.

EXAMPLE 4.27

Find the inverse of

$$A=\begin{bmatrix}1 & 2 & -1\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}.$$

Solution. The given matrix is upper triangular matrix. The readers may prove that inverse of an upper triangular matrix is also upper triangular matrix. Similarly, the inverse of a lower-triangular matrix is again a lower-triangular matrix. So, let

$$\begin{bmatrix}a & b & c\\ 0 & d & e\\ 0 & 0 & f\end{bmatrix}$$

be the inverse of the given matrix. Then, by definition of the inverse, we must have

$$\begin{bmatrix}1 & 2 & -1\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}a & b & c\\ 0 & d & e\\ 0 & 0 & f\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

or

$$\begin{bmatrix}a & b+2d & c+2e-f\\ 0 & d & e+3f\\ 0 & 0 & f\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Equating corresponding entries, we get

$a=1,\ d=1,\ f=1$

$b+2d=0$ so that $b=-2d=-2$

$e+3f=0$ so that $e=-3f=-3$

$c+2e-f=0$ so that $c=f-2e=1+6=7.$

Hence,

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

Remark 4.3 We can also find the inverse of a lower triangular matrix by the above method.

3. Method of Matrix Equation.

Let

$$\begin{aligned}a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2\\ \cdots\quad\cdots\quad&\cdots\quad\cdots\\ \cdots\quad\cdots\quad&\cdots\quad\cdots\\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n&=b_n\end{aligned}$$

be a set of n equations in n variables $x_1, x_2, \ldots, x_n$ In matrix form, we can represent these equations by

$$AX=B,$$

where

$$A=\begin{bmatrix}a_{11} & a_{12} & \ldots & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & \ldots & a_{2n}\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ a_{n1} & a_{n2} & \ldots & \ldots & a_{nn}\end{bmatrix},$$

$$X=\begin{bmatrix}x_1\\ x_2\\ \ldots\\ \ldots\\ x_n\end{bmatrix},\quad B=\begin{bmatrix}b_1\\ b_2\\ \ldots\\ \ldots\\ b_n\end{bmatrix}$$

and A is called the *coefficient matrix*. If A is non-singular matrix, then A^{-1} exists. Premultiplying the matrix equation by A^{-1} we get

$$A^{-1}(AX)=A^{-1}B$$

or

$$(A^{-1}A)X=A^{-1}B$$

or

$$IX=A^{-1}B$$

or

$$X=A^{-1}B.$$

Hence, if we can represent $x_1, x_2, \ldots, x_n$ in terms of $b_1, b_2, \ldots, b_n$, then the coefficient matrix of this system will be the inverse of A.

EXAMPLE 4.28

Find the inverse of

$$A=\begin{bmatrix}1 & 0 & -4\\ 0 & -1 & 2\\ -1 & 2 & 1\end{bmatrix}.$$

Solution. We observe that $|A|=-1\neq 0$. Thus, A is non-singular and so the inverse of A exists. We consider the matrix equation

$$\begin{bmatrix}1 & 0 & -4\\ 0 & -1 & 2\\ -1 & 2 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix},$$

which yields

$$\begin{aligned} x_1 - 4x_3 &= b_1 \\ -x_2 + 2x_3 &= b_2 \\ -x_1 + 2x_2 + x_3 &= b_3. \end{aligned}$$

Solving these equations for x_1, x_2, and x_3, we have

$$\begin{aligned} x_1 &= 5b_1 + 8b_2 + 4b_3 \\ x_2 &= 2b_1 + 3b_2 + 2b_3 \\ x_3 &= b_1 + 2b_2 + \ b_3. \end{aligned}$$

In matrix form, we have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 & 8 & 4 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

that is

$$X = A^{-1}B.$$

Hence

$$A^{-1} = \begin{bmatrix} 5 & 8 & 4 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

4. Method of Elementary Transformation (Gauss-Jordan Method)

The following transformations are called *elementary transformation* of a matrix:

(a) Interchanging of rows (columns).

(b) Multiplication of a row (column) by a non-zero scalar.

(c) Adding/subtracting k multiple of a row (column) to another row (column).

Definition 4.67 A matrix B is said to be *row* (column) *equivalent* to a matrix A if it is obtained from A by applying a finite number of elementary row (column) transformations. In such case, we write B ~ A. In Gauss–Jordan Method, we perform the sequence of elementary row transformations on A and I simultaneously, keeping them side-by-side.

EXAMPLE 4.29

Using elementary row transformations, find A^{-1} if

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 5 \end{bmatrix}.$$

Solution. Consider the augmented matrix

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 5 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & -3 & -4 & 0 & 1 \end{array}\right] \begin{array}{l} R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 4R_1 \end{array}$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 1 & -3 & -4 & 0 & 1 \end{array}\right] R_2 \to -R_2$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & -4 & -6 & 1 & 1 \end{array}\right] R_3 \to R_3 - R_2$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & \frac{3}{2} & -\frac{1}{4} & -\frac{1}{4} \end{array}\right] R_3 \to -\frac{1}{4}R_3$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{4} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{3}{2} & -\frac{1}{4} & -\frac{1}{4} \end{array}\right] R_2 \to R_2 - R_3$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{4} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{3}{2} & -\frac{1}{4} & -\frac{1}{4} \end{array}\right] R_1 \to R_1 - 2R_3$$

Hence

$$A^{-1} = \begin{bmatrix} -2 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{3}{4} & \frac{1}{4} \\ \frac{3}{2} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix}.$$

EXAMPLE 4.30

Using elementary row transformation, find the inverse of the matrix

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}.$$

Solution. Consider the augmented matrix

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 3 & 3 & 1 & 0 & 0 \\ 1 & 4 & 3 & 0 & 1 & 0 \\ 1 & 3 & 4 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 3 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & -1 \\ 1 & 3 & 4 & 0 & 0 & 1 \end{array}\right] R_2 \to R_2 - R_3$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 3 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & -1 \\ 1 & 3 & 1 & -1 & 0 & 1 \end{array}\right] R_3 \to R_3 - R_1$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 6 & 1 & -3 & 3 \\ 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right] R_1 \to R_1 - 3R_2$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 6 & 1 & -3 & 3 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right] R_2 \to R_2 + R_3$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 7 & -3 & -3 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right] R_1 \to R_1 - 6R_3.$$

Hence

$$A^{-1} = \begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

EXAMPLE 4.31

Find the inverse of the matrix

$$A = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & -3 \\ -2 & -4 & -4 \end{bmatrix},$$

by using elementary transformations.

Solution. Write $A = I_3 A$, that is,

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

Now we reduce the matrix A to identity matrix I_3 by elementary row transformation keeping in mind that each such row transformation will apply to the prefactor I_3 on the right hand side.

Performing $R_2 \to R_2 - R_1$ and $R_3 \to R_3 + 2R_1$, we get

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

Performing $R_2 \to \frac{1}{2}R_2$, we get

$$\begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & -3 \\ 0 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 2 & 0 & 1 \end{bmatrix} A.$$

Performing $R_1 \to R_1 - R_2$ and $R_3 \to R_3 + R_2$, we get,

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

Performing $R_3 \to -\frac{1}{4}R_3$, we get

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

Performing $R_1 \to R_1 + 6R_3$ and $R_2 \to R_2 + 3R_3$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & \frac{3}{2} \\ -\frac{5}{4} & -\frac{1}{4} & -\frac{3}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix} A.$$

Thus,

$$I_3 = \begin{bmatrix} 3 & 1 & \frac{3}{2} \\ -\frac{5}{4} & -\frac{1}{4} & -\frac{3}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix} A.$$

Hence

$$A^{-1} = \begin{bmatrix} 3 & 1 & \frac{3}{2} \\ -\frac{5}{4} & -\frac{1}{4} & -\frac{3}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix}.$$

4.13 RANK OF A MATRIX

Definition 4.68 A matrix is said to be of *rank r* if it has at least one non-singular submatrix of order r but has no non-singular submatrix of order more than r.

Rank of a matrix A is denoted by $\rho(A)$.

A matrix is said to be of *rank zero* if and only if all its elements are zero.

EXAMPLE 4.32

Find the rank of the matrix

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

Solution. The matrix A is of order 3×4. Therefore, $\rho(A) \leq 3$. We note that

$$|A_1| = \begin{vmatrix} 1 & 3 & 4 \\ 2 & 4 & 6 \\ -1 & 5 & 4 \end{vmatrix} = 0,$$

$$|A_2| = \begin{vmatrix} 1 & 3 & 2 \\ 2 & 4 & 2 \\ -1 & 5 & 6 \end{vmatrix} = 0,$$

$$|A_3| = \begin{vmatrix} 1 & 4 & 2 \\ 2 & 6 & 2 \\ -1 & 4 & 6 \end{vmatrix} = 0,$$

$$|A_4| = \begin{vmatrix} 3 & 4 & 2 \\ 4 & 6 & 2 \\ 5 & 4 & 6 \end{vmatrix} = 0.$$

Therefore, $\rho(A) \neq 3$. But, we have submatrix $B = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, whose determinant is equal to $-2 \neq 0$. Hence, by definition, $\rho(A) = 2$.

EXAMPLE 4.33

Find the rank of the matrix

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

Solution. Since $|A| = 0$, $\rho(\text{A}) \leq 2$. But, we note that $\begin{vmatrix} 2 & 1 \\ 0 & 3 \end{vmatrix} = 6 \neq 0$. Hence, $\rho(A) = 2$.

Remark 4.4 The rank of a matrix is, of course, uniquely defined when the elements are all explicitly given numbers, but not necessarily otherwise. For example, consider the matrix

$$A = \begin{bmatrix} 4-x & 2\sqrt{5} & 0 \\ 2\sqrt{5} & 4-x & \sqrt{5} \\ 0 & \sqrt{5} & 4-x \end{bmatrix}.$$

We have

$$|A| = (4-x)^3 - 25(4-x) = 0, \text{if } x = 9, 4 \text{ or } -1.$$

When $x = 9$, we have the singular matrix

$$A = \begin{bmatrix} -5 & 2\sqrt{5} & 0 \\ 2\sqrt{5} & -5 & \sqrt{5} \\ 0 & \sqrt{5} & -5 \end{bmatrix},$$

which has non-singular submatrix

$$\begin{bmatrix} -5 & \sqrt{5} \\ \sqrt{5} & -5 \end{bmatrix}.$$

Thus, for $x = 9$ the rank of A is 2. Similarly, the rank is 2 when $x = 4$ or $x = -1$. For other values of x, $|A| \neq 0$ and so the rank of A is 3.

Theorem 4.18 Let A be an $m \times n$ matrix. Then $\rho(A) = \rho(A^T)$.

Proof: Suppose $\rho(A) = r$. Then, there is at least one square submatrix R of A of order r whose determinant is non-zero. If R^{T} is transpose of R, then it is submatrix of A^{T}. Since, the value of a determinant does not alter by interchanging the rows and columns, $|R^{\text{T}}| = |\text{R}| \neq 0$. Therefore, $\rho(A^T) \geq r$.

Now if A^{T} contains a square submatrix S of order $\text{r}+1$, then corresponding to S, S^{T} is a submatrix of A of order $r+1$. But $\rho(A) = r$. Therefore, $|S| = |S^T| = 0$. Thus, A^{T} cannot contain an $(r+1)$ rowed square submatrix with non-zero determinant. Thus, $\rho(A^{\text{T}}) \leq r$. Hence, $\rho(A^{\text{T}}) = r$.

Theorem 4.19 The rank of a matrix does not alter under elementary row (column) transformations.

Proof: Let $\text{A} = [a_{ij}]$ be an $m \times n$ matrix of rank r. We prove the theorem only for elementary row transformation. The proof for column transformation is similar.

Case I. Interchange of a pair of row does not alter the rank.

Let s be the rank of the matrix B obtained from the matrix A of rank r by elementary transformation R_p R_q. Let B_0 be any $(r+1)$ rowed square submatrix of B. The $(r+1)$ rows of B_0 are also the rows of some uniquely determined submatrix A_0 of A. The identical rows of A_0 and B_0 may occur in the same or in different relative positions. Since, the interchange of two rows of a determinant changes only the sign, we have

$$B_0| = |A_0| \quad or \quad |B_0| = -|A_0|.$$

Since $\rho(A)=r$, every $(r+1)$-rowed minor of A vanishes, that is, $|A_0|$ =0. Hence, $|\ B_0\ |=0$. Therefore, every $(r+1)$ rowed minor of B vanishes. Hence, $s=\rho(B)\leq r=\rho(A)$. But A can also be obtained from B by interchanging its rows. Therefore, $r\leq s$. Hence $r=s$.

Case II. Multiplication of the elements of a row by a non-zero number does not alter the rank.

Let s be the rank of the matrix B obtained from the matrix A of rank r by the elementary transformation $R_p \to kR_p$ $(k\neq 0)$ If B_0 is any $(r+1)$-rowed submatrix of B, then there exists a uniquely determined submatrix A_0 of A such that $|B_0|=|A_0|$ (when pth row of B is one of those rows which are deleted to obtain B_0 from B) or $|\ B_0\ |=$k$|\ A_0\ |$ (when pth row of B is retained while obtaining B_0 from B). Since $\rho(A)=r$, every $(r+1)$-rowed submatrix has zero determinant, that is $|A_0|=0$. Hence, $|\ B_0\ |=0$. Thus every $(r+1)$-rowed submatrix of B vanishes. Hence $\rho(B)\leq r$, that is, $s\leq r$. On the other hand, A can be obtained from B by elementary transformation $R_p \to \frac{1}{k}R_p$. Therefore, we have $r\leq s$. Hence $r=s$.

Case III. Addition to the elements of a row, the product by any number k of the corresponding elements of any other row, does not alter the rank.

Let s be the rank of the matrix B obtained from the matrix A by elementary transformation $R_p \to R_p+kR_q$. Let B_0 be any $(r+1)$-rowed square submatrix of B and A_0 be the corresponding placed submatrix of A. The transformation $R_p \to R_p+kR_q$ has changed only the pth row of the matrix A. We know that the value of the determinant does not change if we add to the elements of any row the corresponding elements of any other row multiplied by some number. Therefore, if no row of the submatrix A_0 is a part of the pth row or if two rows of A_0 are parts of the pth and qth rows of A, then $|\ B_0\ |=|\ A_0\ |$. Since $\rho(A)=r$, we have $|\ A_0\ |=0$ and consequently $|\ B_0\ |=0$.

Again, if a row of A_0 is a part of the pth row of A, but no row is part of qth row, then

$$|B_0| = |A_0| + k|C_0|,$$

where C_0 is an $(r+1)$-rowed square matrix which can be obtained from A_0 by replacing the elements of A_0 in the row which corresponds to the pth row of A by the corresponding elements in the qth row of A. All the $(r+1)$ rows of the matrix C_0 are exactly the same as the rows of some $(r+1)$-rowed square submatrix of A, though arranged in some different order. Therefore, $|\ C_0\ |=\pm$ times some $(r+1)$-rowed minor of A. Since the rank of A is r, every $(r+1)$-rowed minor of A is also zero, so that $|\ A_0\ |=0$, $|\ C_0\ |=0$, and so in turn $|\ B_0\ |=0$. Thus, every $(r+1)$-rowed square matrix of B has zero determinant. Hence, $s\leq r$. Also, since, A can be obtained from B by an elementary transformation, $R_p \to R_p+kR_p$ Therefore, as stated, $r\leq s$. Hence $r=s$.

EXAMPLE 4.34

Find the rank of the matrix

$$A = \begin{bmatrix} 3 & 2 & -1 \\ 4 & 2 & 6 \\ 7 & 4 & 5 \end{bmatrix}.$$

Solution. We have

$$A = \begin{bmatrix} 3 & 2 & -1 \\ 4 & 2 & 6 \\ 7 & 4 & 5 \end{bmatrix}$$

$$\sim \begin{bmatrix} -1 & 0 & -7 \\ 4 & 2 & 6 \\ 7 & 4 & 5 \end{bmatrix} R_1 \to R_1 - R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 7 \\ 4 & 2 & 6 \\ 7 & 4 & 5 \end{bmatrix} R_1 \to -R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 7 \\ 0 & 2 & -22 \\ 0 & 4 & -44 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 4R_1 \\ R_3 \to R_3 - 7R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & -11 \\ 0 & 4 & -44 \end{bmatrix} R_2 \to \frac{1}{2} R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & -11 \\ 0 & 0 & 0 \end{bmatrix} R_3 \to R_3 - 4R_2.$$

Thus, $|A| = 0$. Therefore $\rho(A) \neq 3$. But, since

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \neq 0,$$

it follows that $\rho(A) = 2$.

4.14 ELEMENTARY MATRICES

Definition 4.69 A matrix obtained from a unit matrix by a single elementary transformation is called an *elementary matrix*.

For example,

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

is the elementary matrix obtained from I_3 by subjecting it to $C_1 \leftrightarrow C_3$. The matrix

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is the elementary matrix obtained from I_3 by subjecting it to $R_1 \to 4R_1$, whereas the matrix

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is the elementary matrix obtained from I_3 by subjecting it to $R_1 \to R_1 + 2\,R_1$. The elementary matrix obtained by interchanging the ith and jth row of a unit matrix I is denoted by E_{ij}. Since, we obtain same matrix by interchanging ith and jth row or ith and jth column, E_{ij} will also denote the elementary matrix obtained from A by interchanging ith and jth column. $E_i\ (k)$ denotes the elementary matrix obtained by multiplying the ith row or ith column of a unit matrix by k.

Similarly, $E_{ij}(m)$ denotes the elementary matrix obtained by adding to the elements of the ith row (column) of a unit matrix the m multiple of the corresponding elements of the jth row (column).

We note that $|E_{ij}| = -1$, $|E_i(k)| = k \neq 0$ $|\,E_{ij}(m)\,| = 1$. It follows, therefore, that *all the elementary matrices are non-singular and, hence, possess inverse.*

Theorem 4.20 Every elementary row (column) transformation of a matrix can be obtained by pre-multiplication (post-multiplication) with corresponding elementary matrix.

Proof: Let B be the matrix obtained from an $m \times n$ matrix A by row transformation. If E is elementary matrix obtained from I_m by the same row transformation, it is sufficient to show that $B = EA$.

Let

$$M = \begin{bmatrix} R_1 \\ R_2 \\ \dots \\ \dots \\ R_m \end{bmatrix}, \quad N = [C_1 C_2 \ \dots \ C_n].$$

Then

$$MN = \begin{bmatrix} R_1C_1 & R_1C_2 & \dots & \dots & R_1C_n \\ R_2C_1 & R_2C_2 & \dots & \dots & R_2C_n \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ R_mC_1 & R_mC_2 & \dots & \dots & R_mC_n \end{bmatrix}.$$

Clearly, a row transformation applied to M will be the row transformation applied to MN. Hence, elementary row transformation of a product MN of two matrices M and N can be obtained by subjecting the prefactor M to the same elementary row transformation.

Similarly, every elementary column transformation of a product MN can be obtained by subjecting the post-factor N to the same elementary column transformation.

Now, A is an $m \times n$ matrix and I_m is an identity matrix of order m. Therefore, $A = I_m A$ Hence, by the preceding arguments, if we apply a row transformation to A to get a matrix B, then this can be done by applying the same row transformation to I_m. Thus, if B is obtained from A by applying a row transformation and E is obtained from I_m by using the same row transformation, then $B = EA$.

Similarly, if B is obtained from A by subjecting it to a column transformation and E is obtained from I by subjecting it to the same column transformation then $B = AE$.

EXAMPLE 4.35

Let

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 1 & 3 \\ 5 & 3 & 2 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 1 & 3 & 4 \\ 4 & 2 & 6 \\ 5 & 3 & 2 \end{bmatrix}$$

Thus, B has been obtained from A by the row transformation $R_2 \to 2\ R_2$. Now, if, E is the elementary matrix obtained from I_3 by $R_2 \to 2\ R_2$ then

$$EA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 4 \\ 2 & 1 & 3 \\ 5 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 4 \\ 4 & 2 & 6 \\ 0 & 0 & 1 \end{bmatrix} = B.$$

4.15 ROW REDUCED ECHELON FORM AND NORMAL FORM OF MATRICES

Definition 4.70 A matrix is said to be in *row-reduced echelon* form if

(i) The first non-zero entry in each non-zero row is 1.
(ii) The rows containing only zeros occur below all the non-zero rows.
(iii) The number of zeros before the first non-zero element in a row is less than the number of such zeros in the next row.

The rank of a matrix in row reduced echelon form is equal to the number of non-zero rows of the matrix. For example, the matrix,

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

is in the row reduced echelon form and its rank is 2 (the number of non-zero rows).

Theorem 4.21 Every non-zero $m \times n$ matrix of rank r can be reduced, by a sequence of elementary transformation, to the form

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

(normal form or first canonical form), where I_r is the identity matrix of order r.

Proof: Let $A = [a_{ij}]_{m \times n}$ be a matrix of rank r. Since A is non-zero, it has at least one element different from zero. Suppose $a_{ij} \neq 0$. Interchanging the first and ith row and then first and jth column we obtain a matrix B whose leading element is non-zero, say k.

Multiplying the elements of the first row of the matrix B by $\frac{1}{k}$, we obtain a matrix

$$C = \begin{bmatrix} 1 & c_{12} & c_{13} & \dots & \dots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \dots & \dots & c_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & c_{m3} & \dots & \dots & c_{mn} \end{bmatrix},$$

whose leading element is equal to 1. Subtracting suitable multiples of the first column of C from the remaining columns, and suitable multiples of first row from the remaining rows, we obtain a matrix

$$D = \begin{bmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & d_{22} & d_{23} & \dots & \dots & d_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & d_{m2} & d_{m3} & \dots & \dots & d_{mn} \end{bmatrix},$$

in which all elements of the first row and first column except the leading element are equal to zero. If

$$\begin{bmatrix} d_{22} & d_{23} & \dots & \dots & \dots & d_{2n} \\ d_{32} & d_{34} & \dots & \dots & \dots & d_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ d_{m2} & d_{m3} & \dots & \dots & \dots & d_{mn} \end{bmatrix} \neq 0,$$

we repeat the above process for this matrix and get a matrix

$$E = \begin{bmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & e_{33} & \dots & \dots & e_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 3_{m3} & \dots & \dots & e_{mn} \end{bmatrix}.$$

Continuing this process, we obtain a matrix

$$N = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}.$$

The rank of N is k. Since, the matrix N has been obtained from A by elementary transformations, $\rho(N) = \rho(A)$, that is, $k = r$. Hence, every non-zero matrix can be reduced to the form

$$\begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$$

by a finite chain of elementary transformations.

Corollary 4.2 The rank of an $m \times n$ matrix A is r if and only if it can be reduced to the normal form by a sequence of elementary transformations.

Proof: If $\rho(\text{A}) = r$, then by the above theorem it can be reduced to normal form by a sequence of elementary transformations.

Conversely, let the matrix A has been reduced to normal form $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ by elementary transformations. Now the rank of $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ is r and we know that rank of a matrix is not altered by elementary transformation. Therefore, rank of A is also r.

Corollary 4.3 If A is an $m \times n$ matrix of rank r, there exist non-singular matrices P and Q such that

$$\text{PAQ} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Proof: Since A is an $m \times n$ matrix of rank r, it can be reduced to normal form $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ using a sequence of elementary transformations. Further, since the elementary row (column) transformations are equivalent to pre-(post) multiplication by the corresponding elementary matrices, we have

$$P_s P_{s-1} \dots P_1 \, A Q_1 \, Q_2 \dots Q_t = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Now, since, each elementary matrix is non-singular and the product of non-singular matrices is again non-singular, it follows that $P_s \, P_{s-1} \dots P_1$ and $Q_1 \, Q_2 \dots \, Q_t$ are non-singular matrices, say P and Q. Hence

$$PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

where P and Q are non-singular matrices.

4.16 EQUIVALENCE OF MATRICES

Definition 4.71 Two matrices whose elements are real or complex numbers are said to be *equivalent* if and only if each can be transformed into the other by means of elementary transformations.

If the matrix A is equivalent to the matrix B, then we write A $\sim$ B

The relation of equivalence '$\sim$' in the set of all $m \times n$ matrices is an *equivalence relation*, that is, $\sim$ is reflexive, symmetric, and transitive.

Theorem 4.22 If A and B are equivalent matrices, then $\rho(A) = \rho(B)$.

Proof: If $A \sim B$, then B can be obtained from A by a finite number of elementary transformations. But elementary transformation do not alter the rank of a matrix. Hence $\rho(A) = \rho(\text{B})$.

Theorem 4.23 If two matrices A and B have the same size and the same rank, they are equivalent.

Proof: Let A and B be two $m \times n$ matrices of the same rank r. Then they can be reduced to normal form by elementary transformations. Therefore,

$$A \sim \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B \sim \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

or, by symmetry of the relation of equivalence of matrices

$$A \sim \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \text{and} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \sim B.$$

Using transitivity of the relation, '$\sim$' we have $A \sim B$.

Theorem 4.24 If A and B are equivalent matrices, there exist non-singular matrices P and Q such that $B = PAQ$.

Proof: If $A \sim B$, then B can be obtained from A by a finite number of elementary transformations of A. But elementary row (column) transformations are equivalent to pre (post) multiplication by the corresponding elementary matrices. Therefore, there are elementary matrices $P_1, P_2, \ldots, P_s$ $Q_1, Q_2, \ldots, Q_t$ such that

$$P_s\, P_{s-1} \ldots P_1 A \quad Q_1\, Q_2 \ldots Q_t = B.$$

Since, each elementary matrix is non-singular and the product of non-singular matrices is non-singular, we have,

$$PAQ = B,$$

where $P = P_s P_{s-1} \ldots P_1$ and $Q = Q_1 Q_2 \ldots Q_t$ are non-singular matrices.

Theorem 4.25 Any non-singular matrix of explicitly given numbers may be factored into the product of elementary matrices.

Proof: Any non-singular matrix A of order n and the identity matrix I_n have the same order and same rank. Hence $A \sim I_n$. Therefore, by the Theorem 4.41, there exist elementary matrices P_j and Q_j such that

$$A = P_s P_{s-1} \ldots P_1 I_n Q_1 Q_2 \ldots Q_t.$$

EXAMPLE 4.36

Reduce the matrix

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

to normal form and, hence, find its rank.

Solution. We observe that

$$A \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 4 & 5 & -8 & 14 \\ 0 & 3 & 0 & 4 \\ 0 & 1 & 0 & 2 \end{bmatrix} \begin{matrix} C_2 \to C_2 + C_1 \\ C_3 \to C_3 - 2C_1 \\ C_4 \to C_4 + 3C_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & -8 & 14 \\ 0 & 3 & 0 & 4 \\ 0 & 1 & 0 & 2 \end{bmatrix} R_2 \to R_2 - 4R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 3 & 0 & 4 \\ 0 & 5 & -8 & 14 \end{bmatrix} R_2 \leftrightarrow R_4$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & -2 \\ 0 & 5 & -8 & 4 \end{bmatrix} C_4 \to C_4 - 2C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & -8 & 4 \end{bmatrix} \begin{matrix} R_3 \to R_3 - 3R_2 \\ R_4 \to R_4 - 5R_2 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 4 & -8 \end{bmatrix} C_3 \leftrightarrow C_4$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix} \begin{matrix} C_3 \to -\frac{1}{2}C_3 \\ C_4 \to -\frac{1}{8}C_4 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} R_4 \to R_4 + 2R_3$$

$$= I_4.$$

Hence, $\rho(A) = 4$.

EXAMPLE 4.37

Reduce the matrix

$$\begin{bmatrix} 3 & 2 & -1 \\ 4 & 2 & 6 \\ 7 & 4 & 5 \end{bmatrix}$$

to the normal form and, hence, find its rank.

Solution. We note that

$$A = \begin{bmatrix} 3 & 2 & -1 \\ 4 & 2 & 6 \\ 7 & 4 & 5 \end{bmatrix}$$

$$\sim \begin{bmatrix} -1 & 0 & -7 \\ 4 & 2 & 6 \\ 7 & 4 & 5 \end{bmatrix} R_1 \to R_1 - R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 7 \\ 4 & 2 & 6 \\ 7 & 4 & 5 \end{bmatrix} R_1 \to -R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 7 \\ 4 & 2 & -22 \\ 0 & 4 & -44 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 4R_1 \\ R_3 \to R_3 - 7R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & -11 \\ 0 & 4 & -44 \end{bmatrix} R_2 \to \frac{1}{2} R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & -11 \\ 0 & 0 & 0 \end{bmatrix} R_3 \to R_3 - 4R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -11 \\ 0 & 0 & 0 \end{bmatrix} C_3 \to C_3 - 7C_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} C_3 \to C_3 + 11C_2$$

$$\sim \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence $\rho(A) = 2$.

EXAMPLE 4.38

For the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 3 & 1 & 1 \end{bmatrix},$$

find the non-singular matrices P and Q such that PAQ is in the normal form. Hence, find the rank of the matrix A.

Solution. We write

$A = I_3 A I_3$.

Thus,

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We shall apply elementary transformations on A until it is reduced to normal form, keeping in mind that each row transformation will also be applied to the pre-factor I_3 of the product on the right and each column transformation will also be applied to the post-factor I_3 of the product on the right.

Performing $R_2 \to R_2 - R_1$, $R_3 \to R_3 - 3R_1$, we get,

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

Performing $C_2 \to C_2 - C_1$, $C_3 \to C_3 - C_1$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & -2 \\ 0 & -2 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_2 \to -\frac{1}{2} R_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ -3 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_3 \to R_3 + 2R_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ -2 & -1 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Last, performing $C_3 \to C_3 - C_2$, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ -2 & -1 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

or

$$\begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix} = PAQ,$$

where

$$P = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ -2 & -1 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since A is equivalent to $\begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}$, we have $\rho(A) = 2$.

EXAMPLE 4.39

For the matrix

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 3 & 1 & 1 \end{bmatrix},$$

find non-singular matrices P and Q such that PAQ is in the normal form. Hence find the rank of A.

Solution. Write $A = I_3\ A\ I_3$ that is,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

As, in the above example, we shall reduce A to normal form subjecting it to elementary transformations. Performing $R_2 \to R_2 - R_1$, $R_3 - 3R_1$, we have

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

Performing $c_2 \to c_2 + c_1$, $c_3 \to c_3 - c_1$ we have

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

Performing $R_2 \to \frac{1}{2}R_2$, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -3 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_3 \to R_3 - 4R_2$, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -1 & -2 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Performing $C_3 \to -\frac{1}{2}C_3$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -1 & -2 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}.$$

Hence,

$$I_3 = PAQ,$$

where

$$P = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -1 & -2 & 1 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

Since $A \sim I_3$, $\rho(A) = \rho(I_3) = 3$.

Theorem 4.26 The rank of the product of two matrices cannot exceed the rank of either matrix.

Proof: Let A be $m \times n$ and B be $n \times p$ matrices with rank r_A and r_B, respectively. Then, by Corollary 4.9, there exist non-singular matrices C and D of order m and n, respectively, such that

$$CAD = \begin{bmatrix} I_{r_A} & 0 \\ 0 & 0 \end{bmatrix},$$

where $\begin{bmatrix} I_{r_A} & 0 \\ 0 & 0 \end{bmatrix}$ denotes the normal form of A. Thus

$$A = C^{-1} \begin{bmatrix} I_{r_A} & 0 \\ 0 & 0 \end{bmatrix} D^{-1}$$

so that

$$AB = \left[C^{-1} \begin{bmatrix} I_{r_A} & 0 \\ 0 & 0 \end{bmatrix} D^{-1} \right] B$$

$$= C^{-1} \left(\begin{bmatrix} I_{r_A} & 0 \\ 0 & 0 \end{bmatrix} (D^{-1}B) \right).$$

Since C^{-1} is non-singular, AB has same rank as $\begin{bmatrix} I_{r_A} & 0 \\ 0 & 0 \end{bmatrix}(D^{-1}B)$. But $\begin{bmatrix} I_{r_A} & 0 \\ 0 & 0 \end{bmatrix}$ has zeros in the last $m - r_A$ rows and, hence, $\begin{bmatrix} I_{r_A} & 0 \\ 0 & 0 \end{bmatrix}(D^{-1}B)$ also has only zeros in the last $m - r_A$ rows. Hence, the rank of $\begin{bmatrix} I_{rA} & 0 \\ 0 & 0 \end{bmatrix}(D^{-1}B)$ is at most r_A. It follows, therefore, that

$$\rho(AB) \le r_A$$

Also

$$\rho(AB) = \rho\left((AB)^T\right)$$
$$= \rho[B^T A^T]$$

But, as proved earlier,

$$\rho(B^T A^T) \le \rho(B^T) = \rho(B) = r_B.$$

Hence,

$$\rho(AB) \le r_B.$$

This completes the proof of the theorem.

EXAMPLE 4.40

Let A be any non-singular matrix and B a matrix such that AB exists. Show that

$$\rho(AB) = \rho(B).$$

Solution. Let $C = AB$. Since A is non-singular, therefore $B = A^{-1}C$. Since rank of the product of two matrices does not exceed the rank of either matrix, we have,

$$\rho(C) = \rho(AB) \le \rho(B)$$

and

$$\rho(B) = \rho(A^{-1}C) \le \rho(C).$$

Hence

$$\rho(C) = \rho(AB) \le \rho(B) \le \rho(C),$$

which yields

$$\rho(B) = \rho(C) = \rho(AB).$$

4.17 ROW AND COLUMN EQUIVALENCE OF MATRICES

Definition 4.72 A matrix A is said to be *row* (*column*) *equivalent* to B if B is obtainable from A by a finite number of elementary row (column) transformations of A.

Row equivalence of the matrices A and B is denoted by $A \overset{R}{\sim} B$ and column equivalence of A and B is denoted by $A \overset{C}{\sim} B$.

Theorem 4.27 Let A be an $m \times n$ matrix of rank r. Then there exists a non-singular matrix P such that

$$PA = \begin{bmatrix} G \\ 0 \end{bmatrix},$$

where G is an $r \times n$ matrix of rank r and 0 is $(m - r)$ n matrix.

Proof: Since A is an $m \times n$ matrix of rank r, therefore there exist non-singular matrices P and Q such that

$$PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

But every non-singular matrix can be expressed as product of elementary matrices. So,

$$Q = Q_1 Q_2 \ldots Q_t,$$

where Q_1, $Q_2 \ldots Q_t$ are all elementary matrices. Thus,

$$PAQ_1 \; Q_2 \ldots Q_t = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Since, elementary column transformation of a matrix is equivalent to post-multiplication with the corresponding elementary matrix, we post-multiply the left hand side of the above expression by the elementary matrices $Q_t^{-1}, Q_{t-1}^{-1}, \ldots, Q_2^{-1}, Q_1^{-1}$ successively and effect the corresponding column transformations in the right hand side, we get a relation of the form

$$PA = \begin{bmatrix} G \\ 0 \end{bmatrix}.$$

Since elementary transformations do not alter the rank

$$\rho(PA) = \rho(A) = r \text{ and so } \rho\begin{bmatrix} G \\ 0 \end{bmatrix} = r,$$

which implies that $\rho(G) = r$ since G has r rows and last $m \times r$ rows of $\begin{bmatrix} G \\ 0 \end{bmatrix}$ consist of zero elements only.

Theorem 4.28 Every non-singular matrix is row equivalent to a unit matrix.

Proof: Suppose that the matrix A is of order 1. Then $A = [a_{11}]$ which is clearly row equivalent to a unit matrix. We shall prove our result by induction on the order of the matrix. Let A be of order n. Since the result is true for non-singular matrix of order 1, we assume that the result is true for all matrices of order $n - 1$.

Let $A = [a_{ij}]$ be an $n \times n$ non-singular matrix. The first column of the matrix A has at least one non-zero element, otherwise $|A| = 0$, which contradicts the fact that A is non-singular. Let $a_{11} = k \ne 0$. By interchanging (if necessary) the pth row with the first row, we obtain a matrix B whose leading coefficient is $k \ne 0$. Multiplying the elements of the first row by $\frac{1}{k}$, we get the matrix.

$$C = \begin{bmatrix} 1 & c_{12} & c_{13} & \cdots & \cdots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \cdots & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & c_{n3} & \cdots & \cdots & c_{nn} \end{bmatrix}.$$

Using elementary row transformation, we get

$$D = \begin{bmatrix} 1 & d_{12} & d_{13} & \cdots & d_{1n} \\ 0 & & & & \\ \cdots & & & \mathbf{A}_1 & \\ \cdots & & & & \\ 0 & & & & \end{bmatrix},$$

where A_1 is $(n-1)\times(n-1)$ matrix. The matrix A_1 is non-singular otherwise $|A_1|=0$ and so $|D|=0$. Since $A\sim D$, this will imply $|A|=0$ contradicting the fact that A is non-singular. By induction hypothesis, A_1 can be transformed to I_{n-1} by elementary row transformations. Thus, we get a matrix M such that

$$M=\begin{bmatrix} 1 & d_{12} & d_{13} & \dots & \dots & d_{1n} \\ 0 & 1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & 1 \end{bmatrix}.$$

Further, use of elementary row transformation reduces M to the matrix.

$$I_n=\begin{bmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & \dots & 1 \end{bmatrix},$$

which completes the proof of the theorem.

Corollary 4.4 Let A be a on-singular matrix of order n. Then there exists elementary matrices $E_1, E_2, \dots, E_t$ such that

$$E_tE_{t-1}\dots E_2E_1A=I_n.$$

Proof: By the Theorem 4.28, non-singular matrix A can be reduced to I_n by finite number of elementary row transformations. Since elementary row transformation is equivalent to pre-multiplication by the elementary matrix, therefore, there exists elementary matrices $E_1, E_2, \dots, E_t$ such that $E_t, E_{t-1} \dots E_2 E_1 A = I_n$.

Corollary 4.5 Every non-singular matrix is a product of elementary matrices.

Proof: Let A be a non-singular matrix. Then, by Corollary 4.4, there exist elementary matrices $E_1, E_2,\dots, E_t$ such that

$$E_tE_{t-1}\dots E_2E_1A=I_n.$$

Pre-multiplying both sides by $(E_tE_{t-1}\dots E_2E_1)^{-1}$, we get

$$A=E_1^{-1}E_2^{-1}\dots E_t^{-1}.$$

Since, inverse of an elementary matrix is also an elementary matrix, it follows that non-singular matrix can be expressed as a product of elementary matrices.

Corollary 4.6 The rank of a matrix does not alter by pre-multiplication or post-multiplication with a non-singular matrix.

Proof: Every non-singular matrix can be expressed as a product of elementary matrices. Also we know that elementary row (column) transformations are equivalent to pre-(post) multiplication with the corresponding elementary matrices. But elementary transformations do not alter the rank of a matrix. Hence, the rank of a matrix remains unchanged by pre-multiplication or post-multiplication with a non-singular matrix.

4.18 ROW RANK AND COLUMN RANK OF A MATRIX

Definition 4.73 Let A be any $m\times n$ matrix. Then the maximum number of linearly independent rows (columns) of A is called the *row rank* (*column rank*) of A.

The following theorem (stated without proof) shall be used in the sequel.

Theorem 4.29 The row rank, the column rank and the rank of a matrix are equal.

4.19 SOLUTION OF SYSTEM OF LINEAR EQUATIONS

Let

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n&=b_1\\ a_{21}x_1+a_{22}x_2+\dots+a_{2n}x_n&=b_2\\ \dots\quad \dots\quad \dots\quad &\dots\quad \dots\\ \dots\quad \dots\quad \dots\quad &\dots\quad \dots\\ a_{m1}x_1+a_{m2}x_2+\dots+a_{mn}x_n&=b_m \end{aligned}\right\}\qquad(1)$$

be a system of m linear equations in n unknown $x_1, x_2,\dots, x_n$ The matrix form of this system is

$$AX=B,$$

where

$$\text{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & \dots & a_{mn} \end{bmatrix}$$

is called *coefficient matrix* of the system,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix}$$

is the *column matrix of unknowns*, and

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ \dots \\ b_m \end{bmatrix}$$

is *column matrix of known numbers* or the *matrix of constants*. We call the system (1) as the system of *non-homogenous equations*.

Any set of values of $x_1, x_2, \dots, x_n$ from a scalar field which simultaneously satisfy (1) is called a *solution*, over that field, of the system. When such a system has one or more solutions, it is said to be *consistent*, otherwise it is called *inconsistent.*

4.20 SOLUTION OF NON-HOMOGENOUS LINEAR SYSTEM OF EQUATIONS

(A) Matrix Inversion Method.

Consider the non-homogeneous system of linear equations $AX = B$, where A is non-singular $n \times n$ matrix. Since A is non-singular, A^{-1} exists. Pre-multiplication of $AX = B$ by A^{-1} yields

$$A^{-1}(AX) = A^{-1}B$$

or

$$\left(A^{-1}A\right)X = A^{-1}B$$

or

$$IX = A^{-1}B$$

or

$$X = A^{-1}B.$$

Thus if A is non-singular, then the given system of equation can be solved using inverse of A. This method is called the Matrix Inversion Method.

EXAMPLE 4.41

Solve

$$\begin{aligned} x + 2y - 3z &= -4 \\ 2x + 3y + 2z &= 2 \\ 3x - 3y - 4z &= 11 \end{aligned}$$

by Matrix Inversion Method.

Solution. The matrix form of the system is $AX = B$, where

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 3 & 2 \\ 3 & -3 & -4 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ and } B = \begin{bmatrix} -4 \\ 2 \\ 11 \end{bmatrix}.$$

We note that

$$|A| = 1(-6) - 2(-14) - 3(-15) = 67 \neq 0.$$

Thus A is non-singular. Hence the required solution is given by

$$X = A^{-1}B. \qquad (2)$$

The cofactor matrix of A is

$$[A_{ij}] = \begin{bmatrix} -6 & 14 & -15 \\ 17 & 5 & 9 \\ 13 & -8 & -1 \end{bmatrix}$$

and so

$$\text{adj } A = [A_{ij}]^T = \begin{bmatrix} -6 & 17 & 13 \\ 14 & 5 & -8 \\ -15 & 9 & -1 \end{bmatrix}.$$

Hence

$$A^{-1} = \frac{1}{|A|}\text{adj}\, A = \frac{1}{67}\begin{bmatrix} -6 & 17 & 13 \\ 14 & 5 & -8 \\ -15 & 9 & -1 \end{bmatrix}.$$

Substituting A^{-1} in (2), we get

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{67}\begin{bmatrix} -6 & 17 & 13 \\ 14 & 5 & -8 \\ -15 & 9 & -1 \end{bmatrix}\begin{bmatrix} -4 \\ 2 \\ 11 \end{bmatrix}$$

$$= \frac{1}{67}\begin{bmatrix} 201 \\ -134 \\ 67 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}.$$

Hence $x = 3$, $y = -2$, and $z = 1$.

B. Cramer's Rule.

If $|A| \neq 0$, then $AX = B$ has exactly one solution $x_j = \frac{|A_j|}{|A|}$, $j = 1, 2, \dots, n$, where A_j is the matrix obtained from A by replacing the jth column of A by the column of b's.

Consider the matrix form $AX = B$ of the system of linear equations. Again Suppose that A is non-singular. Then, pre-multiplication of $AX = B$ by A^{-1} yields

$$X = A^{-1}B = \frac{1}{|A|}\begin{bmatrix} A_{11} & A_{21} & \dots & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & \dots & A_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & \dots & A_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ b_n \end{bmatrix}$$

or

$$\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} b_1A_{11} + b_2A_{21} + \dots + b_nA_{n1} \\ b_1A_{12} + b_2A_{22} + \dots + b_nA_{n2} \\ \dots \\ \dots \\ b_1A_{1n} + b_2A_{2n} + \dots + b_nA_{nn} \end{bmatrix}.$$

Therefore,

$$x_1 = \frac{1}{|A|}(b_1A_{11} + b_2A_{21} + \dots + b_nA_{n1}) = \frac{|A_1|}{|A|}$$

$$x_2 = \frac{1}{|A|}(b_1A_{12} + b_2A_{22} + \dots + b_nA_{n2}) = \frac{|A_2|}{|A|}$$

$$\dots \quad \dots \quad \dots \quad \dots$$

$$\dots \quad \dots \quad \dots \quad \dots$$

$$x_n = \frac{1}{|A|}(b_nA_{1n} + b_nA_{2n} + \dots + b_nA_{nn}) = \frac{|A_n|}{|A|},$$

where A_j is the matrix obtained from A by replacing the jth column of A by the column of b's.

EXAMPLE 4.42

Solve the system of linear equations

$$3x + y + 2z = 3$$
$$2x - 3y - z = -3$$
$$x + 2y + z = 4$$

by Cramer's rule.

Solution. Let A be the coefficient matrix. Then

$$|A| = \begin{vmatrix} 3 & 1 & 2 \\ 2 & -3 & -1 \\ 1 & 2 & 1 \end{vmatrix} = 8$$

and so A is non-singular. Thus the Cramer's rule is applicable and we have

$$x = \frac{1}{|A|}\begin{vmatrix} 3 & 1 & 2 \\ -3 & -3 & -1 \\ 4 & 2 & 1 \end{vmatrix} = \frac{8}{8} = 1,$$

$$y = \frac{1}{|A|}\begin{vmatrix} 3 & 3 & 2 \\ 2 & -3 & -1 \\ 1 & 4 & 1 \end{vmatrix} = \frac{16}{8} = 2,$$

$$z = \frac{1}{|A|}\begin{vmatrix} 3 & 1 & 3 \\ 2 & -3 & -3 \\ 1 & 2 & 4 \end{vmatrix} = \frac{-8}{8} = -1.$$

Hence, the required solution is $x = 1, y = 2$, and $z = -1$.

Remark 4.5 The above two methods are applicable only when A is non-singular.

4.21 CONSISTENCY THEOREM

Definition 4.74 Let $AX = B$ be the matrix form of a given system of equations. Then the matrix

$$[A{:}B] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}$$

is called the *augmented matrix* of the given system of equations.

Definition 4.75 If a system of linear equations has one or more solution, it is said to be *consistent;* otherwise it is called *inconsistent.*

Theorem 4.30 (Consistency Theorem). The system of linear equations $AX = B$ is consistent if and only if the coefficient matrix A and the augmented matrix $[A{:}B]$ are of the same rank.

Proof: Let $C_1, C_2, \dots, C_n$ denote the column vectors of the matrix A. Then the equation $AX = B$ is equivalent to

$$[C_1C_2 \dots C_n]\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix} = B$$

or

$$x_1C_1 + x_2C_2 + \dots + x_nC_n = B. \qquad (3)$$

Let r be the rank of the matrix A. Then A has r linearly independent columns. Without loss of generality, we assume that $C_1, C_2, \dots, C_r$ form a linearly independent set and so each of the remaining $n - r$ columns is a linear combination of these r columns $C_1, C_2, \dots, C_r$

Suppose the given system of linear equations is consistent. Therefore, there exist n scalar $k_1, k_2, \dots, k_n$ such that

$$k_1C_1 + k_2C_2 + \dots + k_nC_n = B. \qquad (4)$$

Now since each of $n - r$ columns $C_{r+1}, C_{r+2}, \dots, C_n$ is a linear combination of first r columns $C_1, C_2, \dots, C_r$, it follows from (4) that B is also a linear combination of $C_1, C_2, \dots, C_r$ Thus, the maximum number of linearly independent columns of the matrix $[A{:}B]$ is also r. Therefore, the matrix $[A{:}B]$ is also of rank r. Hence, rank of A and the augmented matrix $[A:B]$ is the same.

Conversely, suppose that the matrices A and $[A:B]$ are of the same rank r. Then the maximum number of linearly independent columns of the matrix $[A:B]$ is r. But the first r columns $C_1, C_2, \ldots, C_r$ of the matrix $[A: B]$ had already formed a linearly independent set. Therefore, the column B should be expressed as a linear combination of $C_1, C_2, \ldots, C_r$ Hence, there are scalars $k_1, k_2, \ldots, k_r$, such that

$$k_1C_1 + k_2C_2 + \ldots + k_rC_r = B$$

or

$$k_1C_1 + k_2C_2 + \ldots + k_rC_r + 0C_{r+1}0C_{r+2} + \ldots + 0C_n = B \quad (5)$$

Comparing (3) and (5), we get

$$x_1 = k, x_2 = k_2, \ldots, x_r = k_r,\ x_{r+1} = 0,$$
$$x_{r+2} = 0, \ldots, x_r = 0$$

as the solution of the equation $AX=B$. Hence, the given system of linear equations is consistent. This completes the proof of the theorem.

If the system of linear equations is consistent, then the following cases arises:

Case I. $m \geq n$, that is, number of equations is more than the number of unknowns. In such a case

(i) if $\rho(A)=\rho([A:B])=n$, then the system of equations has a unique solution

(ii) if $\rho(A)=\rho([A:B])=r < n$ then the $(n - r)$ unknowns are assigned arbitrary values and the remaining r unknowns can be determined in terms of these $(n - r)$ unknowns.

Case II. $m < n$, that is, the number of equations is less than the number of unknowns. In such a case

(i) if $\rho(A)=\rho([A:B])=m$, then $n-m$ unknowns can be assigned arbitrary values and the values of the remaining m unknowns can be found in terms of these $n-m$ unknowns, which have already been assigned values

(ii) if $\rho(A)=\rho([A:B])=r < m$, then the $(n - r)$ unknowns can be assigned arbitrary values and the values of remaining r unknowns can be found in terms of these $(n - r)$ unknowns, which have already been assigned values.

EXAMPLE 4.43

Show that the system

$$x+y+z=-3$$
$$3x+y-2z=-2$$
$$2x+4y+7z=7$$

of linear equations is not consistent.

Solution. The matrix form of the system is

$$AX = B$$

and the augmented matrix is

$$[A:B] = \begin{bmatrix} 1 & 1 & 1 & -3 \\ 3 & 1 & -2 & -2 \\ 2 & 4 & 7 & 7 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & -3 \\ 0 & -2 & -5 & 7 \\ 0 & 2 & 5 & 13 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 3R_1 \\ R_3 \to R_3 - 2R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & -3 \\ 0 & -2 & -5 & 7 \\ 0 & 0 & 0 & 20 \end{bmatrix} R_3 \to R_3 + R_2$$

Thus the number of non-zero rows in Echelon form of the matrix $[A: B]$ is 3. But

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

and so $\rho(A)=2$.

Thus,

$$\rho(A) \neq \rho([A:B]).$$

Hence, the given system of equation is inconsistent.

EXAMPLE 4.44

Show that the equations

$$x+2y-z=3$$
$$3x-y+2z=1$$
$$2x-2y+3z=2$$
$$x-y+z=-1$$

are consistent. Also solve them.

Solution. In matrix form, we have

$$AX = \begin{bmatrix} 1 & 2 & -1 \\ 3 & -1 & 2 \\ 2 & -2 & 3 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \\ -1 \end{bmatrix}.$$

The augmented matrix is

$$[A:B] = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 3 & -1 & 2 & 1 \\ 2 & -2 & 3 & 2 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -7 & 5 & -8 \\ 0 & -6 & 5 & -4 \\ 0 & -3 & 2 & -4 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 3R_1 \\ R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 - R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -4 \\ 0 & -6 & 5 & -4 \\ 0 & -3 & 2 & -4 \end{bmatrix} R_2 \to R_2 - R_3$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -4 \\ 0 & 0 & 5 & 20 \\ 0 & 0 & 2 & 8 \end{bmatrix} \begin{matrix} R_3 \to R_3 - 6R_2 \\ R_4 \to R_4 - 3R_2 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} R_3 \to \frac{1}{5}R_3 \\ R_4 \to \frac{1}{2}R_4 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} R_4 \to R_4 - R_3$$

The number of non-zero rows in the echelon form is 3. Hence $\rho([A:B]) = 3$. Also

$$A \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Clearly, $\rho(A) = 3$. Thus, $\rho(A) = \rho([A, B])$ and so the given system is consistent. Further, $r = n = 3$. Therefore, the given system of equation has a unique solution. Rewriting the equation from the augmented matrix, we have

$$x + 2y - z = 3$$
$$-y = -4$$
$$z = 4$$

and so $x = -1, y = 4$ and $z = 4$ is the required solution.

EXAMPLE 4.45

For what values of λ and μ, the system of equations

$$x + y + z = 6$$
$$x + 2y + 3z = 10$$
$$x + 2y + \lambda z = \mu$$

has (i) no solution (ii) a unique solution, and (iii) an infinite number of solutions.

Solution. The matrix form of the given system is

$$AX = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & \lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \\ \mu \end{bmatrix} = B.$$

Therefore, the augmented matrix is

$$[A:B] = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 10 \\ 1 & 2 & \lambda & \mu \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 1 & \lambda - 1 & \mu - 6 \end{bmatrix} \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & \lambda - 3 & \mu - 10 \end{bmatrix} R_3 \to R_3 - R_2.$$

If $\lambda \neq 3$, then $\rho(A) = 3$ and $\rho([A:B]) = 3$. Hence, the given system of equations is consistent. Since $\rho(A)$ is equal to the number of unknowns, therefore, the given system of equations possesses a unique solution for any value of μ.

If $\lambda = 3$ and $\mu \neq 10$, then $\rho(A) = 2$ and $\rho([A:B]) = 3$. Therefore, the given system of equations is inconsistent and so has no solution.

If $\lambda = 3$ and $\mu = 10$ then $\rho(A) = \rho([A:B]) = 2$. Thus, the given system of equation is consistent. Further, $\rho(A)$ is less than the number of unknowns, therefore, in this case the given system of equations possesses an infinite number of solutions.

EXAMPLE 4.46

Determine the value of λ for which the system of equations

$$x_1 + x_2 + x_3 = 2$$
$$x_1 + 2x_2 + x_3 = -2$$
$$x_1 + x_2 + (\lambda - 5)x_3 = \lambda$$

(i) has no solution
(ii) has a unique solution.

Solution. The matrix form of the given system is

$$AX = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & \lambda - 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ \lambda \end{bmatrix} = B.$$

Therefore, the augmented matrix is

$$[A{:}B] = \begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & 2 & 1 & -2 \\ 1 & 1 & \lambda - 5 & \lambda \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & \lambda - 6 & \lambda - 2 \end{bmatrix} \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}.$$

If $\lambda = 6$, then $\rho(A) = 2$ and $\rho([A{:}B]) = 3$. Therefore, the system is inconsistent and so possesses no solution.

If $\lambda \neq 6$, then $\rho(A) = \rho([A{:}B]) = 3$. Hence, the system is consistent in this case. Since $\rho(A)$ is equal to the number of unknowns, the system has a unique solution in this case.

EXAMPLE 4.47

Determine the value of λ for which the system of equations

$$x + y + z = 1$$
$$x + 2y + 4z = \lambda$$
$$x + 4y + 10z = \lambda^2$$

possesses a solution and, hence, find its solution.

Solution. The given system of equations is expressed in the matrix form as

$$AX = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 10 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \end{bmatrix} = B.$$

Therefore, the augmented matrix is

$$[A{:}B] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & \lambda \\ 1 & 4 & 10 & \lambda^2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \lambda - 1 \\ 0 & 3 & 9 & \lambda^2 - 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \lambda - 1 \\ 0 & 0 & 0 & \lambda^2 - 3\lambda + 2 \end{bmatrix}.$$

We note that

$$\rho(A) = \rho([A{:}B]) \text{ if } \lambda^2 - 3\lambda + 2 = 0.$$

Thus, the given equation is consistent if $\lambda^2 - 3\lambda + 2 = 0$, that is if $(\lambda - 2)\,(\lambda - 1) = 0$, that is, if $\lambda = 2$ or $\lambda = 1$. If $\lambda = 2$, then we have

$$[A{:}B] \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and so the given system of equations is equivalent to

$$x + y + z = 1$$
$$y + 3z = 1.$$

These equations yields $y = 1 - 3z$, and $x = 2z$. Therefore, if $z = k$, an arbitrary constant, then x $2k$, $y = 1 - 3k$, and $z = k$ constitute the general solution of the given equation.

If $\lambda = 1$ then, we have

$$[A{:}B] \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and so the given system of equations is equivalent to

$$x + y + z = 1$$
$$y + 3z = 0.$$

These equations yields $y = -3z$, $x = 1 + 2z$. Thus, if c is an arbitrary constant, then $x = 1 + 2c$, $y = -3c$, and $z = c$, constitute the general solution of the given system of equations.

EXAMPLE 4.48

Find the value of λ and μ for which the system of equations

$$3x + 2y + z = 6$$
$$3x + 4y + 3z = \mu$$
$$6x + 10y + \lambda z = \mu$$

has (i) unique solution, (ii) no solution, and (iii) infinite number of solutions.

Solution. The given system of equations is expressed by the matrix equation

$$AX = \begin{bmatrix} 3 & 2 & 1 \\ 3 & 4 & 3 \\ 6 & 10 & \lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \\ \mu \end{bmatrix} = B.$$

Therefore, the augmented matrix is

$$[A{:}B] = \begin{bmatrix} 3 & 2 & 1 & 6 \\ 3 & 4 & 3 & 14 \\ 6 & 10 & \lambda & \mu \end{bmatrix}$$
$$\sim \begin{bmatrix} 3 & 2 & 1 & 6 \\ 0 & 2 & 2 & 8 \\ 0 & 6 & \lambda-2 & \mu-12 \end{bmatrix} \quad \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - 2R_1 \end{matrix}$$
$$\sim \begin{bmatrix} 3 & 2 & 1 & 6 \\ 0 & 2 & 2 & 8 \\ 0 & 0 & \lambda-8 & \mu-36 \end{bmatrix} \quad R_3 \to R_3 - 3R_2.$$

If $\lambda \neq 8$, then $\rho(A) = \rho([A{:}B]) = 3$ and so in this case the system is consistent. Further, since $\rho(A)$ is equal to number of unknowns, the given system has a unique solution.

If $\lambda = 8$, $\mu \neq 36$, then $\rho(A) = 2$ and $\rho([A{:}B]) = 3$. Hence, the system is inconsistent and has no solution.

If $\lambda = 8$, $\mu = 36$, then $\rho(A) = \rho([A{:}B]) = 2$. Therefore, the given system of equation is consistent. Since rank of A is less than the number of unknowns, the given system of equation has infinitely many solutions.

EXAMPLE 4.49

Using consistency theorem, solve the equation

$$x + y + z = 9$$
$$2x + 5y + 7z = 52$$
$$2x + y - z = 0.$$

Solution. The matrix form of the given system of equations is

$$AX = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & 7 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ 52 \\ 0 \end{bmatrix} = B.$$

Therefore, the augmented matrix is

$$[A{:}B] = \begin{bmatrix} 1 & 1 & 1 & 9 \\ 2 & 5 & 7 & 52 \\ 2 & 1 & -1 & 0 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 1 & 1 & 9 \\ 0 & 3 & 5 & 34 \\ 0 & -1 & -3 & -18 \end{bmatrix} \quad \begin{matrix} R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 2R_1 \end{matrix}$$
$$\sim \begin{bmatrix} 1 & 1 & 1 & 9 \\ 0 & -1 & -3 & -18 \\ 0 & 3 & 5 & 34 \end{bmatrix} \quad R_2 \leftrightarrow R_3$$
$$\sim \begin{bmatrix} 1 & 1 & 1 & 9 \\ 0 & -1 & -3 & -18 \\ 0 & 0 & -4 & -20 \end{bmatrix} \quad R_3 \to R_3 - 3R_2.$$

Thus we get echelon form of the matrix $[A{:}B]$. The number of non-zero rows in this form is 3. Therefore $\rho([A{:}B]) = 3$. Further, since

$$A \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -3 \\ 0 & 0 & -4 \end{bmatrix}.$$

Therefore $\rho(A) = 3$. Hence, $\rho(A) = \rho([A{:}B]) = 3$. This shows that the given system of equations is consistent. Also, since $\rho(A)$ is equal to the number of unknowns, the solution of the given system is unique. To find the solution, we note that the given system of equation is equivalent to

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -3 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ -18 \\ -20 \end{bmatrix}$$

and so

$$x + y + z = 9, \; -y - 3z = -18, \; -4z = -20,$$

which yields $z = 5, y = 3,$ and $x = 1$ as the required solution.

4.22 HOMOGENEOUS LINEAR EQUATIONS

Consider the following system of m homogeneous equations in n unknowns $x_1, x_2, \ldots, x_n$:

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0$$
$$\ldots$$
$$\ldots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0.$$

The matrix form of this system is

$$AX = 0,$$

where

$$\begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & \dots & a_{mn} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix}, \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ \dots \\ \dots \\ 0 \end{bmatrix}.$$

It is evident that $x_1 = 0, x_2 = 0, \dots, x_n = 0$, that is, $X = 0$ is a solution of the given system of equations. This solution is called *trivial solution* of the given system.

Let X_1 and X_2 be two solution of $AX = \mathbf{0}$. Then $AX_1 = \mathbf{0}$ and $AX_2 = \mathbf{0}$ and so for arbitrary numbers k_1, k_2, we have

$$A(k_1X_1 + k_2X_2) = k_1(AX_1) + k_2(AX_2)$$
$$= \mathrm{k}_1 0 + \mathrm{k}_2 0 = 0.$$

If follows, therefore, that linear combination of two solutions of $AX = \mathbf{0}$ is also a solution. Hence, the collection of all solutions of the equation $AX = \mathbf{0}$ form a subspace of the vector space V_n.

Theorem 4.31 Let the rank of a matrix A be r. Then the number of linearly independent solutions of m homogeneous linear equations in n variables, $AX = \mathbf{0}$ is $(n - r)$.

Solution. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & \dots & a_{mn} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix}$$

Since $\rho(A) = r$, it has r linearly independent columns. Without loss of generality, suppose that the first r columns of the matrix A are linearly independent. We write

$$A = [C_1 \; C_2 \dots C_n],$$

where $C_1, C_2, \dots, C_n$ are column vectors of A. Therefore, $AX = \mathbf{0}$ can be written as vector equation.

$$x_1C_1 + x_2C_2 + \dots + x_rC_r$$
$$+ x_{r+1}C_{r+1} + \dots + x_nc_n = 0. \qquad (6)$$

Since each of the vector $C_{r+1}, C_{r+2}, \dots, C_n$ is a linear combination of vectors $C_1, C_2, \dots, C_r$, therefore

$$C_{r+1} = p_{11}C_1 + p_{12}C_2 + \dots + p_{1r}C_r,$$
$$C_{r+2} = p_{21}C_1 + p_{22}C_2 + \dots + p_{2r}C_r,$$
$$\dots$$
$$\dots$$
$$C_n = p_{k1}C_1 + p_{k2}C_2 + \dots + p_{kr}C_r, \qquad (7)$$

where $k = n - r$. The expression (7) can be written as

$$p_{11}C_1 + p_{12}C_2 + \dots + p_{1r}C_r - 1.C_{r+1} + 0C_{r+2} + \dots + 0C_n = 0$$
$$p_{21}C_1 + p_{22}C_2 + \dots + p_{2r}C_r + 0.C_{r+1} - 1C_{r+2} + \dots + 0C_n = 0$$
$$\dots$$
$$\dots$$
$$p_{k1}C_1 + p_{k2}C_2 + \dots + p_{kr}C_r + 0.C_{r+1} + 0C_{r+2} + \dots - 1C_n = 0. \qquad (8)$$

Comparing (6) and (8), we note that

$$X_1 = \begin{bmatrix} p_{11} \\ p_{12} \\ \dots \\ \dots \\ p_{1r} \\ -1 \\ 0 \\ \dots \\ \dots \\ 0 \end{bmatrix}, X_2 = \begin{bmatrix} p_{21} \\ p_{22} \\ \dots \\ \dots \\ p_{2r} \\ 0 \\ -1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \dots, X_{n-r} = \begin{bmatrix} p_{k1} \\ p_{k2} \\ \dots \\ \dots \\ p_{kr} \\ 0 \\ 0 \\ \dots \\ 0 \\ -1 \end{bmatrix}$$

are $(n - r)$ solutions of $AX = \mathbf{0}$. Suppose now that

$$c_1X_1 + c_2X_2 + \dots + c_{n-r}X_{n-r} = \mathbf{0}.$$

Comparing $(r+1)$th, $(r+2)$th,…, nth component on both sides, we get

$$-c_1 = 0, -c_2 = 0, \dots, c_{n-r} = 0.$$

Hence $X_1, X_2, \dots, X_{n-r}$ are linearly independent. Suppose that the vector X, with components $x_1, x_2, \dots, x_n$ is any solution of the equation $AX = \mathbf{0}$. We assert that X is linear combination of $x_1, x_2, \dots, x_{n-r}$. To prove it, we note that the vector

$$X + x_{r+1}X_1 + x_{r+2}X_2 + \dots + x_nX_{n-r} \qquad (9)$$

being linear combination of solutions is also a solution. Then the last $n - r$ components of the vector (9) are all zero. Let $z_1, z_2, \dots, z_r$ be the first r components of the vector (28). Then the vector whose components are $(z_1, z_2, \dots, z_r, 0, 0, \dots, 0)$ is a solution of the equation $AX = \mathbf{0}$. Therefore from (6), we have

$$z_1C_1 + z_2C_2 + \dots + z_rC_r = 0.$$

But the vector $C_1, C_2, \dots, C_r$ are linearly independent. Hence $z_1 = z_2 = z_r = 0$. Hence (9) is a zero vector, that is,

$$X + x_{r+1}X_1 + x_{r+2}X_2 + \ldots + x_nX_{n-r} = 0$$

or

$$X = -x_{r+1}X_1 - x_{r+2}X_2 - \ldots - X_nX_{n-r}.$$

Thus, every solution is a linear combination of the $n-r$ linearly independent solution $X_1, X_2, \ldots, X_{n-r}$. It follows, therefore, that the set of solution $[X_1, X_2, \ldots, X_{n-r}]$ form a basis of vector space of all the solutions of the system of equations $AX = \mathbf{0}$.

Remark 4.6 Suppose we have a system of m linear equations in n unknowns. Thus, the coefficient matrix A is of order $m \times n$. Let r be the rank of A. Then, $r \leq n$ (number of column of A).

If $r = n$, then $AX = \mathbf{0}$ possesses $n - n = 0$ number of independent solutions. In this case, we have simply the trivial solution (which forms a linearly dependent system).

If $r < n$, then there are $n - r$ linearly independent solutions. Further any linear combination of these solutions will also be a solution of $AX = \mathbf{0}$. Hence, in this case, the equation $AX = \mathbf{0}$ has infinite number of solutions.

If $m < n$, then since $r \leq m$, we have $r < n$. Hence the system has a non-zero solution. The number of solutions of the equation $AX = \mathbf{0}$ will be infinite.

Theorem 4.32 A necessary and sufficient condition that a system of n homogeneous linear equations in n unknowns have non-trivial solutions is that coefficient matrix be singular.

Proof:

The condition is necessary. Suppose that the system of n homogeneous linear equations in n unknowns have a non-trivial solution. We want to show that $|A| = 0$. Suppose, on the contrary, $|A| \neq 0$. Then rank of A is n. Therefore, number of linearly independent solution is $n - n = 0$. Thus, the given system possesses no linearly independent solution. Thus, only trivial solution exists for the given system. This contradicts the fact that the given system of equation has non-trivial solution. Hence $|A| = 0$.

The condition is sufficient. Suppose $|A| = 0$. Therefore, $\rho(A) < n$. Let r be the rank of A. Then the given equation has $(n - r)$ linearly independent solutions. Since a linearly independent solution can never be zero, therefore, the given system must have a non-zero solution.

EXAMPLE 4.50

Solve

$$\begin{aligned} x + 3y - 2z &= 0 \\ 2x - y + 4z &= 0 \\ x - 11y + 14z &= 0. \end{aligned}$$

Solution. The matrix form of the given system of homogeneous equations is $AX = \mathbf{0}$, where

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & -1 & 4 \\ 1 & -11 & 14 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We note that

$$|A| = \begin{vmatrix} 1 & 3 & -2 \\ 2 & -1 & 4 \\ 1 & -11 & 14 \end{vmatrix} = 30 - 72 + 42 = 0.$$

Therefore A is singular, that is $\rho(A) < n$. Thus, the given system has a non-trivial solution and will have infinite number of solutions.

The given system is

$$\begin{bmatrix} 1 & 3 & -2 \\ 2 & -1 & 4 \\ 1 & -11 & 14 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

$$\sim \begin{bmatrix} 1 & 3 & -2 \\ 0 & -7 & 8 \\ 0 & -14 & 16 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0, \quad \begin{matrix} R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - R_1 \end{matrix},$$

$$\sim \begin{bmatrix} 1 & 3 & -2 \\ 0 & -7 & 8 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0, \ R_3 \to R_3 - 2R_2$$

and so we have

$$\begin{aligned} x + 3y - 2z &= 0 \\ -7y + 8z &= 0. \end{aligned}$$

These equations yield $y = \frac{8}{7}z, x = \frac{-10}{7}z$. Giving different values to z, we get infinite number of solutions.

EXAMPLE 4.51

Solve

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ x_1 + 2x_2 - x_3 &= 0 \\ 2x_1 + x_2 + 3x_3 &= 0. \end{aligned}$$

Solution. In matrix form, we have $AX = \mathbf{0}$, where

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 2 & -1 \\ 2 & 1 & 3 \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \ \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We note that $|A| = 9 \neq 0$. Thus A is non-singular. Hence, the given system of homogeneous equation has only trivial solution $x_1 = x_2 = x_3 = 0$.

EXAMPLE 4.52

Solve

$$\begin{aligned} 2x - 2y + 5z + 3w &= 0 \\ 4x - y + z + w &= 0 \\ 3x - 2y + 3z + 4w &= 0 \\ x - 3y + 7z + 6w &= 0. \end{aligned}$$

Solution. The matrix form of the given system of homogeneous equation is

$$AX = \begin{bmatrix} 2 & -2 & 5 & 3 \\ 4 & -1 & 1 & 1 \\ 3 & -2 & 3 & 4 \\ 1 & -3 & 7 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = 0.$$

Performing row elementary transformations to get echelon form of A, we have

$$A = \begin{bmatrix} 2 & -2 & 5 & 3 \\ 4 & -1 & 1 & 1 \\ 3 & -2 & 3 & 4 \\ 1 & -3 & 7 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 7 & 6 \\ 4 & -1 & 1 & 1 \\ 3 & -2 & 3 & 4 \\ 2 & -2 & 5 & 3 \end{bmatrix} R_1 \leftrightarrow R_4$$

$$\sim \begin{bmatrix} 1 & -3 & 7 & 6 \\ 4 & 11 & -27 & -23 \\ 0 & 7 & -18 & -14 \\ 0 & 4 & -9 & -9 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 4R_1 \\ R_3 \to R_3 - 3R_1 \\ R_4 \to R_4 - 2R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & -3 & 7 & 6 \\ 0 & 4 & -9 & -9 \\ 0 & 7 & -18 & -14 \\ 0 & 4 & -9 & -9 \end{bmatrix} R_2 \to R_2 - R_3$$

$$\sim \begin{bmatrix} 1 & -3 & 7 & 6 \\ 0 & 4 & -9 & -9 \\ 0 & 28 & -72 & -56 \\ 0 & 4 & -9 & -9 \end{bmatrix} R_3 \to 4R_3$$

$$\sim \begin{bmatrix} 1 & -3 & 7 & 6 \\ 0 & 4 & -9 & -9 \\ 0 & 0 & -9 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} R_3 \to R_3 - 7R_2 \\ R_4 \to R_4 - R_2 \end{matrix}.$$

The above echelon form of A suggests that rank of A is equal to the number of non-zero rows. Thus $\rho(A) = 3$. The number of unknowns is 4. Thus $\rho(A) < n$. Hence, the given system possesses non-trivial solution. The number of independent solution will be $(n - r) = 4 - 3 = 1$.

Further, the given system is equivalent to

$$\begin{bmatrix} 1 & -3 & 7 & 6 \\ 0 & 4 & -9 & -9 \\ 0 & 0 & -9 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = 0$$

and so, we have

$$\begin{aligned} x - 3y + 7z + 6w &= 0 \\ 4y - 9z - 9w &= 0 \\ -9z + 7w &= 0 \end{aligned}$$

These equations yield $z = \frac{7}{9}w$, $y = 4w$, $x = \frac{5}{9}w$. Thus taking $w = t$, we get $x = \frac{5}{9}t$, $y = 4t$, $z = \frac{7}{9}t$, $w = t$ as the general solution of the given equations.

EXAMPLE 4.53

Determine the value of λ for which the following equations have non-zero solutions:

$$\begin{aligned} x + 2y + 3z &= \lambda x \\ 3x + y + 2z &= \lambda y \\ 2x + 3y + z &= \lambda z. \end{aligned}$$

Solution. The matrix form of the given equation is

$$AX = \begin{bmatrix} 1-\lambda & 2 & 3 \\ 3 & 1-\lambda & 2 \\ 2 & 3 & 1-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ x \end{bmatrix} = 0.$$

The given system will have non-zero solution only if $|\ A\ | = 0$, that is, if rank of A is less than 3. Thus for the existence of non-zero solution, we must have

$$\begin{vmatrix} 1-\lambda & 2 & 3 \\ 3 & 1-\lambda & 2 \\ 2 & 3 & 1-\lambda \end{vmatrix} = 0$$

or

$$\begin{vmatrix} 6-\lambda & 6-\lambda & 6-\lambda \\ 3 & 1-\lambda & 2 \\ 2 & 3 & 1-\lambda \end{vmatrix} = 0 \text{ using } R_1 \to R_1 + R_2 + R_3$$

or

$$(6-\lambda) \begin{vmatrix} 1 & 1 & 1 \\ 3 & 1-\lambda & 2 \\ 2 & 3 & 1-\lambda \end{vmatrix} = 0$$

or

$$(6-\lambda)\begin{vmatrix} 1 & 0 & 0 \\ 3 & -2-\lambda & -1 \\ 2 & 1 & -1-\lambda \end{vmatrix} = 0, \quad \begin{array}{l} C_2 \to C_2 - C_1 \\ C_3 \to C_3 - C_1 \end{array}$$

or

$$(6-\lambda)[\lambda^2 + 3\lambda + 3] = 0,$$

which yields

$$\lambda = 6 \text{ and } \frac{-3 \pm \sqrt{9-12}}{2}.$$

Thus, the only real value of λ for which the given system of equation has a solution is 6.

4.23 CHARACTERISTIC ROOTS AND CHARACTERISTIC VECTORS

Let A be a matrix of order n, λ a scalar and $X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$ a column vector.

Consider the equation

$$AX = \lambda X \tag{10}$$

Clearly $X = \mathbf{0}$ is a solution of (10) for any value of λ. The question arises whether there exist scalar λ and non-zero vector X, which simultaneously satisfy the equation (10). This problem is known as *characteristic value problem*. If I_n is unit matrix of order n, then (10) may be written in the form

$$(A - \lambda I_n)X = \mathbf{0}. \tag{11}$$

The equation (11) is the matrix form of a system of n homogeneous linear equations in n unknowns. This system will have a non-trivial solution if and only if the determinant of the coefficient matrix $A - \lambda I_n$ vanishes, that is, if

$$|A - \lambda I_n| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots\ \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots\ \dots & a_{2n} \\ \dots & \dots & \dots\ \dots & \dots \\ \dots & \dots & \dots\ \dots & \dots \\ a_{n1} & a_{n2} & \dots\ \dots & a_{nn}-\lambda \end{vmatrix} = 0.$$

The expansion of this determinant yields a polynomial of degree n in λ, which is called the *characteristic polynomial* of the matrix A.

The equation $|A - \lambda I_n| = 0$ is called the *characteristic equation* or *secular equation* of A.

The n roots of the characteristic equation of a matrix A of an order n are called the *characteristic roots, characteristic values, proper values, eigenvalues,* or *latent roots* of the matrix A.

The set of the eigenvalues of a matrix A is called the *spectrum* of A.

If λ is an eigenvalue of a matrix A of order n, then a non-zero vector X such that $AX = \lambda X$ is called a *characteristic vector, eigen vector, proper vector,* or *latent vector* of A corresponding to the characteristic root λ.

Theorem 4.33 The equation $AX = \lambda X$ has a non-trivial solution if and only if λ is a characteristic root of A.

Proof: Suppose first that λ is a characteristic root of the matrix A. Then $|A - \lambda I_n| = 0$ and consequently the matrix $A - \lambda I$ is singular. Therefore, the matrix equation $(A - \lambda I)X = \mathbf{0}$ possesses a non-zero solution. Hence, there exists a non-zero vector X such that $(A - \lambda I)X = \mathbf{0}$ or $AX = \lambda X$.

Conversely, suppose that there exists a non-zero vector X such that $AX = \lambda X$ or $(A - \lambda I)X = \mathbf{0}$. Thus, the matrix equation $(A - \lambda I)X = \mathbf{0}$ has a non-zero solution. Hence $A - \lambda I$ is singular and so $|A - \lambda I| = 0$. Hence, λ is a characteristic root of the matrix A.

Theorem 4.34 Corresponding to a characteristics value λ, there correspond more than one characteristic vectors.

Proof: Let X be a characteristic vector corresponding to a characteristic root λ. Then, by definition, $X \neq 0$ and $AX = \lambda X$. If k is any non-zero scalar, then $kX \neq 0$. Further,

$$A(kX) = k(AX) = k(\lambda X) = \lambda(kX).$$

Therefore, kX is also a characteristic vector of A corresponding to the characteristic root λ.

Theorem 4.35 If X is a proper vector of a matrix A, then X cannot correspond to more than one characteristic root of A.

Proof: Suppose, on the contrary, X be a characteristic vector of a matrix A corresponding to two characteristic roots λ_1 and λ_2. Then, $AX = \lambda_1 X$ and $AX = \lambda_2 X$ and so $(\lambda_1 - \lambda_2) X = 0$. Since $X \neq 0$, this implies $\lambda_1 - \lambda_2 = 0$ or $\lambda_1 = \lambda_2$. Hence the result follows.

Theorem 4.36 Let $X_1, X_2, \dots, X_n$ be non-zero characteristic vectors associated with distinct characteristic roots $\lambda_1, \lambda_2, \dots \lambda_n$ of a matrix A. Then $X_1, X_2, \dots, X_n$ are linearly independent.

Proof: Let $c_1, c_2, \ldots, c_n$ the constants such that

$$c_1X_1 + c_2X_2 + \ldots + c_nX_n = 0 \qquad (12)$$

Multiplying throughout by A and using the fact that $AX_i = \lambda_i X_i$, we get

$$c_1\lambda_1X_1 + c_2\lambda_2X_2 + \ldots + c_n\lambda_nX_n = 0 \qquad (13)$$

Repeating this process, we obtain successively

$$\left.\begin{array}{c} c_1\lambda_1^2X_1 + c_2\lambda_2^2X_2 + \ldots + c_n\lambda_n^2X_n = 0 \\ c_1\lambda_1^3X_1 + c_2\lambda_2^3X_2 + \ldots + c_n\lambda_n^3X_n = 0 \\ \cdots \\ \cdots \\ c_1\lambda_1^{k-1}X_1 + c_2\lambda_2^{k-1}X_2 + \ldots + c_n\lambda_n^{k-1}X_n = 0 \end{array}\right\} \qquad (14)$$

The k equations (12) through (14) in vector unknowns $X_1, X_2, \ldots, X_n$ can be written in the form

$$[c_1X_1\, c_2X_2 \ldots \ldots c_nX_n]\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{k-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \lambda_n & \lambda_n^2 & \ldots & \lambda_n^{k-1} \end{bmatrix} = 0.$$

Since $\lambda_1, \lambda_2, \ldots, \lambda_n$ are distinct, the right factor is a *non-singular Vander-monde matrix*. Since it is non-singular, its inverse exists. Post-multiplication by its inverse yields

$$[c_1X_1\; c_2X_2 \;\ldots\;\ldots\; c_nX_n] = 0$$

Since $X_1, X_2, \ldots, X_n$ are all non-zero, it follows that $c_1 = c_2 = c_n = 0$. Thus, the relation (12) implies c_1 $_1 = c_2 = c_n = 0$. Hence, $X_1, X_2, \ldots, X_n$ are linearly independent.

Let $\phi(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n$ be the characteristic polynomial of a matrix A. Thus

$$\begin{aligned} |A - \lambda I| &= \phi(\lambda) \\ &= a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n. \end{aligned}$$

If we put $\lambda = 0$, then we get $|A| = a_n$. The diagonal term of $|A - \lambda I|$ is $(a_{11} - \lambda)(a_{22} - \lambda)\ldots(a_{nn} - \lambda)$ and this is the only product yielding λ^n and λ^{n-1}. Expanding the product, we obtain $(-1)^n\ \lambda^n$ and $(-1)^{n-1}\sum a_{ii}\lambda^{n-1}$ as the first two terms of $\phi(\lambda)$. Hence

$$a_0 = (-1)^n \text{ and } a_1 = (-1)^{n-1}\sum a_{ii}$$

In $\phi(\lambda)$, the coefficient of λ^{n-1}, namely $(-1)^{n-1}(a_{11} + a_{22} + \ldots + a_{nn})$ is of special interest. As we know, the term $a_{11} + a_{22} + \ldots + a_{nn}$ is called the *trace* or *spur* of the matrix A. It follows from the above discussion that the *sum of the eigenvalues of a matrix is equal to its trace* and the product of the eigenvalues of a matrix A is its determinant | A |.

Theorem 4.37 The characteristic roots of a Hermitian matrix are real.

Proof: Let λ be a characteristic root of a Hermitian matrix. Then there exists a non-zero vector X such that

$$AX = \lambda X. \qquad (15)$$

Taking transpose conjugate, we get

$$X^\theta A^\theta = \bar{\lambda}X^\theta. \qquad (16)$$

Pre-multiplying (15) by X^θ and post-multiplying (16) by X, we get

$$X^\theta AX = \lambda X^\theta X, \text{ and} \qquad (17)$$

$$X^\theta A^\theta X = \bar{\lambda}X^\theta X. \qquad (18)$$

Since A is Hermitian, $A^\theta = A$ and, therefore, (17) and (18) imply

$$\lambda X^\theta X = \bar{\lambda}X^\theta X \Rightarrow (\lambda - \bar{\lambda})X^\theta X = 0.$$

Since $X \neq \mathbf{0}$, we have $\lambda - \bar{\lambda} = 0$ and so $\lambda = \bar{\lambda}$. Hence λ is real.

Corollary 4.7 The characteristic roots of a real symmetric matrix are all real.

Proof: Since a real symmetric matrix is Hermitian, it follows from Theorem 4.37 that the characteristic roots of a real symmetric matrix are all real.

Corollary 4.8 The characteristic roots of a Skew-Hermitian matrix are either pure imaginary or zero.

Proof: Let A be a Skew-Hermitian matrix. Then iA is Hermitian. Let λ be the characteristic root of A. Then $AX = \lambda X$, $X \neq \mathbf{0}$ or $(iA)X = (i\lambda)X$. Thus $i\lambda$ is a characteristic root of iA. But iA is Hermitian and characteristic roots of Hermitian matrix are real. Thus $i\lambda$ is real, which is possible only if λ is zero or pure imaginary. This proves the result.

Corollary 4.9 The characteristic roots of a skew symmetric matrix are either pure imaginary or zero.

Proof: Since a Skew-Symmetric matrix is Skew-Hermitian, the result follows from corollary 4.8.

EXAMPLE 4.54

Find the characteristic vectors of the matrix

$$A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}.$$

Solution. The characteristic equation of the given matrix is

$$|A-\lambda I| = \begin{vmatrix} 3-\lambda & 1 & 0 \\ 0 & 3-\lambda & 1 \\ 0 & 0 & 3-\lambda \end{vmatrix} = 0,$$

which yields $(3-\lambda)^3=0$. Thus 3 is the only distinct characteristic root of A. The characteristic vectors are given by non-zero solutions of the equation $(A-3I)X=0$, that is,

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The coefficient matrix of the equation is of rank 2. Therefore, number of linearly independent solution is $n-r=1$. The above equation yields $x_2=0$, $x_3=0$. Therefore, $x_1=1$, $x_2=0$, $x_3=0$ is a non-zero solution of the above equation. Thus,

$$X = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

is an eigenvector of A corresponding to the eigenvalue 3. Also any non-zero multiple of this vector shall be an eigenvector of A corresponding to $\lambda=3$.

EXAMPLE 4.55

Find the eigenvalues and the corresponding eigenvectors of the matrix

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}.$$

Solution. The characteristic equation of the given matrix is

$$|A-\lambda I| = \begin{vmatrix} 6-\lambda & -2 & 2 \\ -2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{vmatrix} = 0$$

or

$$\begin{vmatrix} 6-\lambda & -2 & 0 \\ -2 & 3-\lambda & 2-\lambda \\ 2 & -1 & 2-\lambda \end{vmatrix} = 0, \quad C_3 \to C_3 + C_2$$

or

$$(2-\lambda)\begin{vmatrix} 6-\lambda & -2 & 0 \\ -2 & 3-\lambda & 1 \\ 2 & -1 & 1 \end{vmatrix} = 0$$

or

$$(2-\lambda)(\lambda-2)(\lambda-8) = 0.$$

Thus, the characteristic roots of A are $\lambda=2, 2, 8$. The eigenvector of A corresponding to the eigenvalue 2 is given by $(A-2I)X=\mathbf{0}$ or

$$\begin{bmatrix} 4 & -2 & 2 \\ -2 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$\begin{bmatrix} -2 & 1 & -1 \\ 4 & -2 & 2 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad R_1 \leftrightarrow R_2$$

or

$$\begin{bmatrix} -2 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{matrix} R_2 \to R_2 + 2R_1 \\ R_3 \to R_3 + R_1 \end{matrix}$$

The coefficient matrix is of rank 1. Therefore, there are $n-r=3-1=2$ linearly independent solution. The above equation is

$$-2x_1 + x_2 - x_3 = 0.$$

Clearly,

$$X_1 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \quad \text{and} \quad X_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

are two linearly independent solutions of this equation. Then X_1 and X_2 are two linearly independent eigenvectors of A corresponding to eigenvalue 2. If k_1, k_2 are scalars not both equal to zero, then $k_1X_1 + k_2X_2$ yields all the eigenvectors of A corresponding to the eigenvalue 2.

The characteristic vectors of A corresponding to the characteristic root 8 are given by $(A-8I)X=\mathbf{0}$ or by

$$\begin{bmatrix} 6-8 & -2 & 2 \\ -2 & -8 & -1 \\ 2 & -1 & 3-8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} -2 & -2 & 2 \\ 0 & -3 & -3 \\ 0 & -3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 + R_1 \end{matrix}$$

$$\sim \begin{bmatrix} -2 & -2 & 2 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad R_3 \to R_3 - R_2$$

The coefficient matrix is of rank 2. Therefore, number of linearly independent solution is $n-r=3$ $2=1$. The above equations give

$$-2x_1 - 2x_2 + 2x_3 = 0$$

$$-3x_2 - 3x_3 = 0.$$

Hence $x_2 = -x_3$. Taking $x_2 = -1$, $x_3 = 1$, we get $x_1 = 2$. Therefore $X_3 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

is an eigenvector of A corresponding to $\lambda = 8$. Further, every non-zero multiple of X_3 is an eigenvector of A corresponding to the eigenvalue 8.

EXAMPLE 4.56

If A is non-singular, show that the eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A.

Solution. Let λ be a characteristic root of the matrix A. Therefore, there exists non-zero vector X such that

$$AX = \lambda X$$
$$\Rightarrow \; A^{-1}AX = \lambda A^{-1}X$$
$$\Rightarrow \; \frac{1}{\lambda}X = A^{-1}X.$$

Hence $\frac{1}{\lambda}$ is a characteristic root of A^{-1} and X is the corresponding characteristic vector.

EXAMPLE 4.57

Show that the characteristic roots of a triangular matrix are just the diagonal elements of the matrix.

Solution. Let

$$\begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ 0 & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & a_{nn} \end{bmatrix}$$

be a triangular matrix of order n. Then

$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots \dots & a_{1n} \\ 0 & a_{22} - \lambda & \dots \dots & a_{2n} \\ \dots & \dots & \dots \dots & \dots \\ \dots & \dots & \dots \dots & \dots \\ 0 & 0 & \dots \dots & a_{nn} - \lambda \end{vmatrix}$$
$$= (a_{11} - \lambda)(a_{22} - \lambda)\dots(a_{nn} - \lambda).$$

Hence, the roots of the characteristic equation $|A - \lambda I| = 0$ are $a_{11}, a_{22}, \dots, a_{nn}$ which are the diagonal element of A.

EXAMPLE 4.58

Show that 0 is an eigenvalue of a matrix A if and only if A is singular.

Solution. If $\lambda = 0$ is an eigenvalue, it satisfies the characteristic equation $|A - \lambda I| = 0$ and so we have $|A| = 0$. Thus A is singular. Conversely if A is singular, then $|A| = 0$. Thus $\lambda = 0$ satisfy the equation $|A - \lambda I| = 0$ and so it is an eigenvalue.

4.24 THE CAYLEY-HAMILTON THEOREM

Let

$$\phi(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n$$

be the characteristic polynomial of a matrix A. Then

$$\phi(A) = a_0A^n + a_1A^{n-1} + \dots + a_{n-1}A + a_nI_n,$$
$$a_0 = (-1)^n$$

is called the *characteristic function* of the matrix A. Concerning this function, we have the following famous theorem.

Theorem 4.38 (Cayley-Hamilton Theorem). Every square matrix A satisfies its characteristic equation $\phi(A) = 0$.

Proof: The characteristic matrix of A is $A - \lambda I_n$. Since the elements of $A - \lambda I_n$ are at most of the first degree in λ, the elements, (cofactor) of the adjoint matrix of $A - \lambda I_n$ are of degree utmost $n-1$ in λ. Therefore, we may represent adj $(A - \lambda I_n)$ as a matrix polynomial

$$\text{adj}\,(A - \lambda I_n) = B_0\lambda^{n-1} + B_1\lambda^{n-2} + \dots + B_{n-2}\lambda + B_{n-1},$$

where B_k is the matrix whose elements are the coefficients of λ^k in the corresponding elements of adj $(A - \lambda I_n)$. But

$$(A - \lambda I_n) \;\text{adj}\; (A - \lambda I_n) = |A - \lambda I_n|\, I_n,$$

that is,

$$A \;\text{adj}\; (A - \lambda I_n) - \lambda \;\text{adj}\; (A - \lambda I_n) = \phi(\lambda)\, I_n.$$

Substituting the expansion of adj $(A - \lambda I_n)$ from above and $\phi(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n$, we get

$$A\left(B_0\lambda^{n-1} + B_1\lambda^{n-2} + \dots + B_{n-2}\lambda + B_{n-1}\right)$$
$$- \lambda\left(B_0\lambda^{n-1} + B_1\lambda^{n-2} + \dots + B_{n-1}\right)$$
$$= \left(a_0\lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n\right) I_n.$$

Comparing coefficients of like powers of λ on both sides, we get

$$-IB_0 = a_0I_n$$
$$AB_0 - IB_1 = a_1I_n$$
$$AB_1 - IB_2 = a_2I_n$$
$$\dots\dots\dots$$
$$\dots\dots\dots$$
$$AB_{n-1} = a_nI_n.$$

Multiplying these successively by A^n, A^{n-1},...,I_n and adding, we get,

$$0 = a_0A^n + a_1A^{n-1} + \ldots + a_{n-1}A + a_nI,$$

that is,

$$\phi(A) = \mathbf{0}.$$

This completes the proof of the theorem.

Corollary 4.10 If A is non-singular, then

$$A^{-1} = \frac{-a_0}{a_n}A^{n-1} - \frac{a_1}{a_n}A^{n-2} - \ldots - \frac{a_{n-1}}{a_n}I$$
$$= \frac{-1}{a_n}\left(a_0A^{n-1} + a_1A^{n-2} + \ldots + a_{n-1}I\right)$$

Proof: By Cayley-Hamilton theorem, we have

$$a_0A^n + a_1A^{n-1} + \ldots + a_{n-1}A + a_nI = 0.$$

Pre-multiplication with A^{-1} yields

$$a_0A^{n-1} + a_1A^{n-2} + \ldots + a_{n-1} + a_nA^{-1} = 0.$$

or

$$A^{-1} = -\frac{a_0}{a_n}A^{n-1} - \frac{a_1}{a_n}A^{n-2} - \ldots - \frac{a_{n-1}}{a_n}I$$
$$= -\frac{1}{a_n}\left(a_0A^{n-1} + a_1A^{n-2} + \ldots + a_{n-1}I\right).$$

Remark 4.7 It follows from above that

$$A^n = -\frac{1}{a_0}\left(a_1A^{n-1} + a_1A^{n-2} + \ldots + a_nI\right).$$

Thus higher powers of a matrix can be obtained using lower powers of A.

EXAMPLE 4.59

Verify Cayley-Hamilton theorem for the matrix

$$A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

and hence find A^{-1}.

Solution. We have

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & -1 & 1 \\ -1 & 2-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{vmatrix}$$
$$= -\lambda^3 + 6\lambda^2 - 9\lambda + 4.$$

Thus, the characteristic equation of the matrix A is

$$\lambda^3 - 6\lambda^2 + 9\lambda - 4 = 0.$$

To verify Cayley-Hamilton theorem, we have to show that

$$A^3 - 6A^2 + 9A - 4I = 0. \qquad (19)$$

We have

$$A^2 = \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix},$$
$$A^3 = \begin{bmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{bmatrix}.$$

Then, we note that

$$A^3 - 6A^2 + 9A - 4I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0.$$

Further, pre-multiplying (19) by A^{-1}, we get

$$A^2 - 6A + 9I - 4A^{-1} = \mathbf{0}$$

and so

$$A^{-1} = \frac{1}{4}\left(A^2 - 6A + 9I\right)$$
$$= \frac{1}{4}\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}.$$

4.25 ALGEBRAIC AND GEOMETRIC MULTIPLICITY OF AN EIGENVALUE

Definition 4.76 If λ is an eigenvalue of order m of matrix A, then m is called the *algebraic multiplicity* of λ.

Definition 4.77 If s is the number of linearly independent eigenvectors corresponding to the eigenvalue λ, then s is called the *geometric multiplicity* of λ.

If r is the rank of the coefficient matrix of $(A - \lambda I)X = \mathbf{0}$, then $s = n - r$, where n is the number of unknowns.

The geometric multiplicity of an eigenvalue cannot exceed its algebraic multiplicity.

4.26 MINIMAL POLYNOMIAL OF A MATRIX

Definition 4.78 A polynomial in x in which the coefficient of the highest power of x is unity is called a *monic polynomial*.

For example, $x^4 - x^3 + 2x^2 + x + 4$ is a monic polynomial of degree 4 over the field of real numbers.

Definition 4.79 The monic polynomial $m(x)$ of lowest degree such that $m(A) = 0$ is called the *minimal polynomial* of the matrix A.

If $m(x)$ is the minimal polynomial of a matrix A, then the equation $m(x)=0$ is called the *minimal equation* of the matrix A.

Theorem 4.39 The minimal polynomial of a matrix is unique.

Proof: Suppose that the minimal polynomial of a matrix A is of degree r. Therefore, for no non-zero polynomial of degree less than r, we can have $m(A)=0$. Let $m_1(x)$ and $m_2(x)$ be two minimal polynomial of A. Then

$$m_1(A) = A^r + a_1A^{r-1} + \ldots + a_rI = 0$$

$$m_2(A) = A^r + b_1A^{r-1} + \ldots + b_rI = 0\ .$$

Subtracting, we have

$$(b_1 - a_1)\,A^{r-1} + \ldots + (b_r - a_r)\ I = 0.$$

Thus, we have a polynomial $f(x)$ of degree $r-1$ such that $f(A)=\mathbf{0}$. Since its degree is less than r, this should be a zero polynomial. Hence

$$b_1 - a_1 = 0, \ldots, b_r - a_r = 0$$

and so

$$a_1 = b_1, \ldots,\ b_r = a_r$$

proving that $m_1(A)=m_2(A)$. Hence, minimal polynomial of A is unique.

Theorem 4.40 Every polynomial $p(\lambda)$ such that $p(A)=0$ is exactly divisible by the minimal polynomial $m(\lambda)$.

Proof: Let $q(\lambda)$ be the quotient when $p(\lambda)$ is divided by $m(\lambda)$ and let $r(\lambda)$ be the remainder, which is of degree less than the degree of $m(\lambda)$. Then, by division algorithm, we have,

$$p(\lambda) = m(\lambda)q(\lambda) + r(\lambda)$$

so

$$0 = p(A) = m(A)q(A) + r(A) = 0.q(\text{A}) + r(\text{A}),$$

which yields $r(A)=0$. Since $r(\lambda)$ is of degree less than the degree of $m(\lambda)$, it follows that $m(\lambda)$ is not a minimal polynomial unless $r(\lambda)=0$. Thus,

$$p(\lambda) = m(\lambda)q(\lambda)$$

and hence $m(\lambda)$ divides $p(\lambda)$.

Corollary 4.11 The minimal polynomial of a matrix is a divisor of the characteristic polynomial of that matrix.

Proof: Let $\phi(\lambda)$ be the characteristic polynomial of a matrix A. Then by Cayley-Hamilton theorem, $\phi(A)=\mathbf{0}$. Let $m(\lambda)$ be the minimal polynomial of A. Then, by Theorem 4.57, $m(\lambda)$ divides $\phi(\lambda)$.

Corollary 4.12 Every root of the minimal equation of a matrix is also a characteristic root of the matrix.

Proof: Let $\phi(\lambda)$ be the characteristic polynomial of a matrix A and $m(\lambda)$ be its minimal polynomial. Then, by Corollary 4.17, $m(\lambda)$ divides $\phi(\lambda)$. Therefore, there exists a polynomial $q(\lambda)$ such that

$$\phi(\lambda) = m(\lambda)q(\lambda).$$

Now, suppose μ is a root of the equation $m(\lambda)=0$. Therefore, $m(\mu)=0$ and so

$$\phi(\mu) = m(\mu)q(\mu) = 0.$$

Hence μ is a root of the characteristic equation $\phi(\lambda)=0$ and so μ is a characteristic root of A.

Theorem 4.41 If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the characteristic roots, distinct or not, of a matrix A of order n and if $g(A)$ is any polynomial function of A, then the characteristic roots of $g(A)$ are $g(\lambda_1), g(\lambda_2), \ldots, g(\lambda_n)$.

Proof: We have

$$|A - \lambda I_n| = (-1)^n\ (\lambda - \lambda_1)\ (\lambda - \lambda_2)\ \ldots\ (\lambda - \lambda_n).$$

We want to show that

$$|g(A) - \lambda I_n| = (-1)^n\ (\lambda - g(\lambda_1)) \times (\lambda - g(\lambda_2))\ \ldots\ (\lambda - g(\lambda_n)).$$

Suppose $g(x)$ is of degree r in x and that for a fixed value of λ, the roots of $g(x) - \lambda = 0$ are $x_1, x_2, \ldots, x_r$. Then

$$g(x) - \lambda = \alpha(x - x_1)(x - x_2)\ldots(x - x_r),$$

where α is the coefficient of x^r in $g(x)$. Hence

$$g(A) - \lambda I_n = \alpha(A - x_1I_n)\ (A - x_2I_n)\ \ldots\ (A - x_rI_n).$$

Therefore if $\phi(\lambda)$ is the characteristic polynomial of A, then

$$\begin{aligned}
&|g(A) - \lambda I_n| \\
&= \alpha^n\,|(A - x_1I_n)|\ |A - x_2I_n|\ \ldots\ |A - x_rI_n| \\
&= \alpha^n\phi(x_1)\,\phi(x_2)\ \ldots\ \phi(x_r) \\
&= \alpha^n(-1)^n\,(x_1 - \lambda_1)\ (x_1 - \lambda_2)\ldots(x_1 - \lambda_n) \\
&\quad \ldots\ (-1)^n\,(x_r - \lambda_1)\ (x_r - \lambda_2)\ \ldots(x_r - \lambda_n) \\
&= \alpha\ (\lambda_1 - x_1)\ (\lambda_1 - x_2)\ \ldots\ (\lambda_1 - x_r) \\
&\quad \ldots\ \alpha\ (\lambda_n - x_1)\ (\lambda_n - x_2)\ \ldots\ (\lambda_n - x_r) \\
&= (g(\lambda_1) - \lambda)\ \ (g(\lambda_2) - \lambda)\ldots\ (g(\lambda_n) - \lambda) \\
&= (-1)^n\ (\lambda - g(\lambda_1))(\lambda - g(\lambda_2))\ldots(\lambda - g(\lambda_n)).
\end{aligned}$$

Hence, $g(\lambda_1), g(\lambda_2), \ldots, g(\lambda_n)$ are the characteristic roots of $g(A)$.

Theorem 4.42 Every root of the characteristic equation of a matrix is also a root of the minimal equation of the matrix.

Proof: Suppose $m(x)$ is the minimal polynomial of a matrix A. Then $m(A)=0$. Let λ be a characteristic root of A. Then, by Theorem 4.58, $m(\lambda)$ is the characteristic root of $m(A)$. But $m(A)=0$ and so each of its characteristic root is zero. Hence $m(\lambda)=0$, which implies that λ is a root of the equation $m(x)=0$. This proves that every characteristic root of a matrix A is also a root of the minimal equation $m(x)=0$.

Corollary 4.12 and Theorem 4.42 combined together yield:

Theorem 4.43 A scalar λ is a characteristic root of a matrix if and only if it is a root of the minimal equation of that matrix.

Definition 4.80 An n-rowed matrix is said to be *derogatory* or *non-derogatory* according as the degree of its minimal equation is less than or equal to n.

It follows from the definition that a matrix is non-derogatory if the degree of its minimal polynomial is equal to the degree of its characteristic polynomial.

Theorem 4.44 If the roots of the characteristic equation of a matrix are all distinct, then the matrix is non-derogatory.

Proof: Let A be a matrix of order n having n distinct characteristic roots. By Theorem 4.60, each of these roots is also a root of the minimal polynomial of A. Therefore, the minimal polynomial of A is of degree n. Hence, by definition, A is non-derogatory.

EXAMPLE 4.60

Show that the matrix

$$A = \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}$$

is derogatory.

Solution. We have

$$|A-\lambda I| = \begin{vmatrix} 7-\lambda & 4 & -1 \\ 4 & 7-\lambda & -1 \\ -4 & -4 & 4-\lambda \end{vmatrix} = -(\lambda-12)(3-\lambda)^2.$$

Therefore, roots of the characteristic equation $|A-\lambda I| = 0$ are $\lambda = 3, 3, 12$.

Since each characteristic root of a matrix is also a root of its minimal polynomial, therefore, $(x-3)$ and $(x-12)$ shall be factors of $m(x)$. Let.

$$g(x) = (x-3)\,(x-12) = x^2 - 15x + 36.$$

We have

$$A^2 = \begin{bmatrix} 69 & 60 & -15 \\ 60 & 69 & -15 \\ -60 & -60 & 24 \end{bmatrix}$$

Then, we observe that

$$g(A) = A^2 - 15A + 36I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus $g(x)$ is the monomic polynomial of lowest degree such that $g(A)=0$. Hence $g(x)$ is minimal polynomial of A. Since its degree is less than the order of the matrix A, the given matrix A is derogatory.

4.27 ORTHOGONAL, NORMAL, AND UNITARY MATRICES

Definition 4.81 Let

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix} \text{ and } Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ \dots \\ y_n \end{bmatrix}$$

be two complex n-vectors. The inner product of X and Y denoted by (X, Y), is defined as

$$(X, Y) = X^{\theta} Y = [\bar{x}_1\, \bar{x}_2 \dots \bar{x}_n] \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ \dots \\ y_n \end{bmatrix}$$

$$= \bar{x}_1 y_1 + \bar{x}_2 y_2 + \dots + \bar{x}_n y_n.$$

If X and Y are real, then their product becomes

$$(X, Y) = X^T Y = [x_1 \;\; x_2 \;\dots\; x_n] \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ \dots \\ y_n \end{bmatrix}$$

$$= x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Definition 4.82 Let X be a complex n-vector. Then the positive square root of the inner product of X with itself is called the *length* or *norm* of X. It is denoted by $\|X\|$.

For example, if

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix}$$

then

$$\|X\| = \sqrt{(X,X)} = \sqrt{X^\theta X}$$
$$= \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}.$$

Obviously, the length of a vector is zero if and only if the vector is a zero vector.

Definition 4.83 A vector X is called a *unit vector* if $\|X\| = 1$.

Definition 4.84 Two complex n-vectors X and Y are said to be *orthogonal* if

$$(X, Y) = X^\theta Y = 0.$$

Obviously, *zero is the only vector which is orthogonal to itself.*

Definition 4.85 A set S of complex n-vectors X_1, X_2, …, X_n is said to be an orthogonal set if any two distinct vectors in S are orthogonal.

Definition 4.86 A set S of complex n-vectors X_1, X_2, …, X_n is said to be an orthonormal set if

(i) S is an orthogonal set

(ii) each vector in S is a unit vector.

Thus the set X_1, X_2, …, X_n is orthonormal if

$$(X_i,\ X_j) = \delta_{ij}, i,j = 1,\ 2, \dots, n,$$

where δ_{ij} (called Kronecker delta) is defined as

$$\delta_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j. \end{cases}$$

Theorem 4.45 An orthogonal set of non-zero vectors is linearly independent.

Proof: Let $S = [X_1, X_2, \dots, X_n]$ be an orthogonal set of non-zero vectors. Let $c_1, c_2, \dots, c_n$ be scalars such that

$$c_1X_1 + c_2X_2 + \dots + c_nX_n = 0 \qquad (20)$$

Let $1 \leq m \leq n$. Then inner product of (20) with X_m is

$$(X_m,\ c_1X_1 + c_2X_2 + \dots + c_nX_n) = (X_m,\ \mathbf{0})$$

or

$$c_1(X_m,X_1) + c_2(X_m,X_2) + \dots + c_m(X_m,X_m) + \dots + c_n(X_m,X_n) = 0.$$

Since $(X_m, X_n) = 0$ for $m \neq n$, the above relation yields

$$c_m(X_m,X_m) = 0.$$

Since $X_m \neq 0$, the inner product $(X_m, X_m) \neq 0$. Hence $c_m = 0$, $m = 1, 2, \dots, n$. Thus, (39) implies $c_1 = c_2 = \dots = c_n = 0$. Hence, X_1, X_2, …, X_n are linearly independent.

Corollary 4.13 Every orthonormal set of vectors is linearly independent.

Proof: Since for every vector X_n, $(X_n, X_n) = 1$, the result follows from Theorem 4.45.

Definition 4.87 A square matrix U with complex element is said to be *unitary* if $U^\theta U = I$.

If U is unitary, then

$$U^\theta U = I$$
$$\Rightarrow |U^\theta U| = |I|$$
$$\Rightarrow |U^\theta|\ |U| = |I| = 1$$
$$\Rightarrow |U| \neq 0.$$

Hence, U is non-singular and so invertible. Thus, U^θ is the inverse of U and we have

$$U^\theta U = I = UU^\theta.$$

Hence, a matrix U is unitary if and only if

$$U^\theta U = UU^\theta = I.$$

If U is a unitary matrix, then the transformation $Y = UX$ is called a *unitary transformation*.

Theorem 4.46 The eigenvalues of a unitary matrix are of unit modulus.

Proof: Let λ be an eigenvalue of a unitary matrix. Therefore, there exists non-zero vector X such that

$$AX = \lambda X \qquad (21)$$

Therefore, taking transposed conjugate of (21), we get

$$(AX)^\theta = (\lambda X)^\theta$$
$$\Rightarrow X^\theta A^\theta = \bar{\lambda} X^\theta \qquad (22)$$

By (21) and (22), we have

$$\begin{aligned} & X^\theta A^\theta AX = \bar{\lambda}\lambda X^\theta X \\ \Rightarrow\ & X^\theta X = \bar{\lambda}\lambda X^\theta X \\ \Rightarrow\ & (1-\bar{\lambda}\lambda)X^\theta X = 0 \\ \Rightarrow\ & (1-\bar{\lambda}\lambda) = 0 \text{ since } X^\theta X \neq 0 \\ \Rightarrow\ & \bar{\lambda}\lambda = 1 \\ \Rightarrow\ & |\lambda| = 1. \end{aligned}$$

Theorem 4.47 (i) If U is unitary matrix, then absolute value of $|U| = 1$

(ii) Any two eigenvectors corresponding to the distinct eigenvalues of a unitary matrix are orthogonal.

Proof: (i) We have

$$|U^\theta| = \left|(\overline{U})^T\right| = |\overline{U}| = \overline{|U|}$$

Therefore

$$\begin{aligned} |U|^2 = \overline{|U|}\,.\,|U| &= |U^\theta|\,|U| = |U^\theta U| \\ &= |I| = 1. \end{aligned}$$

Hence, absolute value of determinant of a unitary matrix is 1.

(ii) Let λ_1 and λ_2 be two distinct eigenvalues of a unitary matrix U and let X_1, X_2 be the corresponding eigenvectors. Then

$$UX_1 = \lambda_1 X_1 \tag{23}$$

$$UX_2 = \lambda_2 X_2 \tag{24}$$

Taking conjugate transpose of (24), we get

$$X_2^\theta U^\theta = \bar{\lambda}_2 X_2^\theta \tag{25}$$

Post-multiplying both sides of (25) by UX_1, we get

$$\begin{aligned} & X_2^\theta U^\theta UX_1 = \bar{\lambda}_2 X_2^\theta UX_1 \\ \Rightarrow\ & X_2^\theta X_1 = \bar{\lambda}_2 X_2^\theta \lambda_1 X_1 \text{ since } U^\theta U = I \\ & \text{and } UX_1 = \lambda_1 X_1 \\ \Rightarrow\ & X_2^\theta X_1 = \bar{\lambda}_2 \lambda_1 X_2^\theta X_1 \\ \Rightarrow\ & (1-\bar{\lambda}_2\lambda_1)\, X_2^\theta X_1 = 0 \end{aligned} \tag{26}$$

But eigenvalues of a unitary matrix are of unit modulus.

Therefore $\bar{\lambda}_2\lambda_2 = 1$, that is, $\bar{\lambda}_2 = \frac{1}{\lambda_2}$. Thus (26) reduces to

$$\begin{aligned} & \left(1-\frac{\lambda_1}{\lambda_2}\right) X_2^\theta X_1 = 0 \\ \Rightarrow\ & \left(\frac{\lambda_2-\lambda_1}{\lambda_2}\right) X_2^\theta X_1 = 0 \\ \Rightarrow\ & X_2^\theta X_1 = 0 \text{ since } \lambda_1 \neq \lambda_2. \end{aligned}$$

Hence, X_1 and X_2 are orthogonal vectors

Theorem 4.48 The product of two unitary matrices of the same order is unitary.

Proof: Let A and B be two unitary matrices of order n. Then

$$AA^\theta = A^\theta A = I \text{ and } BB^\theta = B^\theta B = I.$$

We have

$$\begin{aligned} (AB)^\theta(AB) &= \left(B^\theta A^\theta\right)\,(AB) \\ &= B^\theta\left(A^\theta A\right)B \\ &= B^\theta\, I\, B \\ &= B^\theta B = I. \end{aligned}$$

Hence, AB is an unitary matrix of order n. Similarly,

$$\begin{aligned} (BA)^\theta(BA) &= \left(A^\theta B^\theta\right)\,(BA) \\ &= A^\theta\left(B^\theta B\right)A \\ &= A^\theta\, I\, A \\ &= A^\theta A = I \end{aligned}$$

and so BA is unitary.

Theorem 4.49 The inverse of a unitary matrix of order n is an unitary matrix.

Proof: Let U be an unitary matrix. Then

$$\begin{aligned} & UU^\theta = I \\ \Rightarrow\ & \left(UU^\theta\right)^{-1} = I^{-1} = I \\ \Rightarrow\ & \left(U^\theta\right)^{-1} U^{-1} = I \\ \Rightarrow\ & \left(U^{-1}\right)^\theta U^{-1} = I \end{aligned}$$

Hence, U^{-1} is also a unitary matrix.

Remark 4.8 It follows from Theorem 4.66 that the set of unitary matrices is a group under the binary operation of multiplication. This group is called *unitary group*.

Definition 4.88 A square matrix P is said to be *orthogonal* if $P^TP = I$.

Thus, a real unitary matrix is called an orthogonal matrix.

If P is orthogonal, then

$$\begin{aligned} & P^TP = I \\ \Rightarrow\ & \left|P^TP\right| = |I| = 1 \\ \Rightarrow\ & \left|P^T\right|\,|P| = 1 \\ \Rightarrow\ & |P|^2 = 1 \\ \Rightarrow\ & |P| \neq 0. \end{aligned}$$

Thus P is invertible and have inverse as P^T. Hence $P^TP = I = PP^T$.

If P is an orthogonal matrix, then the transformation $Y=PX$ is called *orthogonal transformation*.

Theorem 4.50 The product of two orthogonal matrices of order n is an orthogonal matrix of order n.

Proof: Let A and B be orthogonal matrices of order n. Therefore, A and B are invertible. Further both AB and BA are matrices of order n. But

$$|AB| = |A|\,|B| \neq 0 \text{ and } |BA| = |B|\,|A| \neq 0$$

Therefore, AB and BA are invertible. Now

$$\begin{aligned}(AB)^T\ (AB) &= \left(B^TA^T\right)\ (AB)\\ &= B^T\left(A^TA\right)B\\ &= B^TI\ B\\ &= B^TB = I.\end{aligned}$$

Hence AB is orthogonal. Similarly BA is also orthogonal.

Theorem 4.51 If a matrix P is orthogonal, then P^{-1} is also orthogonal.

Proof: Since P is orthogonal, we have

$$\begin{aligned}&PP^T = I\\ \Rightarrow\ &\left(PP^T\right)^{-1} = I^{-1} = I\\ \Rightarrow\ &\left(P^T\right)^{-1}P^{-1} = I\\ \Rightarrow\ &\left(P^{-1}\right)^TP^{-1} = I.\end{aligned}$$

Hence, P^{-1} is also orthogonal.

Remark 4.9 The above results show that the set of orthogonal matrices form a multiplication group called *orthogonal group*.

Theorem 4.52 Eigenvalues of an orthogonal matrix are of unit modulus.

Proof: Since an orthogonal matrix is a real unitary matrix, the result follows from Theorem 4.46.

Remark 4.10 It follows from Theorem 4.46 that ±1 can be the only real characteristic roots of an orthogonal matrix.

Definition 4.89 A matrix A is said to be *normal* if and only if $A^\theta A = AA^\theta$.

For example, unitary, Hermitian, and Skew-Hermitian matrices are normal. Also, the diagonal matrices with arbitrary diagonal elements are normal.

Theorem 4.53 If U is unitary, then A is normal if and only if $U^\theta AU$ is normal

Proof: We have

$$\begin{aligned}\left(U^\theta AU\right)^\theta\ \left(U^\theta AU\right) &= \left(U^\theta A^\theta U\right)\ \left(U^\theta AU\right)\\ &= U^\theta A^\theta\left(UU^\theta\right)AU\\ &= U^\theta A^\theta I\,AU\\ &= U^\theta A^\theta AU \qquad (27)\end{aligned}$$

and similarly,

$$\left(U^\theta AU\right)\ \left(U^\theta AU\right)^\theta = U^\theta AA^\theta U \qquad (28)$$

From (27) and (28), we note that $A^\theta A=AA^\theta$ if and only if

$$\left(U^\theta AU\right)^\theta\ \left(U^\theta AU\right) = \left(U^\theta AU\right)\ \left(U^\theta AU\right)^*.$$

Hence, A is normal if and only if $U^\theta AU$ is normal.

4.28 SIMILARITY OF MATRICES

Definition 4.90 Let A and B be matrices of order n. Then B is said to be *similar* to A if there exists a non-singular matrix P such that $B=P^{-1}AP$.

It can be seen easily that the relation of similarity of matrices is an equivalence relation.

If B is similar to A, then

$$\begin{aligned}|B| &= \left|P^{-1}AP\right| = \left|P^{-1}\right|\ |A|\ |P|\\ &= \left|P^{-1}\right|\ |P|\ |A|\\ &= \left|P^{-1}P\right|\ |A|\\ &= |I|\ |A| = |A|.\end{aligned}$$

Therefore it follows that similar matrices have the same determinant.

Theorem 4.54 Similar matrices have the same characteristic polynomial and hence the same characteristic roots.

Proof: Suppose A and B are similar matrices. Then there exists an invertible matrix P such that $B=P^{-1}AP$. Since

$$\begin{aligned}B - xI &= P^{-1}AP - xI\\ &= P^{-1}AP - P^{-1}(xI)P\\ &= P^{-1}(A - xI)P,\end{aligned}$$

we have

$$\begin{aligned}|B - xI| &= \left|P^{-1}(A - xI)\,P\right|\\ &= \left|P^{-1}\right|\ |P|\,|A - xI|\\ &= \left|P^{-1}P\right|\,|A - xI|\\ &= |A - xI|.\end{aligned}$$

Thus A and B have the same characteristic polynomial and so they have same characteristic roots. Further if λ is characteristic root of A, then,

$$AX = \lambda X,\ X \neq 0$$

and so

$$B(P^{-1}X) = (P^{-1}AP)\,P^{-1}X = P^{-1}AX = P^{-1}(\lambda X) = \lambda(P^{-1}X)$$

This shows that $(P^{-1}X)$ is an eigenvector of B corresponding to its eigenvalue λ.

Corollary 4.14 If a matrix A is similar to a diagonal matrix D, the diagonal elements of D are the eigenvalues of A.

Proof: Since A and D are similar, they have same eigenvalues. But the eigenvalues of the diagonal matrix D are its diagonal elements. Hence the eigenvalues of A are the diagonal elements of D.

4.29 DIAGONALIZATION OF A MATRIX

Definition 4.91 A matrix A is said to be *diagonalizable* if it is similar to a diagonal matrix.

Theorem 4.55 *A matrix of order n is diagonalizable if and only if it possesses n linearly independent eigenvectors.*

Proof: Suppose first that A is diagonalizable. Then A is similar to a diagonal matrix

$$D = \text{diag}[\lambda_1 \lambda_2 \ldots \lambda_n].$$

Therefore, there exists an invertible matrix $P = [X_1\ X_2\ \ldots\ X_n]$ such that $P^{-1}AP = D$, that is, $AP = PD$ and so

$$A[X_1\ X_2 \ldots X_n] = [X_1\ X_2 \ldots X_n]\ \text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_n]$$

or

$$[AX_1,\ AX_2 \ldots AX_n] = [\lambda_1 X_1\ \lambda_2 X_2 \ldots \lambda_n X_n].$$

Hence

$$AX_1 = \lambda_1 X_1, AX_2 = \lambda_2 X_2, \ldots, AX_n = \lambda_n X_n.$$

Thus, $X_1, X_2, \ldots, X_n$ are eigenvectors of A corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Since P is non-singular, its column vectors $X_1, X_2, \ldots, X_n$ are linearly independent. Hence A has n linearly independent eigenvectors.

Conversely suppose that A possesses n linearly independent eigenvectors $X_1, X_2, \ldots, X_n$ and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues. Then

$$AX_1 = \lambda_1 X_1,\quad AX_2 = \lambda_2 X_2, \ldots, AX_n = \lambda_n X_n.$$

Let

$$P = [X_1, X_2, \ldots, X_n] \text{ and } D = diag\,[\lambda_1\ \lambda_2 \ldots \lambda_n].$$

Then

$$AP = [AX_1\ AX_2 \ldots\ AX_n] = [\lambda_1 X_1\ \lambda_2 X_2 \ldots\ \lambda_n X_n] = [X_1, X_2 \ldots\ X_n]\ diag\ [\lambda_1\ \lambda_2 \ldots \lambda_n] = PD.$$

Since the column vectors $X_1, X_2, \ldots, X_n$ of the matrix P are linearly independent, so P is invertible and P^{-1} exists.

Therefore,

$$AP = PD \Rightarrow P^{-1}AP = P^{-1}PD$$

$$\Rightarrow P^{-1}AP = D$$

$\Rightarrow$ A is similar to D. $\Rightarrow$A is diagonalizable.

Theorem 4.56 If the eigenvalues of a matrix of order n are all distinct, then it is always similar to a diagonal matrix.

Proof: Suppose that a square matrix of order n has n distinct eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_n$. As eigenvectors of a matrix corresponding to distinct eigenvalues are linearly independent, A has n linearly independent eigenvectors and so, by the above theorem, it is similar to a diagonal matrix.

The following result is very useful in diagonalization of a given matrix.

Theorem 4.57 The necessary and sufficient condition for a square matrix to be similar to a diagonal matrix is that geometric multiplicity of each of its eigenvalues coincide with the algebraic multiplicity.

EXAMPLE 4.61

Show that the matrix

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

is not similar to diagonal matrix.

Solution. The characteristic equation of A is

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 3 & 4 \\ 0 & 2-\lambda & -1 \\ 0 & 0 & 1-\lambda \end{vmatrix} = 0$$

and so

$$(2-\lambda)\ (2-\lambda)\ (1-\lambda) = 0.$$

Hence the eigenvalues of A are 2, 2, and 1. The eigenvector X of A corresponding to $\lambda=2$ is given by $(A-2I)\ X = 0$, that is, by

$$\begin{bmatrix} 0 & 3 & 4 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0 & 3 & 4 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ R_3 \rightarrow R_3 - R_2.$$

The coefficient matrix is of rank 2. Hence number of linearly independent solution is $n-r=1$. Thus geometric multiplicity of 2 is 1. But its algebraic multiplicity is 2. Therefore, geometric multiplicity is not equal to algebraic multiplicity. Hence A is not similar to a diagonal matrix.

EXAMPLE 4.62

Give an example to show that not every square matrix can be diagonalized by a non-singular transformation of coordinates.

Solution. Consider the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

The characteristic equation of A is

$$|A-\lambda I| = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = 0$$

or

$$(1-\lambda)^2 = 0,$$

which yields the characteristic roots as $\lambda = 1,1$.

The characteristic vector corresponding to $\lambda = 1$, is given by $(A-I)X=0$, that is, by

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The rank of the coefficient matrix is 1 and so that number of linearly independent solution is $n-r$ $2-1=1$. Thus the geometric multiplicity of characteristic root is 1, whereas algebraic multiplicity of the characteristic root is 2. Hence, the given matrix is not diagonalizable.

EXAMPLE 4.63

Show that the matrix

$$A = \begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}$$

is diagonalizable. Hence, find the transforming matrix and the diagonal matrix.

Solution. The roots of the characteristic equation

$$|A-\lambda I| = \begin{vmatrix} 8-\lambda & -8 & -2 \\ 4 & -3-\lambda & -2 \\ 3 & -4 & 1-\lambda \end{vmatrix} = 0$$

are 1, 2, 3. Since the eigenvalues are all distinct, A is similar to a diagonal matrix. Further, algebraic multiplicity of each eigenvalues is 1. So there is only one linearly independent eigenvector of A corresponding to each eigenvalues. Now the eigenvector corresponding to $\lambda=1$ is given by $(A-I)X=0$, that is, by

$$\begin{bmatrix} 7 & -8 & -2 \\ 4 & -4 & -2 \\ 3 & -4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\sim \begin{bmatrix} 7 & -8 & -2 \\ -3 & 4 & 0 \\ 3 & -4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, R_2 \to R_2 - R_1$$

$$\sim \begin{bmatrix} 7 & -8 & -2 \\ -3 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, R_3 \to R_3 + R_2.$$

We note that rank of the coefficient matrix is 2. Therefore, there is only one linearly independent solution. Hence geometric multiplicity of the eigenvalues 1 is 1. The equation can be written as

$$7x_1 - 8x_2 - 2x_3 = 0$$

$$-3x_1 + 4x_2 = 0.$$

The last equation yields $x_1 = \frac{4}{3}x_2$. So taking $x_2=3$, we get $x_1=4$. Then the first equation yields $x_3=2$. Hence, the eigenvector corresponding to $\lambda=1$ is

$$X_1 = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}.$$

Similarly, eigenvectors corresponding to $\lambda=2$ and 3 are found to be

$$X_2 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \text{ and } X_3 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

Therefore, the transforming matrix is

$$P = \begin{bmatrix} 4 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix},$$

and so the diagonal matrix is

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

EXAMPLE 4.64

Diagonalize the matrix

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}.$$

Solution. The characteristic equation of the given matrix is

$$|A-\lambda I| = \begin{vmatrix} 1-\lambda & 0 & -1 \\ 1 & 2-\lambda & 1 \\ 2 & 2 & 3-\lambda \end{vmatrix} = 0$$

or

$$-\lambda^3 + 6\lambda^2 - 11\lambda + 6 = 0.$$

The characteristic roots are $\lambda = 1, 2, 3$. Since the characteristic roots are distinct, the given matrix is diagonalizable and the diagonal elements shall be the characteristic roots 1, 2, 3.

The characteristic vectors corresponding to $\lambda = 1$ are given by $(A-I)\,X=0$, that is, by

$$\begin{bmatrix} 0 & 0 & -1 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad R_1 \leftrightarrow R_2$$

$$\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad R_3 \to R_3 - 2R_1.$$

The rank of the coefficient matrix is 2. Therefore, there is only $3-2=1$ linearly independent solution. The above equation yields,

$$x_1 + x_2 + x_3 = 0$$

$$-x_3 = 0.$$

Hence, the characteristic vector corresponding to $\lambda=1$ is

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

The characteristic vector corresponding to $\lambda=2$ is given by $(A-2I)\,X=0$, that is, by

$$\begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{matrix} R_2 \to R_2 + R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$\sim \begin{bmatrix} -1 & 0 & -1 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad R_2 \leftrightarrow R_3.$$

The rank of the coefficient matrix is 2. Therefore, there is only $3-2=1$linearly independent solution. The equation implies

$$-x_1 - x_3 = 0$$

$$x_1 - 2x_3 = 0$$

which yields $x_1 = 2$, $x_2 = -1$, $x_3 = -2$. Therefore, the characteristic vector is

$$\begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}.$$

The characteristic vector corresponding to $\lambda=3$ is given by

$$\begin{bmatrix} -2 & 0 & -1 \\ 1 & -1 & 1 \\ 2 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} -2 & 0 & -1 \\ -1 & -1 & 0 \\ 2 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad R_2 \to R_2 + R_1.$$

The rank of coefficient matrix is 2. and so there is $3-2=1$ independent solution. The equation yields,

$$-2x_1 - x_3 = 0$$

$$-x_1 - x_2 = 0.$$

and so the corresponding characteristic vector is

$$\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}.$$

Thus, the transforming matrix is

$$P = \begin{bmatrix} 1 & 2 & 1 \\ -1 & -1 & -1 \\ 0 & -2 & -2 \end{bmatrix}.$$

We have $|P| = -2$ and the cofactors of P are

$$A_{11} = 0, \quad A_{12} = -2, \quad A_{13} = 2,$$

$$A_{21} = 2, \quad A_{22} = -2, \quad A_{23} = 2,$$

$$A_{31} = -1, \quad A_{32} = 0, \quad A_{33} = 1.$$

Therefore,

$$\text{adj } P = \begin{bmatrix} 0 & 2 & -1 \\ -2 & -2 & 0 \\ 2 & 2 & 1 \end{bmatrix},$$

and so

$$P^{-1} = \frac{1}{|P|}\text{adj } P = \begin{bmatrix} 0 & -1 & \frac{1}{2} \\ 1 & 1 & 0 \\ -1 & -1 & -\frac{1}{2} \end{bmatrix}.$$

Then we observe that

$$P^{-1}AP = \begin{bmatrix} 0 & -1 & \frac{1}{2} \\ 1 & 1 & 0 \\ -1 & -1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 1 \\ -1 & -1 & -1 \\ 0 & -2 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -1 & \frac{1}{2} \\ 1 & 1 & 0 \\ -1 & -1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 4 & 3 \\ -1 & -2 & -3 \\ 0 & -4 & -6 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \text{diag}\,[1 \quad 2 \quad 3]\ .$$

Definition 4.92 Let A and B be square matrices of order n. Then B is said to be *unitarily similar* to A if there exists a unitary matrix U such that $B = U^{-1}AU$.

Theorem 4.58 (Existence Theorem). If A is Hermitian matrix, then there exists a unitary matrix U such that $U^{\theta}AU$ is a diagonal matrix whose diagonal elements are the characteristic roots of A, that is,

$$U^{\theta}AU = \text{diag}[\lambda_1 \ \lambda_2 \dots \ \lambda_n].$$

Proof: We shall prove Theorem 4.75 by induction on the order of A. If $n = 1$, then the theorem is obviously true. We assume that the theorem is true for all Hermitian matrices of order $n - 1$. We shall establish that the theorem holds for all Hermitian matrices of order n.

Let λ_1 be an eigenvalue of A. Thus λ_1 is real. Let X_1 be the eigenvector corresponding to the eigenvalues λ_1. Therefore $AX_1 = \lambda_1 X_1$. We choose an orthonormal basis of the complex vector space V_n having X_1 as a member. Therefore, there exists a unitary matrix S with X_1 as its first column. We now consider the matrix $S^{-1}AS$. Since X_1 is the first column of S, the first column of $S^{-1}AS$ is $S^{-1}AX_1 = S^{-1}\lambda_1 X_1 = \lambda_1 S^{-1}X_1$. But $S^{-1}X_1$ is the first column of $S^{-1}S = \text{I}$. Therefore, the first column of $S^{-1}AS$ is $[\lambda_1 \ 0 \dots 0 \dots 0]^T$. Since S is unitary, $S^{-1} = S^{\theta}$ and so

$$\left(S^{-1}AS\right)^{\theta} = S^{\theta}A^{\theta}\left(S^{-1}\right)^{\theta} = S^{\theta}A^{\theta}S = S^{-1}AS.$$

Hence $S^{-1}AS$ is Hermitian. Therefore, the first row of $S^{-1}AS$ is $[\lambda_1 \ 0 \dots 0 \dots 0]$. Thus,

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & 0 \\ 0 & B \end{bmatrix},$$

where B is a square matrix of order $n-1$. Therefore, by induction hypothesis, there exists a unitary matrix V such that

$$V^{-1}BV = D_1,$$

where D_1 is a diagonal matrix of order $n - 1$.

Let $R = \begin{bmatrix} I & 0 \\ 0 & V \end{bmatrix}$ be a matrix of order n. Then R is invertible and $R^{-1} = \begin{bmatrix} I & 0 \\ 0 & V^{-1} \end{bmatrix}$. Now since V is unitary, $V^{\theta} = V^{-1}$ and so

$$R^{\theta} = \begin{bmatrix} I & 0 \\ 0 & V^{\theta} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & V^{-1} \end{bmatrix} = R^{-1}.$$

Hence, R is uniatary. Since R and S are unitary matrices of order n, SR is also unitary of order n. Let $SR = U$. Then

$$\begin{aligned} U^{-1}AU &= (SR)^{-1}A(SR) \\ &= \left(R^{-1}S^{-1}\right)A(SR) \\ &= R^{-1}\left(S^{-1}AS\right)R \\ &= \begin{bmatrix} I & 0 \\ 0 & V^{-1} \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & V \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & V^{-1}B \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & V \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & V^{-1}BV \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & D_1 \end{bmatrix} = \text{diag}[\lambda_1 \lambda_2 \dots \lambda_n]. \end{aligned}$$

As an immediate consequence of this theorem, we have

Corollary 4.15 If A is a real symmetric matrix, there exists an orthogonal matrix U such that U^TAU is a diagonal matrix, whose diagonal elements are the characteristic roots of A.

Theorem 4.59 If λ is an m-fold eigenvalue of Hermitian matrix A, then rank of A-λI_n is $n - m$.

Proof: By Theorem 4.58, there exists a unitary matrix U such that

$$U^{*}AU = \text{diag}[\lambda\lambda \dots \lambda\lambda_{m+1}\lambda_{m+2} \dots \lambda_n],$$

where λ occurs m times and $\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_n$ are all distinct from λ. Since U is unitary, subtracting λI_n from both sides of the above equation, we get

$$U^{*}[A - \lambda I_n]U = diag[00 \dots 0(\lambda_{m+1} - \lambda) \\ (\lambda_{m+2} - \lambda) \dots (\lambda_n - \lambda)].$$

Since U is non-singular, it follows that the rank of $A - \lambda I_n$ is same as that of the diagonal matrix on the right-hand side. But the rank of the matrix on the right-hand side is $n - m$ because $(\lambda_{m+1} - \lambda)$, $(\lambda_{m+2} - \lambda)$, …, $(\lambda_n - \lambda)$ are all non-zero.

Corollary 4.16 If λ is m-fold eigenvalues of a Hermitian matrix A, then there exists m linearly independent vectors of A associated with λ, that is, with λ_1 there is associated an m-dimensional space of characteristic vectors.

Theorem 4.60 With every Hermitian matrix A we can associate an orthonormal set of n characteristic vectors.

Proof: The eigenvectors associated with a given eigenvalue of A form a vector space for which we can construct an orthonormal basis by Gram-Schmidt process. For each A, there are n vectors in the basis so constructed. Also, the eigenvectors associated with distinct eigenvalues of a Hermitian matrix are orthogonal. It follows, therefore, that these n basis vectors constitute orthonormal set.

Theorem 4.60 indicates how the diagonalization process may be effected. In fact, we have the following theorem.

Theorem 4.61 *If $U_1, U_2, \ldots, U_n$ is an orthonormal system of eigenvectors associated respectively with the eigenvalues $\lambda_1 \ldots \lambda_n$ of Hermitian matrix A and if U is the unitary matrix $[U_1\ U_2\ \ldots\ U_n]$, then*

$$U^* A U = diag[\lambda_1 \lambda_2 \ldots \lambda_n].$$

(The vectors $U_1, U_2, \ldots, U_n$ are often called a set of *principal axes* of A and the transformation with matrix U used to diagonalize A is called *principal axis transformation*).

Proof: We have $AU_j = \lambda_j U_j$, $j = 1, 2, \ldots, n$, where λ_j is the eigenvalue associated with U_j. Thus, if $U = [U_1\ U_2\ \ldots\ U_n]$, then

$$[AU_1 AU_2 \ldots AU_n] = [\lambda_1 U_1\ \lambda_2 U_2 \ldots \lambda_n U_n],$$

that is,

$$AU = U\ \text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_n].$$

Since, U is unitary $U^{-1} = U^\theta$ and so pre-multiplication by U^θ yields

$$\begin{aligned} U^\theta A U &= U^\theta U\ \text{diag}\ [\lambda_1 \lambda_2 \ldots \lambda_n] \\ &= U^{-1} U\ \text{diag}\ [\lambda_1 \lambda_2 \ldots \lambda_n] \\ &= \text{diag}\ [\lambda_1 \lambda_2 \ldots \lambda_n]. \end{aligned}$$

EXAMPLE 4.65

Diagonalize the matrix

$$A = \begin{bmatrix} 2 & 1-2i \\ 1+2i & -2 \end{bmatrix}.$$

Solution. The characteristic equation of A is

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 1-2i \\ 1+2i & -2-\lambda \end{vmatrix} = 0$$

or

$$\lambda^2 - 9 = 0,$$

which yields the characteristic roots as $\lambda = -3, 3$. The eigenvectors corresponding to the eigenvalue -3 is given by $(A + 3I)X = \mathbf{0}$, that is, by

$$\begin{bmatrix} 5 & 1-2i \\ 1+2i & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which yields

$$5x_1 - (1-2i)x_2 = 0$$
$$(1+2i)x_2 + x_2 = 0.$$

Solving these equations, we get $x_1 = 1 - 2i$, $x_2 = -5$. Hence, $X_1 = \begin{bmatrix} 1-2i \\ -5 \end{bmatrix}$ is the eigenvector corresponding to $\lambda = -3$.

The eigenvector corresponding to $\lambda = 3$ is given by $(A - 3I)X = \mathbf{0}$ that is, by

$$\begin{bmatrix} -1 & 1-2i \\ 1+2i & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which yields

$$-x_1 + (1-2i)x_2 = 0$$
$$(1+2i)x_1 - 5x_2 = 0.$$

Solving these equations, we get $x_1 = 5$, $x_2 = 1 + 2i$. Thus the required eigenvector is $X_2 = \begin{bmatrix} 5 \\ 1+2i \end{bmatrix}$.

We note that

$$\begin{aligned} X_2^\theta X_1 &= [5 \quad 1-2i] \begin{bmatrix} 1-2i \\ -5 \end{bmatrix} \\ &= 5(1-2i) - 5(1-2i) = 0. \end{aligned}$$

Thus $\{X_1, X_2\}$ is an orthogonal set. Now

$$\begin{aligned} \text{Norm of } X_1 &= \sqrt{|1-2i|^2 + |-5|^2} = \sqrt{5+25} \\ &= \sqrt{30} \end{aligned}$$

$$\begin{aligned} \text{Norm of } X_2 &= \sqrt{|5|^2 + |1+2i|^2} \\ &= \sqrt{25+5} \\ &= \sqrt{30}. \end{aligned}$$

Therefore, normalized characteristic vectors are

$$U_1 = \begin{bmatrix} \frac{1-2i}{\sqrt{30}} \\ \frac{-5}{\sqrt{30}} \end{bmatrix}, \quad U_2 = \begin{bmatrix} \frac{5}{\sqrt{30}} \\ \frac{1+2i}{\sqrt{30}} \end{bmatrix}.$$

Hence the transforming unitary matrix is

$$U = \begin{bmatrix} \frac{1-2i}{\sqrt{30}} & \frac{5}{\sqrt{30}} \\ \frac{-5}{\sqrt{30}} & \frac{1+2i}{\sqrt{30}} \end{bmatrix}.$$

We then note that

$$U^{\theta}AU = U^{-1}AU = \begin{bmatrix} -3 & 0 \\ 0 & 3 \end{bmatrix} = \text{diag } [-3\ 3].$$

EXAMPLE 4.66

Diagonalize the Hermitian matrix

$$A = \begin{bmatrix} 5 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 5 & -2 \\ 0 & 0 & -2 & 2 \end{bmatrix}.$$

Solution. The characteristic equation of A is

$$|A-\lambda I| = \begin{vmatrix} 5-\lambda & 2 & 0 & 0 \\ 2 & 2-\lambda & 0 & 0 \\ 0 & 0 & 5-\lambda & -2 \\ 0 & 0 & -2 & 2-\lambda \end{vmatrix} = 0.$$

The characteristic roots are 1, 1, 6, 6. The characteristic vector corresponding to $\lambda = 1$ are given by $(A - I)X = \mathbf{0}$, that is, by

$$\begin{bmatrix} 4 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

which yields

$$4x_1 + 2x_2 = 0$$
$$2x_1 + x_2 = 0$$
$$4x_3 - 2x_4 = 0$$
$$2x_3 + x_4 = 0$$

with the complete solution as

$$X_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \end{bmatrix}, X_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}.$$

These vectors are already orthogonal. The normalized vectors are

$$U_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \\ 0 \\ 0 \end{bmatrix} \text{ and } U_2 = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}.$$

Similarly, the normalized vectors corresponding to $\lambda = 6$ are

$$U_3 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \\ 0 \end{bmatrix} \text{ and } U_4 = \begin{bmatrix} 0 \\ 0 \\ -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}.$$

Hence, the transforming unitary matrix is

$$U = \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} & 0 \\ 0 & \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ 0 & \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \end{bmatrix} = 0.$$

and $U^{\theta}\, AU = \text{diag } [1\ 1\ 6\ 6]$.

4.30 TRIANGULARIZATION OF AN ARBITRARY MATRIX

Not every matrix can be reduced to diagonal form by a unitary transformation. But it is always possible to reduce a square matrix to a triangular form. In this direction, we have the following result.

Theorem 4.62 (Jacobi-Thoerem). Every square matrix A over the complex field can be reduced by a unitary transformation to upper triangular form with the characteristic roots on the diagonal.

Proof: We shall prove the theorem by induction on the order n of the matrix A. If $n = 1$, the theorem is obviously true. Suppose that the result holds for all matrices of order $n - 1$. Let λ_1 be the characteristic root of A and U_1 denote the corresponding unit characteristic vector. Then $AU_1 = \lambda_1 U_1$. Let $\{U_1, U_2, \ldots, U_n\}$ be an orthonormal set, that is, $U = [U_1, U_2, \ldots, U_n]$. Then

$$U^{\theta}AU = \begin{bmatrix} U_1^{\theta} \\ U_2^{\theta} \\ \cdots \\ \cdots \\ U_n^{\theta} \end{bmatrix} [AU_1\ AU_2 \ldots AU_n]$$

$$= \begin{bmatrix} U_1^{\theta} \\ U_2^{\theta} \\ \cdots \\ \cdots \\ U_n^{\theta} \end{bmatrix} [\lambda_1 U_1\ AU_2 \ldots AU_n]$$

Since $U_1^\theta U_1 = I$ and $U_2^\theta U_1 = U_3^\theta U_1 = \ldots = U_n^\theta U_1 = 0$, we have

$$U^*AU = \begin{bmatrix} \lambda_1 & U_1^\theta AU_2 & \ldots & \ldots & U_1^\theta AU_n \\ 0 & U_2^\theta AU_2 & \ldots & \ldots & U_2^\theta AU_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & U_n^\theta AU_2 & \ldots & \ldots & U_n^\theta AU_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & B \\ 0 & C \end{bmatrix}.$$

Now, by induction hypothesis, the matrix

$$C = \begin{bmatrix} U_2^\theta AU_2 & \ldots & \ldots & U_2^\theta AU_n \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ U_n^\theta AU_2 & \ldots & \ldots & U_n^\theta AU_n \end{bmatrix},$$

which is of order $n-1$, is triangularizable. Thus there exists a unitary matrix W of order $n-1$ which triangularize C, that is, $W^\theta CW$ is triangular. Let $V = \begin{bmatrix} I & 0 \\ 0 & W \end{bmatrix}$. Then $V^{-1} = \begin{bmatrix} I & 0 \\ 0 & W^{-1} \end{bmatrix}$ and

$$V^\theta = \begin{bmatrix} I & 0 \\ 0 & W^\theta \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & W^{-1} \end{bmatrix} = V^{-1}.$$

Hence V is unitary and

$$V^\theta \left(U^\theta AU\right) V = \begin{bmatrix} I & 0 \\ 0 & W^\theta \end{bmatrix} \begin{bmatrix} \lambda_1 & B \\ 0 & C \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & W \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & B \\ 0 & W^\theta C \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & W \end{bmatrix} = \begin{bmatrix} \lambda_1 & BW \\ 0 & W^\theta CW \end{bmatrix},$$

where $W^\theta CW$ is upper triangular. Thus, we have

$$(UV)^\theta A(UV) = \begin{bmatrix} \lambda_1 & BW \\ 0 & W^\theta CW \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & b_{12} & \ldots & \ldots & b_{1n} \\ 0 & \lambda_2 & \ldots & \ldots & b_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & \lambda_n \end{bmatrix}.$$

Since UV is unitary, the characteristic roots of the triangular matrix are the same as that of A. Thus, diagonal elements of triangular matrix are characteristic roots of A.

Theorem 4.63 A matrix A over the complex field can be diagonalized by a unitary transformation if and only if A is normal.

Proof: Suppose first that U is unitary and A can be diagonalized, that is, $U^\theta AU = \text{diag}[\lambda_1\ \lambda_2\ \ldots\ \lambda_n]$. Then $A = U\ \text{diag}[\lambda_1\ \lambda_2\ \ldots\ \lambda_n]U^\theta$ and so

$$A^\theta A = U\ (\text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_n])^\theta\ (\text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_n])\ U^\theta$$

and

$$AA^\theta = U\ (\text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_n])\ (\text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_n])^\theta\ U^\theta$$

But

$\text{diag}[\lambda_1\ \lambda_2\ \ldots\ \lambda_n]\ (\text{diag}[\lambda_1\ \lambda_2\ \ldots\ \lambda_n])^\theta = (\text{diag}[\lambda_1\ \lambda_2\ \ldots\ \lambda_n])^\theta\ \text{diag}[\lambda_1\ \lambda_2\ \ldots\ \lambda_n]$.

Hence $A^\theta A = AA^\theta$ and so A is normal.

Conversely, suppose A is normal. Then, by Theorem 4.70, there exists unitary matrix U such that $U^\theta AU = B$, where B is upper triangular. But $U^\theta AU$ is normal and so B is normal. Suppose that the upper triangular matrix B is

$$B = \begin{bmatrix} \lambda_1 & b_{12} & \ldots & \ldots & b_{1n} \\ 0 & \lambda_2 & \ldots & \ldots & b_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & \lambda_n \end{bmatrix}$$

Since $B^\theta B = BB^\theta$, we have

$$B = \begin{bmatrix} \bar{\lambda}_1 & 0 & b_{13} & \ldots & \ldots & 0 \\ \bar{b}_{12} & \bar{\lambda}_2 & b_{23} & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \bar{b}_{1n} & \bar{b}_{2n} & 0 & \ldots & \ldots & \bar{\lambda}_n \end{bmatrix} \times \begin{bmatrix} \lambda_1 & b_{12} & b_{13} & \ldots & \ldots & b_{1n} \\ 0 & \lambda_2 & b_{23} & \ldots & \ldots & b_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \ldots & \lambda_n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & b_{12} & b_{13} & \ldots & \ldots & b_{1n} \\ 0 & \lambda_2 & b_{23} & \ldots & \ldots & b_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \ldots & \lambda_n \end{bmatrix} \times \begin{bmatrix} \bar{\lambda}_1 & 0 & b_{13} & \ldots & \ldots & 0 \\ \bar{b}_{12} & \bar{\lambda}_2 & b_{23} & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \bar{b}_{1n} & \bar{b}_{2n} & 0 & \ldots & \ldots & \bar{\lambda}_n \end{bmatrix}$$

Comparison of 1-1 entries on both sides, we have

$$\bar{\lambda}_1\lambda_1 = \lambda_1\bar{\lambda}_1 + b_{12}\bar{b}_{12} + b_{13}\bar{b}_{13} + \ldots + b_{1n}\bar{b}_{1n}$$

or

$$0 = |b_{12}|^2 + |b_{13}|^2 + \ldots + |b_{1n}|^2,$$

which implies that $b_{12} = b_{13} = \ldots = b_{1n} = 0$.
Similarly, comparison of 2-2 entries, we get

$$b_{23} = b_{24} = \ldots = b_{2n} = 0$$

and so on, Hence B is diagonal, that is, $U^{\theta}AU = \text{diag}\,[\lambda_1\ \lambda_2\ \ldots\ \lambda_n]$.

4.31 QUADRATIC FORMS

Definition 4.93 A homogeneous polynomial of the type

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\ x_i\ x_j,$$

where a_{ij} are elements of a field F is called a *quadratic form* in n variables $x_1, x_2, \ldots, x_n$ over the field F.

If a_{ij} are real, then the quadratic form is called *real quadratic form*.

For example, $x_1^2 - 3x_1x_2 + x_2^2 + x_1x_3$ is a real quadratic form.

Theorem 4.64 Every quadratic form over a field F in n variables $x_1, x_2, \ldots, x_n$ can be expressed in the form of X^TBX, where B is a symmetric matrix of order n over F and X is a column vector $[x_1, x_2, \ldots, x_n]^T$.

Proof: Let

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\ x_i\ x_j,$$

be a quadratic form over the field F in n variables $x_1, x_2, \ldots, x_n$. Since x_i, x_j are scalars, we have $x_i\,x_j = x_j\,x_i$. Therefore, the coefficient of $x_i\,x_j$ is $a_{ij} + a_{ji}$. Thus, we assign half of the coefficient to x_{ij} and half to x_{ji}. Let b_{ij} be another set of scalars such that $b_{ii} = a_{ii}$ and $b_{ij} = \frac{1}{2}\left(a_{ij} + a_{ji}\right)$ for $i \neq j$. Then

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\ x_i\ x_j = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\ x_i\ x_j.$$

Since, $b_{ij} = b_{ji}$, the matrix $B = [b_{ij}]_{n\times n}$ is symmetric. We further note that if

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ \cdots \\ x_n \end{bmatrix},$$

then

$$X^TBX = [x_1 x_2 \ldots_n] \begin{bmatrix} b_{11} & b_{12} & \cdots & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & \cdots & b_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ \cdots \\ x_n \end{bmatrix}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\,x_i\,x_j = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\,x_i\,x_j.$$

The symmetric matrix B is called the *matrix of the quadratic form*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\ x_i\ x_j.$$

EXAMPLE 4.67

Find the matrix of the quadratic form $x_1^2 - 3x_1x_2 + x_2^2 + x_1x_3$.

Solution. The given quadratic form can be written as

$$x_1^2 - \frac{3}{2}x_1x_2 - \frac{3}{2}x_2x_1 + x_2^2 + \frac{1}{2}x_1x_3 + \frac{1}{2}x_3x_1.$$

Therefore, the matrix of the given quadratic form is

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

where

$$a_{11} = 1, \quad a_{12} = -\frac{3}{2} \quad a_{13} = \frac{1}{2}$$

$$a_{21} = -\frac{3}{2}, \quad a_{22} = 1 \quad a_{23} = 0$$

$$a_{31} = \frac{1}{2}, \quad a_{32} = 0 \quad a_{33} = 0.$$

Hence

$$A = \begin{bmatrix} 1 & -\frac{3}{2} & \frac{1}{2} \\ -\frac{3}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix},$$

which is symmetric.

EXAMPLE 4.68

Find the quadratic form corresponding to the symmetric matrix.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 3 \\ 3 & 3 & 1 \end{bmatrix}$$

Solution. The required symmetric form is

$$X^TAX = [x_1x_2x_3]\begin{bmatrix}1 & 2 & 3\\ 2 & 0 & 3\\ 3 & 3 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

$$= [x_1x_2x_3]\begin{bmatrix}x_1+2x_2+3x_3\\ 2x_1+3x_2\\ 3x_1+3x_2+x_3\end{bmatrix}$$

$$= x_1(x_1+2x_2+3x_3)+x_2(2x_1+3x_3) + x_3(3x_1+3x_2+x_3)$$

$$= x_1^2+x_3^2+4x_1x_2+6x_1x_3+6x_2x_3.$$

4.32 DIAGONALIZATION OF QUADRATIC FORMS

We know that for every real symmetric matrix A there exists an orthogonal matrix U such that

$$U^T\ AU = \text{diag}\ [\lambda_1\ \lambda_2 \ldots \lambda_n],$$

where $\lambda_1, \lambda_2,\ldots, \lambda_n$ are characteristic roots of A. Applying the orthogonal transformation $X = UY$ to the quadratic form X^TAX, we have

$$X^TAX = \lambda_1y_1^2+\lambda_2y_2^2+\ldots+\lambda_ny_n^2.$$

If the rank of A is r, then n-r characteristic roots are zero and so

$$X^TAX = \lambda_1y_1^2+\lambda_2y_2^2+\ldots+\lambda_ry_r^2,$$

where $\lambda_1, \lambda_2,\ldots, \lambda_n$ are non-zero characteristic roots.

Definition 4.94 A square matrix B of order n over a field F is said to be *congruent* to another square matrix A of order n over F, if there exists a non-singular matrix P over F such that $B=P^TAP$.

The relation of "congruence of matrices" is an equivalence relation in the set of all $n\times n$ matrices over a field F. Further, let A be symmetric matrix and let B be congruent to A. Therefore, there exists a non-singular matrix P such that $B=P^TAP$. Then

$$B^T = (P^T\ AP)^T = P^TA^TP$$

$$= P^TAP, \text{since } A \text{ is symmetric}$$

$$= B.$$

Hence, every *matrix congruent to a symmetric matrix is a symmetric matrix*.

Theorem 4.65 (Congruent reduction of a symmetric matrix). If A is any n rowed non-zero symmetric matrix of rank r over a field F, then there exists an n rowed non-singular matrix P over F such that

$$P^TAP = \begin{bmatrix}A_1 & 0\\ 0 & 0\end{bmatrix},$$

where A_1 is a non-zero singular diagonal matrix of order r over F and each **0** is a null matrix of a suitable size.

Proof: We prove the theorem by induction. When $n = 1$, $r = 1$ also. The quadratic form is simply $a_{11}{x_1}^2$, $a_{11} \neq 0$ and the identity transformation $y_1 = x_1$ is the non-singular transformation. Suppose that the theorem is true for all symmetric matrices of order $n - 1$, then we first show that there exists a matrix $B = [b_{ij}]_{n\times n}$ over F congruent to A such that $b_{11} \neq 0$. We take up the following cases.

Case I. If $a_{11} \neq 0$, then we take $B = A$.

Case II. If $a_{11} = 0$, but some diagonal element of A, say $a_{ii} \neq 0$. Then using R_i R_1, C_i C_1 to A, we obtain a matrix B congruent to A such that $b_{11} = a_{ii} \neq 0$.

Case III. Suppose that each diagonal element of A is zero. Since A is non-zero, there exists, non-zero element a_{ij} such that $a_{ij} = a_{ji} \neq 0$. Applying the congruent operation $R_i \to R_i + R_j$, $C_i \to C_i + C_j$ to A, we obtain a matrix $D = [d_{ij}]_{n\times n}$ congruent to A such that $d_{ii} = a_{ij} + a_{ji} = 2a_{ij} \neq 0$. Now, applying the congruent operation $R_i \to R_1$, C_i C_1 to D, we obtain a matrix $B = [b_{ij}]_{n\times n}$ congruent to D and therefore also congruent to A such that $b_{11} = d_{ii} \neq 0$. Hence, there exists a matrix $B = [b_{ij}]$ congruent to a symmetric matrix such that the leading element of B is non-zero. Since B is congruent to a symmetric matrix, therefore, B itself is symmetric. Since $b_{11} \neq 0$, all elements in the first row and first column except the leading element can be made zero by suitable congruent operation. Thus we have a matrix

$$C = \begin{bmatrix}a_{11} & 0 & \ldots & 0\\ 0 & & & \\ \ldots & & B_1 & \\ 0 & & & \end{bmatrix}$$

congruent to B and, therefore, congruent to A such that B is a square matrix of order $n-1$. Further C is congruent to a symmetric matrix and so C is also symmetric. Consequently B_1 is also a symmetric matrix. By induction hypothesis, B_1 can be reduced to a diagonal matrix by congruent operation. So C can be reduced to a diagonal matrix by congruent operations. Thus, A is congruent to a diagonal matrix, say, diag $[\lambda_1\ \lambda_2 \ldots \lambda_k \ldots 0\,0\,0\,0\,]$. Thus there exists a non-singular matrix P such that

$$P^T AP = \text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_k \ldots 0\ 0\ 0\ 0].$$

Since $\rho(A) = r$ and we know that rank does not alter by multiplying by a non-singular matrix, therefore, rank of $P^T AP = \text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_k \ldots 0\ 0\ 0\ 0]$ is also r. So r elements of diag $[\lambda_1\ \lambda_2 \ldots \lambda_k \ldots 0\ 0\ 0\ ]$ are non-zero. Thus, $k = r$ and so

$$P^T AP = \text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_r \ldots 0\ 0\ 0\ 0].$$

Corollary 4.17 Corresponding to every quadratic form $X^T AX$ over a field F, there exists a non-singular linear transformation $X = PY$ over F such that the form $X^T AX$ transforms to

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_r y_r^2,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_r$ are scalars in F and r is the rank of the matrix A.

Definition 4.95 The rank of the symmetric matrix A is called the *rank of the quadratic form $X^T AX$.*

EXAMPLE 4.69

Find a non-singular matrix P such that $P^T AP$ is a diagonal matrix, where

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}.$$

Find the quadratic form and its rank.

Solution. Write $A = IAI$, that is,

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

Using congruent operations, we shall reduce A to diagonal form. Performing congruent operations $R_2 \to R_2 + \frac{1}{3}R_1, C_2 \to C_2 + \frac{1}{3}C_1$ and $R_3 \to R_3 - \frac{1}{3}R_1, C_3 \to C_3 - \frac{1}{3}C_1$, we have

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & \frac{7}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & \frac{7}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ -\frac{1}{3} & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now performing congruent operation $R_3 \to R_3 + \frac{1}{7}R_2, C_3 \to C_3 + \frac{1}{7}C_2$, we have

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & \frac{7}{3} & 0 \\ 0 & 0 & \frac{16}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ -\frac{2}{7} & \frac{1}{7} & 1 \end{bmatrix} A \begin{bmatrix} 1 & \frac{1}{3} & -\frac{2}{7} \\ 0 & 1 & \frac{1}{7} \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus

$$diag\begin{bmatrix} 6 & \frac{7}{3} & \frac{16}{7} \end{bmatrix} = P^{-1}\ AP,$$

where

$$P = \begin{bmatrix} 1 & \frac{1}{3} & -\frac{2}{7} \\ 0 & 1 & \frac{1}{7} \\ 0 & 0 & 1 \end{bmatrix}.$$

The quadratic form corresponding to the matrix A is

$$X^T AX = 6x_1^2 + 3x_2^2 + 3x_3^2 - 4x_1x_2 - 2x_2x_3 + 4x_3x_1. \tag{48}$$

The non-singular transformation $X = PY$ corresponding to the matrix P is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{3} & -\frac{2}{7} \\ 0 & 1 & \frac{1}{7} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix},$$

which yields

$$x_1 = y_1 + \frac{1}{3}y_2 - \frac{2}{7}y_3$$

$$x_2 = y_2 + \frac{1}{7}y_3$$

$$x_3 = y_3.$$

Substituting these values in (48), we get

$$(PY)^T A(PY) = 6y_1^2 + \frac{7}{3}y_2^2 + \frac{16}{7}y_3^2.$$

It contains a sum of *three* squares. Thus, the rank of the quadratic form is 3.

Theorem 4.66 Let A be any n-rowed real symmetric matrix of rank r. Then there exists a real non-singular matrix P such that

$$P^T AP = \text{diag}[1\ 1 \ldots\ 1\ \ -1\ \ -1\ \ -1\ \ldots\ -1\ 0\ 0\ 0\ \ldots\ 0],$$

where 1 appears p times and -1 appears r - p times.

Proof: Since A is a symmetric matrix of rank r, there exists a non-singular real matrix Q such that

$$Q^T AQ = \text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_r\ \ldots 0\ 0\ 0\ 0].$$

Suppose p of the non-zero diagonal elements are positive and $r - p$ are negative. Then by using congruence operations $R_i\ R_j, C_i\ C_j$, we can assume that first p elements $\lambda_1, \lambda_2, \ldots, \lambda_p$ are positive and $\lambda_{p+1}, \lambda_{p+2}, \ldots, \lambda_r$ are negative. Let

$$S = \text{diag}\left[\frac{1}{\sqrt{\lambda_1}}\frac{1}{\sqrt{\lambda_2}}\cdots\frac{1}{\sqrt{\lambda_p}}\frac{1}{\sqrt{-\lambda_{p+1}}}\cdots\frac{1}{\sqrt{-\lambda_r}}111\right].$$

Then S is non-singular and $S^T = S$. Let $P = QS$. Then P is also real non-singular matrix and we have

$$\begin{aligned}P^TAP &= (QS)^TA(QS) = S^TQ^TAQS\\ &= S^T(\text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_r\ 0 \ldots 0])S\\ &= S(\text{diag}[\lambda_1\ \lambda_2 \ldots \lambda_r\ 0 \ldots 0])S\\ &= \text{diag}[1\ 1 \ldots 1\ \ -1\ \ -1\ \ldots -1\ 0 \ldots 0]\end{aligned}$$

so that 1 appears p times and –1 appears $r - p$ times.

Corollary 4.18 If X^TAX is a real quadratic form of rank r in n variables, then there exists a real non-singular linear transformation $X = PY$ which transform X^TAX to the form

$$Y^TP^TAPY = y_1^2 + y_2^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2,$$

which is called *canonical form* or *normal form* of a real quadratic form.

The number of positive terms in the normal form of X^TAX is called the *index* of the quadratic form, whereas $p - (r - p) = 2p - r$ is called the *signature* of the quadratic form and is usually denoted by s.

A quadratic form X^TAX with a non-singular matrix A of order n is called *positive definite* if $n = r = p$, that is, if n = rank = index. A quadratic form is called *positive semi-definite* if $r < n$ and $r = p$. Similarly a quadratic form is called *negative definite* if its index is zero and $n = r$ and called *negative semi-definite* if $r < n$ and its index is zero.

EXAMPLE 4.70

Find the rank, index, and signature of the quadratic form $x^2 - 2y^2 + 3z^2 - 4yz + 6zx$.

Solution. The matrix of the given quadratic form is

$$A = \begin{bmatrix}1 & 0 & 3\\ 0 & -2 & -2\\ 3 & -2 & 3\end{bmatrix}.$$

Write $A = IAI$, that is,

$$\begin{bmatrix}1 & 0 & 3\\ 0 & -2 & -2\\ 3 & -2 & 3\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} A \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Performing congruence operations $R_3 \to R_3 - 3R_1,\ C_3 \to C_3 - 3C_1$, we get

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

Performing congruence operations $R_3 \to R_3 - R_2, C_3 \to C_3 - C_2$, we have,

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

Performing $R_2 \to \frac{1}{\sqrt{2}}R_2, C_2 \to \frac{1}{\sqrt{2}}C_2$, and $R_3 \to \frac{1}{\sqrt{4}}R_3, C_3 \to \frac{1}{\sqrt{4}}C_3$, we get,

$$\begin{bmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 0 & \frac{1}{\sqrt{2}} & 0\\ -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2}\end{bmatrix} A \begin{bmatrix}1 & 0 & -\frac{3}{2}\\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2}\\ 0 & 0 & \frac{1}{2}\end{bmatrix}.$$

Hence $X = PY$ transforms the given quadratic form to $y_1^2 - y_2^2 - y_3^2$.

The rank of the quadratic form is 3 (the number of non-zero terms in the normal form.)

The number of positive terms is 1. Hence, the index of the quadratic form is 1.

We note that $2p - r = 2 - 3 = -1$. Therefore, signature of the quadratic form is –1.

4.33 MISCELLANEOUS EXAMPLES

EXAMPLE 4.71

Compute the inverse of $\begin{bmatrix}2 & 1 & -1\\ 0 & 2 & 1\\ 5 & 2 & -3\end{bmatrix}$ using elementary transformations.

Solution. Write $A = I_3A$, that is,

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

We reduce the matrix on the L.H.S of the equation to identity matrix by elementary row transformations, keeping in mind that each row transformation will apply to I_3 on the right hand side.

Interchanging R_1 and R_3, we get

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

Performing $R_1 \to R_1 - 2R_3$, we get

$$\begin{bmatrix}1 & 0 & -1\\ 0 & 2 & 1\\ 2 & 1 & -1\end{bmatrix} = \begin{bmatrix}-2 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{bmatrix} A$$

Performing $R_3 \to R_3 - 2R_1$, we have

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

Now performing $R_2 \to R_2 - R_3$, we get

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

Now performing $R_3 \to R_3 - R_2$, we have

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 1 \\ -5 & 1 & 2 \\ 10 & -1 & -4 \end{bmatrix} A$$

Lastly, performing $R_1 \to R_1 + R_3$, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 8 & -1 & -3 \\ -5 & 1 & 2 \\ 10 & -1 & -4 \end{bmatrix} A. = A^{-1}A.$$

Hence

$$A^{-1} = \begin{bmatrix} 8 & -1 & -3 \\ -5 & 1 & 2 \\ 10 & -1 & -4 \end{bmatrix}.$$

EXAMPLE 4.72

Using Cayley-Hamilton theorem, find A^{-1}, given the matrix

$$A = \begin{bmatrix} 13 & -3 & 5 \\ 0 & 4 & 0 \\ -15 & 9 & -7 \end{bmatrix}.$$

Solution. Proceeding as in Example 4.59, the characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 13-\lambda & -3 & 5 \\ 0 & 4-\lambda & 0 \\ -15 & 9 & -7-\lambda \end{vmatrix}$$

or

$$(13-\lambda)(4-\lambda)(-7-\lambda) + 75(4-\lambda) = 0$$

or

$$\lambda^3 - 10\lambda^2 + 8\lambda + 64 = 0.$$

By Cayley's Hamilton theorem, we have

$$A^3 - 10A^2 + 8A + 64I = 0$$

or

$$A^{-1} = -\frac{1}{64}[A^2 - 10A + 8I]$$

$$= -\frac{1}{64}\left\{\begin{bmatrix} 94 & -6 & 30 \\ 0 & 16 & 0 \\ -90 & 18 & -26 \end{bmatrix} - 10\begin{bmatrix} 13 & -3 & 5 \\ 0 & 4 & 0 \\ -15 & 9 & -7 \end{bmatrix} + 8\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\}$$

$$= -\frac{1}{64}\begin{bmatrix} 94-130+8 & -6+30 & 30-50 \\ 0 & 16-40+8 & 0 \\ -90+150 & 18-90 & -26+70+8 \end{bmatrix}$$

$$= -\frac{1}{64}\begin{bmatrix} -28 & 24 & -20 \\ 0 & -16 & 0 \\ 60 & -72 & 52 \end{bmatrix}$$

$$= -\frac{1}{16}\begin{bmatrix} -7 & 6 & -5 \\ 0 & -4 & 0 \\ 15 & -18 & 13 \end{bmatrix}.$$

EXAMPLE 4.73

For the matrix:

$$A = \begin{bmatrix} 2 & 1 & -3 & -6 \\ 3 & -3 & 1 & 2 \\ 1 & 1 & 1 & 2 \end{bmatrix},$$

find non-singular matrices P and Q such that PAQ is in the normal form. Hence find the rank of A.

Solution. Write

$$A = I_3 A I_4,$$

that is,

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

Performing elementary transformation $R_1 \leftrightarrow R_3$, we get

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

Performing $R_2 \to R_2 - 3R_1$ and $R_3 \to R_3 - 2R_1$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -6 & -2 & -4 \\ 0 & -1 & -5 & -10 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -3 \\ 1 & 0 & -2 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Performing $C_2 \to C_2 - C_1, C_3 \to C_3 - C_1, C_4 \to C_4 - 2C_1$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -6 & -2 & -4 \\ 0 & -1 & -5 & -10 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -3 \\ 1 & 0 & -2 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 & -2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now performing $R_3 \leftrightarrow R_2$, we have

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -5 & -10 \\ 0 & -6 & -2 & -4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & -3 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 & -2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_2 \to -R_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 10 \\ 0 & -6 & -2 & -4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & -3 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 & -2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now performing $R_3 \to R_3 + 6R_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 10 \\ 0 & 0 & 28 & 56 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 2 \\ -6 & 1 & 9 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 & -2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Performing $C_3 \to C_3 - 5C_2$ and $C_4 \to C_4 - 10C_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 28 & 56 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 2 \\ -6 & 1 & 9 \end{bmatrix} A \begin{bmatrix} 1 & -1 & 4 & 8 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_4 \to \frac{1}{28} R_4$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 2 \\ -\frac{3}{14} & \frac{1}{28} & \frac{9}{28} \end{bmatrix} A \begin{bmatrix} 1 & -1 & 4 & 8 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Performing $C_4 \to C_4 - 2C_3$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 2 \\ -\frac{3}{14} & \frac{1}{28} & \frac{9}{28} \end{bmatrix} A \begin{bmatrix} 1 & -1 & 4 & 8 \\ 0 & 1 & -5 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

or

$$[I_3 \quad 0] = PAQ,$$

where

$$P = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 2 \\ -\frac{3}{14} & \frac{1}{28} & \frac{9}{28} \end{bmatrix} \text{ and}$$

$$Q = \begin{bmatrix} 1 & -1 & 4 & 0 \\ 0 & 1 & -5 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Also $\rho(A) = 3$

EXAMPLE 4.74

(a) Find the rank of the matrix $\begin{bmatrix} 3 & -1 & 2 \\ -6 & 2 & 4 \\ -3 & 1 & 2 \end{bmatrix}$ by reducing it to the normal form

(b) For the matrix $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix}$,find non-singular matrices P and Q such that PAQ is in the normal form. Hence find the rank of A.

(c) Reduce the following matrix to column echelon and find its rank:

$$A = \begin{bmatrix} 1 & 1 & -1 & 1 \\ -1 & 1 & -3 & -3 \\ 1 & 0 & 1 & 2 \\ 1 & -1 & 3 & 3 \end{bmatrix}.$$

(d) Find all values of μ for which rank of the matrix

$$A = \begin{bmatrix} \mu & -1 & 0 & 0 \\ 0 & \mu & -1 & 0 \\ 0 & 0 & \mu & -1 \\ -6 & 11 & -6 & 1 \end{bmatrix}$$

is equal to 3.

Solution. (a) We have

$$A = \begin{bmatrix} 3 & -1 & 2 \\ -6 & 2 & 4 \\ -3 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 2 \\ -10 & 2 & 4 \\ -5 & 1 & 2 \end{bmatrix} C_1 \to C_1 - C_3$$

$$\sim \begin{bmatrix} 1 & -1 & 2 \\ 0 & -8 & 24 \\ 0 & -4 & 12 \end{bmatrix} \begin{matrix} R_2 \to R_2 + 10R_1 \\ R_3 \to R_3 + 5R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -3 \\ 0 & -4 & 12 \end{bmatrix} R_2 \to -\frac{1}{8}R_2$$

$$\sim \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix} R_3 \to R_3 + 4R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} C_2 \to C_2 + C_1 \\ C_3 \to C_3 - C_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} C_3 \to C_3 + 3C_2$$

$$= \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix} \text{ (normal form)}$$

Hence $\rho(A) = 2$.

(b) Expressing the given matrix in the form $A = I_3 A I_3$, we have

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

Using the elementary transformation $R_2 \to R_2 - R_1$, we get

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

Using the elementary column transformations $C_2 \to C_2 - C_1$ and $C_3 \to C_3 - 2C_1$, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Operating $R_3 \to R_3 + R_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Now operating $C_3 \to C_3 - C_2$, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

or

$$\begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix} = PAQ,$$

where

$$P = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since elementary transformations do not alter the rank of a matrix,

$$\rho(A) = \rho\begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix} = 2.$$

(c) A Matrix is said to be in column echelon form if

(i) The first non-zero entry in each non-zero column is 1.

(ii) The column containing only zeros occurs next to all non-zero columns.

(iii) The number of zeros above the first non-zero entry in each column is less than the number of such zeros in the next column.

The given matrix is

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

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & -4 & -2 \\ 1 & -1 & 2 & 1 \\ 1 & -2 & 4 & 2 \end{bmatrix} \begin{matrix} C_2 \to C_2 - C_1 \\ C_3 \to C_3 + C_1 \\ C_4 \to C_4 - C_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & -4 & 0 \\ 1 & -1 & 2 & 0 \\ 1 & -2 & 4 & 0 \end{bmatrix} C_4 \to C_4 + C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -2 & 0 & 0 \end{bmatrix} C_3 \to C_3 + 2C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 1 & -\frac{1}{2} & 0 & 0 \\ 1 & -1 & 0 & 0 \end{bmatrix} R_2 \to \frac{1}{2}R_2,$$

which is column echelon form. The number of non-zero column is two and therefore $\rho(A) = 2$.

(d) Similar to Remark 4.4

We are given that

$$A = \begin{bmatrix} \mu & -1 & 0 & 0 \\ 0 & \mu & -1 & 0 \\ 0 & 0 & \mu & -1 \\ -6 & 11 & -6 & 1 \end{bmatrix}$$

Therefore

$$|A| = \mu\begin{vmatrix} \mu & -1 & 0 \\ 0 & \mu & -1 \\ -6 & 11 & -6 \end{vmatrix} + 1\begin{vmatrix} 0 & -1 & 0 \\ 0 & \mu & -1 \\ -6 & -6 & 1 \end{vmatrix}$$

$$= \mu^3 - 6\mu^2 + 11\mu - 6$$

$$= 0 \text{ if } \mu = 1,\ 2,\ 3.$$

For $\mu = 3$, we have the singular matrix

$$\begin{bmatrix} 3 & -1 & 0 & 0 \\ 0 & 3 & -1 & 0 \\ 0 & 0 & 3 & -1 \\ -6 & 11 & -6 & 1 \end{bmatrix},$$

which has non-singular sub-matrix

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

Thus for $\mu = 3$, the rank of the matrix A is 3. Similarly, the rank is 3 for $\mu = 2$ and $\mu = 1$. For other values of μ, we have $|A| \neq 0$ and so $\rho(A) = 4$ for other values of μ.

EXAMPLE 4.75

Solve the system of equations :

$$x + y + z = 6$$
$$x - y + 2z = 5$$
$$3x + y + z = 8$$
$$2x - 2y + 3z = 7$$

Solution. The augmented matrix is

$$[A : B] = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 1 & -1 & 2 & 5 \\ 3 & 1 & 1 & 8 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & -2 & 1 & -1 \\ 0 & -2 & -2 & -10 \end{bmatrix} \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - 3R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 6 \\ 0 & -2 & 1 & -1 \\ 0 & -2 & -2 & -10 \end{bmatrix} \begin{matrix} C_2 \to C_2 - C_1 \\ C_3 \to C_3 - C_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 6 \\ 0 & 1 & -2 & -1 \\ 0 & -2 & -2 & -10 \end{bmatrix} C_2 \leftrightarrow C_3$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 6 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & -6 & -12 \end{bmatrix} R_3 \to R_3 + 2R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 6 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & -9 \end{bmatrix} R_3 \to R_3 + 3R_2$$

It follows that $\rho(A) = 2$ and $\rho[A : B] = 3$. Hence the given equation is inconsistent.

EXAMPLE 4.76

Discuss the consistency of the system of equations:

$$2x - 3y + 6z - 5w = 3,\ y - 4z + w = 1,$$
$$4x - 5y + 8z - 9w = \lambda$$

for various values of λ. If consistent, find the solution.

Solution. The matrix equation is $AX = B$, where

$$A = \begin{bmatrix} 2 & -3 & 6 & -5 \\ 0 & 1 & -4 & 1 \\ 4 & -5 & 8 & 9 \end{bmatrix},\ X = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \text{ and}$$

$$B = \begin{bmatrix} 3 \\ 1 \\ \lambda \end{bmatrix}.$$

The augmented matrix is

$$[A : B] = \begin{bmatrix} 2 & -3 & 6 & -5 & 3 \\ 0 & 1 & -4 & 1 & 1 \\ 4 & -5 & 8 & 9 & \lambda \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & -3 & 6 & -5 & 3 \\ 0 & 1 & -4 & 1 & 1 \\ 0 & 1 & -4 & 1 & \lambda - 6 \end{bmatrix} R_3 \to R_3 - 2R_1$$

$$\sim \begin{bmatrix} 2 & -3 & 6 & -5 & 3 \\ 0 & 1 & -4 & 1 & 1 \\ 0 & 0 & 0 & 0 & \lambda - 7 \end{bmatrix} R_3 \to R_3 - R_2$$

We note that $\rho(A) = \rho[A : B]$ if $\lambda - 7 = 0$, that is, if $\lambda = 7$. Thus the given equation is consistent if $\lambda = 7$. Thus if $\lambda = 7$, than we have

$$[A : B] = \begin{bmatrix} 2 & -3 & 6 & -5 & 3 \\ 0 & 1 & -4 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and so the given system of equations is equivalent to

$$2x - 3y + 6z - 5\omega = 3$$
$$y - 4z + \omega = 1.$$

Therefore if $w = k_1,\ z = k_2$, then $y = 1 + 4k_2 - k_1$ and $x = 3 + 3k_2 + k_1$. Hence the general solution of the system is $x = 3 + 3k_2 + k_1,\ y = 1 + 4k_2 - k_1,\ z = k_2,\ w = k_1$.

EXAMPLE 4.77

Test for consistency the following set of equations and solve if it is consistent: $5x + 3y + 7z = 4$, $3x + 26y + 2z = 9, 7x + 2y + 10z = 5$.

Solution. The augmented matrix is

$$[A : B] = \begin{bmatrix} 5 & 3 & 7 & 4 \\ 3 & 26 & 2 & 9 \\ 7 & 2 & 10 & 5 \end{bmatrix}$$

$$\sim \begin{bmatrix} 15 & 9 & 21 & 12 \\ 15 & 130 & 10 & 45 \\ 7 & 2 & 10 & 5 \end{bmatrix} \begin{matrix} R_1 \to 3R_1 \\ R_2 \to 5R_2 \end{matrix}$$

$$\sim \begin{bmatrix} 15 & 9 & 21 & 12 \\ 0 & 121 & -11 & 33 \\ 7 & 2 & 10 & 5 \end{bmatrix} R_2 \to R_2 - R_1$$

$$\sim \begin{bmatrix} 35 & 21 & 49 & 28 \\ 0 & 11 & -1 & 3 \\ 35 & 10 & 50 & 25 \end{bmatrix} \begin{matrix} R_1 \to \frac{7}{3}R_1 \\ R_2 \to \frac{1}{4}R_2 \\ R_3 \to 5R_3 \end{matrix}$$

$$\sim \begin{bmatrix} 35 & 21 & 49 & 28 \\ 0 & 11 & -1 & 3 \\ 0 & -11 & 1 & -3 \end{bmatrix} R_3 \to R_3 - R_1$$

$$\sim \begin{bmatrix} 35 & 21 & 49 & 28 \\ 0 & 11 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} R_3 \to R_3 + R_2.$$

We observe that

$$\rho(A) = 2, \rho([A : B]) = 2,$$

and so $\rho(A) = \rho([A : B])$. Hence the given system of equation is consistent. Further, the given system is equivalent to

$$35x + 21y + 49z = 28$$
$$11y - z = 3,$$

which yield $y = \frac{3+z}{11}$ and $x = \frac{7}{11} - \frac{16}{11}z$.

Taking $z = 0$, we get a particular solution as

$$x = \frac{7}{11},\ y = \frac{3}{11},\ z = 0.$$

EXAMPLE 4.78

(a) Find the value of λ for which the equations

$$(\lambda-1)x+(3\lambda+1)y+2\lambda z=0,$$
$$(\lambda-1)x+(4\lambda-2)y+(\lambda+3)z=0,$$
$$2x+(3\lambda+1)y+3(\lambda-1)z=0.$$

are consistent, and find the ratios of $x:y:z$ when λ has the smallest of these values. What happens when λ has the greater of these values?

(b) Determine b such that the system of homogeneous equations

$$2x+y+2z=0,\quad x+y+3z=0 \quad \text{and}$$
$$4x+3y+bz=0$$

has (i) Trivial solution (ii) Non-trivial solution. Also find the non-trivial solution using matrix method.

Solution. (a) For consistency, the coefficient matrix A, in the matrix equation $AX=0$, should be singular. Therefore, we must have

$$\begin{vmatrix} \lambda-1 & 3\lambda+1 & 2\lambda \\ \lambda-1 & 4\lambda-2 & \lambda+3 \\ 2 & 3\lambda+1 & 3(\lambda-1) \end{vmatrix}=0$$

$$\sim \begin{vmatrix} \lambda-1 & 3\lambda+1 & 2\lambda \\ 0 & \lambda-3 & 3-\lambda \\ 2 & 3\lambda+1 & 3(\lambda-1) \end{vmatrix}=0$$

$$\sim \begin{vmatrix} \lambda-1 & 3\lambda+1 & 5\lambda1+1 \\ 0 & \lambda-3 & 0 \\ 2 & 3\lambda+1 & 6\lambda-2 \end{vmatrix}=0$$

$$\sim (\lambda-3)\begin{vmatrix} \lambda-1 & 5\lambda+1 \\ 2 & 6\lambda-2 \end{vmatrix}=0$$

$$\sim \; 2(\lambda-3)\;[(\lambda-1)(3\lambda-1)-(5\lambda+1)]=0$$

or

$$6\lambda(\lambda-3)^2=0, \text{ which yields } \lambda=0 \text{ or } \lambda=3.$$

When $\lambda=0$, the given system of equations reduces to

$$-x+y=0,$$
$$-x-2y+3z=0,$$
$$2x+y-3z=0.$$

The last two equations yield

$$\frac{x}{3}=\frac{y}{3}=\frac{z}{3} \text{ and so } x=y=z.$$

When $\lambda=3$, all the three equations become identical.

(b) The given system of equation is

$$x+y+3z=0,$$
$$2x+y+2z=0,$$
$$4x+3y+bz=0.$$

The system in matrix form is

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 2 \\ 4 & 3 & b \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This homogenous system will have a non-trivial solution only if $|A|=0$. Thus for non-trivial solution

$$\begin{vmatrix} 1 & 1 & 3 \\ 2 & 1 & 2 \\ 4 & 3 & b \end{vmatrix}=0$$

or

$$1(b-6)-1(2b-8)+3(6-4)=0$$

or

$$-b+8=0, \text{ which yields } b=8.$$

Thus for non-trivial solution, $b=8$. The coefficient matrix for non-trivial solution is

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 2 \\ 4 & 3 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 3 \\ 0 & -1 & -4 \\ 0 & -1 & -4 \end{bmatrix} \begin{matrix} R_2 \to R_2-2R_1 \\ R_3 \to R_3-4R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 3 \\ 0 & -1 & -4 \\ 0 & 0 & 0 \end{bmatrix} R_3 \to R_3-R_2.$$

The last matrix is of rank 2. Thus the given system is equivalent to

$$x+y+3z=0$$
$$-y-4z=0.$$

Hence $y=-4z$ and then $x=z$. Taking $z=t$ the general solution is

$$x=t,\; y=-4t,\;\; z=t.$$

EXAMPLE 4.79

Prove that the sum of the eigenvalues of a matrix A is the sum of the elements of the principal diagonal.

Solution. If $A=[a_{ij}]$ be the matrix of order n, then the characteristic equation of the matrix A is

$$|A-\lambda I|=\lambda^n-\lambda^{n-1}\left(\sum_{i=1}^{n} a_{ii}\right)+\ldots\ldots=0.$$

Form the theory of equations, the sum of the roots $\lambda_1, \lambda_2, ..., \lambda_n$ is equal to negative of the coefficient of λ^{n-1}. Hence

$$\lambda_1 + \lambda_2 + \ ... \ + \lambda_n = a_{11} + a_{22} + \ ... \ + a_{nn}$$
$$= \text{Trace } A.$$

EXAMPLE 4.80

(a) Find the eigen values of A^{-1} if the matrix A is $\begin{bmatrix} 2 & 5 & -1 \\ 0 & 3 & 2 \\ 0 & 0 & 4 \end{bmatrix}$.

(b) Find the eigenvalues and the corresponding vectors of the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}.$$

Solution. (a) By example 4.57, the eigenvalues of triangular matrix are the diagonal elements. Hence the eigenvalues of A are 2, 3 and 4. Since the eigenvalues of A^{-1} are multiplicative inverses of the eigenvalues of the matrix A, the eigenvalues of A^{-1} are $\frac{1}{2}, \frac{1}{3}$ and $\frac{1}{4}$.

(b) We have

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}.$$

The characteristic equation of A is

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & 2-\lambda & 1 \\ 2 & 0 & 3-\lambda \end{vmatrix} = 0.$$

or

$$\lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0,$$

which yields $\lambda = 1, \ 2, \ 3$. Hence the characteristic roots are 1, 2 and 3.

The eigenvector corresponding to $\lambda = 1$ is given by $(A - I)X = 0$, that is, by

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Thus, we have

$$x_2 + x_3 = 0,$$
$$2x_1 + 2x_3 = 0.$$

Hence $x_1 = x_2 = -x_3$. Taking $x_3 = -1$, we get the vector

$$X_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

The eigenvector corresponding to the eigenvalue 2 is given by $(A - 2I)X = 0$, that is, by

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This equation yields

$$X_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ as one of the vector.}$$

Similarly, the eigenvector corresponding to $\lambda = 3$ is given by $(A - 3I)X = 0$ or by

$$\begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 1 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which yields $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ as one of the solution. Hence $X_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

EXAMPLE 4.81

Find the sum and product of the eigen values of the matrix:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 5 & 6 \\ 7 & 4 & 3 & 2 \\ 4 & 3 & 0 & 5 \end{bmatrix}$$

Solution. The given matrix is

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 5 & 6 \\ 7 & 4 & 3 & 2 \\ 4 & 3 & 0 & 5 \end{bmatrix}.$$

The sum of the eigenvalues is the trace (spur) of the matrix and so the sum is $1 + 1 + 3 + 5 = 10$.

Product of the eigenvalues is equal to $|A|$. Expanding $|A|$, we get the product as 262.

EXAMPLE 4.82

One of the eigenvalues of $\begin{bmatrix} 7 & 4 & -4 \\ 4 & -8 & -1 \\ 4 & -1 & -8 \end{bmatrix}$ *is* -9.

Find the other two eigenvalues.

Solution. The characteristic equation of the given matrix is

$$|A-\lambda I| = \begin{vmatrix} 7-\lambda & 4 & -4 \\ 4 & -8-\lambda & -1 \\ 4 & -1 & -8-\lambda \end{vmatrix} = 0.$$

or

$$(7-\lambda)\left[\lambda^2+16\lambda+63\right]-4[-28-4\lambda] - 4[28+4\lambda]=0$$

or

$$\lambda^3+9\lambda^2-49\lambda-441=0.$$

Clearly $\lambda=-9$ satisfies this equation. Then, by synthetic division, the reduced equation is

$$\lambda^2-49=0,$$

Which yields $\lambda=\pm 7$. Thus the eigenvalues of the given matrix are $-9, 7, -7$. The sum of the eigenvalues is $-9+7-7=-9$, which is equal to the trace of the given matrix.

EXAMPLE 4.83

Verify that the following set of vectors in $\mathbb{R}^3$ is linearly dependent: $(1,0,1),(1,1,1),(1,1,2)$ and $(1,2,1)$: Also find the number of linearly independent vectors.

Solution. The vectors is $\mathbb{R}^3$ are given to be

$$v_1=(1,\ 0,\ 1),\ v_2=(1,\ 1,\ 1),$$
$$v_3=(1,\ 1,\ 2),\ v_4=(1,\ 2,\ 1).$$

Let

$$\lambda_1v_1+\lambda_2v_2+\lambda_3v_3+\lambda_4v_4=0, \qquad (29)$$

This gives

$$\lambda_1(1,\ 0,\ 1)+\lambda_2(1,\ 1,\ 1)+\lambda_3(1,\ 1,\ 2) + \lambda_4(1,\ 2,\ 1)=0$$

or

$$\lambda_1+\lambda_2+\lambda_3+\lambda_4=0$$
$$0\lambda_1+\lambda_2+\lambda_3+2\lambda_4=0$$
$$\lambda_1+\lambda_2+2\lambda_3+\lambda_4=0.$$

We have four variable and three equations. Thus there is one degree of freedom. We have

$$\frac{\lambda_1}{-1}=\frac{\lambda_2}{2}=\frac{\lambda_3}{0}=\frac{\lambda_4}{1}=k.$$

Therefore

$$\lambda_1=-k,\ \lambda_2=2k,\ \lambda_3=0,\ \lambda_4=-k.$$

Putting these values of λ_i is (29), we get

$$-kv_1+2kv_2+0v_3-kv_4=0$$

or

$$v_1-2v_2+0v_3+v_4=0.$$

Thus (29) is satisfied for $\lambda_1=1,\ \lambda_2=-2,\ \lambda_3=0$ and $\lambda_4=1$. Since not all of λ_i are zeros, it follows that v_1, v_2, v_3, v_4 are linearly dependent.

EXAMPLE 4.84

What do you mean by an orthogonal matrix? Verify that the following matrix is orthogonal:

$$\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

Solution. A square matrix P is said to be orthogonal if $P^TP=PP^T=I$. If

$$P=\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix},$$

then

$$PP^T=\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$$
$$=\begin{bmatrix} \cos^2\theta+\sin^2\theta & 0 & \begin{matrix}-\sin\theta\cos\theta \\ +\sin\theta\cos\theta\end{matrix} \\ 0 & 1 & 0 \\ \begin{matrix}-\sin\theta\cos\theta \\ +\sin\theta\cos\theta\end{matrix} & 0 & \cos^2\theta+\sin^2\theta \end{bmatrix}$$
$$=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}=I.$$

Hence P is orthogonal.

EXAMPLE 4.85

Show that the transformation

$$y_1=\frac{1}{3}x_1+\frac{2}{3}x_2+\frac{2}{3}x_3,$$
$$y_2=\frac{2}{3}x_1+\frac{1}{3}x_2-\frac{2}{3}x_3,$$
$$y_3=\frac{2}{3}x_1-\frac{2}{3}x_2+\frac{1}{3}x_3 \text{ is orthogonal.}$$

Solution. In matrix form, we have

$$Y=PX,$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \; P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \text{ and}$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

The transformation $Y = PX$ will be orthogonal if $P^T P = I$. To show it, we observe that

$$P^T P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I.$$

Hence $Y = PX$ is orthogonal.

EXAMPLE 4.86

Diagonalise the matrix $A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$ through an orthogonal transformation.

Solution. We shall proceed as in Example 4.66. The characteristic equation of A is

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 0 & 1 \\ 0 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix} = 0$$

or

$$(2-\lambda)(3-\lambda)(2-\lambda) - (3-\lambda) = 0$$
$$(3-\lambda)[(2-\lambda)^2 - 1] = 0$$
$$(3-\lambda)[4 + \lambda^2 - 4\lambda - 1] = 0$$
$$(3-\lambda)(\lambda^2 - 4\lambda + 3] = 0$$
$$\Rightarrow \lambda = 3, \;\; 3, \;\; 1.$$

The ch. vector corresponding to $\lambda = 3$ is given by

$$\left.\begin{aligned} -x_1 + x_3 = 0 \\ x_1 - x_3 = 0 \end{aligned}\right\} \qquad (30)$$

Thus

$$X_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Let $X_2 = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ be another eigenvector of A corresponding to the eigenvalue 3 and orthogonal to X_1. Then

$x - y = 0$, because it satisfies equation (30)

and

$$x + z = 0, \text{ using } X_2^\theta X_1 = 0.$$

Obviously $x = 1$, $z = -1$ is a solution. Therefore

$$X_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

Further eigenvector corresponding to $\lambda = 1$ is given by $(A - I)X = 0$, that is , by

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0.$$

This equation yields

$$x_1 + x_3 = 0, \quad 2x_2 = 0, \quad x_1 + x_3 = 0.$$

Thus $x_1 = 1$, $x_2 = 0$, $x_3 = -1$ and so

$$X_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Length (norm) of the vectors X_1, X_2, X_3 are respectively $\sqrt{2}, \sqrt{3}, \sqrt{2}$. Hence the orthogonal matrix is

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

and

$$P^T AP = dig[3 \quad 3 \quad 1].$$

EXAMPLE 4.87

(a) Show that the matrix $A = \begin{bmatrix} 3 & 1 & -1 \\ -2 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$, is diagonalizable. Hence, find P such that $P^{-1}AP$ is a diagonal matrix. Also obtain the matrix $B = A^2 + 5A + 3I$.

(b) Find a matrix P which diagonalizes the matrix $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$, verify $P^{-1}AP = D$, where D is the diagonal matrix.

(c) Show that the matrix $A = \begin{bmatrix} 3 & 1 & -1 \\ -2 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$ is diagonalizable. Hence, find P such that $P^{-1}AP$ is a diagonal matrix.

Solution. (a) The characteristic equation for the given matrix A is

$$|A - \lambda I| = \begin{vmatrix} 3-\lambda & 1 & -1 \\ -2 & 1-\lambda & 2 \\ 0 & 1 & 2-\lambda \end{vmatrix} = 0,$$

that is,

$$(3-\lambda)(\lambda^2 - 3\lambda) + 4 - 2\lambda + 2 = 0$$

or

$$\lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0.$$

By inspection, $\lambda = 1$ is a root. The reduced equation is $\lambda^2 - 5\lambda + 6 = 0$, which yields $\lambda = 2, 3$.

Since all characteristic roots are distnict, the given matrix A is diagonalizable. To find the non-singular matrix P satisfying $P^{-1}AP = \text{diag}(1, 2, 3)$, we proceed as follows:

The characteristic vectors are given by $(A - \lambda I)X = 0$, that is, by

$$\begin{bmatrix} 3-\lambda & 1 & -1 \\ -2 & 1-\lambda & 2 \\ 0 & 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and so

$$\left.\begin{aligned} (3-\lambda)x_1 + x_2 - x_3 &= 0 \\ -2x_1 + (1-\lambda)x_2 + 2x_3 &= 0 \\ 0x_1 + x_2 + (2-\lambda)x_3 &= 0 \end{aligned}\right\} \qquad (31)$$

For $\lambda = 1$, we get

$$2x_1 + x_2 - x_3 = 0$$

$$-2x_1 + 2x_3 = 0$$

$$x_2 + x_3 = 0$$

We note that $x_1 = 1$, $x_2 = -1$, $x_3 = 1$ satisfy these equations. Hence the eigenvector corresponding to $\lambda = 1$ is $[1 \ -1 \ 1]^T$.

For $\lambda = 2$, we get from (31),

$$x_1 + x_2 - x_3 = 0$$

$$-2x_1 - x_2 + 2x_3 = 0$$

$$0x_1 + x_2 = 0$$

Clearly $x_1 = 1$, $x_2 = 0$, $x_3 = 1$ is a solution. Hence the eigenvector corresponding to $\lambda = 2$ is $[1 \ 0 \ 1]^T$.

For $\lambda = 3$, we have

$$x_2 - x_3 = 0$$

$$-2x_1 - 2x_2 + 2x_3 = 0$$

$$x_2 - x_3 = 0$$

Taking $x_3 = 1$, we get $x_2 = 1$ and $x_1 = 1$. Thus the eigenvector corresponding to $\lambda = 3$ is $[0 \ 1 \ 1]^T$. Hence

$$P = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

and

$$P^{-1}AP = \text{diag}[1 \quad 2 \quad 3] = D, \text{ say} \qquad (32)$$

Premultiplication by P and postmultiplication by P^{-1} reduces (32) to

$$A = PDP^{-1}.$$

Further,

$$A^n = PD^nP^{-1}.$$

Thus

$$A^2 = PD^2P^{-1}.$$

But

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad D^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } P^{-1} = \begin{bmatrix} -1 & -1 & 1 \\ 2 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

Putting these values in $B = A^2 + 5A + 3I$, we get

$$B = A^2 + 5A + 3I = \begin{bmatrix} 25 & 8 & -8 \\ -18 & 9 & 18 \\ -2 & 8 & 19 \end{bmatrix}.$$

(b) We have

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$$

The characteristic equation of A is

$$|A-\lambda I| = \begin{vmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{vmatrix} = 0$$

or

$$(4-\lambda)(3-\lambda)-2=0$$

or

$$\lambda^2-7\lambda+10=0$$

The characteristic roots are $\lambda = \frac{7\pm 3}{2} = 2,\ 5$. Since the eigenvalues are distinct, the matrix A is diagonalizable. The eigenvector corresponding to $\lambda = 2$ is given by $(A-2I)X=0$, that is ,by

$$\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or by

$$2x_1+x_2=0 \quad \text{or} \quad x_1 = -\frac{x_2}{2}.$$

Putting $x_2 = 2$, we get

$$X_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Similarly, eigenvector corresponding to $\lambda = 5$ is given by $(A-5I)X=0$ or by

$$\begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

or by

$$-x_1+x_2=0$$
$$2x_1-2x_2=0$$

and so $x_1 = x_2$. Putting $x_2 = 1$, we get

$$X_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Thus the transforming matrix is

$$P = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}$$

and

$$P^{-1} = \begin{bmatrix} -\frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Then

$$P^{-1}AP = \begin{bmatrix} -\frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} -\frac{2}{3} & \frac{2}{3} \\ \frac{10}{3} & \frac{5}{3} \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}.$$

(c) The characteristic matrix of the given matrix A is

$$|A-\lambda I\,| = \begin{vmatrix} 3-\lambda & 1 & -1 \\ -2 & 1-\lambda & 2 \\ 0 & 1 & 2-\lambda \end{vmatrix} = 0$$

or

$$(3-\lambda)[(1-\lambda)(2-\lambda)-2]-l[-4+2\lambda]-l(-2)$$
$$=0$$

or

$$(3-\lambda)(1-\lambda)(2-\lambda)-6+2\lambda+4-2\lambda+2=0$$

or

$$(3-\lambda)(1-\lambda)(2-\lambda)=0.$$

Hence the given matrix A has distinct characteristic roots $\lambda = 1, 2, 3$. Consequently it is diagonalizable. Now the eigenvector corresponding to $\lambda = 1$ is given by $(A-I)X=0$, that is, by

$$\begin{bmatrix} 2 & 1 & -1 \\ -2 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus

$$2x_1+x_2-x_3=0$$
$$-2x_1+0x_2+2x_3=0$$
$$0x_1+x_2+x_3=0$$

and so $x_1 = x_3 = -x_2$. Taking $x_2 = -1$, we get an eigenvector corresponding to $\lambda = 1$ as

$$X_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Now eigenvector corresponding to $\lambda = 2$ is given by $(A-2I)X=0$, that is, by

$$\begin{bmatrix} 1 & 1 & -1 \\ -2 & -1 & 2 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus

$$x_1 + x_2 - x_3 = 0$$
$$-2x_1 - x_2 + 2x_3 = 0$$
$$x_2 = 0$$

For this system $x_1 = 1,\ x_2 = 0,\ x_3 = 1$ is a solution. Therefore

$$X_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

An eigenvector corresponding to $\lambda = 3$ is given by $(A - 3I)X = 0$, that is, by

$$\begin{bmatrix} 0 & 1 & -1 \\ -2 & -2 & 2 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus

$$x_2 - x_3 = 0$$
$$-2x_1 - 2x_2 + 2x_3 = 0$$
$$x_2 - x_3 = 0,$$

which yields $x_1 = 0,\ x_2 = 1,\ x_3 = 1$ as one of the solution. Thus

$$X_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Therefore the transforming matrix is

$$P = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

and so the diagonal matrix is

$$P^{-1}AP = \begin{bmatrix} -1 & -1 & 1 \\ 2 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & -1 \\ -2 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$
$$\times \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

EXAMPLE 4.88

Reduce the quadratic form $x^2 + 5y^2 + z^2 + 2xy + 6zx + 2yz$ to a canonical form through an Orthogonal transformation.

Solution. The given quadratic form can be written as

$$x^2 + xy + yx + 5y^2 + yz + yz + z^2 + 3zx + 3xz.$$

The matrix of the quadratic form is

$$A = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix}.$$

Write $A = IAI$, that is,

$$\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Using congruent operations $R_2 \to R_2 - R_1,\ C_2 = C_2 - C_1$ and $R_3 \to R_3 - 3R_1,\ C_3 \to C_3 - C_1$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & -2 \\ 0 & -2 & -8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now performing congruent operation $R_3 \to R_3 + \frac{1}{2}R_3,\ C_3 \to C_3 + \frac{1}{2}C_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -\frac{7}{2} & \frac{1}{2} & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & -\frac{7}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus

$$\text{diag}\ [1\ \ 4\ \ -9] = P^{-1}AP,$$

where

$$P = \begin{bmatrix} 1 & -1 & -\frac{7}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence the required canonical form is

$$x^2 + 4y^2 - 9z^2.$$

EXAMPLE 4.89

Reduce the quadratic form $x^2 + y^2 + z^2 - 2xy - 2yz - 2zx$ to canonical form through an orthogonal transformation.

Solution. The matrix of the given quadratic form is

$$A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

The characteristic equation of A is

$$|A-\lambda I| = \begin{vmatrix} 1-\lambda & -1 & -1 \\ -1 & 1-\lambda & -1 \\ -1 & -1 & 1-\lambda \end{vmatrix} = 0$$

or

$$(1-\lambda)\left[\lambda^2 - 2\lambda\right] + \lambda - 2 - 2 + \lambda = 0$$

or

$$\lambda^3 - 3\lambda^2 + 4 = 0,$$

which yields $\lambda = -1,\ 2, 2$.

The eigenvectors will be given by $(A-\lambda I)X = 0$, which implies

$$\begin{bmatrix} 1-\lambda & -1 & -1 \\ -1 & 1-\lambda & -1 \\ -1 & -1 & 1-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$(1-\lambda)\,x - y - z = 0,$$
$$-x + (1-\lambda)y - z = 0 \quad \text{and}$$
$$-x - y + (1-\lambda)z = 0.$$

For $\lambda = -1$, we have

$$2x - y - z = 0, -x + 2y - z = 0 \quad \text{and}$$
$$-x - y + 2z = 0.$$

Solving these equations, we get $x = y = z = 1$ and so the eigenvector is $[1\ 1\ 1]^T$. Its normalized form is $\left[\frac{1}{\sqrt{3}}\ \ \frac{1}{\sqrt{3}}\ \ \frac{1}{\sqrt{3}}\right]^T$.

Corresponding to $\lambda = 2$, we have $-x-y-z = 0$, $-x-y-z = 0$ and $-x-y-z = 0$. We note that $x = -2, y = 1, z = 1$ is a solution. Thus the eigenvector is $[-2\ 1\ 1]^T$. and its normalized form is $\left[\frac{-2}{\sqrt{6}},\ \frac{1}{\sqrt{6}},\ \frac{1}{\sqrt{6}}\right]^T$. To find the second vector, we have

$$-x - y - z = 0$$

and

$$-2x + y + z = 0 \quad \text{using} \quad X_2^{\theta} X_1 = 0.$$

We note that $x = 0$, $y = -1$ and $z = 1$ is a solution. The normalized vector is $\left[0,\ \frac{-1}{\sqrt{2}}\ \ \frac{1}{\sqrt{2}}\right]$.

Hence

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{-2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

and

$$P^T A P = \text{diag}[-1\ 2\ 2].$$

EXERCISES

1. Show that the subset $\{x^2-1,\ x+1,\ x-1\}$ of the vector space of polynomials is linearly independent.
2. Show that the subset $\{(1, 1, 1, 0), (3, 2, 2, 1), (1, 1, 3, -2), (1, 2, 6, -5), (1, 1, 2, 1)\}$ of V_4 is linearly dependent.
3. Show that the subset $\{(0, 0, 1), (1, 0, 1), (1, -1, 1), (3, 0, 1)\}$ is not a basis for V_3.
4. Show that the subset $(1, x, (x-1)x, x(x-1)(x-2)\}$ form a basis for vector space of polynomials of degree 3.
5. Show that

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

form a linearly independent set and describe its linear span geometrically.

Solution. We note that

$$\alpha e_1 + \beta e_2 = 0$$

implies

$$\alpha \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

implies

$$\begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \beta \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

implies

$$\begin{bmatrix} \alpha \\ \beta \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Consequently, $\alpha = \beta = 0$. Hence $\{e_1\ e_2\}$ form a linearly independent set.

Further, the linear span of $\{e_1\ e_2\}$ is the set of all vectors of the form $\begin{bmatrix} \alpha \\ \beta \\ 0 \end{bmatrix}$ which is nothing but (x, y) plane and is a subset of the three dimensional Euclidean space.

6. Show that the vectors

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}, \text{ and } v_3 = \begin{bmatrix} 2 \\ -1 \\ 8 \end{bmatrix}$$

are linearly dependant.

Solution. We note that the relation.

$$\alpha_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} + \alpha_3 \begin{bmatrix} 2 \\ -1 \\ 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

is satisfied if we choose $\alpha_1 = -2, \alpha_2 = -3$ and $\alpha_3 = 1$. Thus $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$ is satisfied by $\alpha_1, \alpha_2, \alpha_3$, where not all of these scalars are zero. Hence the set $\{v_1, v_2, v_3\}$ is linearly dependent.

7. Use the principle of mathematical induction to show that if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, then $A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$ for every positive integer n.

8. Express the matrix

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & -1 & 3 \\ 4 & 6 & 5 \end{bmatrix}$$

as the sum of a symmetric matrix and a Skew symmetric matrix.

Ans. $\begin{bmatrix} 1 & \frac{5}{2} & \frac{9}{2} \\ \frac{5}{2} & -1 & \frac{9}{2} \\ \frac{9}{2} & \frac{9}{2} & 5 \end{bmatrix} + \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{3}{2} \\ -\frac{1}{2} & \frac{3}{2} & 0 \end{bmatrix}$

9. Find the adjoint of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -3 \\ 2 & -1 & 3 \end{bmatrix}$$

and verify the result $A(\text{adj } A) = (\text{adj } A)A = |A|I_n$.

Ans. adj $A = \begin{bmatrix} 3 & -4 & -5 \\ -9 & 1 & 4 \\ -5 & 3 & 1 \end{bmatrix}$

10. Find the inverse of the following matrices:

(i) $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$

(ii) $B = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(iii) $C = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix}$

(iv) $D = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$

(v) $E = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & -3 \\ -2 & -4 & -4 \end{bmatrix}$.

Ans. (i) $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 3 & -1 \\ \frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$, (ii) $\begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

(iii) $\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix}$, (iv) $\begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$,

(v) $-\frac{1}{8}\begin{bmatrix} -24 & -8 & -12 \\ 10 & 2 & 6 \\ 2 & 2 & 2 \end{bmatrix}$

11. Using Gauss Jordan method, find the inverse of the following matrices:

(i) $A = \begin{bmatrix} 3 & -3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{bmatrix}$

(ii) $B = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$

(iii) $C = \begin{bmatrix} 3 & -2 & -1 \\ -4 & 1 & -1 \\ 2 & 0 & 1 \end{bmatrix}$

(iv) $D = \begin{bmatrix} 14 & 3 & -2 \\ 6 & 8 & -1 \\ 0 & 2 & -7 \end{bmatrix}$

Ans. (i) $\begin{bmatrix} 1 & -1 & 0 \\ -2 & 3 & -4 \\ -2 & 3 & -3 \end{bmatrix}$ (ii) $\begin{bmatrix} -1 & 1 & 2 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ -2 & -4 & -5 \end{bmatrix}$ (iv) $-\frac{1}{654}\begin{bmatrix} -54 & 17 & 13 \\ 42 & -98 & 2 \\ 12 & -28 & 94 \end{bmatrix}$

12. Find the rank of the following matrices.

(i) $\begin{bmatrix} 0 & 1 & -3 & -1 \\ 0 & 0 & 1 & 1 \\ 3 & 1 & 0 & 2 \\ 1 & 1 & -2 & 0 \end{bmatrix}$ (ii) $\begin{bmatrix} 2 & 3 & -1 & -1 \\ 1 & -1 & -2 & -4 \\ 3 & 1 & 3 & -2 \\ 6 & 3 & 0 & -7 \end{bmatrix}$

(iii) $\begin{bmatrix} 2 & -1 & 3 & 4 \\ 0 & 3 & 4 & 1 \\ 2 & 3 & 7 & 5 \\ 2 & 5 & 11 & 6 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 3 & 2 & 2 \\ 2 & 4 & 3 & 4 \\ 3 & 7 & 4 & 6 \end{bmatrix}$

(v) $\begin{bmatrix} 1 & 3 & 4 & 3 \\ 3 & 9 & 12 & 9 \\ 1 & 3 & 4 & 1 \end{bmatrix}$

Ans. (i) 3, (ii) 3, (iii) 3, (iv) 3, (v) 2

13. Show that no Skew-symmetric matrix can be of rank 1.

Hint: Diagonal elements are all zeros. If all non-diagonal positive elements are zero, then the corresponding negative elements are also zero and so rank shall be zero, If at least one of the elements is non-zero, then at least one 2-rowed minor is not equal to zero. Hence, rank is greater than or equal to 2.

14. Reduce the following matrices to normal form and, hence, find their ranks.

(i) $\begin{bmatrix} 5 & 3 & 14 & 4 \\ 0 & 1 & 2 & 1 \\ 1 & -1 & 2 & 0 \end{bmatrix}$

(ii) $\begin{bmatrix} 1 & 2 & 1 & 0 \\ -2 & 4 & 3 & 0 \\ 1 & 0 & 2 & -8 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix}$

(iv) $\begin{bmatrix} 2 & -2 & 0 & 6 \\ 4 & 2 & 0 & 2 \\ 1 & -1 & 0 & 3 \\ 1 & -2 & 1 & 2 \end{bmatrix}$

Ans. (i) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, Rank3 (ii) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, Rank3

(iii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ Rank2; (iv) $\begin{bmatrix} I_3 & 0 \\ 0 & 0 \end{bmatrix}$ Rank3

15. Find the inverse of the matrix

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

using elementary operations

Ans. $\begin{bmatrix} -1 & 1 & 2 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}$

16. Using elementary transformation, find the inverse of the matrix

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

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

17. Test for consistency and solve the following system of equations.

(i) $x + y + z = 6$
$x + 2y + 3z = 14$
$x + 4y + 7z = 30.$

Ans. Consistent; rank 2; $x = c - 2$, $y = 8 - 2c$, $z = c$ for arbitrary constant c

(ii) $2x + 6y + 11 = 0$
$6x + 20y + 6z + 3 = 0$
$6y - 18z + 1 = 0$

Ans. Not consistent

(iii) $2x - y + 3z = 8$
$-x + 2y + z = 4$
$3x + y - 4z = 0$

Ans. Consistent, $x = 2$, $y = 2$, $z = 2$

18. Find the values of λ for which the following system of linear equations will have no unique solution

$$3x - y + \lambda z = 1$$
$$2x + y + z = 2$$
$$x + 2y - \lambda z = 1$$

Will this system have any solution for these values of λ?

Ans. $\lambda = -\frac{7}{2}$ and the equations are inconsistent for this value. Hence no solution exists for $\lambda = -\frac{7}{2}$

19. Discuss the existence and nature of solutions for all values of λ for the following system of equations.

$$x + y + 4z = 6$$
$$x + 2y - 2z = 6$$
$$\lambda x + y + z = 6$$

Ans. Unique Solution for $\lambda \neq \frac{7}{10}$. For $\lambda = \frac{7}{10}$, the equations are not consistent.

20. Solve the following equations using matrix method

$$2x - y + 3z = 9$$
$$x + y + z = 6$$
$$x - y + z = 2$$

Ans. Coefficient matrix is non-singular. The unique solution is $x = 1$, $y = 2$, $z = 3$

21. Determine the values of a and b for which the equations

$$x + 2y + 3z = 4$$
$$x + 3y + 4z = 5$$
$$x + 3y + az = b$$

have (i) no solution, (ii) a unique solution, and (iii) an infinite number of solution.

Ans. (i) $a = 4, b \neq 5$ (ii) $a \neq 4$ (iii) $a = 4, b = 5$

22. Solve completely the system of equation

$$x + y + z = 0$$
$$2x - y - 3z = 0$$
$$3x - 5y + 4z = 0$$
$$x + 17y + 4z = 0$$

Ans. Trivial solution $x = y = z = 0$.

23. Solve

$$4x + 2y + z + 3u = 0$$
$$6x + 3y + 4z + 7u = 0$$
$$2x + y + u = 0$$

Ans. $x = c_1$, $u = c_2$, $y = -2c_1 - c_2$, $z = -c_2$

24. Find the eigenvalues of the matrix

$$\begin{bmatrix} a & h & g \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

Ans. a, b, c

25. Find the eigenvalues and the corresponding eigenvectors for the given matrix.

$$\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$$

Ans. $0, 3, 15, c_1 = \begin{bmatrix} \frac{1}{2} \\ 1 \\ 1 \end{bmatrix}, c_2 = \begin{bmatrix} -4 \\ -2 \\ 4 \end{bmatrix}, c_3 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$

26. If the characteristic roots of a matrix A are λ_1, $\lambda_2, \ldots, \lambda_n$, show that the characteristic roots of A^2 are $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$.

Hint: $AX = \lambda X \Rightarrow A(AX) = \lambda(AX) \Rightarrow A^2X = \lambda(\lambda X) = \lambda^2 X$

27. Show that the matrices A and $B^{-1}AB$ have the same characteristic roots

Hint: $\left|B^{-1}AB - \lambda I\right| = \left|B^{-1}AB - B^{-1}\lambda IB\right|$

$= \left|B^{-1}(A - \lambda I)B\right|$

$= \left|B^{-1}\right|\left|A - \lambda I\right|\left|B\right|$

$= \left|A - \lambda I\right|\left|B^{-1}B\right|$

$= \left|A - \lambda I\right|.$

28. Verify Caley-Hamilton theorem for the matrix

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$$

and, hence, find its inverse.

Ans. A satisfies $A^3 - 6A^2 + 7A + 2I = 0$

$$A^{-1} = \frac{1}{2}\left(A^2 - 6A + 7I\right) = \begin{bmatrix} -3 & 0 & 2 \\ -1 & \frac{1}{2} & \frac{1}{2} \\ 2 & 0 & -1 \end{bmatrix}.$$

29. Find the minimal polynomial of the matrix.

$$\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$$

Ans. $x^2 - 3x + 2.$

30. Show that the matrix

$$\begin{bmatrix} a + ic & -b + id \\ b + id & a - ic \end{bmatrix}$$

is unitary if and only if $a^2 + b^2 + c^2 + d^2 = 1$.

31. Show that the matrix

$$A = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

is orthogonal. **Hint:** Show that $A^T A = I$.

32. Show that the matrix

$$\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}$$

is diagonalizable. Obtain the diagonalizing matrix P.

Ans. $P = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$ diag[−1,−1,3]

33. Diagonalize the matrix

$$\begin{bmatrix} 1 & -6 & -4 \\ 0 & 4 & 2 \\ 0 & -6 & -3 \end{bmatrix}.$$

Ans. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

34. Diagonalize the real-symmetric matrix

$$A = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}$$

Ans. $P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix}$, diag [2 3 6]

35. Find a non-singular matrix P such that P^TAP is a diagonal matrix, where

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{bmatrix}$$

Ans. $\begin{bmatrix} 0 & -\frac{1}{2} & -3 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{bmatrix}$

36. Reduce the quadratic form $x^2+4y^2 + 9z^2 + t^2 - 12yz + 6zx = 4xy - 2xt - 6zt$ to canonical form and find its rank and signature

Ans. $y_1^2 - y_2^2 + y_4^2$, Rank:3, Signature:1

37. Reduce the quadratic form $6x_1^2 + 3x_2^2 + 14x_3^2 + 4x_2x_3 + 18x_3x_1 + 4x_1x_2$ to canonical form and find its rank and signature.

Ans. $y_1^2 + y_2^2 + y_3^2$, Rank:3, Signature:3

UNIT III

Integral Calculus

5 Beta and Gamma Functions

The beta and gamma functions, also called Euler's Integrals, are the improper integrals, which are extremely useful in the evaluation of integrals.

5.1 BETA FUNCTION

The integral $\int_0^1 x^{m-1}(1-x)^{n-1}dx$, which converges for $m > 0$ and $n > 0$ is called the *beta function* and is denoted by β (m, n). Thus,

$$\beta(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx, m > 0, n > 0.$$

Beta function is also known as *Eulerian Integral of First Kind*.

As an illustration, consider the integral $\int_0^1 x^{\frac{1}{2}}(1-x)^4 dx$. We can write this integral as

$$\int_0^1 x^{\frac{3}{2}-1}(1-x)^{5-1}dx,$$

which is a beta function, denoted by $\beta\left(\frac{3}{2}, 5\right)$. But, on the other hand, the integral $\int_0^1 x^{\frac{1}{2}}(1-x)^{-3}dx$ is *not* a beta function as $n-1 = -3$ implies $n = -2$ (negative).

5.2 PROPERTIES OF BETA FUNCTION

1. We have

$$\beta(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx$$

$$= \int_0^1 (1-x)^{m-1}(1-(1-x))^{n-1}dx, \text{ since}$$

$$\int_0^a f(x)dx = \int_0^a f(a-x)dx,$$

$$= \int_0^1 x^{n-1}(1-x)^{m-1}dx$$

$$= \beta(n,m), m > 0, \ n > 0.$$

Thus,

$$\beta(m,n) = \beta(n,m), \ m > 0, n > 0.$$

2. We have

$$\beta(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx. \qquad (1)$$

Putting $x = \sin^2\theta$, we get $dx = 2\sin\theta\cos\theta\, d\theta$ and therefore, (1) becomes

$$\beta(m,n) = \int_0^{\frac{\pi}{2}} (\sin^2\theta)^{m-1}(1-\sin^2\theta)^{n-1}$$

$$.2\sin\theta\cos\theta\ d\theta$$

$$= 2\int_0^{\frac{\pi}{2}} \sin^{2m-2}\theta\cos^{2n-2}\theta\sin\theta\cos\theta\ d\theta$$

$$= 2\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\cos^{2n-1}\theta\ d\theta.$$

Thus,

$$\beta(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx$$

$$= 2\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\cos^{2n-1}\theta\ d\theta.$$

3. Let m and n be positive integers. By definition,

$$\beta(m,n)=\int_0^1 x^{m-1}(1-x)^{n-1}dx,\ m>0,\ n>0.$$

Integration by parts yields

$$\beta(m,n)=\left[x^{m-1}\frac{(1-x)^n}{n(-1)}\right]_0^1 - \int_0^1 (m-n)x^{m-2}\left[\frac{(1-x)^n}{n(-1)}\right]dx$$

$$=\frac{m-1}{n}\int_0^1 x^{m-2}(1-x)^n dx$$

$$=\frac{m-1}{n}\beta(m-1,n+1).$$

Similarly,

$$\beta(m-1,n+1)=\frac{m-2}{n+1}\beta(m-2,n+2)$$

$$\beta(2,m+n-2)=\frac{1}{m+n-2}\beta(1,\ m+n-1).$$

Multiplying the preceding equations, we get

$$\beta(m,n)=\frac{(m-1)(m-2).....(2)(1)}{n(n+1)(n+2).....(m+n-2)}\times\beta(1,m+n-1)$$

$$=\frac{(m-1)!\,[1.2.3....(n-1)]}{1.2....(n-1)n(n+1)(n+2)...(m+n-2)}\times\int_0^1 x^{1-1}(1-x)^{m+n-2}dx$$

$$=\frac{(m-1)!\ (n-1)!}{(m+n-2)!}\int_0^1 (1-x)^{m+n-2}dx$$

$$=\frac{(m-1)!\ (n-1)!}{(m+n-2)!}\left[\frac{(1-x)^{m+n-1}}{(m+n-1)(-1)}\right]_0^1$$

$$=\frac{(m-1)!\ (n-1)!}{(m+n-2)!}\cdot\frac{1}{m+n-1}$$

$$=\frac{(m-1)!\ (n-1)!}{(m+n-1)!}.$$

Hence, if m and n are positive integers, then

$$\beta(m,n)=\frac{(m-1)!(n-1)!}{(m+n-1)!}.$$

4. Put $x=\frac{t}{1+t}$ so that $dx=\frac{1}{(1+t)^2}dt$. Then the expression for beta function reduces to

$$\beta(m,n)=\int_0^\infty\left(\frac{t}{1+t}\right)^{m-1}\left(1-\frac{t}{1+t}\right)^{n-1}\cdot\frac{1}{(1+t)^2}dt$$

$$=\int_0^\infty\frac{t^{m-1}}{(1+t)^{m-1}}\cdot\frac{1}{(1+t)^{n-1}}\cdot\frac{1}{(1+t)^2}dt$$

$$=\int_0^\infty\frac{t^{m-1}}{(1+t)^{m+n}}dt=\int_0^\infty\frac{x^{m-1}}{(1+x)^{m+n}}dx.$$

Hence,

$$\beta(m,n)=\int_0^\infty\frac{x^{m-1}}{(1+x)^{m+n}}dx.$$

Since $\beta(m,\ n)=\beta(n,\ m)$, we have

$$\beta(m,n)=\int_0^\infty\frac{x^{m-1}}{(1+x)^{m+n}}dx=\int_0^\infty\frac{x^{n-1}}{(1+x)^{m+n}}dx.$$

5. From the property (4), we have

$$\beta(m,n)=\int_0^\infty\frac{x^{m-1}}{(1+x)^{m+n}}dx$$

$$=\int_0^1\frac{x^{m-1}}{(1+x)^{m+n}}dx+\int_1^\infty\frac{x^{m-1}}{(1+x)^{m+n}}dx$$

$$=I_1+I_2,\ \text{say}. \qquad (2)$$

In I_2, put $x=\frac{1}{t}$ so that $dx=-\frac{1}{t^2}dt$ and so,

$$I_2=\int_1^0\frac{\left(\frac{1}{t}\right)^{m-1}}{\left(1+\frac{1}{t}\right)^{m+n}}\left(-\frac{1}{t^2}\right)dt=\int_0^1\frac{t^{m+n}}{t^{m-1}(t+1)^{m+n}}\cdot\frac{1}{t^2}dt$$

$$=\int_0^1\frac{t^{n-1}}{(1+t)^{m+n}}dt=\int_0^1\frac{x^{n-1}}{(1+x)^{m+n}}dx.$$

Hence, (2) reduces to

$$\beta(m,n)=\int_0^1\frac{x^{m-1}}{(1+x)^{m+n}}dx+\int_0^1\frac{x^{n-1}}{(1+x)^{m+n}}dx$$

$$=\int_0^1\frac{x^{m-1}+x^{n-1}}{(1+x)^{m+n}}dx,\ m>0,\ n>0.$$

EXAMPLE 5.1

Show that

$$\int_0^\infty \frac{x^{m-1}+x^{n-1}}{(1+x)^{m+n}}dx = 2\beta(m,n).$$

Solution. We know that

$$\beta(m,n) = \int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}}dx. \quad (1)$$

Since $\beta(m, n) = \beta(n, m)$, we have

$$\beta(m,n) = \int_0^\infty \frac{x^{n-1}}{(1+x)^{m+n}}dx. \quad (2)$$

Adding (1) and (2), we get

$$2\beta(m,n) = \int_0^\infty \frac{x^{m-1}+x^{n-1}}{(1+x)^{m+n}}dx.$$

EXAMPLE 5.2

Show that

$$\int_0^a x^{m-1}(a-x)^{n-1}dx = a^{m+n-1}\beta(m,n).$$

Solution. Putting $x = ay$, we get

$$\int_0^a x^{m-1}(a-x)^{n-1}dx$$

$$= \int_0^1 (ay)^{m-1}(a-ay)^{n-1}.a\ dy$$

$$= \int_0^1 (ay)^{m-1}a^{n-1}(1-y)^{n-1}.a\ dy$$

$$= \int_0^1 a^{m-1+n-1+1}y^{m-1}(1-y)^{n-1}dy$$

$$= a^{m+n-1}\int_0^1 y^{m-1}(1-y)^{n-1}dy$$

$$= a^{m+n-1}\beta(m,n).$$

EXAMPLE 5.3

Show that

$$\frac{\beta(m+1,n)}{\beta(m,n)} = \frac{m}{m+n}.$$

Solution. We have

$$\beta(m+1,n) = \int_0^1 x^m(1-x)^{n-1}dx$$

$$= \int_0^1 (1-x)^m x^{n-1}dx, \text{ since } \beta(m+1,n)=\beta(n,m+1)$$

$$= \left[(1-x)^m\frac{x^n}{n}\right]_0^1 - \int_0^1 m(1-x)^{m-1}(-1).\frac{x^n}{n}dx$$

$$= \frac{m}{n}\int_0^1 x^{n-1}.x(1-x)^{m-1}dx$$

$$= \frac{m}{n}\int_0^1 x^{n-1}[1-(1-x)](1-x)^{m-1}dx$$

$$= \frac{m}{n}\left[\int_0^1 x^{n-1}(1-x)^{m-1}dx - \int_0^1 x^{n-1}(1-x)^m dx\right]$$

$$= \frac{m}{n}[\beta(n,m)-\beta(n,m+1)]$$

$$= \frac{m}{n}\beta(m,n) - \frac{m}{n}\beta(m+1,n).$$

Thus,

$$\left(1+\frac{m}{n}\right)\beta(m+1,n) = \frac{m}{n}\beta(m,n)$$

or

$$(n+m)\beta(m+1,n) = m\beta(m,n)$$

or

$$\frac{\beta(m+1,n)}{\beta(m,n)} = \frac{m}{m+n}.$$

EXAMPLE 5.4

Prove that

$$\int_0^{\frac{\pi}{2}} \sin^m\theta\cos^n\theta d\theta = \frac{1}{2}\beta\left(\frac{m+1}{2},\frac{n+1}{2}\right),$$

$m > -1$ and $n > -1$.

Solution. We have

$$\int_0^{\frac{\pi}{2}} \sin^m\theta\cos^n\theta\, d\theta$$

$$= \int_0^{\frac{\pi}{2}} \sin^{m-1}\theta\cos^{n-1}\theta.\sin\theta\cos\theta\, d\theta$$

$$= \int_0^{\frac{\pi}{2}} \sin^{m-1}\theta(1-\sin^2\theta)^{\frac{n-1}{2}}\sin\theta\cos\theta\, d\theta.$$

Putting $\sin^2\theta = x$ so that $2\sin\theta\cos\theta\, d\theta = dx$, we get

$$\int_0^{\frac{\pi}{2}} \sin^m\theta\cos^n\theta\, d\theta$$

$$=\frac{1}{2}\int_0^1 x^{\frac{m-1}{2}}(1-x)^{\frac{n-1}{2}}dx=\frac{1}{2}\int_0^1 x^{\frac{m+1}{2}-1}(1-x)^{\frac{n+1}{2}-1}dx$$

$$=\frac{1}{2}\beta\left(\frac{m+1}{2},\frac{n+1}{2}\right), m>-1 \text{ and } n>-1.$$

EXAMPLE 5.5

Show that

$$\beta(m,n) = \beta(m+1,n) + \beta(m,n+1).$$

Solution. By definition,

$$\beta(m+1,n) + \beta(m,n+1)$$

$$= \int_0^1 x^m(1-x)^{n-1}dx + \int_0^1 x^{m-1}(1-x)^n dx$$

$$= \int_0^1 x^{m-1}(1-x)^{n-1}[x+(1-x)]dx$$

$$= \int_0^1 x^{m-1}(1-x)^{n-1}dx = \beta(m,n).$$

EXAMPLE 5.6

Express $\int_0^\infty \frac{x^{m-1}}{(a+bx)^{m+n}}dx,\ m,\ n,\ a,\ b>0$ in terms of beta function.

Solution. Put $bx = ay$ so that $dx = \frac{a}{b}dy$ in the given integral. This gives

$$\int_0^\infty \frac{x^{m-1}}{(a+bx)^{m+n}}dx = \int_0^\infty \frac{\left(\frac{ay}{b}\right)^{m-1}}{(a+ay)^{m+n}}\cdot\frac{a}{b}dy$$

$$= \frac{1}{a^n b^m}\int_0^\infty \frac{y^{m-1}}{(1+y)^{m+n}}dy = \frac{1}{a^n b^m}\beta(m,n),$$

using property (4) of beta function.

EXAMPLE 5.7

Show that $\beta\left(m,\frac{1}{2}\right) = 2^{2m-1}\beta(m,m)$.

Solution. We know [see property (2)] that

$$\beta(m,n) = 2\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\ \cos^{2n-1}\theta\ d\theta. \qquad (1)$$

Putting $n=\frac{1}{2}$, we get

$$\beta\left(m,\frac{1}{2}\right) = 2\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\ d\theta. \qquad (2)$$

Now, putting $n = m$ in (1), we have

$$\beta(m,m) = 2\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\cos^{2m-1}\theta\ d\theta$$

$$= 2\int_0^{\frac{\pi}{2}} (\sin\theta\cos\theta)^{2m-1}d\theta$$

$$= 2\int_0^{\frac{\pi}{2}} \left(\frac{1}{2}\sin 2\theta\right)^{2m-1} d\theta$$

$$= \frac{1}{2^{2m-1}}.2\int_0^{\frac{\pi}{2}} \sin^{2m-1}2\theta\ d\theta$$

$$= \frac{1}{2^{2m-1}}\int_0^{\pi} \sin^{2m-1}\phi\ d\phi,\ \phi = 2\theta$$

$$= \frac{1}{2^{2m-1}} .2 \int_0^{\frac{\pi}{2}} \sin^{2m-1} \phi \, d\phi$$

$$= \frac{1}{2^{2m-1}} \beta\left(m, \frac{1}{2}\right), \text{ using}(2),$$

and so,

$$\beta\left(m, \frac{1}{2}\right) = 2^{2m-1} \beta(m, m).$$

5.3 GAMMA FUNCTION

The *gamma function* is defined as the definite integral

$$\Gamma(n) = \int_0^{\infty} e^{-x} x^{n-1} dx, \; n > 0.$$

The gamma function is also known as *Euler's Integral of Second Kind.*

5.4 PROPERTIES OF GAMMA FUNCTION

1. We have

$$\Gamma(n+1) = \int_0^{\infty} e^{-x} x^n dx = [-x^n e^{-x}]_0^{\infty} + n \int_0^{\infty} x^{n-1} e^{-x} dx$$

$$= n \int_0^{\infty} e^{-x} x^{n-1} dx = n\Gamma(n).$$

Hence,

$$\Gamma(n+1) = n\Gamma(n),$$

which is called the *recurrence formula* for $\Gamma(n)$.

2. Let n be a positive integer. By property (1), we have

$$\begin{aligned}\Gamma(n+1) &= n\Gamma(n) = n(n-1)\Gamma(n-1) \\ &= n(n-1)(n-2)\Gamma(n-2) \\ &= n(n-1)(n-2)...3.2.1 \; \Gamma(1) \\ &= n!\Gamma(1).\end{aligned}$$

But, by definition,

$$\Gamma(1) = \int_0^{\infty} e^{-x} dx = [-e^{-x}]_0^{\infty} = 1.$$

Hence,

$$\Gamma(n+1) = n!, \text{ when } n \text{ is a positive integer.}$$

If we take $n = 0$, then

$$0! = \Gamma(1) = 1,$$

and so, gamma function defines 0!

Further, from the relation $\Gamma(n+1) = n\Gamma(n)$, we deduce that

$$\begin{aligned}\Gamma(2) &= 1.\Gamma(1) = 1!, \\ \Gamma(3) &= 2.\Gamma(2) = 2.1 = 2!, \\ \Gamma(4) &= 3.\Gamma(3) = 3.2.1 = 3!, \text{ and so on.}\end{aligned}$$

Moreover, $\Gamma(0) = \infty$ and $\Gamma(-n) = -\infty$ if $n > 0$.

Also,

$$\begin{aligned}\Gamma(n) &= \frac{\Gamma(n+1)}{n}, n \neq 0 = \frac{(n+1)\Gamma(n+1)}{n(n+1)} \\ &= \frac{\Gamma(n+2)}{n(n+1)}, \; n \neq 0 \text{ and } n \neq -1 \\ &= \frac{(n+2)\Gamma(n+2)}{n(n+1)(n+2)} \\ &= \frac{\Gamma(n+3)}{n(n+1)(n+2)}, \; n \neq 0, \; n \neq -1, \text{ and } n \neq -2 \\ &= \frac{\Gamma(n+k+1)}{n(n+1)(n+2)...(n+k)}, \; n \neq 0, \; n \neq -1, \\ & \quad n \neq -2, \text{ and } n \neq -k.\end{aligned}$$

Thus, $\Gamma(n)$ for $n < 0$ is defined, where k is a least-positive integer such that $n + k + 1 > 0$.

5.5 RELATION BETWEEN BETA AND GAMMA FUNCTIONS

We know that

$$\Gamma(m) = \int_0^{\infty} e^{-t} t^{m-1} dt.$$

Putting $t = x^2$ so that $dt = 2xdx$, we get

$$\Gamma(m) = 2 \int_0^{\infty} e^{-x^2} x^{2m-1} dx. \tag{1}$$

Similarly, we can have

$$\Gamma(n) = 2 \int_0^{\infty} e^{-y^2} y^{2n-1} dy.$$

Therefore,

$$\Gamma(m)\Gamma(n) = 4 \int_0^{\infty} e^{-x^2} x^{2m-1} dx. \int_0^{\infty} e^{-y^2} y^{2n-1} dy$$

$$= 4 \int_0^{\infty} \int_0^{\infty} e^{-(x^2+y^2)} x^{2m-1} y^{2n-1} dx \, dy.$$

Taking $x = r\cos\theta$ and $y = r\sin\theta$, we have $dx\,dy = r d\theta\,dr$.
Therefore,

$$\Gamma(m)\Gamma(n) = 4\int_0^{\frac{\pi}{2}}\int_0^{\infty} e^{-r^2} r^{2(m+n)-1}\cos^{2m-1}\theta \times \sin^{2n-1}\theta\,dr d\theta$$

$$= 2\int_0^{\frac{\pi}{2}}\left[2\int_0^{\infty} e^{-r^2} r^{2(m+n)-1} dr\right] \times \cos^{2m-1}\theta\sin^{2n-1}\theta\,d\theta$$

$$= \Gamma(m+n)\left[2\int_0^{\frac{\pi}{2}}\cos^{2m-1}\theta \times \sin^{2n-1}\theta\,d\theta\right], \text{ using}(1)$$

$= \Gamma(m+n)\beta(m,n)$ using property (2) of beta function.

Hence,

$$\beta(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}.$$

EXAMPLE 5.8

Show that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Solution. We know that

$$\beta(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}.$$

Putting $m = n = \frac{1}{2}$, we get

$$\beta\left(\frac{1}{2},\frac{1}{2}\right) = \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma(1)} = \left[\Gamma\left(\frac{1}{2}\right)\right]^2.$$

Thus,

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^2 = \beta\left(\frac{1}{2},\frac{1}{2}\right) = \int_0^1 x^{\frac{1}{2}-1}(1-x)^{\frac{1}{2}-1}dx$$

$$= \int_0^1 \frac{dx}{\sqrt{x}\sqrt{1-x}} = \int_0^1 \frac{dx}{\sqrt{x-x^2}}$$

$$= \int_0^1 \frac{dx}{\sqrt{\frac{1}{4}-\left(x-\frac{1}{2}\right)^2}} = \left[\sin^{-1}\left(\frac{x-\frac{1}{2}}{\frac{1}{2}}\right)\right]_0^1$$

$$= \sin^{-1}(1) - \sin^{-1}(-1) = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi.$$

Hence,

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Second Method

We know that (see Example 5.4)

$$\int_0^{\frac{\pi}{2}} \sin^m\theta\cos^n\theta d\theta = \frac{1}{2}\beta\left(\frac{m+1}{2},\frac{n+1}{2}\right) = \frac{1}{2}\frac{\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{m+n+2}{2}\right)}.$$

Putting $m = n = 0$, we get

$$\int_0^{\frac{\pi}{2}} d\theta = \frac{\left[\Gamma\left(\frac{1}{2}\right)\right]^2}{2\Gamma(1)} = \frac{\left[\Gamma\left(\frac{1}{2}\right)\right]^2}{2}.$$

Thus,

$$\frac{\pi}{2} = \frac{\left[\Gamma\left(\frac{1}{2}\right)\right]^2}{2}.$$

Hence,

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

EXAMPLE 5.9

Express the integrals $\int_0^{\frac{\pi}{2}}\sqrt{\tan\theta}\,d\theta$ and $\int_0^{\frac{\pi}{2}}\sqrt{\cot\theta}\,d\theta$ in terms of gamma function.

Solution. We have

$$\int_0^{\frac{\pi}{2}}\sqrt{\tan\theta}\,d\theta = \int_0^{\frac{\pi}{2}}\frac{\sin^{\frac{1}{2}}\theta}{\cos^{\frac{1}{2}}\theta}d\theta = \int_0^{\frac{\pi}{2}}\sin^{\frac{1}{2}}\theta\cos^{-\frac{1}{2}}\theta\,d\theta$$

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

$$= \frac{1}{2}\Gamma\left(\frac{3}{4}\right)\Gamma\left(\frac{1}{4}\right).$$

Similarly, we can show that

$$\int_0^{\frac{\pi}{2}}\sqrt{\cot\theta}\,d\theta = \frac{1}{2}\Gamma\left(\frac{3}{4}\right)\Gamma\left(\frac{1}{4}\right).$$

EXAMPLE 5.10

Show that $\Gamma(n)\Gamma(1-n)=\frac{\pi}{\sin n\pi},\ 0<n<1.$

Solution. We know that

$$\beta(m,n)=\int_0^\infty \frac{x^{n-1}}{(1+x)^{m+n}}dx.$$

Also,

$$\beta(m,n)=\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)},\ m>0 \text{ and } n>0.$$

Therefore,

$$\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}=\int_0^\infty \frac{x^{n-1}}{(1+x)^{m+n}}dx.$$

Putting $m=1-n$ so that $m>0$ implies $n<1$, we get

$$\frac{\Gamma(n)\Gamma(1-n)}{\Gamma(1)}=\int_0^\infty \frac{x^{n-1}}{(1+x)}dx$$

or

$$\Gamma(n)\Gamma(1-n)=\int_0^\infty \frac{x^{n-1}}{1+x}dx=\frac{\pi}{\sin n\pi},\ 0<n<1.$$

EXAMPLE 5.11

Show that

$$\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\cos^{2n-1}\theta d\theta=\frac{\Gamma(m)\Gamma(n)}{2\Gamma(m+n)}.$$

Hence, evaluate $\int_0^{\frac{\pi}{2}}\sin^p\theta\, d\theta$ and $\int_0^{\frac{\pi}{2}}\cos^p\theta\, d\theta$.

Solution. We know that

$$2\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\cos^{2n-1}\theta\, d\theta=\beta(m,n)$$

or

$$\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\cos^{2n-1}\theta\, d\theta=\frac{1}{2}.\beta(m,n).$$

But, $\beta(m,n)=\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$. Therefore,

$$\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\cos^{2n-1}\theta\, d\theta=\frac{\Gamma(m)\Gamma(n)}{2\Gamma(m+n)}.$$

If we put $2m-1=p$ and $2n-1=q$, then this result reduces to

$$\int_0^{\frac{\pi}{2}} \sin^p\theta\cos^q\theta\, d\theta=\frac{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{q+1}{2}\right)}{2\Gamma\left(\frac{p+q+2}{2}\right)}. \quad (1)$$

Putting $q=0$ in (1), we get

$$\int_0^{\frac{\pi}{2}} \sin^p\theta\, d\theta=\frac{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{2\Gamma\left(\frac{p+2}{2}\right)}=\frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p+2}{2}\right)}\frac{\sqrt{\pi}}{2}.$$

Similarly, taking $p=0$, we get

$$\int_0^{\frac{\pi}{2}} \cos^q\theta\, d\theta=\frac{\Gamma\left(\frac{(q+1)}{2}\right)}{\Gamma\left(\frac{q+2}{2}\right)}\frac{\sqrt{\pi}}{2}.$$

EXAMPLE 5.12

Show that

$$\Gamma(m)\Gamma\left(m+\frac{1}{2}\right)=\frac{\sqrt{\pi}}{2^{2m-1}}\Gamma(2m)$$

(Duplication Formula).

Solution. In Example 5.7, we have shown that

$$\beta\left(m,\frac{1}{2}\right)=2^{2m-1}\beta(m,m).$$

Converting into gamma function, we get

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

Since $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$, we get

$$\frac{\sqrt{\pi}}{\Gamma\left(m+\frac{1}{2}\right)}=2^{2m-1}\frac{\Gamma(m)}{\Gamma(2m)}$$

or

$$\Gamma(m)\Gamma\left(m+\frac{1}{2}\right)=\frac{\sqrt{\pi}}{2^{2m-1}}\Gamma(2m).$$

EXAMPLE 5.13

Show that

$$\int_0^\infty e^{-ax}x^{n-1}dx=\frac{\Gamma(n)}{a^n},$$

where a and n are positive. Deduce that

(i) $\int_0^\infty e^{-ax}x^{n-1}\cos bx\, dx=\frac{\Gamma(n)}{r^n}\cos n\theta$

(ii) $\int_0^\infty e^{-ax}x^{n-1}\sin bx\, dx=\frac{\Gamma(n)}{r^n}\sin n\theta,$

where $r^2=a^2+b^2$ and $\theta=\tan^{-1}\frac{b}{a}$. Also evaluate $\int_0^\infty e^{-ax}\cos bx\, dx$ and $\int_0^\infty e^{-ax}\sin bx\, dx$.

Solution. Put $ax = z$, so that $adx = dz$, to get

$$\int_0^\infty e^{-ax}x^{n-1}dx = \int_0^\infty e^{-z}\left(\frac{z}{a}\right)^{n-1}\cdot\frac{dz}{a}$$

$$= \frac{1}{a^n}\int_0^\infty e^{-z}z^{n-1}dz = \frac{\Gamma(n)}{a^n}. \quad (1)$$

Replacing a by $a + ib$ in (1), we get

$$\int_0^\infty e^{-(a+ib)x}x^{n-1}dx = \frac{\Gamma(n)}{(a+ib)^n}. \quad (2)$$

But as,

$$e^{-(a+ib)x} = e^{-ax}.e^{-ibx} = e^{-ax}(\cos bx - i\sin bx)$$

and taking $a = r\cos\theta$ and $b = r\sin\theta$, De-Moivre's Theorem implies

$$(a+ib)^n = (r\cos\theta + ir\sin\theta)^n$$
$$= r^n(\cos\theta + i\sin\theta)^n$$
$$= r^n(\cos n\theta + i\sin n\theta).$$

Therefore, (2) reduces to

$$\int_0^\infty [e^{-ax}(\cos bx - i\sin bx)]x^{n-1}dx$$

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

$$= \frac{\Gamma(n)}{r^n}(\cos n\theta - i\sin n\theta).$$

Equating real- and imaginary parts on both sides, we get

$$\int_0^\infty e^{-ax}x^{n-1}\cos bx\, dx = \frac{\Gamma(n)}{r^n}\cos n\theta$$

and

$$\int_0^\infty e^{-ax}x^{n-1}\sin bx\, dx = \frac{\Gamma(n)}{r^n}\sin n\theta.$$

If we put $n = 1$, then

$$\int_0^{\frac{\pi}{2}} e^{-ax}\cos bx\, dx = \frac{\Gamma(1)}{r}\cos\theta = \frac{r\cos\theta}{r^2} = \frac{a}{a^2+b^2}$$

and

$$\int_0^{\frac{\pi}{2}} e^{-ax}\sin bx\, dx = \frac{\Gamma(1)}{r}\sin\theta = \frac{r\sin\theta}{r^2} = \frac{b}{a^2+b^2}.$$

EXAMPLE 5.14

Show that

$$\Gamma\left(n+\frac{1}{2}\right) = \frac{\sqrt{\pi}\Gamma(2n+1)}{2^{2n}\Gamma(n+1)}.$$

Hence, deduce that $\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right) = \pi\sqrt{2}$.

Solution. We know that

$$\beta(m,n) = 2\int_0^{\frac{\pi}{2}} \sin^{2m-1}\theta\cos^{2n-1}\theta\, d\theta.$$

Therefore,

$$\beta\left(n+\frac{1}{2}, n+\frac{1}{2}\right)$$

$$= 2\int_0^{\frac{\pi}{2}} \sin^{2n}\theta\cos^{2n}\theta\, d\theta = 2\int_0^{\frac{\pi}{2}}\left(\frac{\sin 2\theta}{2}\right)^{2n}d\theta$$

$$= \frac{1}{2^{2n-1}}\int_0^{\frac{\pi}{2}}\sin^{2n}2\theta\, d\theta = \frac{1}{2^{2n}}.\int_0^{\pi}\sin^{2n}\phi\, d\phi,\ \phi = 2\theta$$

$$= \frac{1}{2^{2n}}.2\int_0^{\frac{\pi}{2}}\sin^{2n}\phi\, d\phi = \frac{1}{2^{2n-1}}\left[\frac{\Gamma\left(\frac{2n+1}{2}\right)}{\Gamma\left(\frac{2n+2}{2}\right)}\frac{\sqrt{\pi}}{2}\right]$$

(see Example 6.11)

$$= \frac{\sqrt{\pi}\Gamma\left(n+\frac{1}{2}\right)}{2^{2n}\Gamma(n+1)}. \quad (1)$$

Also,

$$\beta\left(n+\frac{1}{2}, n+\frac{1}{2}\right) = \frac{\Gamma\left(n+\frac{1}{2}\right)\Gamma\left(n+\frac{1}{2}\right)}{\Gamma(2n+1)}$$

$$= \frac{\left[\Gamma\left(n+\frac{1}{2}\right)\right]^2}{\Gamma(2n+1)}. \quad (2)$$

From (1) and (2), we have

$$\Gamma\left(n+\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2^{2n}}.\frac{\Gamma(2n+1)}{\Gamma(n+1)}.$$

Further, putting $n = \frac{1}{4}$, we have

$$\Gamma\left(\frac{3}{4}\right) = \frac{\sqrt{\pi}}{\sqrt{2}}.\frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{5}{4}\right)} = \sqrt{\frac{\pi}{2}}\frac{\frac{1}{2}\Gamma\left(\frac{1}{2}\right)}{\frac{1}{4}\Gamma\left(\frac{1}{4}\right)}.$$

Hence,

$$\Gamma\left(\frac{3}{4}\right)\Gamma\left(\frac{1}{4}\right) = \pi\sqrt{2}.$$

EXAMPLE 5.15

Show that

$$\int_0^{\frac{\pi}{2}} \sqrt{\tan\theta}\, d\theta = \int_0^{\frac{\pi}{2}} \sqrt{\cot\theta}\, d\theta = \frac{1}{2}\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right) = \frac{\pi}{\sqrt{2}}.$$

Solution. In Example 5.9, we have proved that

$$\int_0^{\frac{\pi}{2}} \sqrt{\tan\theta}\, d\theta = \int_0^{\frac{\pi}{2}} \sqrt{\cot\theta}\, d\theta = \frac{1}{2}\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right).$$

But,

$$\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right) = \pi\sqrt{2} \text{ (Example 5.14.)}$$

Hence,

$$\int_0^{\frac{\pi}{2}} \sqrt{\tan\theta}\, d\theta = \int_0^{\frac{\pi}{2}} \sqrt{\cot\theta}\, d\theta = \frac{\pi}{\sqrt{2}}.$$

EXAMPLE 5.16

Show that

$$\int_0^1 y^{q-1}\left(\log\frac{1}{y}\right)^{p-1} dy = \frac{\Gamma(p)}{q^p},\ p > 0,\ q > 0.$$

Solution. Putting $\log\frac{1}{y} = x$, we have $\frac{1}{y} = e^x$ or $y = e^{-x}$ and so, $dy = -e^{-x}\,dx$. Therefore,

$$\int_0^1 y^{q-1}\left(\log\frac{1}{y}\right)^{p-1} dy$$

$$= \int_\infty^0 e^{-(q-1)x} x^{p-1}(-e^{-x})dx$$

$$= \int_0^\infty e^{-qx} x^{p-1} dx$$

$$= \int_0^\infty e^{-t}\left(\frac{t}{q}\right)^{p-1} . \frac{dt}{q},\ \text{putting } qx = t$$

$$= \frac{1}{q^p}\int_0^\infty e^{-t} t^{p-1} dt = \frac{\Gamma(p)}{q^p}.$$

EXAMPLE 5.17

Show that

$$\int_0^{\frac{\pi}{2}} \sqrt{\sin\theta}\ d\theta . \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{\sin\theta}} = \pi.$$

Solution. We have

$$\int_0^{\frac{\pi}{2}} \sqrt{\sin\theta}\, d\theta . \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{\sin\theta}}$$

$$= \int_0^{\frac{\pi}{2}} \sin^{\frac{1}{2}}\theta\, d\theta . \int_0^{\frac{\pi}{2}} \sin^{-\frac{1}{2}}\theta\, d\theta.$$

$$= \frac{\Gamma\left(\frac{\frac{1}{2}+1}{2}\right)}{\Gamma\left(\frac{\frac{1}{2}+2}{2}\right)} . \frac{\sqrt{\pi}}{2} . \frac{\Gamma\left(\frac{-\frac{1}{2}+1}{2}\right)}{\Gamma\left(\frac{-\frac{1}{2}+2}{2}\right)} . \frac{\sqrt{\pi}}{2} = \frac{\Gamma\left(\frac{3}{4}\right)\Gamma\left(\frac{1}{4}\right)}{\Gamma\left(\frac{5}{4}\right)\Gamma\left(\frac{3}{4}\right)} . \frac{\pi}{4}$$

(see Example 5.11)

$$= \frac{\Gamma\left(\frac{1}{4}\right)}{\Gamma\left(\frac{5}{4}\right)} . \frac{\pi}{4} = \frac{\Gamma\left(\frac{1}{4}\right)}{\frac{1}{4}\Gamma\left(\frac{1}{4}\right)} . \frac{\pi}{4} = \pi.$$

EXAMPLE 5.18

Prove that

$$\int_0^1 \left(\log\frac{1}{y}\right)^{n-1} dy = \Gamma(n),\ n > 0.$$

Solution. Putting $\log\frac{1}{y} = x$, that is, $\frac{1}{y} = e^x$ or $y = e^{-x}$, we have $dy = -e^{-x}dx$. Hence,

$$\int_0^1 \left(\log\frac{1}{y}\right)^{n-1} dy = -\int_\infty^0 x^{n-1}e^{-x}dx$$

$$= \int_0^\infty e^{-x}x^{n-1}dx$$

$$= \Gamma(n),\ n > 0.$$

EXAMPLE 5.19

Evaluate $\int_0^1 x^4\left(\log\frac{1}{x}\right)^3 dx$.

Solution. Putting $\log\frac{1}{x} = y$, that is, $\frac{1}{x} = e^y$ or $x = e^{-y}$, we have $dx = -e^y\, dy$. Therefore,

$$\int_0^1 x^4\left(\log\frac{1}{x}\right)^3 dx = -\int_\infty^0 e^{-4y}.y^3.e^{-y}dy = \int_0^\infty e^{-5y}.y^3 dy$$

$$= \int_0^\infty e^{-t}.\frac{t^3}{125}.\frac{dt}{5}, \text{ putting } 5y = t$$

$$= \frac{1}{625}\int_0^\infty e^{-t}.t^3 dt = \frac{1}{625}\Gamma(3) = \frac{6}{625}.$$

EXAMPLE 5.20

Show that

$$\int_0^1 x^m(\log x)^n dx = \frac{(-1)^n n!}{(m+n)^{n+1}}.$$

Solution. Putting $\log x = -y$, we have $x = e^{-y}$ and so, $dx = -e^{-y}$. When $x = 1$, $y = 0$ and when $x \to 0$, $y \to \infty$. Therefore,

$$\int_0^1 x^m(\log x)^n dx$$

$$= \int_\infty^0 e^{-my}(-y)^n(-e^{-y})dy = (-1)^n\int_0^\infty e^{-(m+1)y}.y^n dy$$

$$= (-1)^n\int_0^\infty e^{-t}.\left(\frac{t}{m+1}\right)^n\frac{dt}{m+1}, \text{ putting } (m+1)y = t$$

$$= \frac{(-1)^n}{(m+1)^{n+1}}\int_0^\infty e^{-t}t^n dt = \frac{(-1)^n}{(m+1)^{n+1}}\Gamma(n+1)$$

$$= \frac{(-1)^n n!}{(m+1)^{n+1}}.$$

EXAMPLE 5.21

Prove that

$$\int_0^1 \frac{dx}{\sqrt{1+x^4}} = \frac{1}{4\sqrt{2}}\beta\left(\frac{1}{4},\frac{1}{2}\right).$$

Solution. Putting $x^2 = \tan\theta$, we have $2xdx = \sec^2\theta\, d\theta$. Therefore,

$$\int_0^1 \frac{dx}{\sqrt{1+x^4}} = \int_0^{\frac{\pi}{4}} \frac{\sec^2\theta}{\sqrt{2\tan\theta}}\cdot\frac{1}{\sqrt{1+\tan^2\theta}}d\theta$$

$$= \frac{1}{2}\int_0^{\frac{\pi}{4}}\frac{\sec^2\theta}{\sqrt{\tan\theta}}.\frac{1}{\sec\theta}d\theta = \frac{1}{2}\int_0^{\frac{\pi}{4}}\frac{d\theta}{\sin^{\frac{1}{2}}\theta\cos^{\frac{1}{2}}\theta}$$

$$= \frac{1}{2}\int_0^{\frac{\pi}{4}}\frac{d\theta}{(\sin\theta\cos\theta)^{\frac{1}{2}}} = \frac{1}{2}\int_0^{\frac{\pi}{4}}\frac{d\theta}{\left(\frac{\sin 2\theta}{2}\right)^{\frac{1}{2}}}$$

$$= \frac{1}{\sqrt{2}}\int_0^{\frac{\pi}{4}}\frac{d\theta}{\sqrt{\sin 2\theta}} = \frac{1}{\sqrt{2}}\int_0^{\frac{\pi}{2}}\frac{d\phi}{2\sqrt{\sin\phi}}, \quad \phi = 2\theta$$

$$= \frac{1}{2\sqrt{2}}\int_0^{\frac{\pi}{2}}\sin^{-\frac{1}{2}}\phi\cos^0\phi\, d\phi$$

$$= \frac{2}{4\sqrt{2}}\int_0^{\frac{\pi}{2}}\sin^{-\frac{1}{2}}\phi\cos^0\phi\, d\phi$$

$$= \frac{1}{4\sqrt{2}}\beta\left(\frac{1}{4},\frac{1}{2}\right),$$

since $\beta(m,n) = 2\int_0^{\frac{\pi}{2}}\sin^{2m-1}\phi\cos^{2n-1}\phi\, d\phi$.

EXAMPLE 5.22

Show that

$$\int_0^{\frac{\pi}{2}}\left(\sqrt{\tan\theta}+\sqrt{\sec\theta}\right)d\theta$$

$$= \frac{1}{2}\Gamma\left(\frac{1}{4}\right)\left[\Gamma\left(\frac{3}{4}\right)+\frac{\sqrt{\pi}}{\Gamma\left(\frac{3}{4}\right)}\right].$$

Solution. We have

$$\int_0^{\frac{\pi}{2}}\left(\sqrt{\tan\theta}+\sqrt{\sec\theta}\right)d\theta$$

$$= \int_0^{\frac{\pi}{2}}\sqrt{\tan\theta}\, d\theta + \int_0^{\frac{\pi}{2}}\sqrt{\sec\theta\, d\theta}$$

$$= \int_0^{\frac{\pi}{2}} \sin^{\frac{1}{2}}\theta \cos^{-\frac{1}{2}}\theta \, d\theta + \int_0^{\frac{\pi}{2}} \sin^0\theta \cos^{-\frac{1}{2}}\theta \, d\theta$$

$$= \frac{2}{2}\int_0^{\frac{\pi}{2}} \sin^{\frac{1}{2}}\theta \cos^{-\frac{1}{2}}\theta \, d\theta + \frac{2}{2}\int_0^{\frac{\pi}{2}} \sin^0\theta \cos^{-\frac{1}{2}}\theta \, d\theta$$

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

$$= \frac{1}{2}\left[\frac{\Gamma\left(\frac{3}{4}\right)\Gamma\left(\frac{1}{4}\right)}{\Gamma(1)} + \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{4}\right)}{\Gamma\left(\frac{3}{4}\right)}\right]$$

$$= \frac{1}{2}\Gamma\left(\frac{1}{4}\right)\left[\Gamma\left(\frac{3}{4}\right) + \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{3}{4}\right)}\right]$$

$$= \frac{1}{2}\Gamma\left(\frac{1}{4}\right)\left[\Gamma\left(\frac{3}{4}\right) + \frac{\sqrt{\pi}}{\Gamma\left(\frac{3}{4}\right)}\right].$$

5.6 DIRICHLET'S AND LIOUVILLE'S THEOREMS

The following theorems of Dirichlet and Liouville are useful in evaluating multiple integrals.

Theorem 5.1 (Dirichlet). If V is the region, where $x \geq 0$, $y \geq 0$, $z \geq 0$, *and* $x + y + z \leq 1$, then

$$\iiint_V x^{p-1}y^{q-1}z^{r-1}dx\,dy\,dz = \frac{\Gamma(p)\Gamma(q)\Gamma(r)}{\Gamma(p+q+r+1)}.$$

(The Dirichlet's Theorem can be extended to a finite number of variables).

Proof: Since $x + y + z \leq 1$, we have $y + z \leq 1 - x = a$. Therefore,

$$\iiint_V x^{p-1}y^{q-1}z^{r-1}\,dx\,dy\,dz$$

$$= \int_0^1 \int_x^{1-x} \int_0^{1-x-y} x^{p-1}y^{q-1}z^{r-1}\,dx\,dy\,dz$$

$$= \int_0^1 x^{p-1}\left[\int_0^a \int_0^{a-y} y^{q-1}z^{r-1}dz\,dy\right]dx. \qquad (1)$$

Let

$$I = \int_0^a \int_0^{a-y} y^{q-1}z^{r-1}dz\,dy.$$

Putting $y = aY$ and $z = aZ$, this integral reduces to

$$I = \int_D (aY)^{q-1}(aZ)^{r-1}.a^2dZ\,dY,$$

where D is the domain where $X \geq 0$, $Y \geq 0$, and $Y + Z \geq 1$. Thus,

$$I = a^{q+r}\int_0^1 \int_0^{1-Y} Y^{q-1}Z^{r-1}dZ\,dY$$

$$= a^{q+r}\int_0^1 Y^{q-1}\left[\frac{Z^r}{r}\right]_0^{1-Y} dY$$

$$= \frac{a^{q+r}}{r}\int_0^1 Y^{q-1}(1-Y)^r dY$$

$$= \frac{a^{q+r}}{r}\beta(q,\ r+1) = \frac{a^{q+r}}{r}\,\frac{\Gamma(q)\Gamma(r+1)}{\Gamma(q+r+1)}$$

$$= a^{q+r}\frac{\Gamma(q)\Gamma(r)}{\Gamma(q+r+1)}.$$

Hence, (1) yields

$$\iiint_V x^{p-1}y^{q-1}z^{r-1}dx\,dy\,dz$$

$$= \int_0^1 \frac{\Gamma(q)\Gamma(r)}{\Gamma(q+r+1)}x^{p-1}a^{q+r}dx$$

$$= \frac{\Gamma(q)\Gamma(r)}{\Gamma(q+r+1)}\int x^{p-1}(1-x)^{q+r}dx, \text{ since } a = 1-x$$

$$= \frac{\Gamma(q)\Gamma(r)}{\Gamma(q+r+1)}\beta(p, q+r+1)$$

$$= \frac{\Gamma(q)\Gamma(r)}{\Gamma(q+r+1)}.\frac{\Gamma(p)\Gamma(q+r+1)}{\Gamma(p+q+r+1)} = \frac{\Gamma(p)\Gamma(q)\Gamma(r)}{\Gamma(p+q+r+1)}.$$

Remark 5.1. If $x + y + z \leq h$, then by putting $\frac{x}{h} = X$, $\frac{y}{h} = Y$, and $\frac{z}{h} = Z$, we have $X + Y + Z \leq \frac{h}{h} = 1$ and so, the Dirichlet's Theorem takes the form

$$\iiint_V x^{p-1}y^{q-1}z^{r-1}dx\,dy\,dz$$

$$= \frac{\Gamma(p)\Gamma(q)\Gamma(r)}{\Gamma(q+r+1)}.h^{p+q+r}.$$

Theorem 5.2 (Liouville). If x, y, and z are all positive such that $h_1 < x + y + z < h_2$, then

$$\iiint f(x+y+z)x^{p-1}y^{q-1}z^{r-1}dx\,dy\,dz = \frac{\Gamma(p)\Gamma(q)\Gamma(r)}{\Gamma(p+q+r)}\int_{h_1}^{h_2} f(h)h^{p+q+r-1}dh.$$

(Proof, not provided here, is a slight modification of the proof of Dirichlet's Theorem).

EXAMPLE 5.23

Evaluate $\iiint x\ y\ z\ dx\ dy\ dz$ taken throughout the ellipsoid

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\le 1.$$

Solution. Put $\frac{x^2}{a^2}=X$, $\frac{y^2}{b^2}=Y$, and $\frac{z^2}{c^2}=Z$ to get

$$x = aX^{\frac{1}{2}},\ y = bY^{\frac{1}{2}},\ \text{and } z = cZ^{\frac{1}{2}}$$

and

$$xdx = \frac{a^2}{2}dX,\ ydy = \frac{b^2}{2}dY,\ \text{and } zdz = \frac{c^2}{2}dZ.$$

The condition, under this substitution, becomes

$$X+Y+Z\le 1.$$

Therefore, for the first quadrant,

$$\begin{aligned}\iiint xyz\ dx\ dy\ dz &= \iiint (xdx)(ydy)(zdz)\\ &= \iiint \left(\frac{a^2}{2}dX\right)\left(\frac{b^2}{2}dY\right)\left(\frac{c^2}{2}dZ\right)\\ &= \frac{a^2b^2c^2}{8}\iiint X^{1-1}Y^{1-1}Z^{1-1}dX\ dY\ dZ\\ &= \frac{a^2b^2c^2}{8}\ \frac{\Gamma(1)\Gamma(1)\Gamma(1)}{\Gamma(1+1+1+1)},\end{aligned}$$

by Dirichlet's Theorem

$$= \frac{a^2b^2c^2}{8}\cdot\frac{1}{\Gamma(4)} = \frac{a^2b^2c^2}{8.3!} = \frac{a^2b^2c^2}{48}.$$

Therefore, value of the integral for the whole of the ellipsoid is $8\left(\frac{a^2b^2c^2}{48}\right) = \frac{a^2b^2c^2}{6}$.

EXAMPLE 5.24

The plane $\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1$ meets the axes in A, B, and C. Find the volume of the tetrahedron OABC.

Solution. We wish to evaluate

$$\iiint dx\ dy\ dz$$

under the condition $\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1$. Putting $\frac{x}{a}=X$, $\frac{y}{b}=Y$, and $\frac{z}{c}=Z$, we get $X+Y+Z=1$. Also $dx = adX$, $dy = bdY$, and $dz = cdZ$. Therefore, using Dirichlet's Theorem, the required volume of the tetrahedron is

$$\begin{aligned}V &= \iiint dx\ dy\ dz\\ &= \iiint abc\ dX\ dY\ dZ\\ &= abc\iiint X^{1-1}Y^{1-1}Z^{1-1}dX\ dY\ dZ\\ &= abc\frac{\Gamma(1)\Gamma(1)\Gamma(1)}{\Gamma(1+1+1+1)}\\ &= \frac{abc}{\Gamma(4)} = \frac{abc}{3!} = \frac{abc}{6}.\end{aligned}$$

EXAMPLE 5.25

Evaluate $\iiint x^{l-1}\ y^{m-1}\ z^{n-1}\ dx\ dy\ dz$, where $x>0$, $y>0$, and $z>0$ under the condition $\left(\frac{x}{a}\right)^p+\left(\frac{y}{b}\right)^q+\left(\frac{z}{c}\right)^r\le 1$.

Solution. Put $\left(\frac{x}{a}\right)^p = X$, $\left(\frac{y}{b}\right)^q = Y$, and $\left(\frac{z}{c}\right)^r = Z$ so that

$$dx = \frac{a}{p}X^{\frac{1}{p}-1}dX,$$

$$dy = \frac{b}{q}Y^{\frac{1}{q}-1}dY,\ \text{and}$$

$$dz = \frac{c}{r}Z^{\frac{1}{r}-1}dZ.$$

Therefore,

$$\begin{aligned}\iiint x^{l-1}y^{m-1}z^{n-1}dx\ dy\ dz &= \iiint \left(aX^{\frac{1}{p}}\right)^{l-1}\left(bY^{\frac{1}{q}}\right)^{m-1}\left(cZ^{\frac{1}{r}}\right)^{n-1}\\ &\quad\times\frac{abc}{pqr}X^{\frac{1}{p}-1}Y^{\frac{1}{q}-1}Z^{\frac{1}{r}-1}dX\ dY\ dZ\\ &= \frac{a^lb^mc^n}{pqr}\iiint X^{\frac{l}{p}-1}Y^{\frac{m}{q}-1}Z^{\frac{n}{r}-1}dX\ dY\ dZ\\ &= \frac{a^lb^mc^n}{pqr}\cdot\frac{\Gamma\left(\frac{l}{p}\right)\Gamma\left(\frac{m}{q}\right)\Gamma\left(\frac{n}{r}\right)}{\Gamma\left(1+\frac{l}{p}+\frac{m}{q}+\frac{n}{r}\right)}.\end{aligned}$$

EXAMPLE 5.26

Show that $\iint x^{m-1}\, y^{n-1}\, dx\; dy$ over the positive octant of the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ is $\frac{a^m b^n}{2n}\beta\left(\frac{m}{2},\frac{n}{2}+1\right)$.

Solution. Putting $\frac{x^2}{a^2}=X$ and $\frac{y^2}{b^2}=Y$, we get $x=a\sqrt{X}$ and $y=b\sqrt{Y}$ and $dx=\frac{a}{2}X^{-\frac{1}{2}}dX$ and $dy=\frac{b}{2}Y^{-\frac{1}{2}}dY$. Therefore,

$$\begin{aligned}
&\iint x^{m-1}y^{n-1}dxdy\\
&=\iint a^{m-1}X^{\frac{m-1}{2}}b^{n-1}Y^{\frac{n-1}{2}}\frac{a}{2}X^{-\frac{1}{2}}\frac{b}{2}Y^{-\frac{1}{2}}dXdY\\
&=\frac{a^m b^n}{4}\int\int X^{\frac{m}{2}-1}Y^{\frac{n}{2}-1}dX\;dY\\
&=\frac{a^m b^n}{4}\cdot\frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(1+\frac{m}{2}+\frac{n}{2}\right)}=\frac{a^m b^n}{\frac{4n}{2}}\cdot\frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)\frac{n}{2}}{\Gamma\left(\frac{m+n}{2}+1\right)}\\
&=\frac{a^m b^n}{2n}\cdot\frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{m+n}{2}+1\right)}=\frac{a^m b^n}{2n}\beta\left(\frac{m}{2},\frac{n}{2}+1\right).
\end{aligned}$$

5.7. MISCELLANEOUS EXAMPLES

EXAMPLE 5.27

Evaluate $\int_0^\infty \frac{x\,dx}{1+x^6}$ using Beta and Gamma functions.

Solution. Putting $x^6=t$, we have $x=t^{\frac{1}{6}}$ and so $dx=\frac{1}{6}t^{-\frac{5}{6}}$. When $x=0$, $t=0$ and when $x=\infty$, $t=\infty$. Therefore

$$\begin{aligned}
\int_0^\infty \frac{x\,dx}{1+x^6}&=\frac{1}{6}\int_0^\infty \frac{t^{\frac{1}{6}}}{1+t}t^{-\frac{5}{6}}dt\\
&=\frac{1}{6}\int_0^\infty \frac{t^{-\frac{2}{3}}}{1+t}dt=\frac{1}{6}\int_0^\infty \frac{t^{\frac{1}{3}-1}}{(1+t)^{\frac{1}{3}+\frac{2}{3}}}dt\\
&=\frac{1}{6}\beta\left(\frac{1}{3},\frac{2}{3}\right),\quad\text{since}\quad \beta(m,n)\\
&\qquad\qquad=\int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}}dx\\
&=\frac{1}{6}\frac{\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)}{\Gamma\left(\frac{1}{3}+\frac{2}{3}\right)}=\frac{1}{6}\frac{\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)}{\Gamma(1)}\\
&=\frac{1}{6}\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)\\
&=\frac{1}{6}\left(\frac{\pi}{\sin\frac{\pi}{3}}\right),\ \text{using } \Gamma(n)\,\Gamma\left(1-\frac{2}{n}\right)\\
&\qquad=\frac{\pi}{\sin n\pi}(0<n).
\end{aligned}$$

EXAMPLE 5.28

Evaluate $\int_0^a x^4\sqrt{a^2-x^2}dx$ using Gamma function.

Solution. Putting $x=a\sin\,\theta$, we get $dx=a\cos\,\theta\,d\theta$ and $\sqrt{a^2-x^2}=\sqrt{a^2-a^2\sin^2\theta}=a\,\cos\,\theta$. When $x=0$, $\theta=0$ and when $x=a$, $\theta=\frac{\pi}{2}$. Therefore

$$\begin{aligned}
I&=\int_0^a x^4\sqrt{a^2-x^2}dx\\
&=\int_0^{\frac{\pi}{2}} a^4\sin^4\theta(a\cos\theta)(a\cos\theta)d\theta\\
&=a^6\int_0^{\frac{\pi}{2}}\sin^4\theta\cos^2\theta\,d\theta.
\end{aligned}$$

Since $\int_0^{\frac{\pi}{2}}\sin^m\theta\,\cos^n\theta\,d\theta=\frac{1}{2}\frac{\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{m+n+2}{2}\right)}$, we have

$$\begin{aligned}
I&=a^6\left[\frac{\Gamma\left(\frac{5}{2}\right)\,\Gamma\left(\frac{3}{2}\right)}{2\,\Gamma(4)}\right]\\
&=\frac{a^6}{2}\left[\frac{\frac{3}{2}\cdot\frac{1}{2}\sqrt{\pi}.\frac{1}{2}\sqrt{\pi}}{3!}\right]=\frac{\pi a^6}{32}.
\end{aligned}$$

EXAMPLE 5.29

Show that $\int_0^1 \frac{x^n}{\sqrt{1-x^2}}dx=\frac{2.4.\ldots(n-1)}{1.3.5.\ldots n}$, where n is odd integer.

Solution. Putting $x=\sin\theta$, we have $dx=\cos\theta\,d\theta$. Therefore

$$\begin{aligned}
\int_0^1 \frac{x^n}{\sqrt{1-x^2}}dx&=\int_0^{\frac{\pi}{2}}\frac{\sin^n\theta}{\cos\theta}\cos\theta\,d\theta\\
&=\int_0^{\frac{\pi}{2}}\sin^n\theta\,d\theta\\
&=\frac{\Gamma\left(\frac{n+1}{2}\right)}{2\,\Gamma\left(\frac{n+2}{2}\right)}\sqrt{\pi}\quad\text{(see Example 5.11).}
\end{aligned}$$

Since n is odd, we take $n = 2m+1$. Therefore

$$\int_0^1 \frac{x^n}{\sqrt{1-x^2}}dx$$

$$= \frac{\Gamma(m+1)}{2\,\Gamma\left(m+\frac{3}{2}\right)}\frac{\sqrt{\pi}}{2}$$

$$= \frac{m(m-1)(m-2)\ldots 3.2.1}{(m+\frac{1}{2})(m-\frac{1}{2})(m-\frac{3}{2})\ldots\frac{3}{2}\cdot\frac{1}{2}\Gamma(\frac{1}{2})}\frac{\sqrt{\pi}}{2}$$

$$= \frac{2m(2m-2)(2m-4)\ldots 6.4.2}{(2m+1)(2m-1)(2m-3)\ldots 3.1}$$

$$= \frac{(n-1)(n-3)(n-5)\ldots 6.4.2}{n(n-2)\ldots 5.3.1}, \quad (n \text{ odd})$$

which was to be established.

EXAMPLE 5.30

Show that $\int_0^\infty x^m e^{-a^2x^2}dx = \frac{1}{2a^{m+1}}\,\Gamma\left(\frac{m+1}{2}\right)$.

Solution. Putting $ax = \sqrt{z}$, we have $a\,dx = \frac{1}{2}z^{-\frac{1}{2}}dz$. Therefore

$$\int_0^\infty x^m e^{-a^2x^2}dx = \frac{1}{2a}\int_0^\infty e^{-z}\left(\frac{\sqrt{z}}{a}\right)^m z^{-\frac{1}{2}}dz$$

$$= \frac{1}{2a^{m+1}}\int_0^\infty e^{-z}z^{\frac{m-1}{2}}dz$$

$$= \frac{1}{2a^{m+1}}\int_0^\infty e^{-z}z^{\frac{m+1}{2}-1}dz$$

$$= \frac{1}{2a^{m+1}}\,\Gamma\left(\frac{m+1}{2}\right).$$

EXERCISES

1. Show that

   (i) $\beta(2.5, 1.5) = \frac{\pi}{16}$.

   (ii) $\beta\left(\frac{9}{2},\frac{7}{2}\right) = \frac{5\pi}{2048}$.

   (iii) $\Gamma\left(\frac{3}{4}\right)\Gamma\left(\frac{1}{4}\right) = \pi\sqrt{2}$.

2. Show that
$$\int_0^\infty \sqrt{x}e^{-x^3}dx = \frac{\sqrt{\pi}}{3}.$$

3. Show that
$$\int_0^\infty x^{2n-1}e^{-ax^2}dx = \frac{\Gamma(n)}{2a^n}.$$

4. Show that
$$\int_0^\infty \frac{x^a}{a^x}dx = \frac{\Gamma(a+1)}{(\log a)^{a+1}}, \text{ if a} > 1.$$

5. Show that
$$\int_0^{\frac{\pi}{2}} \sin^3 x\cos^{\frac{5}{2}} x dx = \frac{8}{77}.$$

6. Show that
$$\int_a^b (x-a)^m(b-x)^n dx = (b-a)^{m+n+1}\beta(m+1, n+1).$$

7. Prove that
$$\int_0^\infty x^n e^{-a^2x^2}dx = \frac{1}{2a^{n+1}}\Gamma\left(\frac{n+1}{2}\right), \quad n > -1.$$

8. Show that
$$\int_0^1 \frac{x^2dx}{\sqrt{1-x^4}}\cdot\int_0^1 \frac{dx}{\sqrt{1+x^4}} = \frac{\pi}{4\sqrt{2}}.$$

9. Prove that
$$\int_0^1 \frac{xdx}{\sqrt{1-x^5}} = \frac{1}{5}\beta\left(\frac{2}{5},\frac{1}{2}\right).$$

10. Show that

$$\int_0^1 \frac{dx}{\sqrt{1-x^n}} = \frac{\sqrt{\pi}}{n} \cdot \frac{\Gamma\left(\frac{1}{n}\right)}{\Gamma\left(\frac{1}{n}+\frac{1}{2}\right)}.$$

11. Express the integrals in terms of gamma function:

(i) $\int_0^\infty x^{p-1} e^{-kx} dx, k > 0$ **Ans.** (i) $\frac{\Gamma(n)}{k^n}$.

(ii) $\int_0^1 \left(\log \frac{1}{x}\right)^{n-1} dx$ (ii) $\int_0^\infty e^{-y} y^{n-1} dy = \Gamma(n)$.

(iii) $\int_0^\infty e^{-x^2} dx$ (iii) $\frac{1}{2}\int_0^\infty e^{-t} t^{-\frac{1}{2}} dt = \frac{1}{2}\Gamma\left(\frac{1}{2}\right)$.

12. Show that $\int_0^\infty x^3 e^{-x^2} dx = \frac{1}{9}\Gamma\left(\frac{1}{3}\right)$.

13. Show that $\int_0^\infty x^{2n-1} e^{-ax^2} dx = \frac{\Gamma(n)}{2a^n}$.

14. Show that $y\beta(x+1, y) = x\beta(x, y+1)$.

15. Show that $\int_0^\infty \sqrt{x} e^{-x^2} dx . \int_0^\infty \frac{e^{-x^2}}{\sqrt{x}} dx = \frac{\pi}{2\sqrt{2}}$.

16. The plane $\frac{x}{a}+\frac{y}{b}+\frac{z}{c} = 1$ meets the axes in A, B, and C, respectively. Find the mass of the tetrahedron OABC if the density at any point is $\rho = \mu\, xyz$.

Hint: Mass $= \iiint \rho\, dx\, dy\, dz, 0 \le \frac{x}{a}+\frac{y}{b}+\frac{z}{c} \le 1$

$$= \iiint \mu xyz\, dx\, dy\, dz$$

Put $\frac{x}{a} = X$, $\frac{y}{b} = Y$, and $\frac{z}{c} = Z$ and proceed.

Ans. $\frac{\mu a^2 b^2 c^2}{720}$.

17. Show that the volume of the solid bounded by the coordinate planes and the surface $\sqrt{\frac{x}{a}} + \sqrt{\frac{y}{b}} + \sqrt{\frac{z}{c}} = 1$ is $\frac{abc}{90}$.

18. Find the volume of the ellipsoid $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$.

Hint: $V = 8\int\int\int dxdydz$.

Put $\frac{x^2}{a^2} = X, \frac{y^2}{b^2} = Y$, and $\frac{z^2}{c^2} = Z$, and use Dirichlet's Theorem to get $V = \frac{4\pi}{3} abc$.

19. Show that the entire volume of the solid $\left(\frac{x}{a}\right)^{\frac{2}{3}}+\left(\frac{y}{b}\right)^{\frac{2}{3}}+\left(\frac{z}{c}\right)^{\frac{2}{3}}= 1$ is $\frac{4}{35}\pi\, abc$.

6 Multiple Integrals

The aim of this chapter is to study double- and triple integrals along with their applications. Thus, we shall consider here the integrals of the functions of two- and three variables.

6.1 DOUBLE INTEGRALS

The notion of a double integral is an extension of the concept of a definite integral on the real line to the case of two-dimensional space. Let $f(x, y)$ be a continuous function of two independent variables x and y inside and on the boundary of a region R. Divide the region R into subdomains $R_1, R_2, \ldots, R_n$ of areas $\delta R_1, \delta R_2, \ldots, \delta R_n$, respectively. Let (x_i, y_i) be an arbitrary point inside the ith elementary area, δR_i. Consider the sum

$$\begin{aligned} S_n &= f(x_i, y_i)\delta R_1 + f(x_2, y_2)\delta R_2 + \ldots \\ &\quad + f(x_i, y_i)\delta R_i + \ldots + f(x_n, y_n)\delta R_n. \\ &= \sum_{i=1}^{n} f(x_i, y_i)\delta R_i. \end{aligned}$$

When $n \to \infty$, the number of subregions increases indefinitely such that the largest of the areas δR_i approaches zero. The $\lim\limits_{\substack{n\to\infty \\ \delta R_i \to 0}} S_n$, if exists, is called the *double integral* of the function $f(x, y)$ over the region (domain) R and is denoted by

$$\iint_R f(x,y)dR.$$

If the region R is divided into rectangular meshes by a network of lines parallel to the coordinate axes and if dx and dy be the length and breadth of a rectangular mess, then $dxdy$ is an element of area in Cartesian coordinates. In such a case, we have

$$\iint_R f(x,y)dR = \iint_R f(x,y)\ dx\ dy.$$

We now state, without proof, two theorems that provide sufficient conditions for the existence of a double integral over a closed region R.

Theorem 6.1. Let ϕ and ψ be two continuous functions defined on a closed interval $[a, b]$ such that $\phi(x) \le \psi(x)$ for all $x \in [a, b]$. Let f be a continuous function defined over $R = \{(x, y)\colon a \le x \le b, \phi(x) \le y \le \psi(x)\}$. Then, $\iint_R f(x,y)dxdy$ and $\int_a^b \left[\int_{\phi(x)}^{\psi(x)} f(x,y)dy \right] dx$ exist and are equal.

Theorem 6.2. Let ϕ and ψ be two continuous functions defined on a closed interval $[c, d]$ such that $\phi(y) \le \psi(y)$ for $y \in [c, d]$. Let f be a continuous function defined over $R = \{(x, y)\colon c \le y \le d, \phi(y) \le x \le \psi(y)\}$. Then, $\iint_R f(x,y)dxdy$ and $\int_c^d \left[\int_{\phi(y)}^{\psi(y)} f(x,y)dx \right] dy$ exist and are equal.

EXAMPLE 6.1

Show that

$$\int_0^1 \left[\int_0^1 \frac{x-y}{(x+y)^3} dx \right] dy \ne \int_0^1 \left[\int_0^1 \frac{x-y}{(x+y)^3} dy \right] dx.$$

Solution. We have

$$\begin{aligned} &\int_0^1 \left[\int_0^1 \frac{x-y}{(x+y)^3} dx \right] dy \\ &= \int_0^1 \left[\int_0^1 \frac{x+y-2y}{(x+y)^3} dx \right] dy \\ &= \int_0^1 \left[\int_0^1 \left\{ \frac{1}{(x+y)^2} - \frac{2y}{(x+y)^3} \right\} dx \right] dy \end{aligned}$$

$$= \int_0^1 \left[\frac{-1}{x+y} + \frac{2y}{2(x+y)^2} \right]_0^1 dy$$

$$= \int_0^1 \left[\frac{-1}{(1+y)^2} \right] dy = \left[\frac{1}{1+y} \right]_0^1 = -\frac{1}{2}.$$

Similarly, we can show that

$$\int_0^1 \left[\int_0^1 \frac{x-y}{(x+y)^3} dy \right] dx = \frac{1}{2}.$$

Hence,

$$\int_0^1 \left[\int_0^1 \frac{x-y}{(x+y)^3} dx \right] dy \neq \int_0^1 \left[\int_0^1 \frac{x-y}{(x+y)^3} dy \right] dx.$$

The reason is that the function $f(x,y) = \dfrac{x-y}{(x+y)^3}$ is not continuous in $R=\{(x,y): 0 \le x \le 1, 0 \le y \le 1\}$.

6.2 PROPERTIES OF A DOUBLE INTEGRAL

1. Let $K \neq 0$ be any real number. Then,
$$\iint_R Kf(x,y)dx\,dy = K \iint_R f(x,y)dx\,dy.$$

2. The double integral of the algebraic sum of a finite number of functions f_i is equal to the sum of the double integrals taken for each function. Thus,
$$\iint_R [f_1(x,y)+f_2(x,y)+\ldots+f_n(x,y)]dx\,dy$$
$$= \iint_R f_1(x,y)dx\,dy + \iint_R f_2(x,y)dx\,dy$$
$$+\ldots+ \iint_R f_n(x,y)dx\,dy.$$

3. If the region R is partitioned into two regions R_1 and R_2, then
$$\iint_R f(x,y)dx\,dy = \iint_{R_1} f(x,y)dx\,dy$$
$$+ \iint_{R_2} f(x,y)dx\,dy.$$

6.3 EVALUATION OF DOUBLE INTEGRALS (CARTESIAN COORDINATES)

The double integrals can be evaluated using Theorems 6.1 and 6.2. In fact,

(i) If the limits in the inner integral are functions of x, then we evaluate $\int_{\phi(x)}^{\psi(x)} f(x,y)dy$, first taking x as a constant and then evaluate the integrand (function of x), obtained in the first step, integrating it with respect to x between the limits a and b. Thus,

$$\int_a^b \int_{\phi(x)}^{\psi(x)} f(x,y)dx\,dy = \int_a^b \left[\int_{\phi(x)}^{\psi(x)} f(x,y)dy \right] dx.$$

(ii) If the limits in the inner integral are functions of y, then we evaluate $\int_{\phi(y)}^{\psi(y)} f(x,y)dx$, first taking y as a constant and then evaluate the integrand (function of y), obtained as a result of the first step, integrating it with respect to y between the limits c and d. Thus,

$$\int_c^d \int_{\phi(y)}^{\psi(y)} f(x,y)dx\,dy = \int_c^d \left[\int_{\phi(y)}^{\psi(y)} f(x,y)dx \right] dy.$$

EXAMPLE 6.2

Show that

$$\int_1^2 \left[\int_3^4 (xy+e^y)dx \right] dy = \int_3^4 \left[\int_1^2 (xy+e^y)dy \right] dx.$$

Solution. The function $f(x, y) = xy + e^y$ is a continuous function over the rectangle $R = \{(x, y): 1 \le x \le 2, 3 \le y \le 4\}$. Therefore, the values of these integrals are equal. In fact, we note that

$$\int_3^4 (xy+e^y)dx = \frac{7}{2}y + e^y$$

and so,

$$\int_1^2\int_3^4 (xy+e^y)dx\,dy = \int_1^2\left[\frac{7}{2}x+e^y\right]dy$$

$$=\frac{21}{4}+e^2-e.$$

One the other hand,

$$\int_1^2 (xy+e^y)dy = \frac{3}{2}x+e^2-e$$

and so,

$$\int_3^4\left[\int_1^2 (xy+e^y)dy\right]dx = \int_3^4\left[\frac{3}{2}x+e^2-e\right]dx$$

$$=\frac{21}{4}+e^2-e.$$

Hence, the result.

EXAMPLE 6.3

Evaluate $\iint x^2y^2dx\ dy$ over the circle $x^2+y^2\le 1$.

Solution. Since $x^2+y^2\le 1$, it follows that

$$x^2\le 1 \quad\text{and}\quad y^2\le 1-x^2$$

or

$$|x|\le 1 \quad\text{and}\quad |y|\le\sqrt{1-x^2}$$

or

$$-1\le x\le 1 \quad\text{and}\quad -\sqrt{1-x^2}\le y\le\sqrt{1-x^2}.$$

The integrand $f(x, y) = x^2y^2$ is continuous over the region

$$R=\left\{(x,y): -1\le x\le 1, -\sqrt{1-x^2}\le y\le\sqrt{1-x^2}\right\}.$$

Therefore,

$$\iint_R x^2y^2dx\ dy = \int_{-1}^{1}\left[\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} x^2y^2dy\right]dx$$

$$=\int_{-1}^{1}\left[x^2\left\{\frac{y^3}{3}\right\}_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\right]dx = \int_{-1}^{1}\frac{2}{3}x^2(1-x^2)^{\frac{3}{2}}dx$$

$$=\frac{4}{3}\int_0^1 x^2(1-x^2)^{\frac{3}{2}}dx, \text{ since integrand is even}$$

$$=\frac{4}{3}\int_0^{\frac{\pi}{2}}\sin^2\theta\cos^4\theta\ d\theta, \text{ substituting } x=\sin\theta$$

$$=\frac{4}{3}\cdot\frac{1.3.1}{6.4.2}\cdot\frac{\pi}{2}=\frac{\pi}{24}.$$

EXAMPLE 6.4

Evaluate the double integral of the function $f(x, y) = 1 + x + y$ over a region bounded by $y = -x$, $x=\sqrt{y}$, and $y = 2$.

Solution. The region R is bounded by $y = -x$, the parabola $y^2 = x$, and the line $y = 2$. Thus, the limits of integration for x are $x = -y$, $x=\sqrt{y}$, $y = 0$, and $y = 2$. Therefore,

$$\iint_R (1+x+y)dx\,dy = \int_0^2\left[\int_{-y}^{\sqrt{y}} (1+x+y)dx\right]dy$$

$$=\int_0^2\left[x+\frac{x^2}{2}+xy\right]_{-y}^{\sqrt{y}}dy$$

$$=\int_0^2\left[\left(\sqrt{y}+\frac{y}{2}+y\sqrt{y}\right)-\left(-y+\frac{y^2}{2}-y^2\right)\right]dy$$

$$=\int_0^1\left(\frac{y^2}{2}+\frac{3y}{4}+\sqrt{y}+y\sqrt{y}\right)dy$$

$$=\left[\frac{y^3}{2}+\frac{3y^2}{4}+\frac{2}{3}y^{\frac{3}{2}}+\frac{2}{5}y^{\frac{5}{2}}\right]_0^2$$

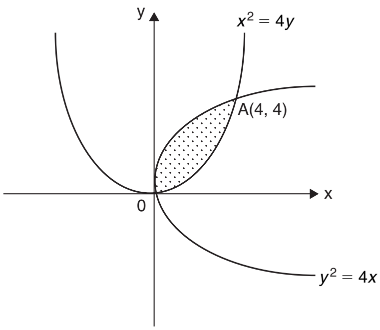
$$=\frac{13}{3}+\frac{44}{15}\sqrt{2}.$$

EXAMPLE 6.5

Evaluate $\iint_R ydx\ dy$, where R is the region bounded by the parabolas $y^2 = 4x$ and $x^2 = 4y$.

Solution. The given parabolas are $y^2 = 4x$ and $x^2 = 4y$. Solving these equations, we get $x = 0$ and $y = 4$. The corresponding values of y are $y = 0$ and $y = 4$. Both the curves pass through the origin and the points of intersection are (0, 0) and (4, 4). Thus, the limits of integration are $x = 0$ to $x = 4$ and $y=\frac{x^2}{4}$ to $y=2\sqrt{x}$. Thus,

$$R=\left\{(x,y):\ 0\le x\le 4;\ \frac{x^2}{4}\le y\le 2\sqrt{x}\right\}.$$



Therefore,

$$\iint_R y\,dx\,dy = \int_0^4 \left[\int_{\frac{x^2}{4}}^{2\sqrt{x}} y\,dy\right] dx$$

$$= \int_0^4 \left[\frac{y^2}{2}\right]_{\frac{x^2}{4}}^{2\sqrt{x}} dx = \int_0^4 \left(\frac{4x}{2} - \frac{x^4}{32}\right) dx$$

$$= \int_0^4 \left(2x - \frac{x^4}{32}\right) dx = \left[x^2 - \frac{x^5}{160}\right]_0^4 = \frac{48}{5}.$$

EXAMPLE 6.6

Calculate the double integral

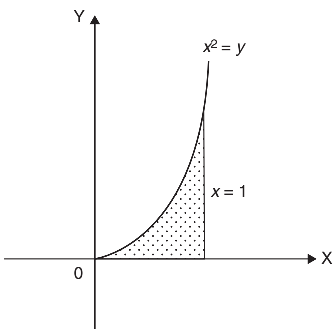
$$\int_0^1 \int_0^{x^2} (x^2 + y^2)\,dx\,dy$$

and determine the region of integration.

Solution. The region of integration is bounded by the lines $x = 0$, $x = 1$, $y = 0$, and the parabola $x^2 = y$. Thus, the region is

$$R = \left\{(x, y): 0 \le x \le 1;\ 0 \le y \le x^2\right\},$$

and is shown in the following figure:



The given integral can be expressed as

$$\int_0^1 \left[\int_0^{x^2} (x^2 + y^2)\,dy\right] dx.$$

So, we evaluate the inner integral first. We have

$$\int_0^{x^2} (x^2 + y^2)\,dy = \left[x^2 y + \frac{y^3}{3}\right]_0^{x^2} = x^4 + \frac{x^6}{3}.$$

Therefore,

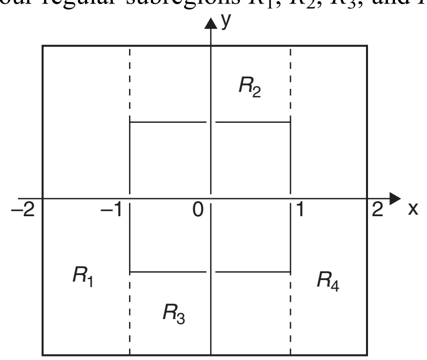
$$\iint_R (x^2 + y^2)\,dx\,dy = \int_0^1 \left(x^4 + \frac{x^6}{3}\right) dx$$

$$= \left[\frac{x^5}{5} + \frac{x^7}{21}\right]_0^1 = \frac{1}{5} + \frac{1}{21} = \frac{26}{105}.$$

EXAMPLE 6.7

Evaluate the double integral $\iint_R e^{x+y}\,dR$ over the region R, which lies between two squares with their center at the origin and with sides parallel to the axes of coordinates, if each side of the inner square is equal to 2 and that of the outer square is 4.

Solution. The region R is irregular. However, the straight lines $x = -1$ and $x = 1$ divide this region into four regular subregions R_1, R_2, R_3, and R_4.



Therefore,

$$\iint_R e^{x+y}\,dR = \iint_{R_1} e^{x+y}\,dx\,dy + \iint_{R_2} e^{x+y}\,dx\,dy$$

$$+ \iint_{R_3} e^{x+y}\,dx\,dy + \iint_{R_4} e^{x+y}\,dx\,dy$$

$$= \int_{-2}^{-1}\left[\int_{-2}^{2} e^{x+y}dy\right]dx + \int_{-1}^{1}\left[\int_{1}^{2} e^{x+y}dy\right]dx$$
$$+ \int_{-1}^{1}\left[\int_{-2}^{-1} e^{x+y}dy\right]dx + \int_{1}^{2}\left[\int_{-2}^{2} e^{x+y}dy\right]dx$$
$$= (e^2 - e^{-2})(e^{-1} - e^{-2}) + (e^2 - e)(e - e^{-1})$$
$$+ \left(e^{-1} - e^{-2}\right)\left(e - e^{-1}\right) + \left(e^2 - e^{-2}\right)$$
$$\times \left(e^2 - e\right)$$
$$= (e^3 - e^{-3})(e - e^{-1}) = 4\sinh 3 \sinh 1.$$

EXAMPLE 6.8

Evaluate $\iint xy\ dx\ dy$ over the positive quadrant of the circle $x^2 + y^2 = a^2$.

Solution. The region of integration is

$$R = \left\{(x, y):\ 0 \le x \le a;\ 0 \le y \le \sqrt{a^2 - x^2}\right\}.$$

The integrand $f(x, y) = xy$ is continuous over R. Therefore,

$$\iint_R xy\ dx\ dy = \int_0^a \left[\int_0^{\sqrt{a^2-x^2}} xy\ dy\right]dx$$
$$= \int_0^a \left[x\frac{y^2}{2}\right]_0^{\sqrt{a^2-x^2}} = \frac{1}{2}\int_0^a x[a^2 - x^2]dx$$
$$= \frac{1}{2}\left[\int_0^a a^2x\ dx - \int_0^a x^3dx\right]$$
$$= \frac{1}{2}\left[a^2\frac{x^2}{2}\right]_0^a - \frac{1}{2}\left[\frac{x^4}{4}\right]_0^a = \frac{a^4}{4} - \frac{a^4}{8} = \frac{a^4}{8}.$$

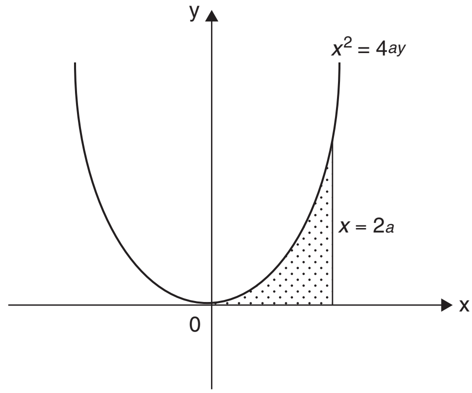
EXAMPLE 6.9

Evaluate $\iint_R xy\ dx\ dy$, where R is the domain bounded by the x-axis, ordinate $x = 2a$, and the curve $x^2 = 4ay$.

Solution. The region of integration is

$$R = \left\{(x, y):\ 0 \le x \le 2a;\ 0 \le y \le \frac{x^2}{4a}\right\}.$$

The region is bounded by $y = 0$, $x = 2a$, and the parabola $x^2 = 4ay$.



Therefore,

$$\iint_R xy\ dx\ dy = \int_0^{2a}\left[\int_0^{\frac{x^2}{4a}} xy\ dy\right]dx$$
$$= \int_0^{2a} x\left\{\left[\frac{y^2}{2}\right]_0^{\frac{x^2}{4a}}\right\}dx = \int_0^{2a}\frac{x^5}{32a^2}dx$$
$$= \frac{1}{32a^2}\left[\frac{x^6}{6}\right]_0^{2a} = \frac{1}{32a^2}\left[\frac{64a^6}{6}\right] = \frac{a^4}{3}.$$

EXAMPLE 6.10

Evaluate $\int_0^1 \int_x^{\sqrt{x}} (x^2 + y^2)dy\ dx$.

Solution. The region of integration is bounded by $x = 0$, $x = 1$, $y = x$, and the parabola $y^2 = x$. Thus,

$$R = \left\{(x, y):\ 0 \le x \le 1;\ x \le y \le \sqrt{x}\right\}.$$

Therefore,

$$\int_0^1\int_x^{\sqrt{x}} \left(x^2 + y^2\right)dy\ dx = \int_0^1\left[\int_x^{\sqrt{x}} \left(x^2 + y^2\right)dy\right]dx$$
$$= \int_0^1 \left\{\left[x^2y + \frac{y^3}{3}\right]_x^{\sqrt{x}}\right\}dx$$
$$= \int_0^1 \left[\left(x^{\frac{5}{2}} + \frac{x^{\frac{3}{2}}}{3}\right) - \left(x^3 + \frac{x^3}{3}\right)\right]dx$$

$$= \int_0^1 \left[x^{\frac{5}{2}} + \frac{x^{\frac{3}{2}}}{3} - \frac{4x^3}{3} \right] dx$$

$$= \left[\frac{x^{\frac{7}{2}}}{\frac{7}{2}} + \frac{x^{\frac{5}{2}}}{\frac{15}{2}} - \frac{4x^4}{12} \right]_0^1 = \frac{2}{7} + \frac{2}{15} - \frac{1}{3} = \frac{9}{105} = \frac{3}{35}.$$

EXAMPLE 6.11

Evaluate $\iint (x+y)^2 dx\ dy$ over the area bounded by the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

Solution. The equation of the ellipse is $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Therefore,

$$\frac{y^2}{b^2} = 1 - \frac{x^2}{a} \quad \text{or} \quad \frac{y}{b} = \pm\sqrt{1 - \frac{x^2}{a^2}} \quad \text{or}$$

$$y = \pm \frac{b}{a}\sqrt{a^2 - x^2}.$$

The ellipse cuts the x-axis at $x = \pm a$. Therefore, the region of integration is

$$R = \{(x,y):\ -a \le x \le a;\ -\frac{b}{a}\sqrt{a^2 - x^2} \le y \le \frac{b}{a}\sqrt{a^2 - x^2}\}.$$

Since $x^2 + y^2$ is an even function of y and xy is an odd function of y, we have

$$\iint_R (x+y)^2 dx\ dy$$

$$= \int_{-a}^{a} \left[\int_{-\frac{b}{a}\sqrt{a^2-x^2}}^{\frac{b}{a}\sqrt{a^2-x^2}} \left(x^2 + y^2 + 2xy\right) dy \right] dx$$

$$= \int_{-a}^{a} \left[\int_{-\frac{b}{a}\sqrt{a^2-x^2}}^{\frac{b}{a}\sqrt{a^2-x^2}} \left(x^2 + y^2\right) dy \right] dx$$

$$+ \int_{-a}^{a} \left[\int_{-\frac{b}{a}\sqrt{a^2-x^2}}^{\frac{b}{a}\sqrt{a^2-x^2}} 2xy\ dy \right] dx$$

$$= \int_{-a}^{a} \left[2 \int_{0}^{\frac{b}{a}\sqrt{a^2-x^2}} \left(x^2 + y^2\right) dy \right] dx + 0$$

$$= 2 \int_{-a}^{a} \left[x^2 y + \frac{y^3}{3} \right]_0^{\frac{b}{a}\sqrt{a^2-x^2}} dx$$

$$= 2 \int_{-a}^{a} \left[\frac{x^2 b}{a}\sqrt{a^2 - x^2} + \frac{b^3}{3a^3}\left(a^2 - x^2\right)^{\frac{3}{2}} \right] dx$$

$$= 4 \int_{0}^{a} \left[\frac{b}{a} x^2 \sqrt{a^2 - x^2} + \frac{b^3}{3a^3}\left(a^2 - x^2\right)^{\frac{3}{2}} \right] dx.$$

Substituting $x = a \sin\theta$, we have $dx = a\cos\theta\ d\theta$ and so,

$$\iint_R (x+y)^2 dx\ dy$$

$$= 4 \int_0^{\frac{\pi}{2}} \left[a^3 b \sin^2\theta \cos^2\theta + \frac{ab^3}{3}\cos^4\theta \right] d\theta$$

$$= 4\left[a^3 b . \frac{1}{4.2} . \frac{\pi}{2} + \frac{ab^3}{3} . \frac{3.1}{4.2} . \frac{\pi}{2} \right]$$

$$= \frac{\pi}{4}\left(a^3 b + ab^3\right) = \frac{\pi ab}{4}\left(a^2 + b^2\right).$$

EXAMPLE 6.12

Evaluate

$$\int_0^1 \int_0^{\sqrt{1+x^2}} \frac{1}{1 + x^2 + y^2} dy\ dx.$$

Solution. We have

$$\int_0^1 \int_0^{\sqrt{1+x^2}} \frac{1}{1 + x^2 + y^2} dy\, dx$$

$$= \int_0^1 \left[\int_0^{\sqrt{1+x^2}} \frac{1}{1 + x^2 + y^2} dy \right] dx$$

$$= \int_0^1 \left[\int_0^{\sqrt{1+x^2}} \frac{1}{(1 + x^2) + y^2} dy \right] dx$$

$$= \int_0^1 \frac{1}{\sqrt{1+x^2}} \left[\tan^{-1} \frac{y}{\sqrt{1+x^2}} \right]_0^{\sqrt{1+x^2}} dx,$$

$$\text{since} \int \frac{dx}{a^2+x^2} = \frac{1}{a}\tan^{-1}\frac{x}{a}$$

$$= \int_0^1 \frac{1}{\sqrt{1+x^2}}[\tan^{-1} 1 - \tan^{-1} 0]dx$$

$$= \frac{\pi}{4}\int_0^1 \frac{dx}{\sqrt{1+x^2}} = \frac{\pi}{4}\left[\log\left(\frac{x+\sqrt{1+x^2}}{1}\right)\right]_0^1$$

$$= \frac{\pi}{4}[\log(1+\sqrt{2}) - \log 1] = \frac{\pi}{4}\log(1+\sqrt{2}).$$

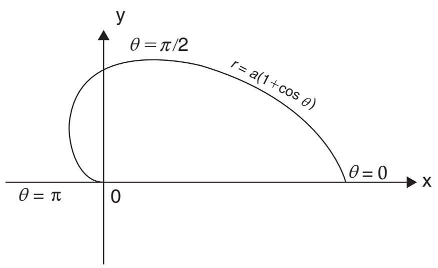
6.4 EVALUATION OF DOUBLE INTEGRALS (POLAR COORDINATES)

We wish to evaluate $\int_{\theta_1}^{\theta_2}\int_{r_1}^{r_2} f(r,\theta)dr\, d\theta$ over the region bounded by the straight lines $\theta = \theta_1$ and $\theta = \theta_2$ and the curves $r = r_1$ and $r = r_2$. To do so, we first integrate $f(r, \theta)$ with respect to r between the limits $r = r_1$ and $r = r_2$. The resulting integrand is then integrated with respect to θ between the limits $\theta = \theta_1$ and $\theta = \theta_2$.

EXAMPLE 6.13

Evaluate $\iint r\sin\theta\, dr\, d\theta$ over the area of the cardioid $r = a(1 + \cos\theta)$ above the initial line.

Solution. To evaluate the given integral above the initial line, we note that the limits of integration for r are $r = 0$ and $r = a(1 + \cos\theta)$, whereas the limits of θ are $\theta = 0$ to $\theta = \pi$.



Therefore,

$$\iint_R r\sin\theta\, dr\, d\theta$$

$$= \int_0^{\pi}\int_0^{a(1+\cos\theta)} r\sin\theta\, dr\, d\theta$$

$$= \int_0^{\pi} \sin\theta \left|\frac{r^2}{2}\right|_0^{a(1+\cos\theta)} dx$$

$$= \frac{1}{2}\int_0^{\pi} a^2(1+\cos\theta)^2 \sin\theta\, d\theta$$

$$= \frac{a^2}{2}\int_0^{\pi}\left(2\cos^2\frac{\theta}{2}\right)^2 2\sin\frac{\theta}{2}\,\cos\frac{\theta}{2}d\theta$$

$$= 4a^2\int_0^{\pi} \cos^5\frac{\theta}{2}\sin\frac{\theta}{2}d\theta$$

$$= 4a^2\int_0^{\frac{\pi}{2}} \cos^5\phi\sin\phi.2d\phi,\ \theta = 2\phi$$

$$= 8a^2\int_0^{\frac{\pi}{2}} \cos^5\phi\sin\phi\, d\phi$$

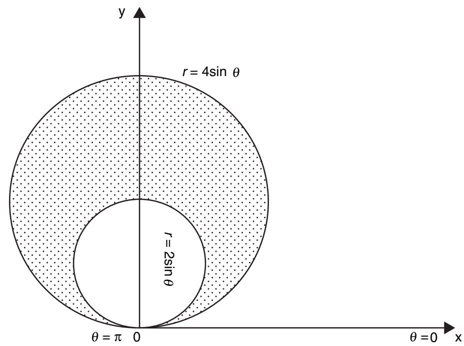
$$= -8a^2\int_0^{\frac{\pi}{2}} \cos^5\phi(-\sin\phi)d\phi$$

$$= -8a^2\left[\frac{\cos^6\phi}{6}\right]_0^{\frac{\pi}{2}} = \frac{4a^2}{3}.$$

EXAMPLE 6.14

Calculate $\iint r^3 dr\, d\theta$ over the area included between the circles $r = 2\sin\theta$ and $r = 4\sin\theta$.

Solution. The region of integration is between the circles as shown in the following figure:



Therefore,

$$\iint r^3 dr\, d\theta = \int_0^{\pi}\int_{2\sin\theta}^{4\sin\theta} r^3 dr\, d\theta = \int_0^{\pi}\left[\frac{r^4}{4}\right]_{2\sin\theta}^{4\sin\theta} d\theta$$

$$= \frac{1}{4}\int_0^{\pi}(256\sin^4\theta - 16\sin^4\theta)d\theta$$

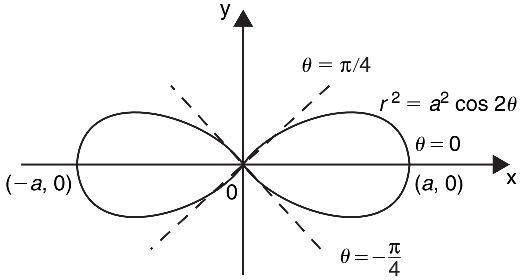
$$= 60\int_0^{\pi} \sin^4\theta\, d\theta = 120\int_0^{\frac{\pi}{2}} \sin^4\phi\, d\,\phi,\ \theta = 2\phi$$

$$= 120.\frac{3.1}{4.2}.\frac{\pi}{2} = 45\frac{\pi}{2}.$$

EXAMPLE 6.15

Evaluate $\iint \frac{r\,dr\,d\theta}{\sqrt{a^2+r^2}}$ over are loop of the lemniscates $r^2 = a^2\cos 2\theta$.

Solution. From the figure of the curve, we note that in the region of integration, r varies from 0 to $a\sqrt{\cos 2\theta}$ and θ varies from $-\frac{\pi}{4}$ to $\frac{\pi}{4}$.



Therefore,

$$\iint \frac{r\,dr\,d\theta}{\sqrt{a^2+r^2}} = \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \left[\int_0^{a\sqrt{\cos 2\theta}} \frac{r}{\sqrt{a^2+r^2}}dr\right]d\theta$$

$$= \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \left[\frac{1}{2}\int_0^{a\sqrt{\cos 2\theta}} 2r(a^2+r^2)^{-\frac{1}{2}}dr\right]d\theta$$

$$= \frac{1}{2}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \left|\frac{(a^2+r^2)^{\frac{1}{2}}}{\frac{1}{2}}\right|_0^{a\sqrt{\cos 2\theta}} d\theta$$

$$= \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} a[(1+\cos 2\theta)^{\frac{1}{2}} - 1]d\theta$$

$$= a\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} a[(2\cos^2\theta)^{\frac{1}{2}} - 1]d\theta$$

$$= a\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} (\sqrt{2}\cos\theta - 1)d\theta$$

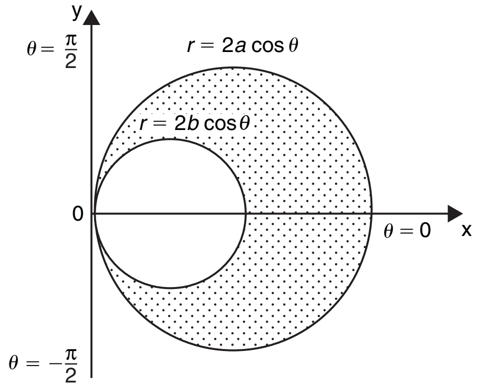
$$= 2a\int_0^{\frac{\pi}{4}} (\sqrt{2}\cos\theta - 1)d\theta$$

$$= 2a\left|\sqrt{2}\sin\theta - \theta\right|_0^{\frac{\pi}{4}} = 2a\left(1 - \frac{\pi}{4}\right).$$

EXAMPLE 6.16

Evaluate $\iint r^3 dr\ d\theta$ over the area included between the circles $r = 2a\cos\theta$, $r = 2b\cos\theta$, where $b < a$.

Solution. The region of integration between the given circles is shown in the following figure:



In the region of integration, r varies from $2b\cos\theta$ to $2a\cos\theta$, whereas θ varies from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$. Therefore,

$$\iint r^3 dr\ d\theta = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left[\int_{2b\cos\theta}^{2a\cos\theta} r^3 dr\right]d\theta$$

$$= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left[\frac{r^4}{4}\right]_{2b\cos\theta}^{2a\cos\theta} d\theta$$

$$= \frac{1}{4}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} [16a^4\cos^4\theta - 16b^4\cos^4\theta]d\theta$$

$$= 4\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (a^4 - b^4)\cos^4\theta\ d\theta$$

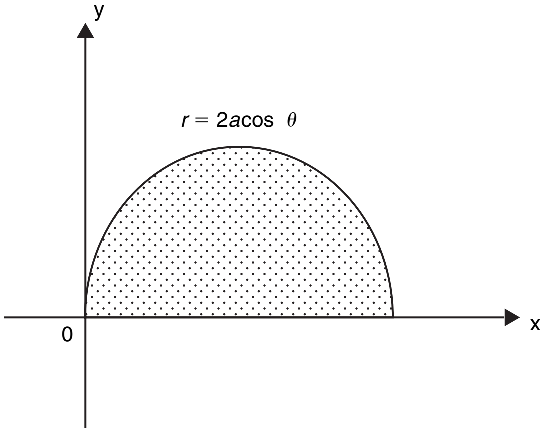
$$= 8\int_0^{\frac{\pi}{2}} (a^4 - b^4)\cos^4\theta\ d\theta$$

$$= 8(a^4 - b^4).\frac{3.1}{4.2}.\frac{\pi}{2} = \frac{3\pi}{2}(a^4 - b^4).$$

EXAMPLE 6.17

Show that $\iint_R r^2 \sin\theta \, dr \, d\theta = \frac{2a^3}{3}$, where R is the region bounded by the semi-circle $r = 2a\cos\theta$, above the initial line.

Solution. The region of integration is shown in the following figure. In this region r varies from 0 to $2a\cos\theta$, whereas θ varies from 0 to $\frac{\pi}{2}$.



Therefore,

$$\iint_R r^2 \sin\theta \, dr \, d\theta = \int_0^{\frac{\pi}{2}} \left[\int_0^{2a\cos\theta} r^2 \sin\theta \, dr \right] d\theta$$

$$= \int_0^{\frac{\pi}{2}} \sin\theta \left| \frac{r^3}{3} \right|_0^{2a\cos\theta} d\theta$$

$$= \frac{8a^3}{3} \int_0^{\frac{\pi}{2}} \sin\theta \cos^3\theta \, d\theta$$

$$= \frac{8a^3}{3} \cdot \frac{2}{4.2} = \frac{2a^3}{3}.$$

6.5 CHANGE OF VARIABLES IN A DOUBLE INTEGRAL

Regarding the change of variables in a double integral, we have the following theorem stated here without proof.

Theorem 6.3. Let D be a domain in $\Re^2$ and $f: D \to \Re$ be integrable over D. Suppose that D be mapped on to a set A of the uv-plane by the transformation $x = \phi(u, v)$ and $y = \psi(u, v)$, where ϕ and ψ have continuous partial derivatives in A and the Jacobian $\frac{\partial(x,y)}{\partial(u,v)} \neq 0$ for all $(u, v) \in A$. Then,

$$\iint_D f(x,y) dx \, dy = \iint_A f(\phi(u,v), \; \psi(u,v)) |J| du \, dv.$$

Deduction. If there is a change in variable from Cartesian- to polar coordinates, then $x = r\cos\theta$, $y = r\sin\theta$, and

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix}$$

$$= r\left(\cos^2\theta + \sin^2\theta\right) = r.$$

Therefore,

$$\iint_D f(x, y) dx \, dy = \iint_A f(r\cos\theta, \; r\sin\theta) \, r \, d\theta \, dr.$$

EXAMPLE 6.18

Calculate the integral

$$\iint \frac{(x-y)^2}{x^2+y^2} dx \, dy$$

over the circle $x^2 + y^2 \le 1$.

Solution. The value of the given integral is four times the value of the integral taken over the positive quadrant of the circle $x^2 + y^2 = 1$. Substituting $x = r\cos\theta$ and $y = r\sin\theta$, the given integral is equal to

$$4 \int_0^1 \int_0^{\frac{\pi}{2}} \frac{1}{r^2} (r\cos\theta - r\sin\theta)^2 \frac{\partial(x,y)}{\partial(r,\theta)} dr \, d\theta$$

$$= 4 \int_0^1 r \int_0^{\frac{\pi}{2}} (\cos^2\theta + \sin^2\theta - 2\sin\theta\cos\theta) dr \, d\theta$$

$$= 4 \int_0^{\frac{\pi}{2}} (1 - \sin 2\theta) \left[\int_0^1 r dr \right] d\theta$$

$$= 2 \int_0^{\frac{\pi}{2}} (1 - \sin 2\theta) d\theta$$

$$= 2 \left[\theta + \frac{\cos 2\theta}{2} \right]_0^{\frac{\pi}{2}} = \pi - 2.$$

EXAMPLE 6.19

Evaluate $\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}dy\,dx$ by changing to polar coordinates. Hence, deduce that $\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$.

Solution. In the given integral, both x and y vary from 0 to ∞. Hence, the region of integration is xy-plane. Changing to polar coordinates by substituting $x = r\cos\theta$ and $y = r\sin\theta$, we get $x^2+y^2 = r^2$; and in the region of integration, r varies from 0 to ∞ and θ varies from 0 to $\frac{\pi}{2}$. Thus,

$$\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}dy\,dx = \int_0^{\frac{\pi}{2}}\int_0^\infty e^{-r^2}\cdot\frac{\partial(x,y)}{\partial(r,\theta)}dr\,d\theta$$

$$= \int_0^{\frac{\pi}{2}}\int_0^\infty re^{-r^2}dr\,d\theta$$

$$= \int_0^{\frac{\pi}{2}}\int_0^\infty e^{-t}.\frac{dt}{2}\,d\theta,\ r^2 = t$$

$$= \frac{1}{2}\int_0^{\frac{\pi}{2}}\left|\frac{e^{-t}}{-1}\right|_0^\infty d\theta$$

$$= \frac{-1}{2}\int_0^{\frac{\pi}{2}}(0-1)d\theta = \frac{\pi}{4}.\quad(1)$$

On the other hand,

$$\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}dy\,dx = \int_0^\infty e^{-x^2}dx\int_0^\infty e^{-y^2}dy$$

$$= \left[\int_0^\infty e^{-r^2}dx\right]^2.$$

Hence, (1) implies

$$\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}.$$

EXAMPLE 6.20

Evaluate $\iint\sqrt{\frac{a^2b^2-b^2x^2-a^2y^2}{a^2b^2+b^2x^2+a^2y^2}}\,dx\,dy$ over the positive quadrant of the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2} = 1$.

Solution. Substituting $\frac{x}{a} = X$ and $\frac{y}{b} = Y$, the problem reduces to the evaluation of $\iint ab\sqrt{\frac{1-X^2-Y^2}{1+X^2+Y^2}}dx\,dy$ over the positive quadrant of the circle $X^2+Y^2 = 1$. Substituting $X = r\cos\theta$ and $Y = r\sin\theta$, we have $dXdY = rdr\,d\theta$. In the region of integration r varies from 0 to 1 and θ varies from 0 to $\frac{\pi}{2}$. Hence, the given integral reduces to

$$ab\int_0^{\frac{\pi}{2}}\int_0^1\sqrt{\frac{1-r^2}{1+r^2}}rdr\,d\theta$$

$$= ab\int_0^1\sqrt{\frac{1-r^2}{1+r^2}}rdr\int_0^{\frac{\pi}{2}}d\theta$$

$$= \frac{ab\pi}{2}\int_0^1\sqrt{\frac{1-r^2}{1+r^2}}rdr$$

$$= \frac{ab\pi}{2}\int_0^{\frac{\pi}{2}}\sqrt{\frac{1-\sin t}{1+\sin t}}.\frac{1}{2}\cos t\,dt,\ r^2 = \sin t$$

$$= \frac{\pi ab}{4}\int_0^{\frac{\pi}{2}}\sqrt{\frac{1-\sin t}{1+\sin t}}.\frac{\sqrt{1-\sin t}}{\sqrt{1-\sin t}}\cos t\,dt$$

$$= \frac{\pi ab}{4}\int_0^{\frac{\pi}{2}}\frac{1-\sin t}{\cos t}\cos t\,dt$$

$$= \frac{\pi ab}{4}\int_0^{\frac{\pi}{2}}(1-\sin t)\,dt$$

$$= \frac{\pi ab}{4}[t+\cos t]_0^{\frac{\pi}{2}} = \frac{\pi ab}{4}\left[\frac{\pi}{2}-1\right]$$

$$= \frac{\pi ab}{8}(\pi-2).$$

Remark 6.1. Substituting $a = b = 1$, the problem reduces to the evaluation of $\iint\sqrt{\frac{1-x^2-y^2}{1+x^2+y^2}}dx\,dy$ over the positive quadrant of the circle $x^2+y^2 = 1$. The value of the integral in that case is $\frac{\pi^2}{8}-\frac{\pi}{4}$.

EXAMPLE 6.21

Evaluate $\iint\limits_{x^2+y^2\le 1} \sin \pi(x^2+y^2)dx\, dy$.

Solution. Changing to polar form yields

$$\iint\limits_{x^2+y^2\le 1} \sin \pi\left(x^2+y^2\right)dx\, dy$$

$$= \int_0^{2\pi}\int_0^1 \sin\left(\pi r^2\right)\, rdr\, d\theta$$

$$= \int_0^{2\pi}\int_0^1 \sin(\pi\ t)\frac{dt}{2}\ d\theta$$

$$= \frac{1}{2}\int_0^{2\pi} \left|-\frac{\cos \pi t}{\pi}\right|_0^1 d\theta$$

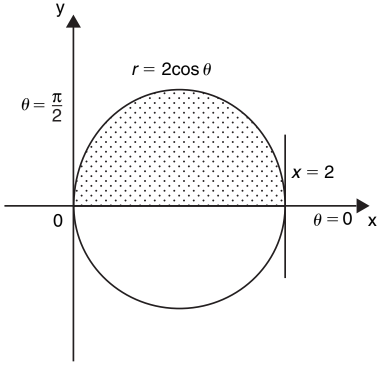
$$= -\frac{1}{2\pi}\int_0^{2\pi} [\cos \pi - \cos 0]d\theta$$

$$= \frac{2}{2\pi}\int_0^{2\pi} d\theta = 2.$$

EXAMPLE 6.22

Evaluate $\int_0^2 \int_0^{\sqrt{2x-x^2}} \frac{x}{\sqrt{x^2+y^2}}dy\, dx$ by changing to polar coordinates.

Solution. The limits of y are from 0 to $\sqrt{2x-x^2}$. But $y=\sqrt{2x-x^2}$ implies $y^2 = 2x - x^2$ or $x^2+y^2 = 2x$. Thus, the region of integration is bounded by $x = 0$, $x = 2$, $y = 0$, and $x^2 + y^2 = 2x$. Changing to polar coordinates $x^2 + y^2 = 2x$ transforms to $r^2 = 2r\cos\theta$ or $r = 2\cos\theta$.



For the region of integration, r varies from 0 to $2\cos\theta$, whereas θ varies from 0 to $\frac{\pi}{2}$. Therefore,

$$\int_0^2 \int_0^{\sqrt{2x-x^2}} \frac{x}{\sqrt{x^2+y^2}}dy\, dx$$

$$= \int_0^{\frac{\pi}{2}}\int_0^{2\cos\theta} \frac{r\cos\theta}{r} r\ dr\ d\theta$$

$$= \int_0^{\frac{\pi}{2}} \cos\theta \int_0^{2\cos\theta} r\ dr\ d\theta$$

$$= \int_0^{\frac{\pi}{2}} \cos\theta \left|\frac{r^2}{2}\right|_0^{2\cos\theta} d\theta$$

$$= 2\int_0^{\frac{\pi}{2}} \cos^3\theta\ d\theta = 2\cdot\frac{2}{3} = \frac{4}{3}.$$

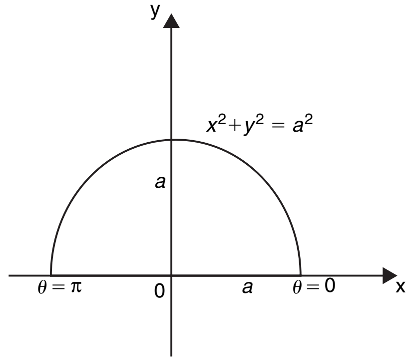
EXAMPLE 6.23

Transform to the Cartesian form and hence, evaluate the integral $\int_0^{\pi}\int_0^{a} r^3 \sin\theta \cos\theta\ dr\ d\theta$.

Solution. We are given that

$$I = \int_0^{\pi}\int_0^{a} r^3 \sin\theta\cos\theta\ dr\ d\theta.$$

Put $x = r\cos\theta$ and $y = r\sin\theta$ so that $dx\, dy = r\, dr\, d\theta$. The region of integration is shown in the following figure:



Therefore, the Cartesian form of the given integral is

$$I = \int_{-a}^{a}\int_0^{a} xy\ dx\ dy.$$

Further,

$$I=\int_0^{\pi}\int_0^{a} r^3\sin\theta\cos\theta\,dr\,d\theta=\int_0^{\pi}\sin\theta\cos\theta\left|\frac{r^4}{4}\right|_0^a d\theta$$

$$=\frac{a^4}{4}\int_0^{\pi}\sin\theta\cos\theta\,d\theta=\frac{a^4}{4}\left|\frac{\sin^2\theta}{2}\right|_0^{\pi}=0.$$

EXAMPLE 6.24

Use the transformation $x+y=u$ and $y=uv$ to show that $\int_0^1\int_0^{1-x} e^{\frac{y}{x+y}}\,dy\,dx=\frac{e-1}{2}$.

Solution. We have

$$x=u-y=u-uv=u(1-v) \text{ and}$$

$$y=uv.$$

Therefore,

$$\frac{\partial(x,y)}{\partial(u,v)}=\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix}=\begin{vmatrix}1-v & -u\\ v & u\end{vmatrix}$$

$$=u(1-v)-(-uv)=u.$$

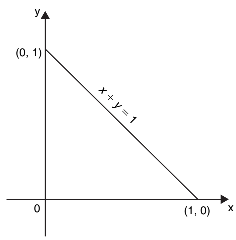
The Jacobian vanishes when $u=0$, that is, when $x=y=0$, but not otherwise. Also the origin (0, 0) corresponds to the whole line $u=0$ of the uv-plane so that the correspondence ceases to be one-to-one. In order to exclude (0, 0), we note that the given integral exists as the limit, when $h\to 0$ of the integral over the region is bounded by

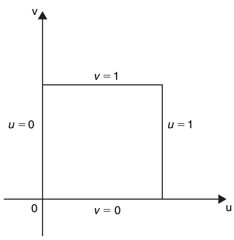
$$x+y=1,\ x=h,\ \text{and } y=0 \text{ where } h>0.$$

The transformed region is then bounded by the lines

$$u=1,\ v=0,\ \text{and } u(1-v)=h.$$

When $h\to 0$, the new region of the uv-plane tends, as its limit, to the square bounded by the lines $u=1$, $v=1$, $u=0$, and $v=0$. Thus, the region of integration in the xy- and uv-planes are as shown in the following figures:





Therefore,

$$\int_0^1\int_0^{1-x} e^{\frac{y}{x+y}}\,dy\,dx=\int_0^1\int_0^1 e^{\frac{uv}{u}}.u\,du\,dv$$

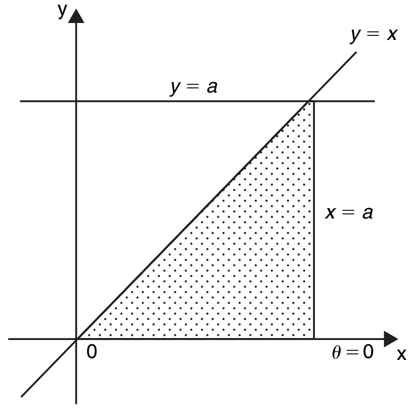
$$=\int_0^1 e^v\int_0^1 u\,du\,dv=\int_0^1 e^v\left[\frac{u^2}{2}\right]_0^1 dv$$

$$=\frac{1}{2}\int_0^1 e^v dv=\frac{1}{2}[e^v]_0^1=\frac{1}{2}(e-1).$$

EXAMPLE 6.25

Evaluate $\int_0^a\int_y^a \frac{x\,dx\,dy}{x^2+y^2}$ by changing into polar coordinates.

Solution. The region of integration is shown in the following figure:



The region is bounded by $x=y$, $x=a$, $y=0$, and $y=a$. Changing to polar coordinates, we have $x=r\cos\theta$, $y=r\sin\theta$, and $dx\,dy=r\,dr\,d\theta$. Further, in the region of integration θ varies from 0 to $\frac{\pi}{4}$. Also, $x=a$ implies $r\cos\theta=a$ or $r=\frac{a}{\cos\theta}$. Therefore, r varies from 0 to $\frac{a}{\cos\theta}$. Hence,

$$\int_0^a\int_y^a \frac{x\,dx\,dy}{x^2+y^2}=\int_0^{\frac{\pi}{4}}\int_0^{\frac{a}{\cos\theta}}\frac{r\cos\theta}{r^2}r\,dr\,d\theta$$

$$=\int_0^{\frac{\pi}{4}}\cos\theta[r]_0^{\frac{a}{\cos\theta}}d\theta=a\int_0^{\frac{\pi}{4}}d\theta=\frac{a\pi}{4}.$$

EXAMPLE 6.26

Evaluate $\iint xy\,(x^2+y^2)^{\frac{n}{2}}dx\,dy$ over the positive octant of the circle $x^2+y^2=4$, supposing $n+3>0$.

Solution. The region of integration is bounded by $x = 0$, $x = 2$, $y = 0$, and $y = 2$. Changing to polar coordinates, we have $x = r\cos\theta$ and $y = r\sin\theta$ and so, $rd\theta\, dr = dxdy$. The limits of integration in the first quadrant of the given circle are now $r = 0$ to $r = 2$ and $\theta = 0$ to $\theta = \frac{\pi}{2}$. Hence,

$$\iint xy\,(x^2+y^2)^{\frac{n}{2}}dx\,dy$$

$$= \int_0^{\frac{\pi}{2}}\int_0^2 r\cos\theta.r\sin\theta\left(r^2\right)^{\frac{n}{2}}.r\,dr\,d\theta$$

$$= \int_0^{\frac{\pi}{2}} \sin\theta\cos\theta\left[\int_0^2 r^{n+3}dr\right]d\theta$$

$$= \int_0^{\frac{\pi}{2}} \sin\theta\cos\theta\left[\frac{r^{n+4}}{n+4}\right]_0^2 d\theta$$

$$= \frac{2^{n+4}}{n+4}\int_0^{\frac{\pi}{2}} \sin\theta\cos\theta\,d\theta = \frac{2^{n+4}}{n+4}.\frac{1.1}{2} = \frac{2^{n+3}}{n+4}.$$

6.6 CHANGE OF ORDER OF INTEGRATION

We have seen that, in a double integration, if the limits of both variables are constant, then we can change the order of integration without affecting the result. But if the limits of integration are variable, a change in the order of integration requires a change in the limits of integration. Some integrals are easily evaluated by changing the order of integration in them.

EXAMPLE 6.27

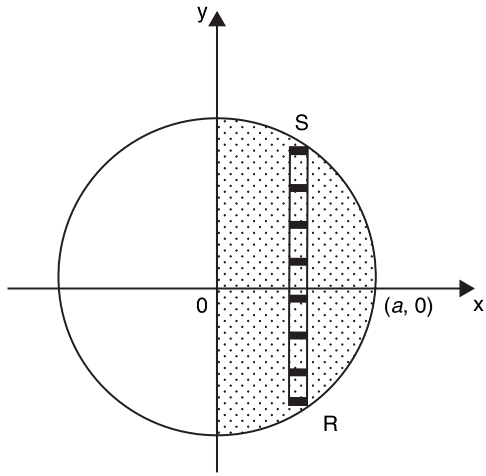
Change the order of integration in the integral

$$I = \int_{-a}^{a}\int_0^{\sqrt{a^2-y^2}} f(x,y)dx\,dy.$$

Solution. The region of integration is bounded by $y = -a$, $y = a$, $x = 0$, and $x^2 + y^2 = a^2$. We have

$$I = \int_{-a}^{a}\left[\int_0^{\sqrt{a^2-y^2}} f(x,y)dx\right]dy.$$

Thus, in the given form, we first integrate with respect to x and then with respect to y.



On changing the order of integration, we first integrate with respect to y, along a vertical ship RS, which extends from $y = -\sqrt{a^2-x^2}$ to $y = \sqrt{a^2-x^2}$. To cover the whole region of integration, we then integrate with respect to x from $x = 0$ to $x = a$. Thus,

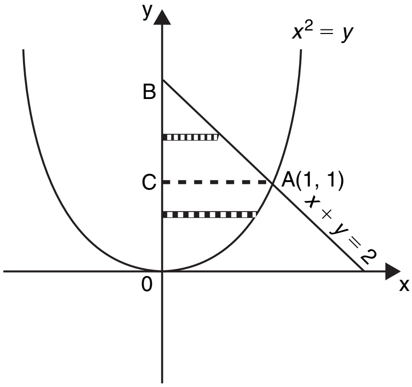
$$I = \int_0^a dx \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} f(x,y)\,dy$$

$$= \int_0^a\left[\int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} f(x,\ y)\,dy\right]dx.$$

EXAMPLE 6.28

Change the order of integration in $I = \int_0^1\int_{x^2}^{2-x} xydx\,dy$ and hence, evaluate the same.

Solution. For a given integral, the region of integration is bounded by $x = 0$, $x = 1$, $y = x^2$, (parabola), and the line $y = 2 - x$. Thus, the region of integration OABO is as shown in the following figure:



In the given form of the integral, we have to integrate first with respect to y and then with respect to x. Therefore, on changing the order of integration, we first integrate the integrand, with respect to x and then, with respect to y. The integration with respect to x requires the splitting-up of the region OABO into two parts OACO and the triangle ABC. For the subregion OACO, the limit of integration are from $x = 0$ to $x = \sqrt{y}$ and $y = 0$ to $y = 1$. Thus, the contribution to the integral I from the region OACO is

$$I_1 = \int_0^1 \left[\int_0^{\sqrt{y}} xy\, dx \right] dy.$$

For the subregion ABC, the limits of integration are from $x = 0$ to $x = 2 - y$ and $y = 1$ to $y = 2$. Thus, the contribution to I from the subregion ABC is

$$I_2 = \int_1^2 \left[\int_0^{2-y} xy\, dx \right] dy.$$

Hence, on changing the order of integration, we get

$$I = \int_0^1 \left[\int_0^{\sqrt{y}} xy\, dx \right] dy + \int_1^2 \left[\int_0^{2-y} xy\, dx \right] dy$$

$$= \int_0^1 \left[y\frac{x^2}{2} \right]_0^{\sqrt{y}} dy + \int_1^2 \left[\frac{yx^2}{2} \right]_0^{2-y} dy$$

$$= \frac{1}{2}\int_0^1 y^2 dy + \frac{1}{2}\int_1^2 y(2-y)^2 dy$$

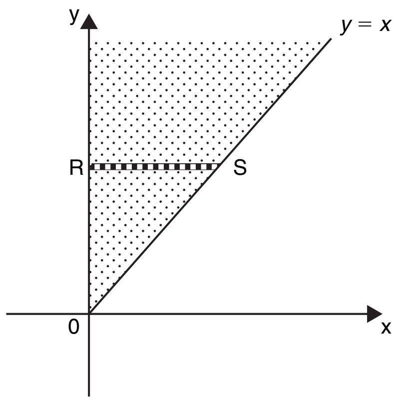
$$= \frac{1}{2}\left[\frac{y^3}{3}\right]_0^1 + \frac{1}{2}\left[\frac{y^4}{4} + 4\frac{y^2}{2} - 4\frac{y^3}{3}\right]_1^2$$

$$= \frac{1}{6} + \frac{5}{24} = \frac{3}{8}.$$

EXAMPLE 6.29

Changing the order of integration, find the value of the integral $\int_0^\infty \int_x^\infty \frac{e^{-y}}{y} dy\, dx$.

Solution. The region of integration is bounded by $x = 0$ and $y = x$. The limits of x are from 0 to ∞ and those of y are from x to ∞. The region of integration is shown in the following figure:



On changing the order of integration, we first integrate the integrand, with respect to x, along a horizontal strip RS, which extends from $x = 0$ to $x = y$. To cover the region of integration, we then integrate, with respect to y, from $y = 0$ to $y = \infty$. Thus,

$$I = \int_0^\infty \left[\int_0^y \frac{e^{-y}}{y} dx \right] dy = \int_0^\infty \frac{e^{-y}}{y} [x]_0^y\, dy$$

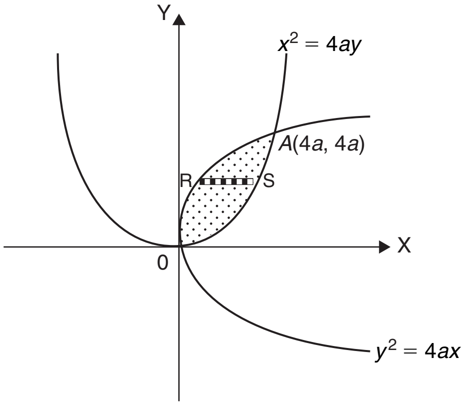
$$= \int_0^\infty e^{-y}\, dy = [-e^{-y}]_0^\infty = -\left[\frac{1}{e^y}\right]_0^\infty$$

$$= -(0 - 1) = 1.$$

EXAMPLE 6.30

Change the order of integration in the integral $\int_0^{4a}\int_{\frac{x^2}{4a}}^{2\sqrt{ax}} dy\,dx$ and evaluate.

Solution. The given integral is $\int_0^{4a}\int_{\frac{x^2}{4a}}^{2\sqrt{ax}} dy\,dx$. The integration is first carried out with respect to y and then with respect to x. The region of integration is bounded by $x = 0$, $x = 4a$, and the parabolas $x^2 = 4ay$ and $y^2 = 4ax$. Thus, the region of integration is as shown in the following figure:



The coordinates at the point of intersection of the parabolas are $A(4a, 4a)$.

On changing the order of integration, we first integrate the integrand, with respect to x, along the horizontal strip RS, which extends from $x = \frac{y_2}{4a}$ to $x = \sqrt{4ay} = 2\sqrt{ay}$. To cover the region of integration, we then integrate with respect to y from $y = 0$ to $y = 4a$. Thus,

$$\int_0^{4a}\int_{\frac{x^2}{4a}}^{2\sqrt{ax}} dy\,dx = \int_0^{4a}\left[\int_{\frac{y^2}{4a}}^{2\sqrt{ay}} dx\right]dy = \int_0^{4a} [x]_{\frac{y^2}{4a}}^{2\sqrt{ay}} dy$$

$$= \int_0^{4a}\left[2\sqrt{ay} - \frac{y^2}{4a}\right]dy = \left[\frac{2\sqrt{a}y^{\frac{3}{2}}}{\frac{3}{2}} - \frac{y^3}{12a}\right]_0^{4a}$$

$$= \frac{32a^2}{3} - \frac{16a^2}{3} = \frac{16a^2}{3}.$$

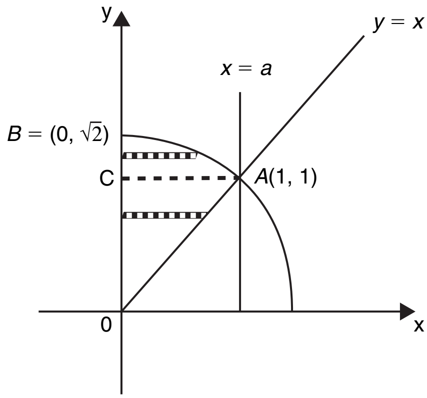
EXAMPLE 6.31

Evaluate the integral

$$\int_0^1\int_x^{\sqrt{2-x^2}} \frac{x\,dy\,dx}{\sqrt{x^2+y^2}}$$

by changing the order of integration.

Solution. For the given integral, the region of integration is bounded by $x = 0$, $x = 1$, $y = x$, and the circle $x^2 + y^2 = 2$. Thus, the region of integration is as shown in the following figure:



The point of intersection of the circle $x^2 + y^2 = 2$ and $x = y$ is $A(1, 1)$. Draw *AC OB*. Thus, the region of integration is divided into two subregions ABCA and *ACO*.

On changing the order of integration, we first integrate with respect to x, along the strips parallel to the x-axis.

In the subregion ABCA, the strip extends from $x = 0$ to $x = \sqrt{2-y^2}$. To cover the subregion, we then integrate with respect to y from $y = 1$ to $y = \sqrt{2}$. Thus, the contribution to the integral due to this subregion is

$$I_1 = \int_1^{\sqrt{2}}\left[\int_0^{\sqrt{2-y^2}} \frac{x}{\sqrt{x^2+y^2}}dx\right]dy.$$

On the other hand, in the subregion ACO, the strip extends from $x = 0$ to $x = y$. To cover this subregion, we then integrate with respect to y from $y = 0$ to $y = 1$. Thus, the contribution to the integral by this subregion is

$$I_2 = \int_0^1 \left[\int_0^y \frac{y}{\sqrt{x^2+y^2}} dx \right] dy.$$

Hence, the given integral is equal to

$$I = I_1 + I_2 = \int_1^{\sqrt{2}} \left[\int_0^{\sqrt{2-y^2}} \frac{x}{\sqrt{x^2+y^2}} dx \right] dy$$

$$+ \int_0^1 \left[\int_0^y \frac{x}{\sqrt{x^2+y^2}} dx \right] dy$$

$$= \int_1^{\sqrt{2}} [(x^2+y^2)^{\frac{1}{2}}]_0^{\sqrt{2-y^2}} dy + \int_0^1 [(x^2+y^2)^{\frac{1}{2}}]_0^y dy$$

$$= \int_1^{\sqrt{2}} \left(\sqrt{2} - y\right) dy + \int_0^1 \left(\sqrt{2}y - y\right) dy$$

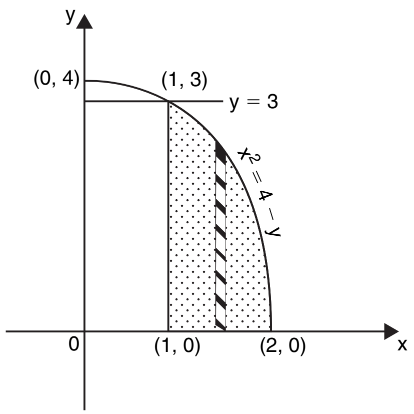
$$= \left[\sqrt{2}y - \frac{y^2}{2}\right]_1^{\sqrt{2}} + \left(\sqrt{2}-1\right)\left[\frac{y^2}{2}\right]_0^1$$

$$= \frac{2-\sqrt{2}}{2} = 1 - \frac{1}{\sqrt{2}}.$$

EXAMPLE 6.32

Evaluate the integral $\int_0^3 \int_1^{\sqrt{4-y}} (x+y) dx\, dy$ by changing the order of integration.

Solution. The region of integration is bounded by $x = 1$, $x^2 = 4 - y$, $y = 0$, and $y = 3$, as shown in the following figure:



On changing the order of integration, we first integrate the integrand, with respect to y, by taking the strip parallel to the axis of y. In the region of integration, y varies from 0 to $4 - x^2$ and x varies from 1 to 2. Therefore,

$$\int_0^3 \int_0^{\sqrt{4-y}} (x+y) dx\, dy$$

$$= \int_1^2 \left[\int_0^{4-x^2} (x+y) dy \right] dx = \int_1^2 \left[xy + \frac{y^2}{2} \right]_0^{4-x^2} dx$$

$$= \int_1^2 \left[4x - x^3 + 8 + \frac{x^4}{2} - 4x^2 \right] dx$$

$$= \left[4\frac{x^2}{2} - \frac{x^4}{4} + 8x + \frac{x^5}{10} - \frac{4x^3}{3} \right]_1^2$$

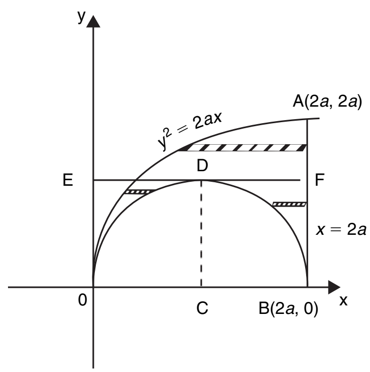
$$= \left(8 - 4 + 16 + \frac{32}{10} - \frac{32}{3} \right)$$

$$- \left(2 - \frac{1}{4} + 8 + \frac{1}{10} - \frac{4}{3} \right) = \frac{241}{60}.$$

EXAMPLE 6.33

Change the order of integration in $I = \int_0^{2a} \int_{\sqrt{2ax-x^2}}^{\sqrt{2ax}} \phi(x,y) dy\, dx.$

Solution. The region of integration is bounded by $x = 0$ and $x = 2a$, the circle $x^2 + y^2 = 2ax$, and the parabola $y^2 = 2ax$. The equation of the circle can be written as $(x - a)^2 + y^2 = a^2$ and so, has the center at $(a, 0)$. The region of integration is as shown in the following figure:



We divide the region of integration into three parts by drawing the line EDF through D parallel to the x-axis. On changing the order of integration, we first integrate the integrand, with respect to x and then integrate the resultant integrand, with respect to y. So, we draw horizontal strips parallel to the x-axis.

In the subregion OEDO, x varies from $\frac{y^2}{2a}$ to $a - \sqrt{a^2 - y^2}$ and y varies from 0 to a. Thus, the contribution to the integral due to this subregion is

$$I_1 = \int_0^a \left[\int_{\frac{y^2}{2a}}^{a-\sqrt{a^2-y^2}} \phi(x, y)dx \right] dy.$$

Similarly, the contribution to the integral due to the subregion DBFD is

$$I_2 = \int_0^a \left[\int_{a+\sqrt{a^2-y^2}}^{2a} \phi(x, y)dx \right] dy,$$

and the contribution to the integral due to the subregion AEFA is

$$I_3 = \int_0^a \left[\int_{\frac{y^2}{2a}}^{2a} \phi(x, y)dx \right] dy.$$

Hence,

$$I = \int_0^a \int_{\frac{y^2}{2a}}^{a-\sqrt{a^2-y^2}} \phi(x, y)dx\, dy + \int_0^a \int_{a+\sqrt{a^2-y^2}}^{2a} \phi(x, y)dx\, dy + \int_0^{2a} \int_{\frac{y^2}{2a}}^{2a} \phi(x, y)dx\, dy.$$

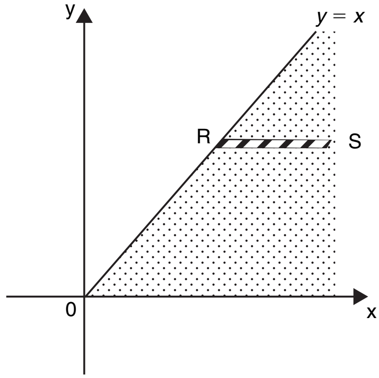
EXAMPLE 6.34

Evaluate $\int_0^\infty \int_0^x x\, e^{-\frac{x^2}{y}} dy\, dx$ by changing the order of integration.

Solution. The region of integration is bounded by the lines

$$x = 0,\ x = \infty,\ y = 0, \text{ and } y = x.$$

Therefore, the region of integration is as shown in the following figure:



On changing the order of integration, we first integrate, with respect to x and then, with respect to y. Thus,

$$\int_0^\infty \int_0^x x\, e^{-\frac{x^2}{y}} dy\, dx = \int_0^\infty \left[\int_y^\infty x\, e^{-\frac{x^2}{y}} dx \right] dy. \quad (1)$$

We first evaluate the inner integral. Substituting $x^2 = t$, we have $2x\, dx = dt$. When $x = y$, $t = y^2$ and when $x = \infty$, $t = \infty$, Therefore,

$$\int_y^\infty x\, e^{-\frac{x^2}{y}} dx = \frac{1}{2} \int_{y^2}^\infty e^{-\frac{t}{y}} dt = \frac{1}{2} \left[\frac{e^{-\frac{t}{y}}}{-\frac{1}{y}} \right]_{y^2}^\infty$$

$$= \frac{1}{2} y\, e^{-y}.$$

Therefore, (1) reduces to

$$\int_0^\infty \int_0^x x\, e^{-\frac{x^2}{y}} dy\, dx = \frac{1}{2} \int_0^\infty y\, e^{-y} dy$$

$$= \frac{1}{2} \left[\frac{y\, e^{-y}}{-1} \right]_0^\infty + \frac{1}{2} \int_0^\infty e^{-y} dy$$

$$= \frac{1}{2} \left[\frac{e^{-y}}{-1} \right]_0^\infty = \frac{1}{2}.$$

6.7 AREA ENCLOSED BY PLANE CURVES (CARTESIAN AND POLAR COORDINATES)

(A) Cartesian Coordinates: The area A of the region

$$R = \{(x, y):\ a \le x \le b;\ f_1(x) \le y \le f_2(x)\}$$

is given by the double integral

$$A = \int_a^b \left[\int_{f_1(x)}^{f_2(x)} dy \right] dx.$$

Similarly, the area A of the region $R = \{(x, y): \ c \le y \le d; \ f_1(y) \le x \le f_2(y)\}$ is given by the double integral

$$A = \int_c^d \left[\int_{f_1(y)}^{f_2(y)} dx \right] dy.$$

(B) Polar Coordinates: The area A of the region

$$R = \{(r, \theta); \ \alpha \le \theta \le \beta; \ f_1(\theta) \le r \le f_2(\theta)\}$$

is given by

$$A = \int_\alpha^\beta \left[\int_{f_1(\theta)}^{f_2(\theta)} r\, dr \right] d\theta.$$

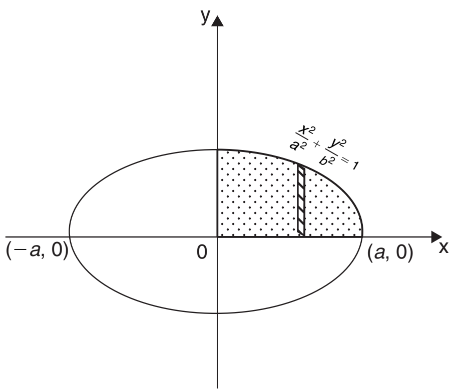
Similarly, the area A of the region $R = \{(r, \theta); \ r_1 \le r \le r_2; \ f_1(r) \le \theta \le f_2(r)\}$ is given by

$$A = \int_{r_1}^{r_2} \left[\int_{f_1(r)}^{f_2(r)} d\theta \right] r\, dr.$$

EXAMPLE 6.35

Find the area of a plate in the form of a quadrant of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Hence, find the area enclosed by the given ellipse.

Solution. From the figure, we note that the required area is bounded by $x = 0$, $x = a$, $y = 0$, and $y = \frac{b}{a}\sqrt{a^2 - x^2}$.



Thus,

$$A = \int_0^a \left[\int_0^{\frac{b}{a}\sqrt{a^2-x^2}} dy \right] dx = \int_0^a [y]_0^{\frac{b}{a}\sqrt{a^2-x^2}} dx$$

$$= \frac{b}{a} \int_0^a \sqrt{a^2 - x^2}\, dx$$

$$= \frac{b}{a} \left[\frac{x\sqrt{a^2 - x^2}}{2} + \frac{a^2}{2} \sin^{-1} \frac{x}{a} \right]_0^a$$

$$= \frac{b}{a} \left[\left(0 + \frac{a^2}{2} \sin^{-1} 1 \right) - 0 \right]$$

$$= \frac{ba}{2} \sin^{-1} 1 = \frac{ba}{2} \cdot \frac{\pi}{2}$$

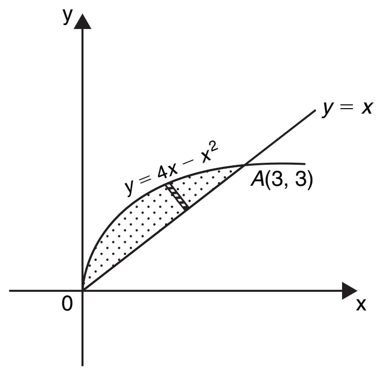
$$= \frac{\pi ab}{4} \text{ sq units.}$$

Hence, the total area enclosed by the given ellipse is four times the area enclosed by the plate in the form of one quadrant $= \pi\, ab$ sq units.

EXAMPLE 6.36

Find the area lying between the parabola $y = 4x - x^2$ and the line $y = x$.

Solution. The parabola passes through the origin. Solving $y = 4x - x^2$ and $y = x$ for x, we get $x = 0$ and $x = 3$. Thus, the curves $y = 4x - x^2$ and $y = x$ intersect at $x = 0$ and $x = 3$. When $0 < x < 3$, $4x - x^2$ is greater than x. Therefore, the region of integration is as shown in the following figure:



Thus, the required area lies between $y=x$, $y=4x-x^2$, $x=0$, and $x=3$. Therefore,

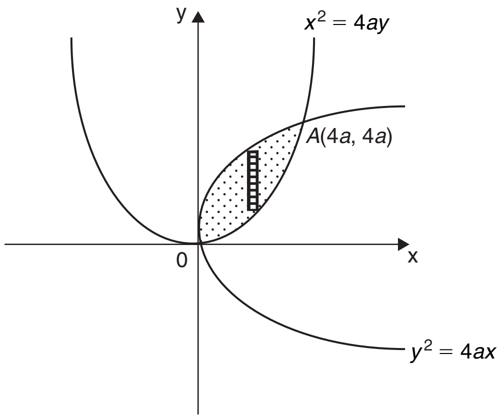
$$A=\int_0^3\left[\int_x^{4x-x^2}dy\right]dx=\int_0^3[y]_x^{4x-x^2}dx=\int_0^3(3x-x^2)dx$$

$$=\left[\frac{3x^2}{2}-\frac{x^3}{3}\right]_0^3=\frac{27}{2}-\frac{27}{3}=\frac{9}{2}.$$

EXAMPLE 6.37

Find the area lying between the parabola $y^2=4ax$ and $x^2=4ay$.

Solution. Solving the equation of the given parabola, we have O(0, 0) and $A(4a, 4a)$ as the points of intersection. The region of integration is shown in the following figure:



Therefore, the required area is

$$A=\int_0^{4a}\left[\int_{\frac{x^2}{4a}}^{2\sqrt{ax}}dy\right]dx=\int_0^{4a}[y]_{\frac{x^2}{4a}}^{2\sqrt{ax}}$$

$$=\int_0^{4a}\left[2\sqrt{ax}-\frac{x^2}{4a}\right]dx$$

$$=2\sqrt{a}\int_0^{4a}\sqrt{x}dx-\frac{1}{4a}\int_0^{4a}x^2dx$$

$$=2\sqrt{a}\left[\frac{x^{\frac{3}{2}}}{\frac{3}{2}}\right]_0^{4a}-\frac{1}{4a}\left[\frac{x^3}{3}\right]_0^{4a}$$

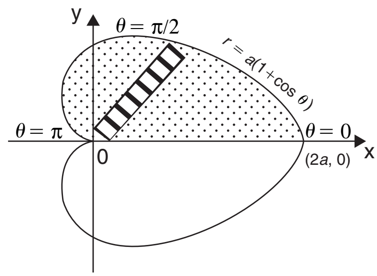
$$=\frac{4}{3}\sqrt{a}\left(8a^{\frac{3}{2}}\right)-\frac{1}{12a}(64a^3)$$

$$=\frac{32a^2}{3}-\frac{16a^2}{3}=\frac{16}{3}a^2.$$

EXAMPLE 6.38

Find the area of the cardioid $r=a(1+\cos\theta)$.

Solution. The curve passes through the origin and cuts the x-axis at $x=2a$. Clearly, θ varies from 0 to π and r varies from 0 to $a(1+\cos\theta)$ in the upper-half part of the integration region.



The required area is given by

$$A=2\int_0^{\pi}\left[\int_0^{a(1+\cos\theta)}r\,dr\right]d\theta=2\int_0^{\pi}\left[\frac{r^2}{2}\right]_0^{a(1+\cos\theta))}d\theta$$

$$=\int_0^{\pi}a^2(1+\cos\theta)^2d\theta$$

$$=4a^2\int_0^{\pi}\left(\cos^2\frac{\theta}{2}\right)^2d\theta$$

$$=4a^2\int_0^{\pi}\cos^4\frac{\theta}{2}d\theta$$

$$=8a^2\int_0^{\frac{\pi}{2}}\cos^4\phi\,d\theta,\ \theta=2\phi$$

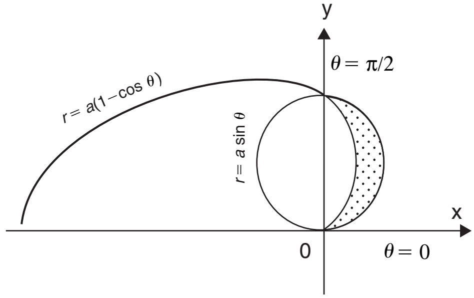
$$=8a^2.\frac{3}{4.2}.\frac{\pi}{2}=\frac{3\pi a^2}{2}.$$

EXAMPLE 6.39

Find the area lying inside the circle $r=a\sin\theta$ and outside the cardioid $r=a(1-\cos\theta)$.

Solution. We have $r=a\sin\theta$ and $r=a(1-\cos\theta)$. Therefore, $a\sin\theta=\text{a}(1-\cos\theta)$, which yields $\sin\theta$

$+\cos\theta = 1$ or $\sin^2\theta + \cos^2\theta + 2\sin\theta\cos\theta = 1$ or $\sin 2\theta = 0$. Hence, $2\theta = 0$ and π and so, $\theta = 0$ or $\frac{\pi}{2}$. Further, from the region of integration, it is clear that r varies from $a(1-\cos\theta)$ to $a\sin\theta$.



Therefore,

$$A = \int_0^{\frac{\pi}{2}} \left[\int_{a(1-\cos\theta)}^{a\sin\theta} r\, dr \right] d\theta = \int_0^{\frac{\pi}{2}} \left[\frac{r^2}{2} \right]_{a(1-\cos\theta)}^{a\sin\theta} d\theta$$

$$= \frac{a^2}{2} \int_0^{\frac{\pi}{2}} [\sin^2\theta - (1-\cos\theta)^2] d\theta$$

$$= \frac{a^2}{2} \int_0^{\frac{\pi}{2}} [\sin^2\theta - \cos^2\theta + 2\cos\theta - 1] d\theta$$

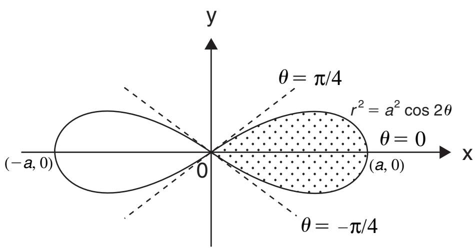
$$= \frac{a^2}{2} \int_0^{\frac{\pi}{2}} (-2\cos^2\theta + 2\cos\theta) d\theta$$

$$= a^2 \left[-\frac{1}{2} \cdot \frac{\pi}{2} + 1 \right] = a^2 \left(1 - \frac{\pi}{4} \right).$$

EXAMPLE 6.40

Find the area of one loop of the lemniscates $r^2 = a^2 \cos 2\theta$.

Solution. The region of integration is shown in the following figure:



The required area is given by

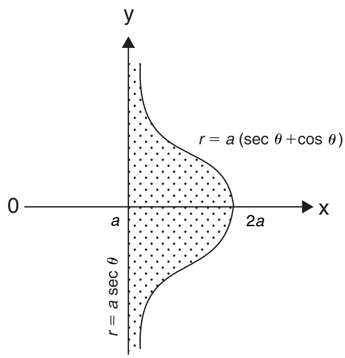
$$A = 2 \int_0^{\frac{\pi}{4}} \left[\int_0^{a\sqrt{\cos 2\theta}} r\, dr \right] d\theta = 2 \int_0^{\frac{\pi}{4}} \left[\frac{r^2}{2} \right]_0^{a\sqrt{\cos 2\theta}} d\theta$$

$$= a^2 \int_0^{\frac{\pi}{4}} \cos 2\theta\, d\theta = \frac{a^2}{2} [\sin 2\theta]_0^{\frac{\pi}{4}} = \frac{a^2}{2}.$$

EXAMPLE 6.41

Find the area included between the curve $r = a(\sec\theta + \cos\theta)$ and its asymptote.

Solution. The curve $r = a(\sec\theta + \cos\theta)$ is symmetrical about the initial line. The equation of the asymptote is $r = a\sec\theta$.



The required area is

$$A = 2 \int_0^{\frac{\pi}{2}} \left[\int_{a\sec\theta}^{a(\sec\theta+\cos\theta)} r\, dr \right] d\theta$$

$$= 2 \int_0^{\frac{\pi}{2}} \left[\frac{r^2}{2} \right]_{a\sec\theta}^{a(\sec\theta+\cos\theta)} d\theta$$

$$= \int_0^{\frac{\pi}{2}} [a^2(\sec\theta + \cos\theta)^2 - a^2\sec^2\theta] d\theta$$

$$= a^2 \int_0^{\frac{\pi}{2}} (\sec^2\theta + \cos^2\theta + 2 - \sec^2\theta) d\theta$$

$$= a^2 \int_0^{\frac{\pi}{2}} (\cos^2\theta + 2) d\theta = a^2 \left[\frac{1}{2} \cdot \frac{\pi}{2} + \frac{2\pi}{2} \right] = \frac{5\pi a^2}{4}.$$

EXAMPLE 6.42

Find the area bounded by the parabolas $y^2 = 4 - x$ and $y^2 = 4 - 4x$.

Solution. The required area is given by

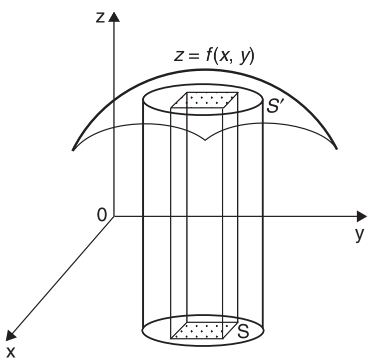
$$A = 2\int_0^2 \left(\int_{\frac{4-y^2}{4}}^{4-y^2} dx \right) dy = 2\int_0^2 \left[4 - \frac{y^2}{4} - \frac{4-y^2}{4} \right] dy$$

$$= 2\int_0^2 \left(3 - \frac{3}{4}y^2 \right) dy = 2\left[3y - \frac{y^3}{4} \right]_0^2$$

$$= 2[6 - 2] = 8.$$

6.8 VOLUME AND SURFACE AREA AS DOUBLE INTEGRALS

(A) Volume as a Double Integral: Consider a surface $z = f(x, y)$. Let the region S be the orthogonal projection of the portion S' of $z = f(x, y)$ on the xy-plane. Divide S into elementary rectangles of area $\delta x\, \delta y$ by drawing lines parallel to the x- and y-axis. On each of these rectangles, erect a prism which has a length parallel to Oz. Then, the volume of the prism between S' and S is $z\, \delta x\, \delta y$.



Therefore, the volume of the solid cylinder with S as base, is composed of these prisms and so,

$$V = \lim_{\substack{\delta x \to 0 \\ \delta y \to 0}} \sum z\, \delta x\, \delta y = \iint_S z\, dx\, dy$$

$$= \iint_S f(x,y)\, dx\, dy.$$

In the polar coordinates, the region S is divided into elements of area $r\, \delta r\, \delta\theta$ and so, the volume in that case is given by

$$V = \iint_S f(r\cos\theta,\ r\sin\theta) r dr\, d\theta.$$

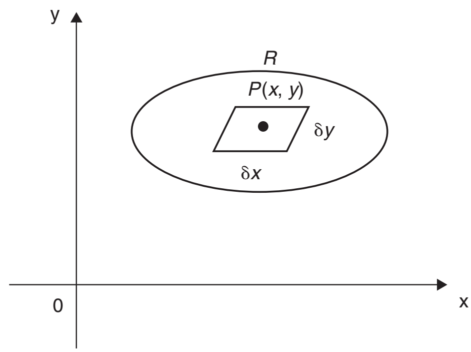
(B) Volumes of Solids of Revolution: Let $P(x, y)$ be a point in a plane area R. Suppose that the elementary area $\delta x\, \delta y$ at $P(x, y)$ revolves about the x-axis. This will generate a ring of radius y. The elementary volume of this ring is $\delta V = 2\pi y\, \delta y\, \delta x$. Hence, the total volume of the solid formed by the revolution of the area R about the x-axis is given by

$$V = 2\pi \iint_R y\, dy\, dx.$$

Changing to polar coordinates, we get

$$V = 2\pi \iint_R r\sin\theta\, r\, dr\, d\theta$$

$$= 2\pi \iint_R r^2 \sin\theta\, dr\, d\theta.$$



Similarly, the volume V of the area R revolved about the y-axis is given by

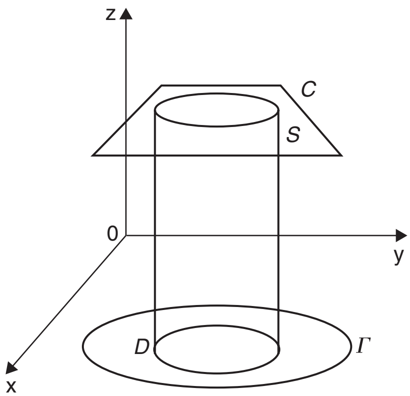
$$V = 2\pi \iint_R x\, dx\, dy.$$

Changing to polar coordinates, we have

$$V = 2\pi \iint_R r\cos\theta\, r\, dr\, d\theta$$

$$= 2\pi \iint_R r^2 \cos\theta\, dr\, d\theta.$$

(C) Surface Area as a Double Integral: Let $z = \psi(x, y)$ be a surface bounded by a curve C. Let the projection of C on the xy-plane be bounded by and let D be the domain on the xy-plane bounded by .



Then, the area of the surface S is given by

$$S = \iint_D \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\ dx\, dy.$$

EXAMPLE 6.43

Find the volume of the sphere $x^2 + y^2 + z^2 = a^2$ using polar coordinates.

Solution. The solid under consideration is bounded above by $z^2 = a^2 - (x^2 + y^2) = a^2 - r^2$. The sphere cuts the xy-plane in the circle $x^2 + y^2 = a^2$or $r^2 = a^2$. Because of symmetry, the required volume is given by

$$V = 2\int_0^{2\pi}\int_0^a \sqrt{a^2 - r^2}\ r\, dr\, d\theta$$

$$= \int_0^{2\pi}\int_0^a \sqrt{a^2 - r^2}.2r\, dr\, d\theta$$

$$= \int_0^{2\pi}\left[\frac{(a^2 - r^2)^{\frac{3}{2}}}{\frac{3}{2}}\right]_0^a d\theta = \frac{2a^3}{3}\int_0^{2\pi} d\theta = \frac{4\pi a^3}{3}.$$

EXAMPLE 6.44

Find the volume of the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

Solution. Due to symmetry, the volume of the given ellipsoid is eight times the volume of the portion of the ellipsoid in the first octant. For the positive octant, the given equation yields

$$z = c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}.$$

The region in this octant is bounded by

$$x = 0,\ x = a,\ y = 0,\ \text{and } y = b\sqrt{1 - \frac{x^2}{a^2}}.$$

Hence, the required volume is given by

$$V = 8\int_0^a \int_0^{b\sqrt{1-\frac{x}{a^2}}} z\, dy\, dx$$

$$= 8\int_0^a \int_0^{b\sqrt{1-\frac{x^2}{a^2}}} c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}\, dy\, dx$$

$$= 8\int_0^a \int_0^{b\sqrt{1-\frac{x^2}{a^2}}} c\left[\left(1 - \frac{x^2}{a^2}\right) - \frac{y^2}{b^2}\right]^{\frac{1}{2}} dy\, dx.$$

Substituting $\frac{y}{b} = \sqrt{1 - \frac{x^2}{a^2}}\sin\theta$, we get $dy = b\sqrt{1 - \frac{x^2}{a^2}}$ $\cos\theta\, d\theta$ (as x is a constant). Therefore,

$$V = 8\int_0^a \int_0^{\frac{\pi}{2}} c\left[\left(1 - \frac{x^2}{a^2}\right) - \left(1 - \frac{x^2}{a^2}\right)\sin^2\theta\right]^{\frac{1}{2}}$$

$$\times b\sqrt{1 - \frac{x^2}{a^2}}\cos\theta\, d\theta$$

$$= 8bc\int_0^a \left(1 - \frac{x^2}{a^2}\right)\int_0^{\frac{\pi}{2}} [1 - \sin^2\theta]^{\frac{1}{2}}\cos\theta\, d\theta$$

$$= 8bc\int_0^a \left(1 - \frac{x^2}{a^2}\right)\int_0^{\frac{\pi}{2}} \cos^2\theta\, d\theta$$

$$= 8bc\int_0^a \left(1 - \frac{x^2}{a^2}\right).\frac{1}{2}.\frac{\pi}{2}$$

$$= 2\pi bc\left[x - \frac{x^3}{3a^2}\right]_0^a = \frac{4}{3}\pi abc.$$

EXAMPLE 6.45

Find the volume contained between the ellipsoid $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ and the cylinder $\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{x}{a}$.

Solution. The equation of the given elliptical cylinder is

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{x}{a}.$$

Substituting $\frac{x}{a}=r\cos\theta$ and $\frac{y}{b}=r\sin\theta$, this equation yields

$$r^2=r\cos\theta \text{ or } r=\cos\theta.$$

The required volume is given by

$$V=4\iint c\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}}\,dy\,dx$$

$$=4abc\int_0^{\frac{\pi}{2}}\int_0^{\cos\theta}\sqrt{1-r^2}\,r\,dr\,d\theta$$

$$=-\frac{4abc}{2}\int_0^{\frac{\pi}{2}}\left[\frac{(1-r^2)^{\frac{3}{2}}}{\frac{3}{2}}\right]_0^{\cos\theta}d\theta$$

$$=-\frac{4abc}{2}\int_0^{\frac{\pi}{2}}(\sin^3\theta-1)\,d\theta$$

$$=-\frac{4abc}{3}\left[\frac{2}{3}-\frac{\pi}{2}\right]=\frac{2}{9}abc[3\pi-4].$$

EXAMPLE 6.46

Find the volume common to a sphere $x^2+y^2+z^2=a^2$ and a circular cylinder $x^2+y^2=ax$.

(particular case of Example 6.45, taking $a=b=c$).

Solution. The required volume is the part of the sphere lying within the cylinder and is given by

$$V=4\iint_R z\,dy\,dx=4\iint_R\sqrt{a^2-x^2-y^2}\,dy\,dx,$$

where R is the half of the circle lying in the first quadrant. Substituting $x=r\cos\theta$ and $y=r\sin\theta$, the equation $x^2+y^2=ax$ yields

$$r^2=a\,r\cos\theta \text{ or } r=a\cos\theta.$$

Thus, the region of integration is bounded by

$$r=0, r=a\cos\theta,\ \theta=0, \text{ and } \theta=\frac{\pi}{2}.$$

Hence,

$$V=4\int_0^{\frac{\pi}{2}}\int_0^{a\cos\theta}\sqrt{a^2-r^2}\,r\,dr\,d\theta$$

$$=\frac{4}{-2}\int_0^{\frac{\pi}{2}}\int_0^{a\cos\theta}\sqrt{a^2-r^2}\,(-2r)\,dr\,d\theta$$

$$=-2\int_0^{\frac{\pi}{2}}\left[\frac{(a^2-r^2)^{\frac{3}{2}}}{\frac{3}{2}}\right]_0^{a\cos\theta}d\theta$$

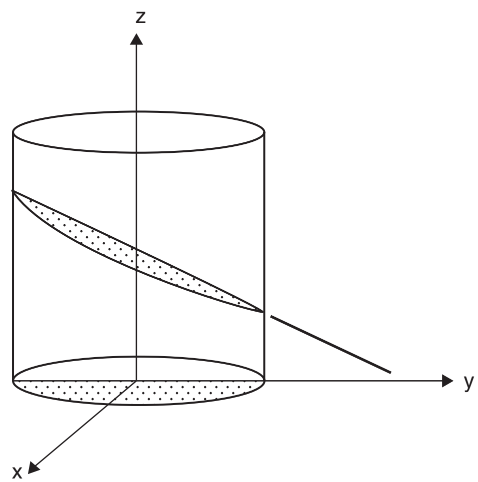
$$=-\frac{4}{3}\int_0^{\frac{\pi}{2}}(a^3\sin^3\theta-a^3)\,d\theta$$

$$=-\frac{4a^3}{3}\left[\frac{2}{3}-\frac{\pi}{2}\right]=\frac{2a^3}{9}(3\pi-4).$$

EXAMPLE 6.47

Find the volume bounded by the cylinder $x^2+y^2=4$ and the planes $y+z=4$ and $z=0$.

Solution. To find the required volume, $z=4-y$ is to be integrated over the circle $x^2+y^2=4$ in the xy-plane.



To cover the area (half of the circle) in the xy-plane, x varies from 0 to $\sqrt{4-y^2}$ and y varies from -2 to 2.

Thus,

$$V = 2\int_{-2}^{2}\left[\int_{0}^{\sqrt{4-y^2}} z\ dx\right] dy$$

$$= 2\int_{-2}^{2}\left[\int_{0}^{\sqrt{4-y^2}} (4-y)\ dx\right] dy$$

$$= 2\int_{-2}^{2} (4-y)[x]_0^{\sqrt{4-y^2}}\ dy$$

$$= 2\int_{-2}^{2} (4-y)\sqrt{4-y^2}\ dy$$

$$= 2\left[4\int_{-2}^{2} \sqrt{4-y^2}\ dy - \int_{-2}^{2} y\sqrt{4-y^2}\ dy\right]$$

$$= 8\int_{-2}^{2} \sqrt{4-y^2}\ dy, \text{ second integrand being odd}$$

$$= 16\int_{0}^{2} \sqrt{4-y^2}\ dy, \text{because of even integrand}$$

$$= 16\left[\frac{y\sqrt{4-y^2}}{2} + \frac{4}{2}\sin^{-1}\frac{y}{2}\right]_0^2$$

$$= 16[2\sin^{-1} 1] = \frac{32\pi}{2} = 16\pi.$$

EXAMPLE 6.48

Find the volume of the solid bounded above by the parabolic cylinder $z = 4 - y^2$ and bounded below by the elliptic paraboloid $z = x^2 + 3y^2$.

Solution. The two surfaces intersect in a space curve, whose projection on the xy-plane is the ellipse $x^2 + 4y^2 = 4$ or $\frac{x^2}{4} + y^2 = 1$. Substituting $x = 2r\cos\theta$ and $y = r\sin\theta$, the ellipse becomes $r^2 = 1$.

Further,

$$z = \phi_1(x) = 4 - y^2 \text{ and}$$
$$z = \phi_2(x) = x^2 + 3y^2.$$

Therefore,

$$\phi_1(x) - \phi_2(x) = 4 - y^2 - x^2 - 3y^2$$
$$= 4 - 4y^2 - x^2 = 4(1 - r^2).$$

Also,

$$J = \frac{\partial\ (x, y)}{\partial\ (r, \theta)} = 2.$$

Since the solid is symmetrical about x- and y-axis, we have

$$V = 4\int_0^{\frac{\pi}{2}}\int_0^1 4(1 - r^2)2r\ dr\ d\theta$$

$$= 32\int_0^{\frac{\pi}{2}}\int_0^1 (r - r^3)\ dr\ d\theta$$

$$= 32\int_0^{\frac{\pi}{2}}\left[\frac{r^2}{2} - \frac{r^4}{4}\right]_0^1 d\theta\ = \frac{32\pi}{8} = 4\pi.$$

EXAMPLE 6.49

Find the volume bounded by xy-plane, the cylinder $x^2 + y^2 = 1$, and the plane $x + y + z = 3$.

Solution. We have to integrate $z = 3 - x - y$ over the circle $x^2 + y^2 = 1$. Substituting $x = r\cos\theta$ and $y = r\sin\theta$, so that $x^2 + y^2 = r^2$, the integrand reduces to $3 - r\cos\theta - r\sin\theta = 3 - r(\cos\theta + \sin\theta)$ and the circle $x^2 + y^2 = 1$ reduces to $r^2 = 1$. Thus, to cover half of the region, r varies from 0 to 1 and θ varies from 0 to $\frac{\pi}{2}$. Hence,

$$V = 4\int_0^{\frac{\pi}{2}}\left[\int_0^1 \{3 - r(\cos\theta + \sin\theta)\}r\ dr\right] d\theta$$

$$= 4\int_0^{\frac{\pi}{2}}\left[3\frac{r^2}{2} - \frac{r^3}{3}(\cos\theta + \sin\theta)\right]_0^1 d\theta$$

$$= 4\int_0^{\frac{\pi}{2}}\left[\frac{3}{2} - \frac{1}{3}(\cos\theta + \sin\theta)\right] d\theta$$

$$= \int_0^{\frac{\pi}{2}} 6d\theta - \frac{4}{3}\int_0^{\frac{\pi}{2}}(\cos\theta + \sin\theta)d\theta$$

$$= \frac{6\pi}{2} - \frac{4}{3}[\sin\theta - \cos\theta]_0^{\frac{\pi}{2}}$$

$$= \frac{6\pi}{2} - \frac{4}{3}[1 + 1] = 3\pi - \frac{8}{3}.$$

EXAMPLE 6.50

Find the volume common to the cylinders $x^2+y^2=a^2$ and $x^2+z^2=a^2$.

Solution. The required volume is given by

$$V = 8\int_0^a \left[\int_0^{\sqrt{a^2-r^2}} \sqrt{a^2-x^2}\, dy\right] dx$$

$$= 8\int_0^a \sqrt{a^2-x^2}\,[y]_0^{\sqrt{a^2-x^2}}\, dx$$

$$= 8\int_0^a (a^2-x^2)\, dx \;= 8\left[a^2x-\frac{x^3}{3}\right]_0^a = \frac{16a^3}{3}.$$

EXAMPLE 6.51

Prove that the volume, enclosed between $x^2+y^2=2ax$ and $z^2=2ax$ is $\frac{128a^3}{15}$.

Solution. To find the required volume, $z=\sqrt{2ax}$ is to be integrated over the curve $x^2+y^2=2ax$ in the xy-plane. Changing to polar coordinates by substituting $x=r\cos\theta$ and $y=r\sin\theta$, the required volume is given by

$$V = 4\int_0^{\frac{\pi}{2}}\int_0^{2a\cos\theta} \sqrt{2ar\cos\theta}\; r\, dr\, d\theta$$

$$= 4\sqrt{2a}\int_0^{\frac{\pi}{2}}\int_0^{2a\cos\theta} \sqrt{r}\sqrt{\cos\theta}\; r\, dr\, d\theta$$

$$= 4\sqrt{2a}\int_0^{\frac{\pi}{2}}\sqrt{\cos\theta}\left[\int_0^{2a\cos\theta} r^{\frac{3}{2}}\, dr\right] d\theta$$

$$= 4\sqrt{2a}\int_0^{\frac{\pi}{2}}\sqrt{\cos\theta}\left[\frac{r^{\frac{5}{2}}}{\frac{5}{2}}\right]_0^{2a\cos\theta}$$

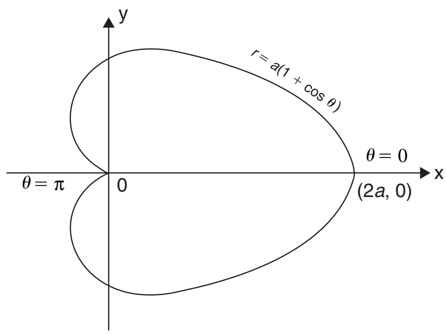
$$= \frac{8\sqrt{2a}}{5}\int_0^{\frac{\pi}{2}} (2a)^{\frac{5}{2}}\sqrt{\cos\theta}(\cos\theta)^{\frac{5}{2}}\, d\theta$$

$$= \frac{8}{5}(2a)^3\int_0^{\frac{\pi}{2}} \cos^3\theta\, d\theta = \frac{64a^3}{5}.\frac{2}{3} = \frac{128}{15}a^3.$$

EXAMPLE 6.52

Find, by double integration, the volume generated by revolving the cardioid $r=a(1+\cos\theta)$ about the initial line.

Solution. We observe that the upper and lower halves of the cardioid $r=a(1+\cos\theta)$ generate the same volume. Therefore, it is sufficient to consider the revolution of the upper-half cardioid only, for which r varies from 0 to $a(1+\cos\theta)$ and θ varies from 0 to π.



Hence,

$$\text{Volume of Revolution} = 2\pi\int_0^{\pi}\int_0^{a(1+\cos\theta)} r^2\sin\theta\, dr\, d\theta$$

$$= 2\pi\int_0^{\pi}\sin\theta\left[\frac{r^3}{3}\right]_0^{a(1+\cos\theta)}$$

$$= \frac{2\pi a^3}{3}\int_0^{\pi}\sin\theta(1+\cos\theta)^3 d\theta$$

$$= -\frac{2\pi a^3}{3}\left[\frac{(1+\cos\theta)^4}{4}\right]_0^{\pi}$$

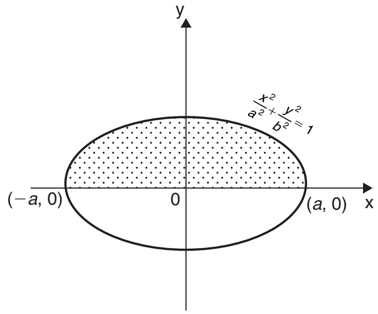
$$= -\frac{2\pi a^3}{3}(-2^4) = \frac{8}{3}\pi a^3.$$

EXAMPLE 6.53

Find the volume of the solid generated by revolving the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ about the x-axis.

Solution. Due to symmetry, it is sufficient to calculate the volume obtained on revolving the upper half of the ellipse. For this, x varies from $-a$ to a

and y varies from 0 to $b\sqrt{1-\frac{x^2}{a^2}}$.



Therefore, the required

$$\text{volume of revolution} = 2\pi \int_{-a}^{a} \int_{0}^{b\sqrt{1-\frac{x^2}{a^2}}} y\, dy\, dx$$

$$= 2\pi \int_{-a}^{a} \left[\frac{y^2}{2}\right]_0^{b\sqrt{1-\frac{x^2}{a^2}}} dy = \pi \int_{-a}^{a} b^2\left(1-\frac{x^2}{a^2}\right) dx$$

$$= \frac{2\pi b^2}{a^2}\int_0^a (a^2-x^2)dx = \frac{2\pi b^2}{a^2}\left[a^2x-\frac{x^3}{3}\right]_0^a$$

$$= \frac{2\pi b^2}{a^2}\int_0^a \left[a^3-\frac{a^3}{3}\right] = \frac{4}{3}\pi ab^2.$$

EXAMPLE 6.54

Find the area of the surface of the paraboloid $x^2+y^2=z$, which lies between the planes $z=0$ and $z=1$.

Solution. The required surface area is given by

$$S = \iint \sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy.$$

But,
$$\frac{\partial z}{\partial x} = 2x \text{ and } \frac{\partial z}{\partial y} = 2y.$$

Therefore,

$$S = \iint \sqrt{1+4(x^2+y^2)}\, dx\, dy$$

$$= \iint \sqrt{1+4r^2}\, r\, dr\, d\theta$$

(changing to polar coordinates).

To find the limits, we see that the projection on the plane $z=1$ is the circle $x^2+y^2=1$ or $r^2=1$ and this circle lies between $\theta=0$ and $\theta=2\pi$. Hence,

$$S = \int_0^{2\pi}\int_0^1 \sqrt{1+4r^2}\, r\ \ d\theta\, dr$$

$$= \frac{1}{8}\int_0^{2\pi}\int_0^1 \sqrt{1+4r^2}\, 8r\ \ d\theta\, dr$$

$$= \frac{1}{8}\int_0^{2\pi}\left[\frac{(1+4r^2)^{\frac{3}{2}}}{\frac{3}{2}}\right]_0^1 d\theta = \frac{1}{12}\int_0^{2\pi}(5\sqrt{5}-1)d\theta$$

$$= \frac{5\sqrt{5}-1}{12}[\theta]_0^{2\pi} = \frac{\pi}{6}[5\sqrt{5}-1].$$

EXAMPLE 6.55

Compute the surface area of the sphere $x^2+y^2+z^2=a^2$.

Solution. The surface area of the sphere is twice the surface area of the upper-half sphere $z=\sqrt{a^2-x^2-y^2}$. We have

$$\frac{\partial z}{\partial x} = -\frac{x}{\sqrt{a^2-x^2-y^2}} \text{ and}$$
$$\frac{\partial z}{\partial y} = -\frac{y}{\sqrt{a^2-x^2-y^2}}.$$

Therefore,

$$S = \iint \sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy.$$

$$= \iint \frac{a}{\sqrt{a^2-x^2-y^2}}\, dx\, dy.$$

The domain of integration is the circle $x^2+y^2=a^2$ on the xy-plane. Therefore,

$$S = 2\int_{-a}^{a}\left[\int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \frac{a}{\sqrt{a^2-x^2-y^2}} dy\right] dx.$$

Changing to polar coordinates, we have

$$S = 2\int_0^{2\pi}\left[\int_0^a \frac{a}{\sqrt{a^2-r^2}} r\,dr\right]d\theta$$

$$= \int_0^{2\pi}\left[\int_0^a \frac{a}{\sqrt{a^2-r^2}} 2r\,dr\right]d\theta \;= 4\pi a^2.$$

EXAMPLE 6.56

Find the area of the spherical surface $x^2 + y^2 + z^2 = a^2$ inside the cylinder $x^2 + y^2 = ax$.

Solution. We have

$$S = 4\iint \sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}\,dx\,dy$$

$$= 4\iint \frac{a dx\,dy}{\sqrt{a^2-x^2-y^2}} \text{ over } x^2+y^2 = ax$$

$$= 4a\int_0^{\frac{\pi}{2}}\int_0^{a\cos\theta} \frac{r\,dr\,d\theta}{\sqrt{a^2-r^2}},\; x = r\cos\theta,\; y = r\sin\theta$$

$$= 2a^2(\pi - 2).$$

EXAMPLE 6.57

Find the area of that part of the cylinder $x^2 + y^2 = a^2$, which is cut off by the cylinder $x^2 + z^2 = a^2$.

Solution. The equation of the surface has the form $y = \sqrt{a^2-x^2}$ so that

$$\frac{\partial y}{\partial x} = -\frac{x}{\sqrt{a^2-x^2}},\; \frac{\partial y}{\partial z} = 0$$

and

$$\sqrt{1+\left(\frac{\partial y}{\partial x}\right)^2+\left(\frac{\partial y}{\partial z}\right)^2} = \frac{a}{\sqrt{a^2-x^2}}.$$

The domain of integration is a quarter circle $x^2 + z^2 = a^2$ where $x \geq 0$ and $z \geq 0$ on the xz-plane. Therefore,

$$S = 8\int_0^a\left[\int_0^{\sqrt{a^2-x^2}} \frac{a}{\sqrt{a^2-x^2}}\,dz\right]dx \;= 8a^2.$$

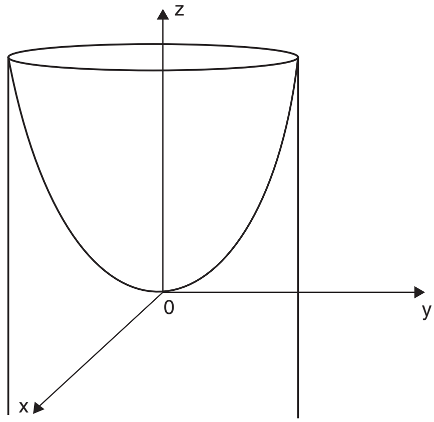
EXAMPLE 6.58

Find the area of the paraboloid $2z = \frac{x^2}{a} + \frac{y^2}{b}$ inside the cylinder $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

Solution. The required area is

$$S = 4\iint \sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}\,dx\,dy,$$

where the integration extends over the positive octant of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.



We have $\frac{\partial z}{\partial x} = \frac{x}{a}$ *and* $\frac{\partial z}{\partial y} = \frac{y}{b}$. Therefore,

$$S = 4\iint\left(1+\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)^{\frac{1}{2}} dx\,dy$$

$$= 4ab\iint\left(1+\xi^2+\eta^2\right)d\xi d\eta,$$

$$x = a\xi,\; y = b\xi,\; \text{so that } \xi^2+\eta^2 = 1$$

$$= 4ab\int_0^{\frac{\pi}{2}}\int_0^1 \left(1+r^2\right) r\,dr\,d\theta,$$

$$\xi = r\cos\theta,\; \eta = r\sin\theta$$

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

6.9 TRIPLE INTEGRALS AND THEIR EVALUATION

Let $f(x, y, z)$ be a continuous function in a finite region V of $\Re^3$. Divide the region V into n sub-regions of respective volumes $\delta V_1, \delta V_2, \ldots, \delta V_n$. If

(x_i, y_j, z_i) be an arbitrary point of the ith subregion, then

$$\lim_{\substack{n\to\infty \\ \delta V_i\to 0}} \sum_{i=1}^{n} f(x_i,\ y_i,\ z_i)\delta V_i,$$

if exists, is called the triple integral of $f(x,y,z)$ over the region V, and is denoted by $\iiint_V f(x,y,z)dV$ or $\iiint_V f(x,y,z)dx\ dy\ dz$.

Evaluation of Triple Integrals

(a) If the region V is specified by the inequalities,

$$a \le x \le b,\ c \le y \le d,\ \text{and } e \le z \le f,$$

and if a, b, c, d, e, and f are constants, then

$$\iiint_V f(x,y,z)dx\ dy\ dz = \int_a^b\int_c^d\int_e^f f(x,y,z)dx\ dy\ dz$$
$$= \int_a^b dx \int_c^d dy \int_e^f f(x,y,z)dz.$$

Since a, b, c, d, e, and f are constant, the order of integration is immaterial, and the integration with respect to any of x, y, and z can be performed first.

(b) If the limits of z are given as functions of x and y, and the limits of y as functions of x while x takes the constant values say from a to b, then

$$\iiint_V f(x,y,z)dx\ dy\ dz$$
$$= \int_a^b \left[\int_{y_1(x)}^{y_2(x)} \left\{\int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)dz\right\}dy\right]dx.$$

Thus, the integration with respect to z is performed first regarding x and y as constants, then the integration with respect to y is performed regarding x as constant and in the last, the integration with respect to x is performed.

EXAMPLE 6.59

Evaluate

$$I = \int_0^{\log 2}\int_0^{x}\int_0^{x+\log y} e^{x+y+z}dz\ dy\ dx.$$

Solution. We have

$$I = \int_0^{\log 2}\int_0^{x} [e^{x+y+z}]_0^{x+\log y}\ dy\ dx$$
$$= \int_0^{\log 2}\int_0^{x} [e^{x+y+x+\log y} - e^{x+y}]\ dy\ dx$$
$$= \int_0^{\log 2}\int_0^{x} [e^{2x}.e^{y}.e^{\log y} - e^{x}.e^{y}]\ dy\ dx$$
$$= \int_0^{\log 2}\int_0^{x} [e^{2x}ye^{y} - e^{x}.e^{y}]\ dy\ dx$$
$$= \int_0^{\log 2}\left[\int_0^{x} e^{2x}ye^{y}dy - \int_0^{x} e^{x}e^{y}dy\right]dx$$
$$= \int_0^{\log 2}\left[e^{2x}\{ye^{y}\}_0^{x} - e^{2x}\int_0^{x} e^{y}dy - e^{x}\int_0^{x} e^{y}dy\right]dx$$
$$= \int_0^{\log 2} [e^{2x}\cdot xe^{x} - e^{2x}(e^{x}-1) - e^{x}(e^{x}-1)]dx$$
$$= \int_0^{\log 2} [xe^{3x} - e^{3x} + e^{x}]dx$$
$$= \int_0^{\log 2} xe^{3x}dx - \int_0^{\log 2} e^{3x}dx + \int_0^{\log 2} e^{x}dx$$
$$= \frac{1}{3}[xe^{3x}]_0^{\log 2} - \frac{1}{3}\int_0^{\log 2} e^{3x}dx - \int_0^{\log 2} e^{3x}dx + \int_0^{\log 2} e^{x}dx$$
$$= \frac{1}{3}\log 2\ \ e^{3\log 2} - \frac{4}{3}\left[\frac{e^{3x}}{3}\right]_0^{\log 2} + [e^{x}]_0^{\log 2}$$
$$= \frac{1}{3}\log 2\ \ e^{\log 8} - \frac{4}{9}\left(e^{\log 8} - 1\right) + \left(e^{\log 2} - 1\right)$$
$$= \frac{8}{3}\log 2 - \frac{4}{9}(8-1) + (2-1)$$
$$= \frac{8}{3}\log 2 - \frac{28}{9} + 1 = \frac{8}{3}\log 2 - \frac{19}{9}.$$

EXAMPLE 6.60

Evaluate

$$I = \int_1^e \int_1^{\log y} \int_1^{e^x} \log z \, dz \, dx \, dy.$$

Solution. We have

$$I = \int_1^e \int_1^{\log y} \left[\int_1^{e^x} \log z. \, 1 dz \right] dx \, dy$$

$$= \int_1^e \int_1^{\log y} \left\{ [z \log z]_0^{e^x} - \int_1^{e^x} z. \frac{1}{z} dz \right\} dx \, dy$$

$$= \int_1^e \int_1^{\log y} [e^x \log e^x - 0 - e^x + 1] dx \, dy$$

$$= \int_1^e \int_1^{\log y} [(x-1)e^x + 1] dx \, dy$$

$$= \int_1^e \left[\int_1^{\log y} (x-1)e^x dx + \int_1^{\log y} dx \right] dy$$

$$= \int_1^e \left[\int_1^{\log y} (x-1)e^x dx + \log y - 1 \right] dy$$

$$= \int_1^e \left\{ [(x-1)e^x]_1^{\log y} - \int_1^{\log y} e^x dx + \log y - 1 \right\} dy$$

$$= \int_1^e [(\log y - 1)e^{\log y} - (e^{\log y} - e) + \log y - 1] \, dy$$

$$= \int_1^e [y(\log y - 1) - (y - e) + \log y - 1] \, dy$$

$$= \int_1^e [(y+1)\log y - 2y + e - 1] \, dy$$

$$= \left[\log y \left(\frac{y^2}{2} + y \right) \right]_1^e - \int_1^e \frac{1}{y} \left(\frac{y^2}{2} + y \right) dy$$

$$- \left[2\frac{y^2}{2} \right]_1^e + (e-1)[y]_1^e$$

$$= \frac{e^2}{2} + e - \int_1^e \left(\frac{y}{2} + 1 \right) dy - (e^2 - 1) + (e-1)^2$$

$$= \frac{e^2}{2} + e - \left(\frac{y^2}{4} + y \right)_1^e - 2e + 2$$

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

$$= \frac{e^2}{4} - 2e + \frac{13}{4} = \frac{1}{4}(e^2 - 8e + 13).$$

EXAMPLE 6.61

Evaluate the integral

$$\int_0^1 \int_0^{\sqrt{1-x^2}} \int_0^{\sqrt{1-x^2-y^2}} xyz \, dz \, dy \, dx.$$

Solution. The given triple integral is

$$I = \int_0^1 xy \int_0^{\sqrt{1-x^2}} \left[\frac{z^2}{2} \right]_0^{\sqrt{1-x^2-y^2}} dy \, dx$$

$$= \int_0^1 \int_0^{\sqrt{1-x^2}} \frac{1}{2} xy(1 - x^2 - y^2) \, dy \, dx$$

$$= \int_0^1 \frac{1}{2} x \left[\int_0^{\sqrt{1-x^2}} (y - x^2 y - y^3) \, dy \right] dx$$

$$= \frac{1}{2} \int_0^1 x \left[\frac{y^2}{2} - x^2 \frac{y^2}{2} - \frac{y^4}{4} \right]_0^{\sqrt{1-x^2}} dx$$

$$= \frac{1}{8} \int_0^1 x(1 - x^2)^2 dx$$

$$= \frac{1}{8} \int_0^{\frac{\pi}{2}} \sin\theta (1 - \sin^2\theta)^2 \cos\theta \, d\theta, \; x = \sin\theta$$

$$= \frac{1}{8} \int_0^{\frac{\pi}{2}} \sin\theta \cos^5\theta \, d\theta = \frac{1}{8} \cdot \frac{1.4.2}{6.4.2} = \frac{1}{48}.$$

EXAMPLE 6.62

Evaluate

$$I = \int_0^a \int_0^{a-x} \int_0^{a-x-y} x^2 dx \, dy \, dz.$$

Solution. We have

$$I = \int_0^a \int_0^{a-x} x^2 \left[\int_0^{a-x-y} dz \right] dx\ dy.$$

$$= \int_0^a \int_0^{a-x} x^2 [z]_0^{a-x-y} dx\ dy$$

$$= \int_0^a x^2 \left[\int_0^{a-x} (a-x-y)dy \right] dx$$

$$= \int_0^a x^2 \left[ay - xy - \frac{y^2}{2} \right]_0^{a-x} dx$$

$$= \int_0^a x^2 \left[a^2 - ax - ax + x^2 - \frac{(a-x)^2}{2} \right] dx$$

$$= \frac{1}{2} \int_0^a (x^2a^2 - 2ax^3 + x^4)dx$$

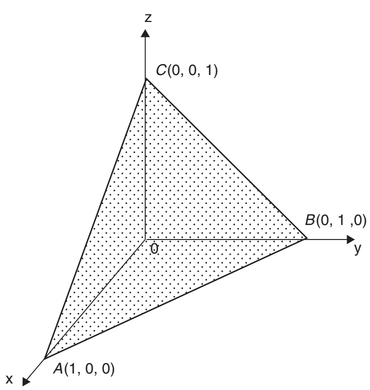
$$= \frac{1}{2} \left[a^2 \frac{x^3}{3} - 2a\frac{x^4}{4} + \frac{x^5}{5} \right]_0^a = \frac{a^5}{60}.$$

EXAMPLE 6.63

Evaluate $\iiint \frac{dx\ dy\ dz}{(x+y+z+1)^3}$ over a tetrahedron bounded by coordinate planes and the plane $x+y+z=1$.

Solution. The region of integration is bounded by the coordinate planes $x=0$, $y=0$, and $z=0$ and the plane $x+y+z=1$. Thus,

$$R = \{(x,y,z);\ x \geq 0,\ y \geq 0,\ z \geq 0,\ x+y+z \leq 1\}$$
$$= \{(x,y,z);\ 0 \leq x \leq 1,\ 0 \leq y \leq 1-x,$$
$$0 \leq z \leq 1-x-y\}.$$



Therefore,

$$\iiint_R \frac{dx\,dy\,dz}{(x+y+z+1)^3}$$

$$= \int_0^1 \int_0^{1-x} \int_0^{1-x-y} \frac{1}{(x+y+z+1)^3} dz\ dy\ dx$$

$$= \int_0^1 \int_0^{1-x} \left[\int_0^{1-x-y} (x+y+z+1)^{-3} dz \right] dy\,dx$$

$$= \int_0^1 \int_0^{1-x} \left[\frac{(x+y+z+1)^{-2}}{-2} \right]_0^{1-x-y} dy\,dx$$

$$= \frac{1}{2} \int_0^1 \int_0^{1-x} \left[\frac{1}{(x+y+1)^2} - \frac{1}{4} \right] dy\,dx$$

$$= \frac{1}{2} \int_0^1 \left[\frac{-1}{(x+y+1)} - \frac{y}{4} \right]_0^{1-x} dx$$

$$= \frac{1}{2} \int_0^1 \left(-\frac{1}{2} - \frac{1-x}{4} + \frac{1}{x+1} \right) dx$$

$$= \frac{1}{2} \int_0^1 \left[-\frac{3}{4} + \frac{x}{4} + \frac{1}{x+1} \right] dx$$

$$= \frac{1}{2} \left[-\frac{3x}{4} + \frac{x^2}{8} + \log(x+1) \right]_0^1$$

$$= \frac{1}{2} \left[-\frac{3}{4} + \frac{1}{8} + \log 2 \right] = \frac{1}{2} \log 2 - \frac{5}{16}.$$

EXAMPLE 6.64

Evaluate $\int_0^a \int_0^x \int_0^{x+y} e^{x+y+z} dz\ dy\ dx.$

Solution. We have

$$\int_0^a \int_0^x \int_0^{x+y} e^{x+y+z} dz\ dy\ dx$$

$$= \int_0^a \int_0^x [e^{x+y+z}]_0^{x+y} dy\ dx$$

$$= \int_0^a \int_0^x [e^{2(x+y)} - e^{x+y}]dy\ dx$$

$$= \int_0^a \left[\int_0^x e^{2(x+y)}dy - \int_0^x e^{x+y}dy \right] dx$$

$$= \int_0^a \left\{ \left[\frac{e^{2(x+y)}}{2} \right]_0^x - [e^{x+y}]_0^x \right\} dx$$

$$= \int_0^a \left[\frac{e^{4x}}{2} - \frac{e^{2x}}{2} - e^{2x} + e^x \right] dx$$

$$= \left[\frac{e^{4x}}{8} - \frac{e^{2x}}{4} - \frac{e^{2x}}{2} + e^x \right]_0^a$$

$$= \frac{e^{4a}}{8} - \frac{3e^{2a}}{4} + e^a.$$

EXAMPLE 6.65

Evaluate the triple integral

$$I = \int_0^a \int_0^{\sqrt{a^2-x^2}} \int_0^{\sqrt{a^2-x^2-y^2}} \frac{1}{\sqrt{a^2 - x^2 - y^2 - z^2}} dz\ dy\ dx.$$

Solution. We have

$$I = \int_0^a \int_0^{\sqrt{a^2-x^2}} \left[\int_0^{\sqrt{a^2-x^2-y^2}} \frac{1}{\sqrt{(a^2 - x^2 - y^2) - z^2}} dz \right] dy\, dx$$

$$= \int_0^a \int_0^{\sqrt{a^2-x^2}} \left[\sin^{-1} \frac{z}{\sqrt{a^2 - x^2 - y^2}} \right]_0^{\sqrt{a^2-x^2-y^2}} dy\ dx$$

$$= \int_0^a \left[\int_0^{\sqrt{a^2-x^2}} [\sin^{-1} 1] dy \right] dx$$

$$= \frac{\pi}{2} \int_0^a [y]_0^{\sqrt{a^2-x^2}} dx = \frac{\pi}{2} \int_0^a \sqrt{a^2 - x^2} dx$$

$$= \frac{\pi}{2} \left[\frac{x\sqrt{a^2 - x^2}}{2} + \frac{a^2}{2} \sin^{-1} \frac{x}{a} \right]_0^a$$

$$= \frac{\pi}{2} \left[0 + \frac{a^2}{2} \sin^{-1} 1 - 0 \right] = \frac{\pi a^2}{4} \sin^{-1} 1$$

$$= \frac{\pi a^2}{4} . \frac{\pi}{2} = \frac{\pi^2 a^2}{8}.$$

Note: The earlier example may be restated as "Evaluate $\iiint \frac{dx\ dy\ dz}{\sqrt{a^2 - x^2 - y^2 - z^2}}$ over the positive octant of the sphere $x^2 + y^2 + z^2 = a^2$."

EXAMPLE 6.66

Evaluate

$$I = \int_{-c}^{c} \int_{-b}^{b} \int_{-a}^{a} \left(x^2 + y^2 + z^2\right) dz\ dy\ dx$$

Solution. We observe that the integrand $x^2 + y^2 + z^2$ is symmetrical in x, y, and z. Therefore, the limits of integration can be assigned as per our preference. We have

$$I = \int_{-c}^{c} \int_{-b}^{b} \int_{-a}^{a} \left(x^2 + y^2 + z^2\right) dx\ dy\ dz$$

$$= 2 \int_{-c}^{c} \int_{-b}^{b} \left[\int_0^a \left(x^2 + y^2 + z^2\right) dx \right] dy\ dz,$$

since $x^2 + y^2 + z^2$ is even in x

$$= 2 \int_{-c}^{c} \int_{-b}^{b} \left[\frac{x^3}{3} + y^2 x + z^2 x \right]_0^a dy\ dz$$

$$= 2 \int_{-c}^{c} \int_{-b}^{b} \left[\frac{a^3}{3} + ay^2 + az^2 \right] dy\ dz$$

$$= 4 \int_{-c}^{c} \left[\int_0^b \left(\frac{a^3}{3} + ay^2 + az^2 \right) dy \right] dz,$$

since integrand is even in y

$$= 4 \int_{-c}^{c} \left[\frac{a^3 y}{3} + \frac{ay^3}{3} + az^2 y \right]_0^b dz$$

$$= 4 \int_{-c}^{c} \left[\frac{ba^3}{3} + \frac{ab^3}{3} + abz^2 \right] dz$$

$$= 8\int_0^c \left[\frac{ba^3}{3} + \frac{ab^3}{3} + abz^2\right]dz,$$
since integrand is even in z
$$= 8\left[\frac{ba^3}{3}z + \frac{ab^3z}{3} + \frac{abz^3}{3}\right]_0^c$$
$$= 8\left[\frac{ba^3c}{3} + \frac{ab^3c}{3} + \frac{abc^3}{3}\right]$$
$$= \frac{8abc}{3}[a^2 + b^2 + c^2].$$

EXAMPLE 6.67

Evaluate $\iiint xyz\,dx\,dy\,dz$ over the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

Solution. The region of integration is bounded by
$$z = -c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} \text{ and } z = c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}.$$
The projection on the xy-plane is the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Hence, the limits of integration for y and x are from $y = -b\sqrt{1 - \frac{x^2}{a^2}}$ to $y = b\sqrt{1 - \frac{x^2}{a^2}}$ and $x = -a$ to $x = a$. Thus,
$$R = \left\{\begin{matrix}(x,y,z); -a \le x \le a, -b\sqrt{1-\frac{x^2}{a^2}} \le y \le b\sqrt{1-\frac{x^2}{a^2}}, \\ -c\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}} \le z \le c\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}}\end{matrix}\right\}.$$
Hence,
$$\iiint_R xyz\,dx\,dy\,dz$$
$$= \int_{-a}^{a}\int_{-b\sqrt{1-\frac{x^2}{a^2}}}^{b\sqrt{1-\frac{x^2}{a^2}}} xy\left[\int_{-c\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}}}^{c\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}}} z\,dz\right]dy\,dx$$
$= 0$, since the integrand z is an odd function.

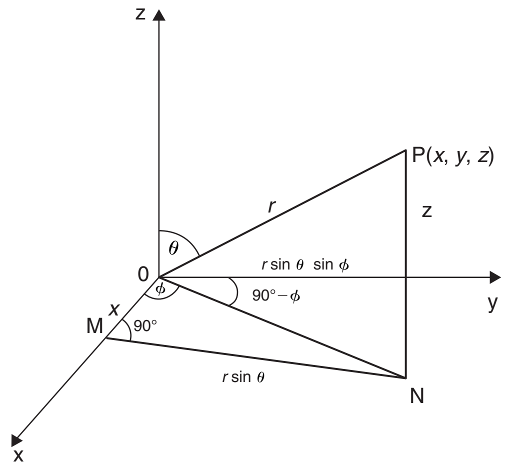
6.10 CHANGE TO SPHERICAL POLAR COORDINATES FROM CARTESIAN COORDINATES IN A TRIPLE INTEGRAL

Let $P(x, y, z)$ be any point in $\Re^3$. Then, the position of this point is determined by the following three numbers:

(i) The distance $r = \sqrt{x^2 + y^2 + z^2}$ of $P(x, y, z)$ from the origin (0, 0, 0).

(ii) The polar distance θ, where θ is the angle between the radius vector OP and the positive direction of z-axis.

(iii) The angle ϕ, which the projection of the radius vector OP on the xy-plane makes with the x-axis.



Then,
$$x = r\ sin\theta\ cos\phi,\ \ y = r\ sin\theta\ sin\phi,\ \ z = r\ cos\theta$$
and so, $x^2 + y^2 + z^2 = r^2$. Under these transformations, the region
$$R = \{(x, y, z);\ x^2 + y^2 + z^2 \le a^2\}$$
is mapped into the region
$R' = \{(r, \theta, \phi); 0 \le r \le a, 0 \le \theta \le \pi, 0 \le \phi \le 2\pi\}$.
Also,
$$\frac{\partial(x,\ y,\ z)}{\partial\ (r, \theta, \phi)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix}$$
$$= \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}$$
$$= r^2\sin\theta.$$

Hence,

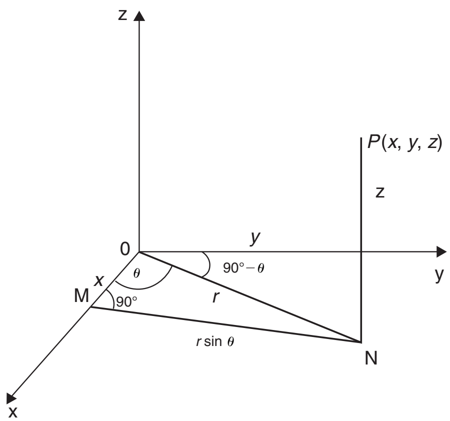
$$I = \iiint_R f(x,\ y,\ z)dx\ dy\ dz$$

$$= \iiint_{R'} f(r\sin\theta\cos\phi,\ r\sin\theta\sin\phi,\ r\cos\theta)$$

$$\cdot\, r^2 \sin\theta\ dr\ d\theta\ d\phi.$$

The polar spherical coordinates are useful when the region of integration is a sphere or a part of it. If the region of integration is a whole sphere, then $0 \le r \le a$, $0 \le \theta \le \pi$, and $0 \le \phi \le 2\pi$. But if the region of integration is the positive octant of the sphere, then $0 \le r \le a$, $0 \le \theta \le \pi$, and $0 \le \phi \le \frac{\pi}{2}$.

Remark 6.2. If the region of integration is a right circular cylinder, then the Cartesian coordinates are changed to *cylindrical polar coordinates* $(r,\ \theta,\ z)$ because the position of $P(x,\ y,\ z)$ is determined by r, θ, and z as shown in the following figure:



Then,

$$x = r\cos\theta,\ y = r\sin\theta,\ \text{and } z = z,$$

and

$$\frac{\partial(x,y,z)}{\partial(r,\theta,z)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$= r\cos^2\theta + r\sin^2\theta = r.$$

Hence,

$$I = \iiint_R f(x,y,z)dx\ dy\ dz$$

$$= \iiint_R f(r\cos\theta,\ r\sin\theta,\ z)\ r\ dr\ d\theta\ dz.$$

EXAMPLE 6.68

Evaluate $I = \iiint z(x^2+y^2)dx\ dy\ dz$ over $x^2 + y^2 \le 1$ and $2 \le z \le 3$.

Solution. The region of integration is

$$V = \{(x,y,z);\ x^2+y^2 \le 1,\ 2 \le z \le 3\}.$$

Using the transformation $x = r\cos\theta$, $y = r\sin\theta$, and $z = z$ (cylindrical polar coordinates), we have

$$I = \int_0^1 \int_0^{2\pi} \left[\int_2^3 zr^2.r\ dz \right] d\theta\ dr = \int_0^1 r^3 \int_0^{2\pi} \left[\frac{z^2}{2}\right]_2^3 d\theta\ dr$$

$$= \int_0^1 r^3 \int_0^{2\pi} \left[\frac{9}{2} - \frac{4}{2}\right] d\theta\ dr = \frac{5}{2}\int_0^1 r^3 [\theta]_0^{2\pi} dr$$

$$= 5\pi \int_0^1 r^3 dr = 5\pi \left[\frac{r^4}{4}\right]_0^1 = \frac{5\pi}{4}.$$

EXAMPLE 6.69

Evaluate $I = \iiint \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2}}\ dx\,dy\,dz$ over the region

$$V = \left\{(x,y,z);\ x \ge 0,\ y \ge 0,\ z \ge 0,\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \le 1\right\}$$

Solution. Substituting $\frac{x}{a} = X$, $\frac{y}{b} = Y$, and $\frac{z}{c} = Z$ so that $dx = adX$, $dy = bdY$, $dz = cdZ$ and hence, $dx\ dy\ dz = abc\ dX\ dY\ dZ$. Therefore,

$$I = abc \iiint (1 - X^2 - Y^2 - Z^2)^{\frac{1}{2}} dx\ dy\ dz,$$

over the region

$$V' = \{(X,Y,Z);\ X \ge 0,\ Y \ge 0,\ Z \ge 0,\ X^2 + Y^2 + Z^2 \le 1\}.$$

Using spherical polar coordinates,

$$X = r\sin\theta\cos\phi,\ Y = r\sin\theta\sin\phi,\ \text{and}$$
$$Z = r\cos\theta,$$

the region of integration becomes

$$V'' = \left\{(r,\theta,\phi);\ 0 \le r \le 1, 0 < 0, \le \frac{\pi}{2},\ 0 \le \phi \le \frac{\pi}{2}\right\}.$$

Hence,

$$I = abc \int_0^1 \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} (1-r^2)^{\frac{1}{2}} r^2 \sin\theta \, dr\, d\theta\, d\phi$$

$$= abc \int_0^1 r^2(1-r^2)^{\frac{1}{2}} \int_0^{\frac{\pi}{2}} \sin\theta \left[\int_0^{\frac{\pi}{2}} d\phi\right] d\theta\, dr$$

$$= abc \int_0^1 r^2(1-r^2)^{\frac{1}{2}} \int_0^{\frac{\pi}{2}} \sin\theta [\phi]_0^{\frac{\pi}{2}}\, d\theta\, dr$$

$$= \frac{abc\pi}{2} \int_0^1 r^2(1-r^2)^{\frac{1}{2}} \left[\int_0^{\frac{\pi}{2}} \sin\theta\, d\theta\right] dr$$

$$= \frac{abc\pi}{2} \int_0^1 r^2(1-r^2)^{\frac{1}{2}} [-\cos\theta]_0^{\frac{\pi}{2}}\, dr$$

$$= \frac{abc\pi}{2} \int_0^1 r^2(1-r^2)^{\frac{1}{2}}\, dr. \qquad (1)$$

But, substituting $r = \sin t$ so that $dr = \cos t\, dt$, we have

$$\int_0^1 r^2\left(1-r^2\right)^{\frac{1}{2}} dr = \int_0^{\frac{\pi}{2}} \sin^2 t \sqrt{1-\sin^2 t}\ \cos t\, dt$$

$$= \int_0^{\frac{\pi}{2}} \sin^2 t \cos^2 t\, dt = \frac{1}{4.2}\cdot\frac{\pi}{2} = \frac{\pi}{16}.$$

Hence (1) reduces to

$$I = \frac{\pi abc}{2}\left(\frac{\pi}{16}\right) = \frac{\pi^2 abc}{32}$$

EXAMPLE 6.70

Evaluate

$$I = \int_0^1 \int_0^{\sqrt{1-x^2}} \int_0^{\sqrt{1-x^2-y^2}} \frac{dz\, dy\, dx}{\sqrt{1-x^2-y^2-z^2}}$$

by changing to spherical polar coordinates.

Solution. The region of integration is

$$V = \left\{(x, y, z);\ 0 \le x \le 1,\ 0 \le y \le \sqrt{1-x^2},\ 0 \le z \le \sqrt{1-x^2-y^2}\right\}.$$

Now, we transform the region by using spherical polar coordinates, by substituting $x = r \sin\theta \cos\phi$, $y = r \sin\theta \sin\phi$, and $z = r\cos\theta$. The transformed region is

$$V' = \left\{(r,\theta,\phi); 0 \le r \le 1, 0 \le \theta \le \frac{\pi}{2}, 0 \le \phi \le \frac{\pi}{2}\right\}.$$

Therefore,

$$I = \int_0^1 \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \frac{r^2 \sin\theta}{\sqrt{1-r^2}}\, dr\, d\theta\, d\phi$$

$$= \int_0^1 \frac{r^2}{\sqrt{1-r^2}} \int_0^{\frac{\pi}{2}} \sin\theta \left[\int_0^{\frac{\pi}{2}} d\phi\right] d\theta\, dr$$

$$= \int_0^1 \frac{r^2}{\sqrt{1-r^2}} \int_0^{\frac{\pi}{2}} \sin\theta [\phi]_0^{\frac{\pi}{2}}\, d\theta\, dr$$

$$= \frac{\pi}{2} \int_0^1 \frac{r^2}{\sqrt{1-r^2}} \left[\int_0^{\frac{\pi}{2}} \sin\theta\, d\theta\right] dr$$

$$= \frac{\pi}{2} \int_0^1 \frac{r^2}{\sqrt{1-r^2}} [-\cos\theta]_0^{\frac{\pi}{2}}\, dr$$

$$= \frac{\pi}{2} \int_0^1 \frac{r^2}{\sqrt{1-r^2}} dr = \frac{\pi}{2} \int_0^1 \frac{1-(1-r^2)}{\sqrt{1-r^2}}\, dr$$

$$= \frac{\pi}{2}\left[\int_0^1 \frac{dr}{\sqrt{1-r^2}} - \int_0^1 \sqrt{1-r^2}\right] dr$$

$$= \frac{\pi}{2}\left\{\left[\sin^{-1} r - \left(\frac{r\sqrt{1-r}}{2} + \frac{1}{2}\sin^{-1} r\right)\right]_0^1\right\}$$

$$= \frac{\pi}{2}\left\{\frac{\pi}{2} - \frac{\pi}{4}\right\} = \frac{\pi}{2}\cdot\frac{\pi}{4} = \frac{\pi^2}{8}.$$

Note: This example is a particular case of Example 6.65 for $a = 1$.

EXAMPLE 6.71

Evaluate $I = \iiint\limits_V (x^2+y^2+z^2)^m dx\, dy\, dz,\ m > 0$ over the region $V = \{(x, y, z);\ x^2 + y^2 + z^2 \le 1\}$.

Solution. The given region of integration is

$$V = \left\{(x, y, z);\ x^2 + y^2 + z^2 \le 1\right\}.$$

Changing to spherical polar coordinates by substituting

$$x = r\sin\theta\cos\phi,\ y = r\sin\theta\sin\phi,\ \text{and } z = r\cos\theta,$$

we get $x^2 + y^2 + z^2 = r^2$ and $\dfrac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = r^2\sin\theta$.

Therefore, the region of integration reduces to

$V' = \{(r,\ \theta,\ \phi); 0 \le r \le 1, 0 \le \theta \le \pi, 0 \le \phi \le 2\pi\}$.

Hence,

$$\begin{aligned}
I &= \int_0^1\int_0^\pi\int_0^{2\pi} r^{2m+2}\sin\theta\ dr\ d\theta\ d\phi \\
&= \int_0^1 r^{2m+2}\int_0^\pi \sin\theta\left[\int_0^{2\pi} d\phi\right] d\theta\ dr \\
&= \int_0^1 r^{2m+2}\int_0^\pi \sin\theta[\phi]_0^{2\pi} d\theta\ dr \\
&= 2\pi\int_0^1 r^{2m+2}\left[\int_0^\pi \sin\theta\ d\theta\right] dr \\
&= 2\pi\int_0^1 r^{2m+2}[-\cos\theta]_0^\pi dr \\
&= 4\pi\int_0^1 r^{2m+2} dr = 4\pi\left[\frac{r^{2m+3}}{2m+3}\right]_0^1 \\
&= \frac{4\pi}{2m+3}.
\end{aligned}$$

6.11 VOLUME AS A TRIPLE INTEGRAL

In Cartesian coordinates, the volume of a region V is given by the triple integral

$$\iiint dx\ dy\ dz,$$

where the limits of integration are chosen to cover the entire region V.

In spherical polar coordinates, the volume of a region V is given by the triple integral

$$\iiint r^2\sin\theta\ dr\ d\theta\ d\phi,$$

where the limits of integration are chosen to cover the entire region V.

In cylindrical coordinates, the volume of a region V is given by the triple integral

$$\iiint r\ dr\ d\theta\ dz,$$

where the limits of integration are chosen to cover the entire region V.

EXAMPLE 6.72

Find the volume of the sphere $x^2 + y^2 + z^2 = r^2$.

Solution. The required volume is given by

$$V = 8\iiint dx\ dy\ dz,$$

taken over the positive octant of the given sphere. Changing to spherical polar coordinates, we put $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, and $z = r\cos\theta$. So, $x^2 + y^2 + z^2 = r^2$. In the positive octant, we have

$$0 \le r \le a,\ 0 \le \theta \le \frac{\pi}{2},\ \text{and } 0 \le \phi \le \frac{\pi}{2}.$$

Therefore,

$$\begin{aligned}
V &= 8\int_0^a\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} r^2\sin\theta\ dr\ d\theta\ d\phi \\
&= 8\int_0^a r^2\int_0^{\frac{\pi}{2}}\sin\theta\left[\int_0^{\frac{\pi}{2}} d\phi\right] d\theta\ dr \\
&= 8\int_0^a r^2\int_0^{\frac{\pi}{2}}\sin\theta[\phi]_0^{\frac{\pi}{2}} d\theta\ dr \\
&= 4\pi\int_0^a r^2\left[\int_0^{\frac{\pi}{2}}\sin\theta\ d\theta\right] dr \\
&= 4\pi\int_0^a r^2[-\cos\phi]_0^{\frac{\pi}{2}}\ dr = 4\pi\int_0^a r^2\ dr \\
&= 4\pi\left[\frac{r^3}{3}\right]_0^a = \frac{4\pi\ a^3}{3}.
\end{aligned}$$

EXAMPLE 6.73

Find the volume of the ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$.

Solution. Substituting $\frac{x}{a} = X,\ \frac{y}{b} = Y,$ and $\frac{z}{c} = Z$, we have
$dx = adX$, $dy = adY$, and $z = adZ$. Therefore, the volume is given by

$$V = abc \iiint dX\ dY\ dZ,$$

taken over $X^2 + Y^2 + Z^2 = 1$. Changing to spherical polar coordinates by substituting $X = r \sin\theta \cos\phi$, $y = r\sin\theta \sin\phi$, and $z = r\cos\theta$, we have

$$V = 8abc \int_0^1 \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} r^2 \sin\theta\ dr\ d\theta\ d\phi$$

$$= 8abc \int_0^1 r^2 \int_0^{\frac{\pi}{2}} \sin\theta \left[\int_0^{\frac{\pi}{2}} d\phi\right] d\theta\ dr$$

$$= 8abc \int_0^1 r^2 \int_0^{\frac{\pi}{2}} \sin\theta [\phi]_0^{\frac{\pi}{2}}\ d\theta\ dr$$

$$= 4\pi abc \int_0^1 r^2 \left[\int_0^{\frac{\pi}{2}} \sin\theta\ d\theta\right] dr$$

$$= 4\pi abc \int_0^1 r^2 [-\cos\theta]_0^{\frac{\pi}{2}}\ dr = 4\pi abc \int_0^1 r^2\ dr$$

$$= 4\pi abc \left[\frac{r^3}{3}\right]_0^1 = \frac{4}{3}\pi abc.$$

EXAMPLE 6.74

Find the volume of the solid bounded by the surface

$$\left(\frac{x}{a}\right)^{\frac{2}{3}} + \left(\frac{y}{b}\right)^{\frac{2}{3}} + \left(\frac{z}{c}\right)^{\frac{2}{3}} = 1.$$

Solution. Substituting $\left(\frac{x}{a}\right)^{\frac{2}{3}} = X,\ \left(\frac{y}{b}\right)^{\frac{2}{3}} = Y,$ and $\left(\frac{z}{c}\right)^{\frac{2}{3}} = Z$, that is, $x = aX^3$, $y = bY^3$, and $z = cZ^3$, we get $dx = 3aX^2\ dX$, $dy = 3bY^2\ dY$, and $dz = 3cZ^2\ dZ$. Then, the required volume is given by

$$V = \iiint_V dx\ dy\ dz$$

$$= \iiint 27abcX^2Y^2Z^2 dX\ dY\ dZ,$$

taken throughout the sphere $X^2 + Y^2 + Z^2 = 1$.

Changing to spherical polar coordinates (r, θ, ϕ), we have

$$V = 8 \iiint_{V'} 27abc\ r^2 \sin^2\theta \cos^2\phi.$$

$$r^2 \sin^2\theta \sin^2\phi\ r^2 \cos^2\theta . r^2 \sin\theta\ dr\ d\theta\ d\phi,$$

where

$$V' = \left\{(r,\ \theta,\ \phi); 0 \le r \le 1, 0 \le \theta \le \frac{\pi}{2}, 0 \le \phi \le \frac{\pi}{2}\right\}.$$

Thus,

$$V = 216abc \int_0^1 r^8 \int_0^{\frac{\pi}{2}} \sin^5\theta \cos^2\theta$$

$$. \left[\int_0^{\frac{\pi}{2}} \sin^2\phi \cos^2\phi\ d\phi\right] d\theta\ dr$$

$$= 216abc \int_0^1 r^8 \left[\int_0^{\frac{\pi}{2}} \sin^5\theta \cos^2\theta . \frac{1}{4.2} . \frac{\pi}{2} d\theta\right] dr$$

$$= \frac{216}{16}\pi abc \int_0^1 r^8 \left[\int_0^{\frac{\pi}{2}} \sin^5\theta \cos^2\theta\ d\theta\right] dr$$

$$= \frac{216}{16}\pi abc \int_0^1 r^8 \left[\frac{4.2.1}{7.5.3.1}\right] dr$$

$$= \frac{216}{210}\pi abc \left[\frac{r^9}{9}\right]_0^1$$

$$= \frac{216\pi abc}{1890}$$

$$= \frac{4}{35}\pi abc.$$

Second Method: For the positive octant, we have

$$V = \iiint dx\ dy\ dz.$$

Substituting $\left(\frac{x}{a}\right)^{\frac{2}{3}} = X,\ \left(\frac{y}{b}\right)^{\frac{2}{3}} = y,$ and $\left(\frac{z}{c}\right)^{\frac{2}{3}} = Z$, we have $dx = \frac{3}{2} aX^{\frac{1}{2}} dX,\ dy = \frac{3}{2} bY^{\frac{1}{2}} dY,\ dz = \frac{3}{2} cZ^{\frac{1}{2}} dZ,$ and $X + Y + Z \le 1$. Therefore, by Dirichlet's Theorem,

$$V = \int \iint \frac{27}{8} abcX^{\frac{1}{2}}Y^{\frac{1}{2}}Z^{\frac{1}{2}}\ dX\ dY\ dZ$$

$$= \frac{27}{8} abc \int \iint X^{\frac{3}{2}-1} Y^{\frac{3}{2}-1} Z^{\frac{3}{2}-1} dX\ dYdZ$$

$$= \frac{27}{8} abc \frac{\Gamma\left(\frac{3}{2}\right)\,\Gamma\left(\frac{3}{2}\right)\,\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(1+\frac{3}{2}+\frac{3}{2}+\frac{3}{2}\right)}$$

$$= \frac{27abc}{8} \frac{\left(\Gamma\left(\frac{3}{2}\right)\right)^3}{\Gamma\left(\frac{9}{2}+1\right)}$$

$$= \frac{27}{8}\ abc \frac{\left(\frac{1}{2}.\sqrt{\pi}\right)^2}{\frac{9}{2}\cdot\frac{7}{2}\cdot\frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi}}$$

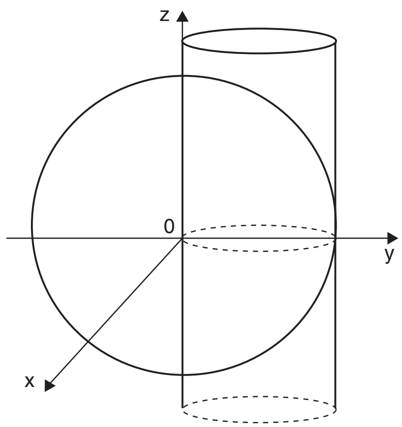
$$= \frac{\pi}{8} abc. \frac{4}{35}.$$

Hence, the total volume is $= 8\,\frac{\pi abc}{8}.\frac{4}{35} = \frac{4\pi abc}{35}$.

EXAMPLE 6.75

Find the volume of the portion cut off from the sphere $x^2 + y^2 + z^2 = a^2$ by the cylinder $x^2 + y^2\ ax$.

Solution. The required volume is



$$V = 4 \int_0^a \int_0^{\sqrt{ax-x^2}} \int_0^{\sqrt{a^2-x^2-y^2}} dx\ dy\ dz$$

$$= 4 \int_0^a \int_0^{\sqrt{ax-x^2}} \left(a^2 - x^2 - y^2\right)^{\frac{1}{2}} dy\ dx$$

$$= 4 \int_0^{\frac{\pi}{2}} \int_0^{a\cos\theta} \sqrt{a^2 - r^2}\ r\ dr\ d\theta,$$

changing to polar coordinates

$$= \frac{4}{2} \int_0^{\frac{\pi}{2}} \int_0^{a\cos\theta} 2r\sqrt{a^2 - r^2}\ dr\ d\theta$$

$$= 2 \int_0^{\frac{\pi}{2}} \left[\frac{-(a^2 - r^2)^{\frac{3}{2}}}{\frac{3}{2}}\right]_0^{a\cos\theta} d\theta$$

$$= \frac{4}{3} \int_0^{\frac{\pi}{2}} [-a^3 \sin^3\theta + a^3] d\theta$$

$$= \frac{4}{3} a^3 \left[-\frac{2}{3} + \frac{\pi}{2}\right]$$

$$= \frac{2}{3} a^3 \left(\pi - \frac{4}{3}\right).$$

EXAMPLE 6.76

Prove that the volume of the wedge intercepted between the cylinder $x^2 + y^2 = 2ax$ and the planes $z = mx$ and $z = nx$ is $\pi(m - n)a^3$.

Solution. The required volume is given by

$$V = 2 \int_0^{2a} \int_0^{\sqrt{2ax-x^2}} \int_{nx}^{mx} dx\ dy\ dz$$

$$= 2 \int_0^{2a} \int_0^{\sqrt{2ax-x^2}} [z]_{nx}^{mx}\ dy\ dx$$

$$= 2 \int_0^{2a} \int_0^{\sqrt{2ax-x^2}} (m - n)x\ dy\ dx$$

$$= 2 \int_0^{2a} (m - n)x[y]_0^{\sqrt{2ax-x^2}}\ dx$$

$$= 2(m - n) \int_0^{2a} x\sqrt{2ax - x^2}\ dx$$

$$= 2(m - n) \int_0^{2a} x^{\frac{3}{2}}\sqrt{2ax - x^2}\ dx.$$

Substituting $x = 2a\sin^2\theta$, we get $dx = 4a\sin\theta\cos\theta\ d\theta$. The limits of integration are $\theta = 0$ to $\theta = \frac{\pi}{2}$. Therefore,

$$V = 2(m - n) \int_0^{\frac{\pi}{2}} 16a^3 \sin^4\theta \cos^2\theta\ d\theta$$

$$= 32(m - n). \frac{3.1}{6.4.2}.\frac{\pi}{2} = \pi(m - n)a^3.$$

EXAMPLE 6.77

The axes of two right circular cylinders of the same radius a, intersect at right angles. Prove that the volume inside both the cylinders is $\frac{16a^3}{3}$.

Solution. Let the equations of the cylinders be

$$x^2 + y^2 = a^2 \text{ and } x^2 + z^2 = a^2.$$

Therefore, the region of integration is defined by

$$-\sqrt{a^2 - x^2} \le z \le \sqrt{a^2 - x^2},$$
$$-\sqrt{a^2 - x^2} \le y \le \sqrt{a^2 - x^2}, \text{ and}$$
$$-a \le x \le a.$$

Hence, the required volume is

$$V = \int_{-a}^{a} \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} dx\, dy\, dz$$

$$= 8 \int_{0}^{a} \int_{0}^{\sqrt{a^2-x^2}} \int_{0}^{\sqrt{a^2-x^2}} dx\, dy\, dz$$

$$= 8 \int_{0}^{a} \int_{0}^{\sqrt{a^2-x^2}} \sqrt{a^2 - x^2}\, dy\, dx$$

$$= 8 \int_{0}^{a} \sqrt{a^2 - x^2} [y]_0^{\sqrt{a^2-x^2}}\, dx$$

$$= 8 \int_{0}^{a} (a^2 - x^2)\, dx$$

$$= 8 \left[a^2 x - \frac{x^3}{3}\right]_0^a = 8\left[a^3 - \frac{a^3}{3}\right] = \frac{16a^3}{3}.$$

EXAMPLE 6.78

Find the volume in the positive octant bounded by the coordinate planes and the plane $x + 2y + 3z = 4$.

Solution. The region of integration is

$$V = \left\{(x,\ y,\ z);\ 0 \le x \le 4,\ 0 \le y \le \frac{4-x}{2},\ 0 \le z \le \frac{4 - x - 2y}{3}\right\}.$$

Therefore, the required volume is

$$V = \int_0^4 \int_0^{\frac{4-x}{2}} \left[\int_0^{\frac{4-x-2y}{3}} dz\right] dy\, dx$$

$$= \frac{1}{3} \int_0^4 \int_0^{\frac{4-x}{2}} (4 - x - 2y)\, dy\, dx$$

$$= \frac{1}{3} \int_0^4 \left(4y - xy - 2\frac{y^2}{2}\right)_0^{\frac{4-x}{2}} dx$$

$$= \frac{1}{3} \int_0^4 \left(8 - 2x - 2x + \frac{x^2}{2} - 4 - \frac{x^2}{4} + 2x\right) dx$$

$$= \frac{1}{3} \int_0^4 \left(4 - 2x + \frac{x^2}{4}\right) dx = \frac{1}{3}\left(4x - x^2 + \frac{x^3}{12}\right)_0^4$$

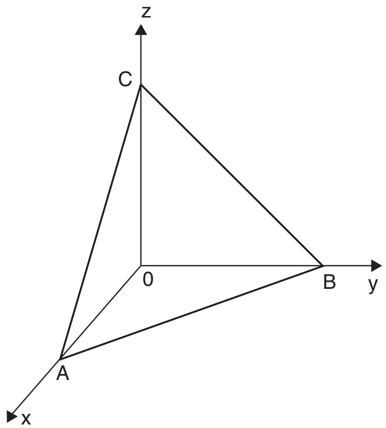
$$= \frac{1}{3}\left[16 - 16 + \frac{64}{12}\right] = \frac{16}{9}.$$

EXAMPLE 6.79

Find the volume of the tetrahedron bounded by the planes $x = 0$, $y = 0$, $z = 0$, and $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$; $a, b, c \ge 0$.

Solution. The region of integration is bounded by four planes $x = 0$, $y = 0$, $z = 0$, and $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$. It is bounded below and above by $z = 0$ and $z = c\left(1 - \frac{x}{a} - \frac{y}{b}\right)$. Its projection on the xy-plane is a triangle bounded by $x = 0$, $y = 0$, and $\frac{x}{a} + \frac{y}{b} = 1$. Therefore, the region is

$$V = \left\{(x, y, z);\ 0 \le x \le a,\ 0 \le y \le b\left(1 - \frac{x}{a}\right),\ 0 \le Z \le c\left(1 - \frac{x}{a} - \frac{x}{b}\right)\right\}.$$



Hence, the volume of the tetrahedron OABC is

$$V = \iiint_V dx\,dy\,dz = \int_0^a \int_0^{b\left(1-\frac{x}{a}\right)} \int_0^{c\left(1-\frac{x}{a}-\frac{y}{b}\right)} dz\,dy\,dx$$

$$= \int_0^a \int_0^{b\left(1-\frac{x}{a}\right)} [z]_0^{c\left(1-\frac{x}{a}-\frac{y}{b}\right)} dy\,dx$$

$$= c\int_0^a \left[\int_0^{b\left(1-\frac{x}{a}\right)} \left(1-\frac{x}{a}-\frac{y}{b}\right)dy\right] dx$$

$$= c\int_0^a \left[\frac{\left(1-\frac{x}{a}-\frac{y}{b}\right)^2}{2\left(-\frac{1}{b}\right)}\right]_0^{b\left(1-\frac{x}{a}\right)} dx = \frac{bc}{2}\int_0^a \left(1-\frac{x}{a}\right)^2 dx$$

$$= \frac{bc}{2}\left[\frac{\left(1-\frac{x}{a}\right)^3}{3\left(-\frac{1}{a}\right)}\right]_0^a - \frac{abc}{6}[0-1]$$

$$= \frac{abc}{6} \text{cu units.}$$

Second Method: Substituting $\frac{x}{a} = u$, $\frac{y}{b} = v$, and $\frac{z}{c} = w$, we have $u \geq o, v \geq 0$, $w \geq 0$, and $u + v + w \leq 1$. Therefore

$$V = \iiint_V dx\,dy\,dz$$

$$= \iiint_{V'} a\,d\,u.b\,d\,u.c\,d\,u,\ u+v+w \leq 1$$

$$= abc \iiint_{V'} u^{1-1}v^{1-1}w^{1-1} du\,dv\,dw$$

$$= abc\frac{\Gamma(1)\,\Gamma(1)\,\Gamma(1)}{\Gamma(1+1+1+1)}, \text{ by Dirichlet's Theorem}$$

$$= \frac{abc}{3!} = \frac{abc}{6}.$$

EXAMPLE 6.80

Find the volume of the portion cut off from a sphere $x^2 + y^2 + z^2 = a^2$ by a cone $x^2 + y^2 = z^2$.

Solution. The origin is the center of the sphere and the vertex of the cone $x^2 + y^2 = z^2$. Therefore, the volume is symmetrical about the plane $z = 0$. Hence,

$$V = 2\iiint dx\,dy\,dz.$$

Changing the coordinates to spherical polar, $x^2 + y^2 + z^2 = a^2$ reduces to $r^2 = a^2$ or $r = a$. Further, $x^2 + y^2 = z^2$ reduces to

$$r^2 \sin^2\theta \cos^2\phi + r^2 \sin^2\theta \sin^2\phi = r^2 \cos^2\theta$$

or

$$r^2 \sin^2\theta(\cos^2\phi + \sin^2\phi) = r^2 \cos^2\theta$$

or

$$\sin^2\theta = \cos^2\theta, \text{ which yields } \theta = \frac{\pi}{4}.$$

Thus, θ varies from 0 to $\frac{\pi}{4}$ and ϕ varies from 0 to π. Therefore,

$$V = 2\int_0^a \int_0^{\frac{\pi}{4}} \int_0^{2\pi} r^2 \sin\theta\,dr\,d\theta\,d\phi$$

$$= 2\int_0^a r^2 \int_0^{\frac{\pi}{4}} \sin\theta\,[\phi]_0^{2\pi}\,d\theta\,dr$$

$$= 4\pi\int_0^a r^2 \int_0^{\frac{\pi}{4}} \sin\theta\,d\theta$$

$$= 4\pi\int_0^a r^2[-\cos\theta]_0^{\frac{\pi}{4}}dr - 4\pi\int_0^a r^2\left(1-\frac{1}{\sqrt{2}}\right)dr$$

$$= 4\pi\left(1+\frac{1}{\sqrt{2}}\right)\left[\frac{r^3}{3}\right]_0^a = \frac{4\pi a^3}{3}\left(1-\frac{1}{\sqrt{2}}\right)$$

$$= \frac{2\pi a^3}{3}(2-\sqrt{2}).$$

6.12 MISCELLANEOUS EXAMPLES

EXAMPLE 6.81

Evaluate $\int_4^3\int_1^2 (x+y)^{-2}dx\,dy$.

Solution.

$$\int_4^3\left[\int_1^2 (x+y)^{-2}dx\right]dy$$

$$= \int_4^3\left[\frac{(x+y)^{-1}}{-1}\right]_0^2 dy$$

$$= -\int_4^3\left[\frac{1}{2+y}-\frac{1}{1+y}\right]dy$$

$$= -[\log(2+y)-\log(1+y)]_4^3$$

$$= \log 4+\log 6-2\log 5$$

$$= \log\frac{24}{25}.$$

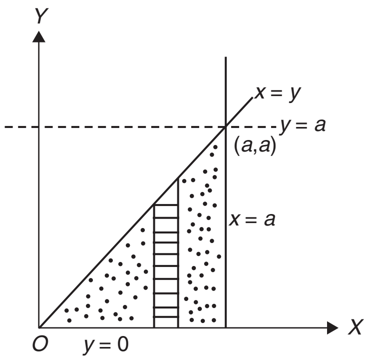
EXAMPLE 6.82

Evaluate the integral $\int_0^a\int_y^a \frac{x}{x^2+y^2}dxdy$ by changing the order of integration.

Solution. The given integral is

$$I = \int_0^a\int_y^a \frac{x}{x^2+y^2}dx\,dy.$$

The region of integration is bounded by the lines $x=y$, $x=a$, $y=0$ and $y=a$. Thus the region of integration is shown in the figure below:



On changing the order of integration, we first integrate with respect to y along the strip parallel to y axis. The strip extends from $y=0$ to $y=x$. To cover the whole region, we then integrate with respect to x from $x=0$ to $x=a$. Hence

$$I = \int_0^a\int_0^x \frac{x}{x^2+y^2}dx\,dy = \int_0^a x\left[\frac{1}{x}\tan^{-1}\frac{y}{x}\right]_0^x dx$$

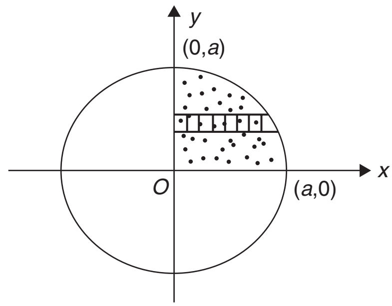
$$= \int_0^a \frac{\pi}{4}dx = \frac{\pi}{4}a.$$

EXAMPLE 6.83

By changing the order of integrations, evaluate

$$\int_0^a\int_0^{\sqrt{a^2-x^2}} (a^2-x^2-y^2)dy\,dx.$$

Solution. The region of integration is bounded by $x=0,\ x=a,\ y=0$ and the circle $x^2+y^2=a^2$.



After changing the order of integration, we have to integrate the integrand first with respect to x and then with respect to y. We take a strip parallel to x-axis. The limit of x varies from 0 to $\sqrt{a^2-y^2}$. To cover the whole region, the limits of y will vary from 0 to a. Hence the given integral is

$$I = \int_0^a\int_0^{\sqrt{a^2-y^2}} \left[(a^2-y^2)-x^2\right]dxdy$$

$$= \int_0^a \left[\frac{x}{2}\sqrt{(a^2-y^2)-x^2} + \frac{a^2-y^2}{2}\sin^{-1}\frac{x}{\sqrt{a^2-y^2}}\right]_0^{\sqrt{a^2-y^2}} dy$$

$$= \int_0^a \left[0 + \frac{\pi(a^2 - y^2)}{4} - 0\right] dy$$

$$= \frac{\pi}{4}\int_0^a (a^2 - y^2)dy = \frac{\pi}{4}\left[a^2 y - \frac{y^3}{3}\right]_0^a$$

$$= \frac{\pi}{4}\left[a^3 - \frac{a^3}{3}\right] = \frac{\pi}{6}a^3.$$

EXAMPLE 6.84

Changing the order of integration of $\int_0^\infty \int_0^\infty e^{-xy} \sin nx \, dx\, dy$, show that $\int_0^\infty \left(\frac{\sin nx}{x}\right) dx = \frac{\pi}{2}$.

Solution. We have

$$I = \int_0^\infty \int_0^\infty e^{-xy} \sin nx \, dx\, dy$$

$$= \int_0^\infty \sin nx \left[\int_0^\infty e^{-xy} dy\right] dx$$

$$= \int_0^\infty \sin nx \left[\frac{e^{-xy}}{-x}\right]_0^\infty dx$$

$$= \int_0^\infty \frac{\sin nx}{x} dx. \qquad (1)$$

On the other hand,

$$I = \int_0^\infty \int_0^\infty e^{-xy} \sin nx \, dx\, dy$$

$$= \int_0^\infty \left[\int_0^\infty e^{-xy} \sin nx \, dx\right] dy$$

$$= \int_0^\infty \left[\frac{e^{-xy}}{n^2 + y^2}(n \cos nx + y \sin nx)\right]_0^\infty dy$$

$$= \int_0^\infty \frac{n}{n^2 + y^2} dy = \left[\tan^{-1}\frac{y}{n}\right]_0^\infty = \frac{\pi}{2} \qquad (2)$$

From (1) and (2), it follows that

$$\int_0^\infty \frac{\sin nx}{x} dx = \frac{\pi}{2}.$$

EXAMPLE 6.85

By transforming into polar co-ordinates, evaluate $\iint \frac{x^2 y^2}{x^2+y^2} dxdy$ over the annular region between the circles $x^2 + y^2 = a^2$ and $x^2 + y^2 = b^2$, where $b > a$.

Solution. Putting $x = r \cos\theta,\ y = r \sin\theta$, we have $dx\, dy = r dr d\theta$. Therefore

$$\iint \frac{x^2 y^2}{x^2 + y^2} dx\, dy = \int_a^b \int_0^{2\pi} \frac{r^4 \sin^2\theta \cos^2\theta}{r^2} r dr\, d\theta$$

$$= \frac{1}{16}(b^4 - a^4) \int_0^{2\pi} \left(\frac{1 - \cos 4\theta}{2}\right) d\theta$$

$$= \frac{b^4 - a^4}{16}\left[\frac{1}{2}\theta - \frac{1}{8}\sin 4\theta\right]_0^{2\pi}$$

$$= \frac{\pi}{16}(b^4 - a^4).$$

EXAMPLE 6.86

Evaluate the integral $\int_0^1 \int_0^{\sqrt{1-x^2}} \int_{\sqrt{x^2+y^2}}^0 \frac{dz\, dy\, dx}{\sqrt{x^2+y^2+z^2}}$.

Solution. We have

$$I = \int_0^1 \int_0^{\sqrt{1-x^2}} \int_{\sqrt{x^2+y^2}}^0 \frac{1}{\sqrt{x^2 + y^2 + z^2}} dz\, dy\, dx$$

$$I = \int_0^1 \int_0^{\sqrt{1-x^2}} \int_{\sqrt{x^2+y^2}}^0 \frac{1}{\sqrt{(\sqrt{x^2 + y^2})^2 + z^2}} dz\, dy\, dx$$

$$= \int_0^1 \int_0^{\sqrt{1-x^2}} \left[\log\left|z + \sqrt{x^2 + y^2 + z^2}\right|\right]_{\sqrt{x^2+y^2}}^0 dx\, dy$$

$$= \int_0^1 \int_0^{\sqrt{1-x^2}} \left[\log\left|\sqrt{x^2 + y^2}\right| - \log\left|\sqrt{x^2 + y^2} + \sqrt{2(x^2 + y^2)}\right|\right] dxdy$$

$$= -\log(\sqrt{2}+1)\int_0^1\int_0^{\sqrt{1-x^2}} dy\,dx$$

$$= -\log(\sqrt{2}+1)\int_0^1 \sqrt{1-x^2}dx$$

$$= -\log(\sqrt{2}+1)\left[\frac{1}{2}x\sqrt{1-x^2}+\frac{1}{2}\sin^{-1}\frac{x}{1}\right]_0^1$$

$$= -\log(\sqrt{2}+1)\left[\frac{1}{2}\sin^{-1}1\right] = -\frac{\pi}{4}\log\left(\sqrt{2}+1\right).$$

EXAMPLE 6.87

Find the area in the first quadrant enclosed by the curve $\left(\frac{x}{a}\right)^{\alpha}+\left(\frac{y}{b}\right)^{\beta}=1$, where $\alpha>0,\ \beta>0$.

Solution. The equation of the curve is

$$\left(\frac{x}{a}\right)^{\alpha}+\left(\frac{y}{b}\right)^{\beta}=1,\quad \alpha,\ \beta>0.$$

The parametric form of the curve is

$$x=a\cos^{\frac{2}{\alpha}}t,\quad y=b\sin^{\frac{2}{\beta}}t.$$

Therefore, the required area is

$$A=\int_{\frac{\pi}{2}}^{0} y\,dx=\int_{\frac{\pi}{2}} y\frac{dx}{dt}dt$$

$$=\int_{\frac{\pi}{2}}^{0}(b\sin^{\frac{2}{\beta}}t)\left(-\frac{2a}{\alpha}\cos^{\left(\frac{2}{\alpha}-1\right)}t\right)\sin t\ dt$$

$$=\frac{2ab}{\alpha}\int_0^{\frac{\pi}{2}}\sin^{\left(\frac{2}{\beta}+1\right)}t\cos^{\left(\frac{2}{\alpha}-1\right)}t\,dt$$

$$=\frac{2ab}{\alpha}\left[\frac{\Gamma\left(\frac{\left(\frac{2}{\beta}+2\right)}{2}\right)\Gamma\left(\frac{\left(\frac{2}{\alpha}-1+1\right)}{2}\right)}{2\Gamma\left(\frac{\frac{2}{\beta}+1+\frac{2}{\alpha}-1+2}{2}\right)}\right]$$

$$=\frac{2ab}{2\alpha\beta}\left[\frac{\Gamma\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{1}{\beta}\right)}{\Gamma\left(\frac{1}{\alpha}+\frac{1}{\beta}+1\right)}\right]$$

$$=\frac{ab}{\alpha+\beta}\frac{\Gamma\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{1}{\beta}\right)}{F\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)}$$

EXERCISES

Evaluation of a Double Integral

1. Evaluate $\iint_D (4-x^2-y^2)\ dx\ dy$ if the region D is bounded by the lines $x=0,\ x=1,\ y=0,$ and $y=\frac{3}{2}$.

   Ans. $\frac{35}{8}$.

2. Evaluate $\int\int e^{2x+3y}dx$ over the triangle bounded by $x=0,\ y=0$, and $x+y=1$.

   Ans. $\frac{1}{6}(2e+1)(e-1)^2$.

3. Evaluate $\int_0^a\int_0^{\sqrt{a^2-y^2}}(a^2-x^2-y^2)^{\frac{1}{2}}dx\ dy$.

   Ans. $\frac{\pi a^3}{6}$.

4. Evaluate $\int_1^a\int_1^b \frac{dx\ dy}{xy}$.

   Ans. $\log b\log a$.

5. Evaluate $\int_0^1\int_0^{\sqrt{1+x^2}}\frac{dx\ dy}{1+x^2+y^2}$.

   Hint: $$I=\int_0^1\frac{1}{\sqrt{1+x^2}}\left[\tan^{-1}\frac{y}{\sqrt{1+x^2}}\right]_0^{\sqrt{1+x^2}}dx$$

   $$=\int_0^1\frac{1}{\sqrt{1+x^2}}[\tan^{-1}1-\tan^{-1}0]dx$$

   $$=\frac{\pi}{4}\int_0^1\frac{1}{\sqrt{1+x^2}}dx$$

   $$=\frac{\pi}{4}[\log\{x+\sqrt{1+x^2}\}]_0^1$$

   $$=\frac{\pi}{4}\log[1+\sqrt{2}].$$

6. Evaluate $\iint x^2y^2dx\ dy$ over the region bounded by $x=0,\ y=0$, and $x^2+y^2=1$.

   Ans. $\int_0^1\int_0^{\sqrt{1-y^2}} x^2y^2dx\ dy=\frac{\pi}{96}$.

7. Evaluate $\iint_R y\, dx\, dy$, where R is the region in the first quadrant bounded by the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$.

Hint: $R=\{(x,y);0 \le x \le a, 0 \le y \le b\sqrt{1-\frac{x^2}{a^2}}\}$

Ans. $\frac{ab^2}{3}$.

8. Evaluate $\iint xy(x+y)dx\, dy$ over the area between $y=x^2$ and $y=x$.

Ans. $\frac{3}{56}$.

9. Evaluate $\iint_D xy\, dx\, dy$, where A is the region common to the circle $x^2+y^2=x$ and $x^2+y^2=y$.

Hint: The points of intersection of two circles are $(0,0)$ and $\left(\frac{1}{2},\frac{1}{2}\right)$. The limits are $x=0$ to $x=\frac{1}{2}$ and $y=\frac{1-\sqrt{1-4x^2}}{2}$ to $y=\sqrt{x-x^2}$.

Ans. $\frac{1}{96}$.

10. Evaluate $\int_0^{\pi}\int_0^{a(1+\cos\theta)} r^2\cos\theta\, d\theta\, dr$.

Ans. $\frac{5\pi a^3}{8}$.

11. Evaluate $\iint r\sin\theta dr\, d\theta$ over the cardioid $r=a\,(1-\cos\,\theta)$ above the initial line.

Hint : $\iint_R r\sin\theta\, dr\, d\theta = \int_0^{\pi}\int_0^{a(1-\cos\theta)} r\sin\theta\, dr\, d\theta$

$$= \int_0^{\pi}\sin\theta\left[\frac{r^2}{2}\right]_0^{a(1-\cos\theta)} d\theta$$

$$= \frac{a^2}{2}\int_0^{\pi}\sin\theta(1-\cos\theta)^2 d\theta$$

$$= \frac{a^2}{2}\left[\frac{(1-\cos\theta)^3}{3}\right]_0^{\pi} = \frac{4a^2}{3}.$$

12. Evaluate $\iint r^2 d\theta\, dr$ over the area of the circle $r=a\cos\,\theta$.

Ans. $\frac{4a^3}{9}$.

13. Evaluate $\int_0^{\frac{\pi}{2}}\int_0^{a\cos\theta} r\sqrt{a^2-r^2}dr\, d\theta$.

Ans. $\frac{a^3}{18}(3\pi-4)$.

14. Evaluate $\iint \frac{r\, dr\, d\theta}{\sqrt{a^2+r^2}}$ over one loop of $r^2=a^2\cos^2 2\theta$.

Ans. $\frac{a}{2}(4-\pi)$.

Change of Variable in a Double Integral

15. Transform the following double integral to polar coordinates and hence, evaluate the same.

$$I=\int_0^{a}\int_0^{\sqrt{a^2-y^2}}\left(a^2-x^2-y^2\right)dx\, dy.$$

Hint: $I=\int_0^{\frac{\pi}{2}}\int_0^{1}\left(a^2-r^2\right) r\, dr\, d\theta$.

Ans. $\frac{\pi a^4}{8}$.

16. Changing to polar coordinates, evaluate $\iint\sqrt{\frac{1-x^2-y^2}{1+x^2+y^2}}dx\, dy$ over the positive quadrant of the circle $x^2+y^2=1$.

Hint: $I=\int_0^{\frac{\pi}{2}}\int_0^{1}\sqrt{\frac{1-r^2}{1+r^2}}r\, dr\, d\theta = \int_0^{\frac{\pi}{2}}\int_0^{1}\frac{1-r^2}{\sqrt{1-r^4}}r\, dr\, d\theta$.

Ans. $\frac{\pi}{4}\left(\frac{\pi}{2}-1\right)$.

17. Using the transformation $x+y=u$ and $y=uv$, show that $\iint[xy(1-x-y)]^{\frac{1}{2}}dx\, dy$, taken over the area of triangle bounded by the lines $x=0$, $y=0$, and $x+y=1$, is $\frac{2\pi}{105}$.

Hint: $x=u-y=u-uv$ and $y=uv$. Therefore, Jacobian $J=u$. So, $dx\, dy=u\, du\, dv$. Further, $0 \le u \le 1$ and $0 \le v \le 1$. Thus,

$$I=\int_0^{1}\int_0^{1} u(1-u)^{\frac{1}{2}}\, v^{\frac{1}{2}}(1-v)^{\frac{1}{2}}u\, du\, dv$$

$$= \int_0^1 u^2(1-u)^{\frac{1}{2}}\,du.\int_0^1 v^{\frac{1}{2}}(1-v)^{\frac{1}{2}}dv$$

$$= \int_0^1 u^{3-1}(1-u)^{\frac{3}{2}-1}du.\int_0^1 v^{\frac{3}{2}-1}(1-v)^{\frac{3}{2}-1}dv$$

$$= \beta\left(3,\frac{3}{2}\right)\,\beta\left(\frac{3}{2},\frac{3}{2}\right), \text{ convert to gamma function.}$$

18. Evaluate $\iint \sqrt{a^2 - x^2 - y^2}\; dx\, dy$ over the semi-circle $x^2 + y^2 = ax$ in the positive quadrant.

    Hint: $I = \int_0^{\frac{\pi}{2}} \int_0^{a\cos\theta} \sqrt{a^2 - r^2}\; r\, dr\, d\theta.$

    Ans. $\frac{a^3}{3}\left(\frac{\pi}{2} - \frac{2}{3}\right).$

19. Evaluate $\iint \sqrt{\dfrac{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}{1 + \frac{x^2}{a^2} + \frac{y^2}{b^2}}}\; dx\, dy$ over the positive quadrant of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$

    Hint: Another from of Example 6.20.

20. Change $\int_0^{4a} \int_{\frac{y^2}{4a}}^{y} \frac{x^2 - y^2}{x^2 + y^2}\, dx\, dy$ in polar coordinates and hence, evaluate the same.

    Ans. $\int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \int_0^{\frac{4a\cos\theta}{\sin^2\theta}} r\; \cos 2\theta\; d\theta\; dr = 8a^2\left(\frac{\pi}{2} - \frac{5}{3}\right).$

21. Changing to polar coordinates, evaluate $\iint xy\left(x^2 + y^2\right)^{\frac{3}{2}} dx\, dy$ over the positive quadrant of the circle $x^2 + y^2 = 1$.

    Ans. $\frac{1}{14}.$

Change of Order of Integration

22. Change the order of integration in

    $$\int_0^1 \int_x^{x(2-x)} f(x,y)dx\; dy.$$

    Ans. $\int_0^1 \int_{1-\sqrt{1-y}}^{y} f(x,y)dy\; dx.$

23. Changing the order of integration, evaluate the integral $\int_0^a \int_{\frac{x^2}{a}}^{2a-x} xy\; dy\; dx.$

    Ans. $\frac{5a^4}{6}.$

24. Changing the order of integration, evaluate $\int_0^a \int_y^a \frac{x\; dx\; dy}{x^2 + y^2}.$

    Ans. $\frac{\pi a}{4}.$

25. Change the order of integration

    $\int_0^a \int_x^{\frac{a^2}{x}} (x + y)\; dx\; dy.$ and hence evaluate it.

    Ans. $\infty.$

26. Evaluate the integral $\int_0^b \int_0^{\frac{a}{b}\sqrt{b^2-y^2}} xy\; dy\; dx.$

    Ans. $\frac{a^2b^2}{8}.$

27. Changing the order of integration, evaluate the integral $\int_0^{\frac{a}{\sqrt{2}}} \int_y^{\sqrt{a^2-y^2}} \log\left(x^2 + y^2\right)dx\; dy,\; a > 0.$

    Ans. $\frac{\pi a^4}{4}\left(\log a - \frac{1}{2}\right).$

Area Enclosed by Curves

28. Find the area bounded by the parabola $y = x^2$ and the line $y = 2x + 3$.

    Ans. $\frac{32}{3}.$

29. Find the area of the region bounded by the lines, $x = -2$ and $x = 2$and the circle, $x^2 + y^2 = 9$.

    Ans. $4\sqrt{5} + 18\sin^{-1}\frac{2}{3}$ sq. units.

30. Find the area of the cardioid $r = a(1 - \cos\theta)$.

    Ans. Area $= 2\int_0^{\pi} \int_0^{a(1-\cos\theta)} r\; dr\; d\theta = \frac{3\pi a^2}{2}$ sq. units.

31. Find the area outside the circle $r = a$ and inside the cardioid. $r = a(1 + \cos\theta)$.

    Ans. $\frac{\pi a^2}{2}.$

32. Find by double integration, the area lying inside the cardioid $r = a(1 + \cos\theta)$ and outside the parabola $r(1+ \cos\theta).= 1$.
Hint: Eliminating r between the two equations, we get $\cos^2\theta+2\cos\theta=0$, which implies $\theta = \pm\frac{\pi}{2}$.
Then Area $= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{\frac{1}{1+\cos\theta}}^{1+\cos\theta} r\,dr\,d\theta = \frac{9\pi+16}{12}$.

33. Find, using double integration, the smaller of the areas bounded by the circle $x^2 + y^2 = 9$ and the line $x + y = 3$.
Ans. $\frac{9}{4}(\pi - 2)$.

Volume and Surface Areas as Double Integrals

34. Find the volume of the solid region under the surface $z = 3 - x^2 - 2y^2$ for $x^2 + y^2 \le 1$.
Ans. $\frac{9\pi}{4}$.

35. Using double integration, find the volume of the tetrahedron bounded by the coordinate planes and the plane $\frac{x}{a}+\frac{y}{b}+\frac{z}{c} = 1$.
Ans. $\int_0^a \left[\int_0^{b\left(1-\frac{x}{a}\right)} c\left(1 - \frac{x}{a} - \frac{y}{b}\right)dy\right] dx = \frac{abc}{6}$.

36. Find the volume of the region bounded by the surfaces $y = x^2$ and $x = y^2$ and the planes $z = 0$ and $z = 3$.
Hint: $V = \int_0^1 \int_{x^2}^{\sqrt{x}} 3\,dy\,dx = 1$.

37. Calculate the volume of the solid bounded by the surfaces $x = 0$, $y = 0$, $z = 0$, and $x + y + z = 1$.
Ans. $\frac{1}{6}$.

38. Find the volume of the cylinder $x^2 + y^2 - 2ax = 0$ intercepted between the paraboloid $x^2 + y^2 = 2az$ and the xy-plane.
Ans. $3\pi a^3$.

39. Find the volume bounded by the xy-plane, the paraboloid $2z = x^2 + y^2 = 2ax$, and the cylinder $x^2 + y^2 = 4$.
Ans. 4π.

40. Find the volume common to the surface $y^2 + z^2 = 4ax$ and $x^2 + y^2 = 2ax$, the axis being rectangular.
Hint:
$$V = \int_0^{2a} \int_{-\sqrt{2ax-x^2}}^{\sqrt{2ax-x^2}} \sqrt{4ax - y^2}\,dy\,dx$$
$$= 2\int_0^{2a} \int_0^{\sqrt{2ax-x^2}} \sqrt{4ax - y^2}\,dy\,dx.$$
Ans. $\frac{2}{3}(3\pi + 8)a^3$.

41. Find the volume of the sphere $x^2 + y^2 + z^2 = 9$.
Ans. 36π.

42. Find the area of the surface $z^2 = 2xy$ included between $x = 0$, $x = a$, $y = 0$, and $y = b$.
Ans. $\frac{2\sqrt{2}}{3}\sqrt{ab}(a + b)$.

43. Find the area of the portion of the sphere $x^2 + y^2 + z^2 = 9$ lying inside the cylinder $x^2 + y^2 = 3y$.
Hint: $z^2 = 9 - x^2 - y^2$. Then $1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 = \frac{9}{9 - x^2 - y^2}$. Change to polar coordinates. Surface area
$$= 4\int_0^{\frac{\pi}{2}} \int_0^{3\sin\theta} \frac{3}{\sqrt{9 - r^2}}\,dr\,d\theta = 18\pi - 36.$$

44. Find the area of the portion of the cylinder $x^2 + y^2 = 4y$ lying inside the sphere $x^2 + y^2 + z^2 = 16$.
Ans. 64 sq. units.

45. Using double integration, find the volume generated by the revolution of the cardioid $r = a(1 - \cos\theta)$ about its axis.
Hint: Volume of revolution
$$= 2\pi\int_0^{\pi} \int_0^{a(1-\cos\theta)} r^2 \sin\theta\,dr\,d\theta = \frac{8}{3}\pi a^3.$$

46. Find the volume generated by revolving the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2} = 1$ about the y-axis.
Ans. $\frac{4}{3}\pi a^2 b$.

Evaluation of Triple Integral

47. Evaluate $\int_0^1 \int_0^{1-x} \int_0^{1-x-y} xyz\,dx\,dy\,dz.$

Ans. $\dfrac{1}{720}.$

48. Evaluate $\iiint z^2 dx\,dy\,dz$ over the sphere. $x^2 + y^2 + z^2 = 1.$

Ans. $\dfrac{4\pi}{15}.$

49. Evaluate $\int_1^3 \int_{\frac{1}{x}}^1 \int_0^{\sqrt{xy}} xyz\,dz\,dy\,dx.$

Ans. $\dfrac{1}{3}\left(\dfrac{13}{3} - \dfrac{1}{2}\log 3\right).$

50. Evaluate $\int_0^{\frac{\pi}{2}} \int_0^{a\cos\theta} \int_0^{\sqrt{a^2-r^2}} r\,dz\,dr\,d\theta.$

Ans. $a^3\left(\dfrac{\pi}{6} - \dfrac{\pi}{9}\right).$

51. Evaluate $\iiint_V dx\,dy\,dz$, where $V = \{(x, y, z);$ $x^2 + y^2 + z^2 \le a^2, 0 \le z \le h\}$, using cylindrical polar coordinates.

Ans. $\int_0^{\frac{\pi}{2}} \int_0^{2\pi} \int_0^{h} r\,dz\,d\theta\,dr = \pi a^2 h.$

Volume as a Triple Integral

52. Find the volume bounded by the surface $x^2 + y^2 = a^2$ and $\dfrac{x^2}{p} + \dfrac{y^2}{q} = 2z,\ p > 0,\ q > 0.$

Ans. $\dfrac{\pi a^4}{8}\left(\dfrac{1}{p} + \dfrac{1}{q}\right).$

53. Find the volume of the paraboloid of revolution $x^2 + y^2 = 4z$ cut off by the plane $z = 4$.

Hint: $V = 4\int_0^4 \int_0^{\sqrt{16-x^2}} \int_{\frac{x^2+y^2}{4}}^{4} dx\,dy\,dz = 32\pi.$

54. Find the volume bounded above by the sphere $x^2 + y^2 + z^2 = 2a^2$ and below by the paraboloid $az = x^2 + y^2$.

Ans. $\pi\,a^3\left(\dfrac{4\sqrt{2}}{3} - \dfrac{7}{6}\right).$

55. Show that the volume enclosed by the cylinder $x^2 + y^2 = 2ax$ and $z^2 = 2ax$ is $\dfrac{128a^3}{15}$.

56. Show that the volume of the wedge intercepted between the cylinder $x^2 + y^2 = 2ax$ and the planes $z = x$ and $z = 2x$ is $\pi\,a^3$.
Hint: See Example 6.76.

Vector Calculus

UNIT IV

7 Vector Calculus

We know that *scalar* is a quantity that is characterized solely by magnitude whereas *vector* is a quantity which is characterized by both magnitude and direction. For example, time, mass, and temperature are scalar quantities whereas displacement, velocity, and force are vector quantities. We represent a vector by an arrow over it. Geometrically, we represent a vector $\vec{a}$ by a directed line segment $\vec{PQ}$, where $\vec{a}$ has direction from P to Q. The point P is called the *initial point* and the point Q is called the *terminal point* of $\vec{a}$. The length $\left|\vec{PQ}\right|$ of this line segment is the *magnitude* of $\vec{a}$. Two vectors $\vec{a}$ and $\vec{b}$ are said to be *equal* if they have the same magnitude and direction. The product of a vector $\vec{a}$ and a scalar m is a vector m $\vec{a}$ with magnitude $|m|$ times the magnitude of $\vec{a}$ with direction, the same or opposite to that of $\vec{a}$, according as $m > 0$ or $m < 0$. In particular, if $m = 0$, then $m\,\vec{a}$ is a null vector $\vec{0}$. A vector with unit magnitude is called a *unit vector*. If $\vec{a}$ is non-zero vector, then $\frac{\vec{a}}{|\vec{a}|} = \frac{\vec{a}}{a}$ is a unit vector having the same direction as that of $\vec{a}$ and is denoted by $\hat{a}$.

If $\vec{a}, \vec{b}$ and $\vec{c}$ are vectors and m and n are scalars (real or complex), then addition and scalar multiplication of vectors satisfy the following properties:

(i) $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ (Commutative law for addition).

(ii) $\vec{a} + \left(\vec{b} + \vec{c}\right) = \left(\vec{a} + \vec{b}\right) + \vec{c}$ (Associative law for addition).

(iii) $\vec{m}\left(\vec{a} + \vec{b}\right) = m\vec{a} + m\vec{b}$ (Distributive law for addition).

(iv) $(m + n)\vec{a} = m\vec{a} + n\vec{a}$ (Distributive law for scalars).

(v) $\vec{a} + \vec{0} = \vec{a} = \vec{0} + \vec{a}$ (Existence of identity for addition).

(vi) $\vec{a} + (-\vec{a}) = \vec{0} = (-\vec{a}) + \vec{a}$ (Existence of inverse for addition).

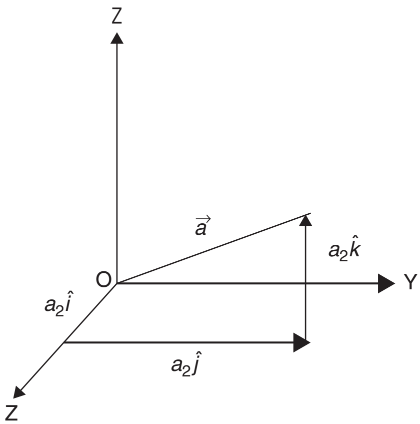
(vii) $|m\vec{a}| = |m|\;|\vec{a}|$.

(viii) $m(n\vec{a}) = (mn)\vec{a}$.

(ix) $n(m\vec{a}) = m(n\vec{a})$.

The unit vectors in the directions of positive x-, y-, and z-axes of a three-dimensional, rectangular coordinate system are called the *rectangular unit vectors* and are denoted, respectively, by $\hat{i}$, $\hat{j}$, and $\hat{k}$.

Let a_1, a_2, and a_3 be the rectangular coordinates of the terminal point of vector $\vec{a}$ with the initial point at the origin O of a rectangular coordinate system in three dimensions. Then, the vectors $a_1\hat{i}$, $a_2\hat{j}$, and $a_3\hat{k}$ are called *rectangular component vectors* or simply *component vectors* of $\vec{a}$ in the x, y, and z directions, respectively.



The resultant (sum) of $a_1\hat{i}$, $a_2\hat{j}$, and $a_3\hat{k}$ is the vector $\vec{a}$ and so,

$$\vec{a} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k}.$$

Further, the magnitude of $\vec{a}$ is

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

In particular, the *radius vector* or *position vector* $\vec{r}$ from O to the point (x, y, z) in a three-dimension space is expressed as

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$$

and

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

The *scalar product* or *dot product* or *inner product* of two vectors $\vec{a}$ and $\vec{b}$ is a *scalar* defined by

$$\vec{a}.\vec{b} = |\vec{a}|\,\left|\vec{b}\right| \cos\theta,$$

where θ is the angle between the vectors $\vec{a}$ and $\vec{b}$ and $0 \le \theta \le \pi$.

The scalar product satisfies the following properties:

(i) $\vec{a}.\vec{b} = \vec{b}.\vec{a}$.

(ii) $\vec{a}.\left(\vec{b}+\vec{c}\right) = \vec{a}.\vec{b} + \vec{a}.\vec{c}$.

(iii) $t\left(\vec{a}.\vec{b}\right) = \left(t\,\vec{a}.\vec{b}\right) = \vec{a}.\left(t\,\vec{b}\right) = \left(\vec{a}.\vec{b}\right)t$, where t is a scalar.

(iv) $\hat{i}.\hat{i} = \hat{j}.\hat{j} = \hat{k}.\hat{k} = 1$ and $\hat{i}.\hat{j} = \hat{j}.\hat{k} = \hat{k}.\hat{i} = 0$.

(v) If $\vec{a} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k}$, then $\vec{a}.\vec{a} = |\vec{a}|^2 = a^2 = a_1^2 + a_2^2 + a_3^2$.

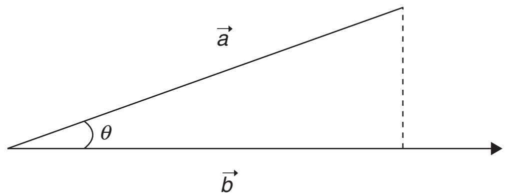
(vi) If $\vec{a} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k}$ and $\vec{b} = b_1\hat{i} + b_2\hat{j} + b_3\hat{k}$, then

$$\vec{a}.\vec{b} = a_1b_2 + a_2b_2 + a_3b_3.$$

(vii) If $\vec{a}.\vec{b} = 0$ and $\vec{a}$ and $\vec{b}$ are nonzero vectors, then $\cos\theta = 0$ and so, $\theta = \frac{\pi}{2}$. Hence, $\vec{a}$ and $\vec{b}$ are perpendicular.

(viii) The *projection* of a vector $\vec{a}$ on a vector $\vec{b}$ is a vector defined by "projection of $\vec{a}$ on $\vec{b} = (a\cos\theta)\vec{e}_b = (\vec{a}.\vec{e}_b)\vec{e}_b$", where θ is the angle between $\vec{a}$ and $\vec{b}$, and $\vec{e}_b$ is a unit vector in the direction of the vector $\vec{b}$.

Let a vector $\vec{a}$ makes angles α, β, and γ, respectively, with positive directions of x, y, and z. Then, the numbers $\cos\alpha$, $\cos\beta$ and $\cos\gamma$ are called the *directions cosines* of $\vec{a}$. Thus, $\cos\alpha = \hat{i}.\vec{e}_a$, $\cos\beta = \hat{j}.\vec{e}_a$, and $\cos\gamma = \hat{k}.\vec{e}_a$, where $\vec{e}_a$ is a unit vector in the direction of $\vec{a}$.



The *vector-* or *cross product* of two vectors $\vec{a}$ and $\vec{b}$ is a vector defined by $\vec{a}\times\vec{b} = |\vec{a}|\left|\vec{b}\right| \sin\theta\,\hat{e} = ab\sin\theta\,\hat{e}$, where θ is the angle between the vectors $\vec{a}$ and $\vec{b}$ such that $0 \le \theta \le \pi$ and $\hat{e}$ is a unit vector perpendicular to both $\vec{a}$ and $\vec{b}$. The direction of $\vec{a}\times\vec{b}$ is perpendicular to the plane of A and B, such that $\vec{a}$, $\vec{b}$, and $\vec{a}\times\vec{b}$ form a right-handed triad of vectors.

In particular, if $\vec{a} = \vec{b}$ or $\vec{a}$ is parallel to $\vec{b}$, then $\vec{a}\times\vec{b} = \vec{0}$.

If $\vec{a} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k}$ and $\vec{b} = b_1\hat{i} + b_2\hat{j} + b_3\hat{k}$, then

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

where $\hat{i}\times\hat{i} = \hat{j}\times\hat{j} = \hat{k}\times\hat{k} = 0$;

$$\hat{i}\times\hat{j} = \hat{k},\ \hat{j}\times\hat{k} = \hat{i},\ \text{and}\ \hat{k}\times\hat{i} = j;\ \text{and}$$

$$\hat{j}\times\hat{i} = -\hat{k},\ \hat{k}\times\hat{j} = -\hat{i},\ \text{and}\ \hat{i}\times\hat{k} = -\hat{j}.$$

The magnitude $\left|\vec{a}\times\vec{b}\right|$ of $\vec{a}\times\vec{b}$ is equal in the area of the parallelogram with sides $\vec{a}$ and $\vec{b}$.

The vector product satisfies the following properties:

(i) $\vec{a}\times\vec{b} = -\vec{b}\times\vec{a}$ (Anti-commutative law).

(ii) $\vec{a}\times\left(\vec{b}+\vec{c}\right) = \vec{a}\times\vec{b} + \vec{a}\times\vec{c}$ (Distributive law over addition).

(iii) $t\left(\vec{a}\times\vec{b}\right) = (t\vec{a})\times\vec{b} = \vec{a}\times\left(t\vec{b}\right) = \left(\vec{a}\times\vec{b}\right)t$, t is a scalar.

The dot- and cross multiplication of three vectors $\vec{a}, \vec{b}$, and $\vec{c}$ follow the following laws:

(i) $\left(\vec{a} \cdot \vec{b}\right)\vec{c} \neq \vec{a}\left(\vec{b} \cdot \vec{c}\right)$.

(ii) $\vec{a} \cdot \left(\vec{b} \times \vec{c}\right) = \vec{b} \cdot (\vec{c} \times \vec{a}) = \vec{c} \cdot \left(\vec{a} \times \vec{b}\right)$.

If $\vec{a} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k}$, $\vec{b} = b_1\hat{i} + b_2\hat{j} + b_3\hat{k}$, and $\vec{c} = c_1\hat{i} + c_2\hat{j} + c_3\hat{k}$, then

$$\vec{a} \cdot \left(\vec{b} \times \vec{c}\right) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

(iii) $\vec{a} \times \left(\vec{b} \times \vec{c}\right) \neq \left(\vec{a} \times \vec{b}\right) \times \vec{c}$.

(iv) $\vec{a} \times \left(\vec{b} \times \vec{c}\right) = (\vec{a} \cdot \vec{c})\vec{b} - \left(\vec{a} \cdot \vec{b}\right)\vec{c}$,

$$\left(\vec{a} \times \vec{b}\right) \times \vec{c} = (\vec{a} \cdot \vec{c})\vec{b} - \left(\vec{b} \cdot \vec{c}\right)\vec{a}.$$

The product $\vec{a}.\left(\vec{b} \times \vec{c}\right)$ is called the *scalar triple product* or *box product* and is denoted by $[abc]$.

The product $\vec{a} \times \left(\vec{b} \times \vec{c}\right)$ is called the *vector triple product*.

7.1 DIFFERENTIATION OF A VECTOR

A vector $\vec{r}$ is said to be a *vector function* of a scalar variable t if to each value of t there corresponds a value of $\vec{r}$.

A vector function is denoted by $\vec{r} = \vec{r}(t)$ or $\vec{r} = \vec{f}(t)$. For example, the position vector $\vec{r}$ of a particle moving along a curved path is a vector function of time t. In rectangular coordinate system, the vector function $\vec{f}$ can be expressed in a component form as

$$\vec{r} = \vec{f}(t) = f_1\hat{i} + f_2\hat{j} + f_3\hat{k},$$

where f_1, f_2, and f_3 are scalar functions of t and are called *components* of $\vec{f}$.

Let $\vec{r} = \vec{f}(t)$ be a vector function of the scalar variable t. If Δt denotes a small increment in t and $\Delta\vec{r}$ the corresponding increment in $\vec{r}$, then

$$\frac{d\vec{r}}{dt} = \lim_{\Delta t \to 0} \frac{\Delta\vec{r}}{\Delta t} = \lim_{\Delta t \to 0} \frac{\vec{f}(t + \Delta t) - \vec{f}(t)}{\Delta t},$$

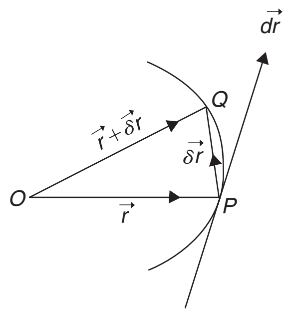
if exists, is called the *ordinary derivative* of $\vec{r}$ with respect to the scalar t.

Since $\frac{d\vec{r}}{dt}$ is itself a vector depending on t, we can further consider its derivative with respect to t. If this derivative exists, it is denoted by $\frac{d^2\vec{r}}{dt^2}$. Similarly, higher derivatives of $\vec{r}$ can be defined.

Geometric Significance of $\frac{d\vec{r}}{dt}$: Let $\vec{r} = \vec{f}(t)$ be the vector equation of a curve C in space. Let P and Q be two neighboring points on C with position vectors $\vec{r}$ and $\vec{r} + \delta\vec{r}$. Then, $OP = \vec{r}$, $OQ = \vec{r} + \delta\vec{r}$ and so,

$$\vec{PQ} = \vec{OQ} - \vec{OP} = \vec{r} + \delta\vec{r} - \vec{r} = \delta\vec{r}.$$

Therefore, $\frac{\delta\vec{r}}{\delta t}$ is directed along the chord PQ. As $\delta t \to 0$, that is, as $Q \to P$, the chord PQ tends to the tangent to the curve C at P. Hence, $\frac{d\vec{r}}{dt} = \lim_{\delta t \Rightarrow 0} \frac{\delta\vec{r}}{\delta t}$ is a vector along the tangent to the curve at P.



Unit Tangent Vector to a Curve: Suppose that we take an arc length s from any point, say A, on the curve C, up to the point P as the parameter, instead of t. Then, $AP = s$, $AQ = s + \delta s$, and so, $PQ = \delta s$. In this case, $\frac{d\vec{r}}{ds}$ will be a vector along the tangent at P. Further,

$$\left|\frac{d\vec{r}}{ds}\right| = \lim_{\delta s \to 0} \left|\frac{\delta\vec{r}}{\delta s}\right| = \lim_{Q \to P} \frac{\text{chord } PQ}{\text{Arc } PQ} = 1.$$

Hence, $\frac{d\vec{r}}{ds}$ is the unit vector $\hat{t}$ along the tangent at P.

Theorem 7.1. *If $\vec{a}, \vec{b}$, and $\vec{c}$ are differentiable vector functions of a scalar t and ϕ is a differentiable scalar function of t, then*

(i) $\frac{d}{dt}\left(\vec{a} \pm \vec{b}\right) = \frac{d\vec{a}}{dt} + \frac{d\vec{b}}{dt}$.

(ii) $\frac{d}{dt}\left(\vec{a} \cdot \vec{b}\right) = \vec{a} \cdot \frac{d\vec{b}}{dt} + \frac{d\vec{a}}{dt} \cdot \vec{b}$.

(iii) $\frac{d}{dt}\left(\vec{a} \times \vec{b}\right) = \vec{a} \times \frac{d\vec{b}}{dt} + \frac{d\vec{a}}{dt} \times \vec{b}$.

(iv) $\frac{d}{dt}(\phi\vec{a}) = \phi\frac{d\vec{a}}{dt} + \frac{d\phi}{dt}\vec{a}$.

(v) $\frac{d}{dt}\left[\vec{a}\cdot\left(\vec{b}\times\vec{c}\right)\right]=\vec{a}\cdot\left(\vec{b}\times\frac{d\vec{c}}{dt}\right)+\vec{a}\cdot\left(\frac{d\vec{b}}{dt}\times\vec{c}\right)$
$+\frac{d\vec{a}}{dt}\cdot\left(\vec{b}\times\vec{c}\right).$

(vi) $\frac{d}{dt}\left[\vec{a}\times\left(\vec{b}\times\vec{c}\right)\right]=\vec{a}\times\left(\vec{b}\times\frac{d\vec{c}}{dt}\right)$
$+\vec{a}\times\left(\frac{d\vec{b}}{dt}\times\vec{c}\right)+\frac{d\vec{a}}{dt}\times\left(\vec{b}\times\vec{c}\right).$

Proof: We prove (i), (iii), and (v). The other parts may similarly be proved by the readers themselves.

(i) $\frac{d}{dt}\left(\vec{a}+\vec{b}\right)$

$$=\lim_{\Delta t\to 0}\frac{\left[(\vec{a}+\Delta\vec{a})+\left(\vec{b}+\Delta\vec{b}\right)\right]-\left(\vec{a}+\vec{b}\right)}{\Delta t}$$

$$=\lim_{\Delta t\to 0}\frac{\Delta\vec{a}+\Delta\vec{b}}{\Delta t}=\lim_{\Delta t\to 0}\frac{\Delta\vec{a}}{\Delta t}+\lim_{\Delta t\to 0}\frac{\Delta\vec{b}}{\Delta t}$$

$$=\frac{d\vec{a}}{dt}+\frac{d\vec{b}}{dt}.$$

(iii) $\frac{d}{dt}\left(\vec{a}\times\vec{b}\right)$

$$=\lim_{\Delta t\to 0}\frac{\left[(\vec{a}+\Delta\vec{a})\times\left(\vec{b}+\Delta\vec{b}\right)\right]-\vec{a}\times\vec{b}}{\Delta t}$$

$$=\lim_{\Delta t\to 0}\frac{\vec{a}\times\vec{b}+\vec{a}\times\Delta\vec{b}+\Delta\vec{a}\times\vec{b}+\Delta\vec{a}\times\Delta\vec{b}-\vec{a}\times\vec{b}}{\Delta t}$$

$$=\lim_{\Delta t\to 0}\frac{\vec{a}\times\Delta\vec{b}+\Delta\vec{a}\times\vec{b}+\Delta\vec{a}\times\Delta\vec{b}}{\Delta t}$$

$$=\lim_{\Delta t\to 0}\left[\vec{a}\times\frac{\Delta\vec{b}}{\Delta t}+\frac{\Delta\vec{a}}{\Delta t}\times\vec{b}+\frac{\Delta\vec{a}}{\Delta t}\times\Delta\vec{b}\right]$$

$$=\vec{a}\times\frac{d\vec{b}}{dt}+\frac{d\vec{a}}{dt}\times\vec{b}+\frac{d\vec{a}}{dt}\times\vec{0}$$

$$=\vec{a}\times\frac{d\vec{b}}{dt}+\frac{d\vec{a}}{dt}\times\vec{b}+\vec{0}=\vec{a}\times\frac{d\vec{b}}{dt}+\frac{d\vec{a}}{dt}\times\vec{b}.$$

(v) Using (ii) and (iii), we have

$$\frac{d}{dt}\left[\vec{a}\cdot\left(\vec{b}\times\vec{c}\right)\right]$$

$$=\vec{a}\cdot\frac{d}{dt}\left(\vec{b}\times\vec{c}\right)+\frac{d\vec{a}}{dt}\cdot\left(\vec{b}\times\vec{c}\right)$$

$$=\vec{a}\cdot\left[\vec{b}\times\frac{d\vec{c}}{dt}+\frac{d\vec{b}}{dt}\times\vec{c}\right]+\frac{d\vec{a}}{dt}\cdot\left(\vec{b}\times\vec{c}\right)$$

$$=\vec{a}\cdot\left[\vec{b}\times\frac{d\vec{c}}{dt}\right]+\vec{a}\cdot\left[\frac{d\vec{b}}{dt}\times\vec{c}\right]+\frac{d\vec{a}}{dt}\cdot\left(\vec{b}\times\vec{c}\right).$$

Theorem 7.2. The derivative of a constant vector is the zero vector.

Proof: We know that a vector is called constant if its magnitude and direction do not change. Let $\vec{c}$ be a constant vector and let $\vec{r}=\vec{c}$. Then, $\vec{r}+\delta\vec{r}=\vec{c}$ and so, $\delta\vec{r}=\vec{0}$. Thus, $\frac{\delta\vec{r}}{\delta t}=\frac{\vec{0}}{\delta t}=\vec{0}$ and hence, $\frac{d\vec{r}}{dt}=\lim_{\delta t\to 0}\frac{\delta\vec{r}}{\delta t}=\lim_{\delta t\to 0}\vec{0}=\vec{0}$.

Theorem 7.3. A vector function $\vec{f}$ of a scalar variable t is constant if and only if $\frac{d\vec{f}}{dt}=\vec{0}$.

Proof: If $\vec{f}$ is a constant vector, then, by Theorem 7.2, $\frac{d\vec{f}}{dt}=\vec{0}$.

Conversely, suppose that $\frac{d\vec{f}}{dt}=\vec{0}$. If f_1, f_2, and f_3 are the components of $\vec{f}$ along x-, y-, and z-axes, then $\vec{f}=f_1\hat{i}+f_2\hat{j}+f_3\hat{k}$. Hence,

$$\vec{0}=\frac{d\vec{f}}{dt}=\frac{df_1}{dt}\hat{i}+\frac{df_2}{dt}\hat{j}+\frac{df_3}{dt}\hat{k}.$$

Therefore, equality of two vectors implies

$$\frac{df_1}{dt}=\frac{df_2}{dt}=\frac{df_3}{dt}=0.$$

Therefore, f_1, f_2, and f_3 are constant scalars, independent of t. Hence, $\vec{f}$ is a constant vector function.

Theorem 7.4. A vector function $\vec{f}$ of a scalar variable t has a constant magnitude if and only if $\vec{f}\cdot\frac{d\vec{f}}{dt}=0$.

Proof: First, suppose that the vector function $\vec{f}$ has a constant magnitude c. Then,

$$\vec{f}\cdot\vec{f}=\left|\vec{f}\right|^2=c^2$$

and so,

$$\frac{d}{dt}\left(\vec{f}\cdot\vec{f}\right)=\frac{d}{dt}(c^2)=0.$$

But,

$$\frac{d}{dt}\left(\vec{f}\cdot\vec{f}\right)=\vec{f}\cdot\frac{d\vec{f}}{dt}+\frac{d\vec{f}}{dt}\cdot\vec{f}=2\vec{f}\cdot\frac{d\vec{f}}{dt}.$$

Hence,

$$2\vec{f}\cdot\frac{d\vec{f}}{dt}=0 \text{ or } \vec{f}\cdot\frac{d\vec{f}}{dt}=0.$$

Conversely, suppose that $\vec{f}.\frac{d\vec{f}}{dt}=0$. Therefore, $2\vec{f}\cdot\frac{d\vec{f}}{dt}=0$ or $\vec{f}\cdot\frac{d\vec{f}}{dt}+\frac{d\vec{f}}{dt}\cdot\vec{f}=0$ *or* $\frac{d}{dt}\left(\vec{f}\cdot\vec{f}\right)=0$, which implies $\vec{f}\cdot\vec{f}$ is constant $=c^2$, say.

Hence, $\left|\vec{f}\right|^2 = c^2$ or $\left|\vec{f}\right| = c$, that is, $\vec{f}$ has a constant magnitude.

Theorem 7.5. The necessary and sufficient condition for a vector function $\vec{f}$ of a scalar variable t to have a constant direction is that $\vec{f} \times \frac{d\vec{f}}{dt} = \vec{0}$.

Proof: Let $\vec{F}$ be a vector function of modulus unity for all t. Let $\left|\vec{f}\right| = f$. Then, $\vec{f} = f\vec{F}$.

The condition is necessary: Suppose that $\vec{f}$ has a constant direction. Since $\vec{f} = f\vec{F}$, it follows that $\vec{f}$ and $\vec{F}$ have the same direction. Thus, $\vec{F}$ has a constant magnitude, equal to unity and a constant direction too and so, is a constant vector. Therefore, $\frac{d\vec{F}}{dt} = 0$. Differentiating $\vec{f} = f\vec{F}$ with respect to t, we have

$$\frac{d\vec{f}}{dt} = \frac{df}{dt}\vec{F} + f\frac{d\vec{F}}{dt}.$$

Now,

$$\begin{aligned}
\vec{f} \times \frac{d\vec{f}}{dt} &= (f\vec{F}) \times \left[\frac{df}{dt}\vec{F} + f\frac{d\vec{F}}{dt}\right] \\
&= f\frac{df}{dt}\vec{F} \times \vec{F} + f^2\vec{F} \times \frac{d\vec{F}}{dt} \\
&= \vec{0} + f^2\vec{F} \times \frac{d\vec{F}}{dt}, \text{ since } \vec{F} \times \vec{F} = \vec{0} \\
&= f^2\vec{F} \times \frac{d\vec{F}}{dt} \\
&= f^2\vec{F} \times \vec{0} = \vec{0}, \text{ since } \frac{d\vec{F}}{df} = \vec{0} \text{ (as shown earlier).}
\end{aligned}$$

The condition is sufficient: Suppose that $\vec{f} \times \frac{d\vec{f}}{dt} = \vec{0}$. Therefore, as shown previously, $f^2\vec{F} \times \frac{d\vec{F}}{dt} = \vec{0}$ and so $\vec{F} \times \frac{d\vec{F}}{dt} = \vec{0}$. Also, since $\vec{F}$ is of constant magnitude, $\vec{F} \cdot \frac{d\vec{F}}{dt} = \vec{0}$. These two facts imply that $\frac{d\vec{F}}{dt} = \vec{0}$. Therefore, $\vec{F}$ is a constant vector. But magnitude of F is constant (unity). Therefore, $\vec{\mathrm{F}}$ has a constant direction. But $\vec{f} = f\vec{F}$. Therefore, direction of $\vec{\mathrm{f}}$ is also constant.

Corollary 7.1: The derivative of a vector function of a scalar variable t having a constant direction is collinear with it.

Proof: Since $\vec{f}$ has a constant direction, $\vec{f} \times \frac{d\vec{f}}{dt} = 0$ and so, $\vec{f}$ and $\frac{d\vec{f}}{dt}$ are collinear. This completes the proof of the corollary.

From Theorems 7.3–7.5, we conclude that

(i) $\frac{d\vec{f}}{dt} = \vec{0}$ if and only if $\vec{f}$ is a constant vector function in both magnitude and direction

(ii) $\vec{f} \cdot \frac{d\vec{f}}{dt} = 0$ if and only if $\vec{f}$ has a constant magnitude.

(iii) $\vec{f} \times \frac{d\vec{f}}{dt} = 0$ if and only if $\vec{f}$ has a constant direction.

Theorem 7.6. If $\vec{f} = f_1\hat{i} + f_2\hat{j} + f_3\hat{k}$ is a vector function of the scalar variable t, then

$$\frac{d\vec{f}}{dt} = f_1'(t)\hat{i} + f_2'(t)\hat{j} + f_3'(t)\hat{k}.$$

Proof: We have

$$\vec{f} = f_1\hat{i} + f_2\hat{j} + f_3\hat{k},$$

where f_1, f_2, and f_3 are scalar functions of t. Therefore,

$$\begin{aligned}
\frac{d\vec{f}}{dt} &= \frac{d}{dt}(f_1\hat{i}) + \frac{d}{dt}(f_2\hat{j}) + \frac{d}{dt}\left(f_3\hat{k}\right) \\
&= f_1\frac{d\hat{i}}{dt} + \frac{df_1}{dt}\hat{i} + f_2\frac{d\hat{j}}{dt} + \frac{df_2}{dt}\hat{j} + f_3\frac{d\hat{k}}{dt} + \frac{df_3}{dt}\hat{k} \\
&= \vec{0} + \frac{df_1}{dt}\hat{i} + \vec{0} + \frac{df_2}{dt}\hat{j} + \vec{0} + \frac{df_3}{dt}\hat{k} \\
&= \frac{df_1}{dt}\hat{i} + \frac{df_2}{dt}\hat{j} + \frac{df_3}{dt}\hat{k}.
\end{aligned}$$

Thus, to differentiate a vector, it is sufficient to differentiate its components.

Velocity and Acceleration: Let $\vec{r}$ be the position vector of a moving particle P, and let $\delta\vec{r}$ be the displacement of the particle in time δt, where t denotes time. Then, the vector $\frac{\delta\vec{r}}{\delta t}$ denotes the average velocity of the particle during the interval δt of time. Therefore, the velocity vector $\vec{v}$ of the particle at P is given by

$$\vec{v} = \lim_{\delta t \to 0}\frac{\delta\vec{r}}{\delta t} = \frac{d\vec{r}}{dt},$$

and its direction is along the tangent at P. Further, if $\delta\vec{v}$ is the change in velocity $\vec{v}$ during the time interval δt, then the rate of change of velocity, that is, $\frac{\delta\vec{v}}{\delta t}$ is the average acceleration of the particle during the interval δt. Thus, the acceleration of the particle at P is

$$\vec{a} = \lim_{\delta t \to 0}\frac{\delta\vec{v}}{\delta t} = \frac{d\vec{v}}{dt} = \frac{d}{dt}\left(\frac{d\vec{r}}{dt}\right) = \frac{d^2\vec{r}}{dt^2}.$$

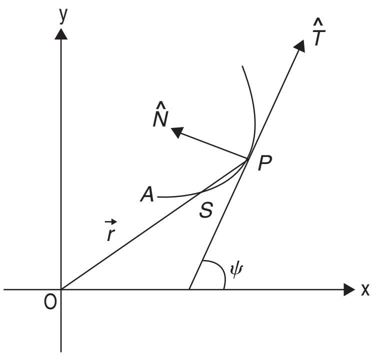
Tangential and Normal Acceleration: Let $\vec{r}$ be the position vector of a point P moving in a plane curve at any time t. Then the velocity of the moving point is given by

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{d\vec{r}}{ds}\cdot\frac{ds}{dt}.$$

But $\frac{d\vec{r}}{ds} = \hat{T}$ is a unit vector along the tangent at P. Therefore, $\vec{v} = \frac{ds}{dt}\hat{T}$. Thus, the acceleration is

$$\frac{d\vec{v}}{dt} = \frac{d}{dt}\left(\frac{ds}{dt}\hat{T}\right) = \frac{d^2s}{dt^2}\hat{T} + \frac{ds}{dt}\cdot\frac{d\hat{T}}{dt}. \qquad (1)$$

But,



$$\frac{d\hat{T}}{dt} = \frac{d\hat{T}}{d\psi}\cdot\frac{d\psi}{dt} = \frac{d\psi}{dt}\hat{N},$$

where $\hat{N}$ is a unit vector along the normal at P. Therefore,

$$\frac{d\hat{T}}{dt} = \frac{d\psi}{ds}\cdot\frac{ds}{dt}\hat{N} = \frac{1}{\rho}\frac{ds}{dt}\hat{N},$$

where ρ is the radius of curvature at P. Hence, (1) reduces to

$$\frac{d\vec{v}}{dt} = \frac{d^2s}{dt^2}\hat{T} + \frac{1}{\rho}\left(\frac{ds}{dt}\right)^2\hat{N} = \frac{dv}{dt}\hat{T} + \frac{v^2}{\rho}\hat{N}.$$

Therefore,

$$\text{Tangential acceleration} = \frac{dv}{dt} = \frac{d^2s}{dt^2}$$

and

$$\text{Normal acceleration} = \frac{v^2}{\rho}.$$

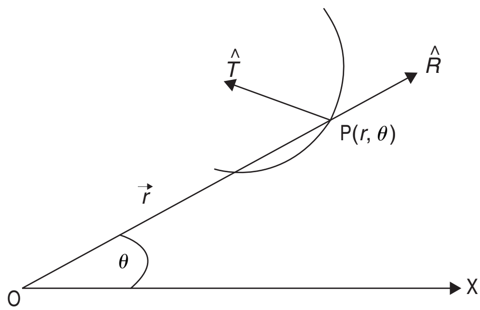
Radial and Transverse Acceleration of a Moving Particle: Let $\vec{r}$ be the position vector of a moving particle $P\,(r, \theta)$. Suppose that $\hat{R}$ and $\hat{T}$ are the unit vectors in radial- and transverse directions, respectively. Then, $\hat{r} = r\hat{R}$ and

$$\text{velocity } \vec{v} = \frac{d\vec{r}}{dt} = \frac{d}{dt}\left(r\hat{R}\right) = \frac{dr}{dt}\hat{R} + r\frac{d\hat{R}}{dt}$$

$$= \frac{dr}{dt}\hat{R} + r\frac{d\hat{R}}{d\theta}\cdot\frac{d\theta}{dt} = \frac{dr}{dt}\hat{R} + r\frac{d\theta}{dt}\hat{T}.$$

Therefore, the components of the velocity in the radial- and transverse directions are

$$v_R = \frac{dr}{dt} \text{ and } v_T = r\frac{d\theta}{dt}.$$



Further, since $\frac{d\hat{R}}{d\theta} = \hat{T}$ and $\frac{d\hat{T}}{d\theta} = -\hat{R}$, we have

acceleration $\vec{a}$

$$= \frac{d\vec{v}}{dt} = \frac{d}{dt}\left(\frac{dr}{dt}\hat{R} + r\frac{d\theta}{dt}\hat{T}\right)$$

$$= \frac{d^2r}{dt^2}\hat{R} + \frac{dr}{dt}\cdot\frac{d\hat{R}}{dt} + \frac{dr}{dt}\cdot\frac{d\theta}{dt}\hat{T} + r\frac{d^2\theta}{dt^2}\vec{T} + r\frac{d\theta}{dt}\cdot\frac{d\hat{T}}{dt}$$

$$= \frac{d^2r}{dt^2}\hat{R} + \frac{dr}{dt}\frac{d\hat{R}}{d\theta}\cdot\frac{d\theta}{dt} + \left(\frac{dr}{dt}\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}\right)\hat{T}$$

$$+ r\frac{d\theta}{dt}\frac{d\hat{T}}{d\theta}\cdot\frac{d\theta}{dt}$$

$$= \frac{d^2r}{dt^2}\hat{R} + \frac{dr}{dt}\cdot\frac{d\theta}{dt}\hat{T} + \left(\frac{dr}{dt}\cdot\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}\right)\hat{T}$$

$$- r\left(\frac{d\theta}{dt}\right)^2\hat{R}$$

$$= \left[\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right]\hat{R} + \left[2\frac{dr}{dt}\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}\right]\hat{T}.$$

Hence,

$$\text{Radial acceleration} = a_R = \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2$$

and

$$\text{Transverse acceleration} = a_T = 2\frac{dr}{dt}\cdot\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}.$$

EXAMPLE 7.1

If $\vec{a} = \sin\theta\,\hat{i} + \cos\theta\,\hat{j} + \theta\hat{k}$, $\vec{b} = \cos\theta\hat{i} - \sin\theta\hat{j} - 3\hat{k}$, and $\vec{c} = 2\vec{i} + 3\vec{j} - \vec{k}$, find $\frac{d}{d\theta}\left(\vec{a}\times(\vec{b}\times\vec{c})\right)$ at $\theta = 0$.

Solution. We are given that
$\vec{a} = \sin\theta\,\hat{i} + \cos\theta\,\hat{j} + \theta\hat{k}$,
$\vec{b} = \cos\theta\hat{i} - \sin\theta\hat{j} - 3\hat{k}$, and $\vec{c} = 2\hat{i} + 3\hat{j} - \hat{k}$.
Therefore,

$$\vec{b}\times\vec{c} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \cos\theta & -\sin\theta & -3 \\ 2 & 3 & -1 \end{vmatrix}$$

$$= (\sin\theta + 9)\hat{i} - (-\cos\theta + 6)\hat{j} + (3\cos\theta + 2\sin\theta)\hat{k}.$$

Then,

$$\vec{a}\times\left(\vec{b}\times\vec{c}\right) = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \sin\theta & \cos\theta & \theta \\ 9+\sin\theta & -\cos\theta + 6 & 3\cos\theta + 2\sin\theta \end{vmatrix}$$

$$= \left(3\cos^2\theta + \sin 2\theta - \theta\cos\theta + 6\theta\right)\hat{i} - \left(\frac{3}{2}\sin 2\theta + 2\sin^2\theta - 9\theta - \theta\sin\theta\right)\hat{j} + (-6\sin\theta - 9\cos\theta)\hat{k}.$$

Therefore,

$$\frac{d}{d\theta}\left[\vec{a}\times\left(\vec{b}\times\vec{c}\right)\right] = (-6\cos\theta\sin\theta + 2\cos 2\theta - \cos\theta + \theta\sin\theta + 6)\hat{i} - (3\cos 2\theta + 4\sin\theta\cos\theta - 9 - \theta\cos\theta - \sin\theta)\hat{j} + (-6\cos\theta + 9\sin\theta)\hat{k}.$$

Putting $\theta = 0$, we get

$$\frac{d}{d\theta}\left[\vec{a}\times\left(\vec{b}\times\vec{c}\right)\right] = (2-1+6)\hat{i} - (3-9)\hat{j} - 6\hat{k} = 7\hat{i} + 6\hat{j} - 6\hat{k}.$$

EXAMPLE 7.2

If $\vec{r} = (\cos nt)\hat{i} + (\sin nt)\hat{j}$, where n is a constant and t varies, show that $\vec{r}\times\frac{d\vec{r}}{dt} = n\hat{k}$.

Solution. We have

$$\vec{r} = (\cos nt)\hat{i} + (\sin nt)\hat{j}.$$

Therefore,

$$\frac{d\vec{r}}{dr} = (-n\sin nt)\hat{i} + (n\cos nt)\hat{j}.$$

Therefore,

$$\vec{r}\times\frac{d\vec{r}}{dt} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \cos nt & \sin nt & 0 \\ -n\sin nt & n\cos nt & 0 \end{vmatrix} = \hat{k}\left(n\cos^2 nt + n\sin^2 nt\right) = n\hat{k}.$$

EXAMPLE 7.3

If $\vec{a}$ and $\vec{b}$ are constant vectors, ω is a constant scalar, and $\vec{r} = \vec{a}\sin\omega t + \vec{b}\cos\omega t$, show that
(i) $\frac{d^2\vec{r}}{dt^2} + \omega^2\vec{r} = \vec{0}$ and (ii) $\vec{r}\times\frac{d\vec{r}}{dt} = -\omega\,\vec{a}\times\vec{b}$.

Solution. (i) Since $\vec{a}$ and $\vec{b}$ are constant vectors, we have

$$\frac{d\vec{a}}{dt} = \vec{0} \text{ and } \frac{d\vec{b}}{dt} = \vec{0}. \qquad (1)$$

Now it is given that $\vec{r} = \vec{a}\sin\omega t + \vec{b}\cos\omega t$. Therefore,

$$\frac{d\vec{r}}{dt} = \sin\omega t\frac{d\vec{a}}{dt} + \frac{d}{dt}(\sin\omega t)\vec{a} + \cos\omega t\frac{d\vec{b}}{dt} + \frac{d}{dt}(\cos\omega t)\vec{b}$$

$$= 0 + \frac{d}{dt}(\sin\omega t)\vec{a} + 0 + \frac{d}{dt}(\cos\omega t)\vec{b}, \text{ using (1).}$$

$$= (\omega\cos\omega t)\vec{a} - (\omega\sin\omega t)\vec{b}$$

and

$$\frac{d^2\vec{r}}{dt^2} = \left(-\omega^2\sin\omega t\right)\vec{a} - \left(\omega^2\cos\omega t\right)\vec{b} = -\omega^2\left(\vec{a}\sin\omega t + \vec{b}\cos\omega t\right) = -\omega^2\vec{r}.$$

Hence,

$$\frac{d^2\vec{r}}{dt^2} + \omega^2\vec{r} = \vec{0}.$$

(ii) Since $\vec{a}\times\vec{a} = \vec{0}$, $\vec{b}\times\vec{b} = 0$, and $\vec{a}\times\vec{b} = -\vec{b}\times\vec{a}$, we have

$$\begin{aligned}\vec{r}\times\frac{d\vec{r}}{dt} &= \left[(\cos\,\omega t)\vec{a} + (\sin\,\omega t)\vec{b}\right]\\ &\quad\times\left[(-\omega\sin\,\omega t)\vec{a} + (\omega\cos\,\omega t)\vec{b}\right]\\ &= \left(\omega\,\cos^2\omega t\right)\left(\vec{a}\times\vec{b}\right) - \left(\omega\,\sin^2\omega t\right)\left(\vec{b}\times\vec{a}\right)\\ &= \left(\omega\,\cos^2\omega t\right)\left(\vec{a}\times\vec{b}\right) + \left(\omega\,\sin^2\omega t\right)\left(\vec{a}\times\vec{b}\right)\\ &= \left[\omega\left(\cos^2\omega t + \sin^2\omega t\right)\right]\left(\vec{a}\times\vec{b}\right)\\ &= \omega\left(\vec{a}\times\vec{b}\right).\end{aligned}$$

EXAMPLE 7.4

Show that

$$\frac{d}{dt}\left(\vec{a}\times\frac{d\vec{b}}{dt} - \frac{d\vec{a}}{dt}\times\vec{b}\right) = \vec{a}\times\frac{d^2\vec{b}}{dt^2} - \frac{d^2\vec{a}}{dt^2}\times\vec{b}.$$

Solution. We have

$$\begin{aligned}&\frac{d}{dt}\left(\vec{a}\times\frac{d\vec{b}}{dt} - \frac{d\vec{a}}{dt}\times\vec{b}\right)\\ &= \frac{d\vec{a}}{dt}\times\frac{d\vec{b}}{dt} + \vec{a}\times\frac{d^2\vec{b}}{dt^2} - \frac{d^2\vec{a}}{dt^2}\times\vec{b} - \frac{d\vec{a}}{dt}\times\frac{d\vec{b}}{dt}\\ &= \vec{a}\times\frac{d^2\vec{b}}{dt^2} - \frac{d^2\vec{a}}{dt^2}\times\vec{b},\end{aligned}$$

which proves our assertion.

EXAMPLE 7.5

Let $\vec{r} = t^2\hat{i} - 3t\hat{j} + (2t+1)\hat{k}$, find at $t = 0$, the value of $\left|\frac{d^2\vec{r}}{dt^2}\right|$.

Solution. Let $\vec{r} = t^2\hat{i} - 3t\hat{j} + (2t+1)\hat{k}$. Then,

$$\begin{aligned}\frac{d\vec{r}}{dt} &= \frac{d}{dt}\left(t^2\right)\hat{i} - \frac{d}{dt}(3t)\hat{j} + \frac{d}{dt}(2t+1)\hat{k}\\ &= 2t\hat{i} - 3\hat{j} + 2\hat{k} \quad \text{and}\end{aligned}$$

$$\frac{d^2\vec{r}}{dt^2} = \frac{d}{dt}(2t)\hat{i} - \frac{d}{dt}(3)\hat{j} + \frac{d}{dt}(2)\hat{k} = 2\hat{i}.$$

When $t = 0$, we have $\frac{d^2\vec{r}}{dt^2} = 2\hat{i}$. Further,

$$\left|\frac{d^2\vec{r}}{dt^2}\right| = \sqrt{2^2 + 0^2 + 0^2} = 2.$$

EXAMPLE 7.6

If $\vec{a} = 5t^2\hat{i} + t\hat{j} - t^2\hat{k}$ and $\vec{b} = \sin\,t\hat{i} - \cos\,t\hat{j}$, find $\frac{d}{dt}\left(\vec{a}.\vec{b}\right)$ and $\frac{d}{dt}\left(\vec{a}\times\vec{b}\right)$.

Solution. Let $\vec{a} = 5t^2\hat{i} + t\hat{j} - t^2\hat{k}$ and $\vec{b} = \sin\,t\hat{i} - \cos\,t\hat{j}$. Then,

$$\vec{a}.\,\vec{b} = 5t^2\sin t - t\cos\,t.$$

Therefore,

$$\begin{aligned}\frac{d}{dt}\left(\vec{a}.\vec{b}\right) &= \frac{d}{dt} = \left(5t^2\sin\,t - t\cos\,t\right)\\ &= 5t^2\cos\,t + 10t\,\sin\,t + t\,\sin\,t - \cos\,t\\ &= \left(5t^2 - 1\right)\cos t + 11\,t\sin t.\end{aligned}$$

Also,

$$\begin{aligned}\vec{a}\times\vec{b} &= \begin{vmatrix}\hat{i} & \hat{j} & \hat{k}\\ 5t^2 & t & -t\\ \sin t & -\cos\,t & 0\end{vmatrix}\\ &= \left(-t^2\cos t\right)\hat{i} - \left(t^3\sin\,t\right)\hat{j}\\ &\quad + \left(-5t^2\cos t - t\,\sin\,t\right)\hat{k}.\end{aligned}$$

Therefore,

$$\begin{aligned}&\frac{d}{dt}\left(\vec{a}\times\vec{b}\right)\\ &= \frac{d}{dt}\left(-t^3\cos t\right)\hat{i} - \frac{d}{dt}\left(t^3\sin\,t\right)\hat{j}\\ &\quad + \frac{d}{dt}\left(-5t^2\cos\,t - t\sin\,t\right)\hat{k}\\ &= \left(t^3\sin\,t - 3t^2\cos\,t\right)\hat{i} - \left(t^3\cos\,t + 3t^2\sin\,t\right)\hat{j}\\ &\quad + \left(5t^2\sin\,t - 11t\cos\,t - \sin t\right)\hat{k}.\end{aligned}$$

EXAMPLE 7.7

Find a unit tangent vector to any point on the curve $x = a\,\cos\,\omega t$, $y = a\,\sin\,\omega t$, and $z = bt$, where a, b, and ω are constants.

Solution. Let $\vec{r}$ be the position vector of any point (x, y, z) on the given curve. Then,

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k} = (a\,\cos\,\omega t)\hat{i} + (a\,\sin\,\omega t)\hat{j} + (bt)\hat{k}.$$

Therefore,

$$\frac{d\vec{r}}{dt} = (-a\omega \sin \omega t)\hat{i} + (a\omega \cos \omega t)\hat{j} + b\hat{k}.$$

The vector $\frac{d\vec{r}}{dt}$ is along the tangent at the point (x, y, z) to the given curve. Hence, unit tangent vector is given by

$$\vec{T} = \frac{\frac{d\vec{r}}{dt}}{\left|\frac{d\vec{r}}{dt}\right|} = \frac{(-a\omega \sin \omega t)\hat{i} + (a\omega \cos \omega t)\hat{j} + b\hat{k}}{\sqrt{a^2\omega^2 + b^2}}.$$

EXAMPLE 7.8

A particle moves along the curve $x = 3t^2$, $y = t^2 - 2t$ and $z = t^3$. Find its velocity and acceleration at $t = 1$ in the direction of $\hat{i} + \hat{j} - \hat{k}$.

Solution. Let $\vec{r}$ be the position vector of any point (x, y, z) on the given curve. Then,

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k} = 3t^2\hat{i} + \left(t^2 - 2t\right)\hat{j} + t^3\hat{k}$$

and so, the velocity and acceleration of the particle are, respectively,

$$\vec{v} = \frac{d\vec{r}}{dt} = 6t\hat{i} + (2t - 2)\hat{j} + 3t^2\hat{k} = 6\hat{i} + 3\hat{k} \text{ at } t = 1$$

and

$$\vec{a} = \frac{d^2\vec{r}}{dt^2} = 6\hat{i} + 2\hat{j} + 6t\hat{k} = 6\hat{i} + 2\hat{j} + 6\hat{k} \text{ at } t = 1.$$

The unit vector in the direction of $\hat{i} + \hat{j} - \hat{k}$ is

$$\hat{n} = \frac{\hat{i} + \hat{j} - \hat{k}}{|\hat{i} + \hat{j} - \hat{k}|} = \frac{\hat{i} + \hat{j} - \hat{k}}{\sqrt{3}}.$$

Therefore, the components of velocity and acceleration in the direction of $\hat{i} + \hat{j} - \hat{k}$ are

$$\vec{v}.\hat{n} = \left(6\hat{i} + 3\hat{k}\right) \cdot \frac{\left(\hat{i} + \hat{j} - \hat{k}\right)}{\sqrt{3}} = \frac{3}{\sqrt{3}} = \sqrt{3}$$

and

$$\vec{a}.\hat{n} = \left(6\hat{i} + 2\hat{j} + 3\hat{k}\right) \cdot \frac{\left(\hat{i} + \hat{j} - \hat{k}\right)}{\sqrt{3}} = \frac{2}{\sqrt{3}} = \frac{2\sqrt{3}}{3}.$$

EXAMPLE 7.9

Find the angle between the tangents to the curve $\vec{r} = t^2\hat{i} + 2t\hat{j} - t^3\hat{k}$ at the points $t = \pm 1$.

Solution. We have

$$\vec{r} = t^2\hat{i} + 2t\hat{j} - t^3\hat{k}.$$

Therefore, the vector along the tangent at any point is

$$\frac{d\vec{r}}{dt} = 2t\hat{i} + 2\hat{j} - 3t^2\hat{k}.$$

Thus, the vectors along the tangents at $t = \pm 1$ are

$$\vec{T}_1 = 2\hat{i} + 2\hat{j} - 3\hat{k} \text{ and } \vec{T}_2 = -2\hat{i} + 2\hat{j} - 3\hat{k}.$$

The angle θ between the tangents is given by

$$\cos\theta = \frac{\vec{T}_1.\vec{T}_2}{|\vec{T}_1||\vec{T}_2|} = \frac{2(-2) + 2(2) - 3(-3)}{\sqrt{4+4-9}.\sqrt{4+4+9}} = \frac{9}{17}.$$

Hence,

$$\theta = \cos^{-1}\left(\frac{9}{17}\right).$$

EXAMPLE 7.10

A particle moves along the curve

$$\vec{r} = \left(t^3 - 4t\right)\hat{i} + \left(t^2 + 4t\right)\hat{j} + \left(8t^2 - 3t^3\right)\hat{k}.$$

where t denotes time. Find the magnitude of acceleration along the tangent and normal at time $t = 2$.

Solution. The curve is

$$\vec{r} = \left(t^3 - 4t\right)\hat{i} + \left(t^2 + 4t\right)\hat{j} + \left(8t^2 - 3t^3\right)\hat{k}.$$

Therefore,

$$\begin{aligned}\vec{v} &= \frac{d\vec{r}}{dt} = \left(3t^2 - 4\right)\hat{i} + (2t + 4)\hat{j} + \left(16t - 9t^2\right)\hat{k} \\ &= 8\hat{i} + 8\hat{j} - 4\hat{k} \text{ at } t = 2\end{aligned}$$

and

$$\begin{aligned}\text{acceleration } \vec{a} &= \frac{d^2\vec{r}}{dt^2} = 6t\hat{i} + 2\hat{j} + (16 - 18t)\hat{k} \\ &= 12\hat{i} + 2\hat{j} - 20\hat{k} \text{ at } t = 2.\end{aligned}$$

The velocity is along the tangent to the curve. Therefore,
Component of $\vec{a}$ along the tangent

$$\begin{aligned}&= \vec{a}.\frac{\vec{v}}{|\vec{v}|} \\ &= \left(12\hat{i} + 2\hat{j} - 20\hat{k}\right)\frac{8\hat{i} + 8\hat{j} - 4\hat{k}}{|8\hat{i} + 8\hat{j} - 4\hat{k}|} \\ &= \left(12\hat{i} + 2\hat{j} - 20\hat{k}\right)\frac{8\hat{i} + 8\hat{j} - 4\hat{k}}{\sqrt{64 + 64 + 16}} \\ &= \frac{96 + 16 + 80}{12} = 16\end{aligned}$$

and

Component of $\vec{a}$ along the normal

$= |\ \vec{a} -$ resolved part of $\vec{a}$ along the tangent $|$

$$= |12\hat{i} + 2\hat{j} - 20\hat{k} - 16\frac{8\hat{i} + 8\hat{j} - 4\hat{k}}{\sqrt{64+64+16}}$$

$$= \frac{1}{3}|\ 4\hat{i} - 26\hat{j} - 44\hat{k}\ |$$

$$= \frac{1}{3}[\sqrt{16+676+1936}\ ] = 2\sqrt{73}.$$

7.2 PARTIAL DERIVATIVES OF A VECTOR FUNCTION

Let $\vec{f}$ be a vector function of x, y, and z. Let $\delta\vec{f}$ be the change in $\vec{f}$ corresponding to a small change δx in x. Then, $\delta\vec{f} = \vec{f}(x+\delta x,\ y,\ z) - \vec{f}(x,y,z)$. The limit

$$\lim_{\delta x \to 0} \frac{\delta\vec{f}}{\delta x} = \lim_{\delta x} \frac{\vec{f}(x+\delta x,\ y,\ z) - \vec{f}(x,y,z)}{\delta x},$$

if it exists, is called the *partial derivative of the vector function* $\vec{f}$ *with respect to* x and is denoted by $\frac{\partial\vec{f}}{\partial x}$ or $\vec{f}_x$.

Similarly, the partial derivatives of $\vec{f}$ with respect to y and z are defined by

$$\vec{f}_y = \lim_{\delta y \to 0} \frac{\vec{f}(x, y+\delta y,\ z) - \vec{f}(x,y,z)}{\delta y}$$

and

$$\vec{f}_z = \lim_{\delta z \to 0} \frac{\vec{f}(x,y,z+\delta z) - \vec{f}(x,y,z)}{\delta z},$$

provided these limits exist.

If $\vec{f}$ and $\vec{g}$ are differentiable vector functions of the independent variables x, y, and z and ϕ is a differentiable scalar function of x, y, and z, then

(i) $\frac{\partial}{\partial x}\left(\vec{f} + \vec{g}\right) = \frac{\partial\vec{f}}{\partial x} + \frac{\partial\vec{g}}{\partial x}$.

(ii) $\frac{\partial}{\partial x}\left(\phi\vec{f}\right) = \phi\frac{\partial\vec{f}}{\partial x} + \frac{\partial\phi}{\partial x}\vec{f}$

(iii) $\frac{\partial}{\partial x}\left(\vec{f}.\vec{g}\right) = \vec{f}.\frac{\partial\vec{g}}{\partial x} + \frac{\partial\vec{f}}{\partial x}.\vec{g}$.

(iv) $\frac{\partial}{\partial x}\left(\vec{f} \times \vec{g}\right) = \vec{f} \times \frac{\partial\vec{g}}{\partial x} + \frac{\partial\vec{f}}{\partial x} \times \vec{g}$.

Similar expressions for partial derivatives with respect to y and z are valid. Higher partial derivatives of $\vec{f}$ may also be defined in the same way. For example,

$$\vec{f}_{xx} = \frac{\partial^2\vec{f}}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial\vec{f}}{\partial x}\right).$$

EXAMPLE 7.11

If $\vec{f} = xyz\hat{i} + xz^2\hat{j} - y^3\hat{k}$, find $\frac{\partial^2 f}{\partial x\partial y}$ at the origin.

Solution. We have

$$\vec{f} = xyz\hat{i} + xz^2\hat{j} - y^3\hat{k}.$$

Therefore,

$$\frac{\partial\vec{f}}{\partial y} = xz\hat{i} + \vec{0} - 3y^2\hat{k} \text{ and}$$

$$\frac{\partial^2\vec{f}}{\partial x\partial y} = z\hat{i} = \vec{0} \text{ at } (0,0,0).$$

EXAMPLE 7.12

If $\vec{a} = x^2yz\hat{i} - 2xz^3\hat{j} + xz^2\hat{k}$ and $\vec{b} = 2z\hat{i} + y\hat{j} - x^2\hat{k}$, find the value of $\frac{\partial^2}{\partial x^2}\left(\vec{a} \times \vec{b}\right)$ at the point (1, 0, 1).

Solution. We have

$\vec{a} = x^2yz\hat{i} - 2xz^3\hat{j} + xz^2\hat{k}$ and $\vec{b} = 2z\hat{i} + y\hat{j} - x^2\hat{k}$.

Therefore,

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x^2yz & -2xz^3 & xz^2 \\ 2z & y & -x^2 \end{vmatrix}$$

$$= \left(2x^3z^3 - xyz^2\right)\hat{i} - \left(-x^4yz - 2xz^3\right)\hat{j} + \left(x^2y^2z + 4xz^4\right)\hat{k}.$$

Hence,

$$\frac{\partial}{\partial x}\left(\vec{a} \times \vec{b}\right) = \left(6x^2z^3 - yz^2\right)\hat{i} - \left(-4x^3yz - 2z^3\right)\hat{j} + \left(2xy^2z + 4z^4\right)\hat{k}$$

and

$$\frac{\partial^2}{\partial x^2}\left(\vec{a} \times \vec{b}\right) = \left(12xz^3\right)\hat{i} - \left(-12x^2yz\right)\hat{j} + \left(2y^2z\right)\hat{k}$$

$$= 12\hat{i} \text{ at } (1,0,1).$$

7.3 GRADIENT OF A SCALAR FIELD

A variable quantity whose value at any point in a region of space depends upon the position of the point is called a *point function*. If for each point $P(x, y, z)$ of a region R, there corresponds a scalar $\phi(x, y, z)$, then ϕ is called a *scalar-point function* for the region R. The region R is then called a *scalar field*. For example, the temperature at any point within or on the surface of the earth is a scalar-point function. Similarly, atmospheric pressure in the space is a scalar-point function. On the other hand, if for each point $P(x, y, z)$, of a region R, there exists a vector $\vec{f}(x,y,z)$, then $\vec{f}$ is called a *vector-point function* and the region R is then called a *vector field*. For example, the gravitational force is a vector-point function.

Let $f(x, y, z)$, be a scalar-point function. Then, the points satisfying an equation of the type $f(x, y, z) = c$ (constant) constitute a family of surface in a three-dimensional space. The surfaces of this family are called *level surfaces*. Since the value of the function f at any point of the surface is the same, these surfaces are also called *iso-f-surfaces*.

The operator ∇, defined by

$$\nabla = \hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z},$$

is called the *vector differential operator* and is read as *del* or *nabla*.

Let ϕ be a scalar function defined and differentiable at each point (x, y, z) in a certain region of space. Then, the vector defined by

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)\phi,$$

$$= \hat{i}\frac{\partial\phi}{\partial x} + \hat{j}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}$$

is called the *gradient of the scalar function* ϕ and is denoted by grad ϕ or $\nabla\phi$.

Thus, grad ϕ is a vector with components $\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}$, and $\frac{\partial\phi}{\partial z}$. We note that ϕ is a scalar-point function, whereas $\nabla\phi$ is a vector-point function.

7.4 GEOMETRICAL INTERPRETATION OF A GRADIENT

Let $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$ be the position vector of a point P through which a level surface $\phi(x, y, z) = c$ (constant) passes. Then, differentiating $\phi(x, y, z) = c$ with respect to t, we get

$$\frac{d\phi}{dt} = 0 \quad \text{or} \quad \frac{\partial\phi}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial\phi}{\partial y}\cdot\frac{dy}{dt} + \frac{\partial\phi}{\partial z}\cdot\frac{dz}{dt} = 0.$$

or

$$\left(\frac{\partial\phi}{\partial x}\hat{i} + \frac{\partial\phi}{\partial y}\hat{j} + \frac{\partial\phi}{\partial z}\hat{k}\right)\cdot\left(\frac{dx}{dt}\hat{i} + \frac{dy}{dt}\hat{j} + \frac{dz}{dt}\hat{k}\right) = 0$$

or

$$\nabla\phi.\frac{d\vec{r}}{dt} = 0.$$

Since $\frac{d\vec{r}}{dt}$ is the vector tangent to the curve at P and since P is an arbitrary point on $\phi(x, y, z) = c$, it follows that $\nabla\phi$ is perpendicular to $\phi(x, y, z) = c$ at every point. Hence, $\nabla\phi$ *is normal to the surface* $\phi(x, y, z) = c$.

7.5 PROPERTIES OF A GRADIENT

The following theorem illustrates the properties satisfied by a gradient.

Theorem 7.7. If ϕ and ψ are two scalar-point functions, and c is a constant, then,

(i) $\nabla(\phi \pm \psi) = \nabla\phi \pm \nabla\psi$.

(ii) $\nabla(\phi\psi) = \phi\nabla\psi + \psi\nabla\phi$.

(iii) $\nabla\left(\frac{\phi}{\psi}\right) = \frac{\psi\nabla\phi - \phi\nabla\psi}{\psi^2}$, provided that $\psi \neq 0$.

(iv) $\nabla(c\phi) = c\nabla\phi$.

(v) $\nabla\phi$ is a constant if and only if ϕ is a constant.

Proof: (i). By the definition of a gradient, we have

$$\begin{aligned}
&\nabla(\phi \pm \psi) \\
&= \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(\phi \pm \psi) \\
&= \hat{i}\frac{\partial}{\partial x}(\phi \pm \psi) + \hat{j}\frac{\partial}{\partial y}(\phi \pm \psi) + \hat{k}\frac{\partial}{\partial z}(\phi \pm \psi) \\
&= \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)\phi \pm \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)\psi \\
&= \nabla\phi \pm \nabla\psi.
\end{aligned}$$

(ii) $\nabla(\phi\psi) = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(\phi\psi)$

$$= \hat{i}\frac{\partial}{\partial x}(\phi\psi) + \hat{j}\frac{\partial}{\partial y}(\phi\psi) + \hat{k}\frac{\partial}{\partial z}(\phi\psi)$$

$$= \hat{i}\left[\phi\frac{\partial\psi}{\partial x} + \psi\frac{\partial\phi}{\partial x}\right] + \hat{j}\left[\phi\frac{\partial\psi}{\partial y} + \psi\frac{\partial\phi}{\partial y}\right] + \hat{k}\left[\phi\frac{\partial\psi}{\partial z} + \psi\frac{\partial\phi}{\partial z}\right]$$

$$= \phi\left[\hat{i}\frac{\partial\psi}{\partial x} + \hat{j}\frac{\partial\psi}{\partial y} + \hat{k}\frac{\partial\psi}{\partial z}\right] + \psi\left[\hat{i}\frac{\partial\phi}{\partial x} + \hat{j}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}\right]$$

$$= \phi\,\nabla\,\psi + \psi\,\nabla\,\phi.$$

(iii) $\nabla\left(\frac{\phi}{\psi}\right) = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)\left(\frac{\phi}{\psi}\right)$

$$= \hat{i}\frac{\partial}{\partial x}\left(\frac{\phi}{\psi}\right) + \hat{j}\frac{\partial}{\partial y}\left(\frac{\phi}{\psi}\right) + \hat{k}\frac{\partial}{\partial z}\left(\frac{\phi}{\psi}\right)$$

$$= \hat{i}\left[\frac{\psi\frac{\partial\phi}{\partial x} - \phi\frac{\partial\psi}{\partial x}}{\psi^2}\right] + \hat{j}\left[\frac{\psi\frac{\partial\phi}{\partial y} - \phi\frac{\partial\psi}{\partial y}}{\psi^2}\right] + \hat{k}\left[\frac{\psi\frac{\partial\phi}{\partial z} - \phi\frac{\partial\psi}{\partial z}}{\psi^2}\right]$$

$$= \psi\,\frac{\left[\hat{i}\frac{\partial\phi}{\partial x} + \hat{j}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}\right]}{\psi^2} - \phi\,\frac{\left[\hat{i}\frac{\partial\psi}{\partial x} + \hat{j}\frac{\partial\psi}{\partial y} + \hat{k}\frac{\partial\psi}{\partial z}\right]}{\psi^2}$$

$$= \frac{\psi\,\nabla\,\phi - \phi\,\nabla\,\psi}{\psi^2},\quad \psi \neq 0.$$

(iv) We have

$$\nabla(c\phi) = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(c\phi)$$

$$= \hat{i}\frac{\partial}{\partial x}(c\phi) + \hat{j}\frac{\partial}{\partial y}(c\phi) + \hat{k}\frac{\partial}{\partial z}(c\phi)$$

$$= c\hat{i}\frac{\partial}{\partial x}\phi + c\hat{j}\frac{\partial}{\partial y}\phi + c\hat{k}\frac{\partial}{\partial z}\phi$$

$$= c\left(\hat{i}\frac{\partial\phi}{\partial x} + \hat{j}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}\right) = c\,\nabla\,\phi.$$

(v) We note that

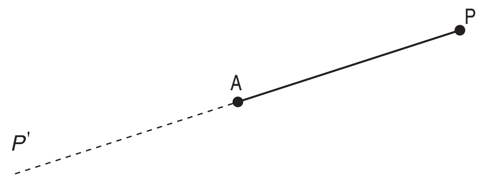
$$\nabla\,\phi = \vec{0} \Leftrightarrow \hat{i}\frac{\partial\phi}{\partial x} + \hat{j}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z} = \vec{0}$$

$$\Leftrightarrow \frac{\partial\phi}{\partial x} = 0,\ \frac{\partial\phi}{\partial y} = 0,\ \frac{\partial\phi}{\partial z} = 0$$

$\Leftrightarrow \phi$ is a constant.

7.6 DIRECTIONAL DERIVATIVES

Let A be any given point in the region of definition of a scalar-point function ϕ. Let P be a point on any line drawn on one side of A. Then $\lim_{P\to A}\frac{\phi(P)-\phi(A)}{AP}$, if exists, is called the *directional derivative of the scalar-point function* ϕ at A in the direction of AP. The length of AP is regarded as positive. The direction derivative in the direction of AP', where P' is a point on the other side of A, is negative of that in the direction of AP.



The *directional derivative of the vector function* $\vec{f}$ at A in the direction of AP is defined as $\lim_{P\to A}\frac{\vec{f}(P)-\vec{f}(A)}{AP}$, provided the limit exists.

7.6.1 Directional Derivatives Along Coordinate Axes

Let $A(x, y, z)$ be a point and let $P(x + \delta x, y, z)$ be a point on a line drawn through A and parallel to the positive direction of x-axis. Then, $AP = \delta x > 0$. Therefore, directional derivative of a scalar-point function at A along AP is defined as

$$\lim_{P\to A}\frac{\phi(P) - \phi(A)}{AP} = \lim_{\delta x\to 0}\frac{\phi(x + \delta x, y, z) - \phi(x, y, z)}{\delta x} = \frac{\partial\phi}{\partial x}.$$

Thus, the directional derivative of a scalar-point function ϕ along the x-axis is the *partial derivative of* ϕ *with respect to x.*

Similarly, directional derivatives of ϕ along y- and z-axis are, respectively, $\frac{\partial\phi}{\partial y}$ and $\frac{\partial\phi}{\partial z}$.

The directional derivatives of a vector-point function $\vec{f}$ along the coordinate axes are similarly $\frac{\partial\vec{f}}{\partial x}$, $\frac{\partial\vec{f}}{\partial y}$, and $\frac{\partial\vec{f}}{\partial z}$, respectively.

Further, if *l, m,* and *n,* are direction cosines of $AP = r$, then the coordinates of P are $x + lr, y + mr$, and $z + nr$ and so, the directional derivative of the scalar-point function ϕ along AP becomes

$$\lim_{P\to A}\frac{\phi(P)-\phi(A)}{AP}$$
$$=\lim_{r\to 0}\frac{\phi(x+lr,y+mr,z+nr)-\phi(x,y,z)}{r}$$
$$=\lim_{r\to 0}\frac{\phi(x,y,z)+\left(lr\frac{\partial\phi}{\partial x}+mr\frac{\partial\phi}{\partial y}+nr\frac{\partial\phi}{\partial z}\right)+\ldots-\phi(x,y,z)}{r}$$
$$=l\frac{\partial\phi}{\partial x}+m\frac{\partial\phi}{\partial y}+n\frac{\partial\phi}{\partial z},$$

by the application of Taylor's Theorem for function of several variables under the assumption that ϕ has a continuous first-order partial derivatives.

Similarly, the directional derivative of a vector-point function $\vec{f}$ along any line with direction cosines l, m, and n is $l\frac{\partial\vec{f}}{\partial x} + m\frac{\partial\vec{f}}{\partial y} + n\frac{\partial\vec{f}}{\partial z}$.

Theorem 7.8. The directional derivative of a scalar-point function ϕ along the direction of unit vector $\hat{a}$ is $\nabla\phi\cdot\hat{a}$.

Proof: The unit vector $\hat{a}$ along a line whose direction cosines are *l*, *m*, and *n* is

$$\hat{a} = l\hat{i} + m\hat{j} + n\hat{k}.$$

Therefore,

$$\nabla\phi.\hat{a} = \left(\hat{i}\frac{\partial\phi}{\partial x}+\hat{j}\frac{\partial\phi}{\partial y}+\hat{k}\frac{\partial\phi}{\partial z}\right)\left(l\hat{i}+m\hat{j}+n\hat{k}\right)$$
$$= l\frac{\partial\phi}{\partial x}+m\frac{\partial\phi}{\partial y}+n\frac{\partial\phi}{\partial z},$$

which is nothing but directional derivative of ϕ in the direction of the unit vector $\hat{a}$.

Theorem 7.9. Grad ϕ is a vector in the direction of which the maximum value of the directional derivative of ϕ occurs. Hence, the directional derivative is maximum along the normal to the surface and the maximum value is

$$|grad\ \phi| = |\nabla\phi|.$$

Proof: Recall that $\vec{a}.\vec{b} = |a||b|\cos\theta$, where θ is the angle between the vectors $\vec{a}$ and $\vec{b}$. Since (grad ϕ). $\hat{a}$ gives the directional derivative in the direction of unit vector $\hat{a}$, that is, the rate of change of $\phi(x, y, z)$ in the direction of the unit vector $\hat{a}$, it follows that the rate of change of $\phi(x, y, z)$ is zero along directions perpendicular to grad $\phi\left(\text{since } \cos\frac{\pi}{2} = 0\right)$ and is maximum along the direction parallel to grad ϕ. Since grad ϕ acts along the normal direction to the level surface of $\phi(x, y, z)$, the directional derivative is maximum along the normal to the surface. The maximum value is | grad ϕ | = $|\nabla\phi|$.

EXAMPLE 7.13

If $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$ and $|\vec{r}| = r$, show that
(i) $\nabla f(r) = f'(r)\nabla r$ and (ii) $\nabla f(r)\times\vec{r} = \vec{0}$.

Solution. (i) By the definition of gradient,

$$\nabla f(r) = \hat{i}\frac{\partial}{\partial x}f(r) + \hat{j}\frac{\partial}{\partial y}f(r) + \hat{k}\frac{\partial}{\partial z}f(r)$$
$$= \hat{i}f'(r)\frac{\partial r}{\partial x} + \hat{j}f'(r)\frac{\partial r}{\partial y} + \hat{k}f'(r)\frac{\partial r}{\partial z}$$
$$= f'(r)\left(\hat{i}\frac{\partial r}{\partial x} + \hat{j}\frac{\partial r}{\partial y} + \hat{k}\frac{\partial r}{\partial z}\right) = f'(r)\nabla r.$$

(ii) As in part (i), we have

$$\nabla f(r) = f'(r)\left(\hat{i}\frac{\partial r}{\partial x} + \hat{j}\frac{\partial r}{\partial y} + \hat{k}\frac{\partial r}{\partial z}\right).$$

Since $r = |\vec{r}| = \sqrt{x^2+y^2+z^2}$, we have

$$\frac{\partial r}{\partial x} = \frac{1}{2(x^2+y^2+z^2)^{\frac{1}{2}}}(2x) = \frac{x}{\sqrt{x^2+y^2+z^2}} = \frac{x}{r},$$

and similarly, $\frac{\partial r}{\partial y} = \frac{y}{r}$ and $\frac{\partial r}{\partial z} = \frac{z}{r}$. Therefore,

$$\nabla f(r) = f'(r)\left(\hat{i}\frac{x}{r} + \hat{j}\frac{y}{r} + \hat{k}\frac{z}{r}\right) = f'(r)\frac{\vec{r}}{r}.$$

Hence,

$$\nabla f(r) \times \vec{r} = \frac{f'(r)}{r}(\vec{r} \times \vec{r}) = \vec{0}.$$

EXAMPLE 7.14

If $f(\vec{r}) = x^2yz^2$, find ∇f at the point (1,2,3). Hence calculate

(i) the directional derivative of $f(\vec{r})$ at (1,2,3) in the direction of the vector (– 2,3, – 6).

(ii) the maximum rate of change of the function at (1,2,3) and its direction.

Solution. Since $\frac{\partial f}{\partial x} = 2xyz^2$, $\frac{\partial f}{\partial y} = x^2z^2$, and $\frac{\partial f}{\partial z} = 2x^2yz$, we have

$$\nabla f = 2xyz^2\hat{i} + x^2z^2\hat{j} + 2x^2yz\hat{k}.$$

Therefore, at the point (1,2,3),

$$\text{grad}\, f = \nabla f = 36\hat{i} + 9\hat{j} + 12\hat{k}.$$

(i) The unit vector $\hat{a}$ in the direction of the vector (– 2,3, –6) is

$$\frac{-2\hat{i} + 3\hat{j} - 6\hat{k}}{\sqrt{4+9+36}} = \frac{-2}{7}\hat{i} + \frac{3}{7}\hat{j} - \frac{6}{7}\hat{k}.$$

Therefore, the directional derivative at (1,2,3) in the direction of the vector (– 2,3, –6) is

$$\nabla f.\hat{a} = \left(36\hat{i} + 9\hat{j} + 12\hat{k}\right).\left(-\frac{2}{7}\hat{i} + \frac{3}{7}\hat{j} - \frac{6}{7}\hat{k}\right)$$

$$= -\frac{72}{7} + \frac{27}{7} - \frac{72}{7} = -\frac{117}{7}.$$

(ii) The maximum rate of change of the function at (1,2,3) occurs along the direction parallel to ∇f at (1,2,3), that is, parallel to $36\hat{i} + 9\hat{j} + 12\hat{k}$. The unit vector in that direction is $\frac{36\hat{i}+9\hat{j}+12\hat{k}}{\sqrt{1296+144+81}} = \frac{36\hat{i}+9\hat{j}+12\hat{k}}{39} = \frac{12\hat{i}+3\hat{j}+4\hat{k}}{13}$.

The maximum rate of change of $f(\vec{r})$ is $|\, grad\, f\,| = \sqrt{1296 + 144 + 81} = 39.$

EXAMPLE 7.15

If $\vec{r}$ is the usual position vector $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$ with $|\vec{r}| = r$, evaluate
(*i*) ∇r, (*ii*) $\nabla\left(\frac{1}{r}\right)$, (*iii*) ∇r^n, and (*iv*) $\nabla\left(\frac{1}{r^2}\right)$.

Solution. Since $|\vec{r}| = r = \sqrt{x^2 + y^2 + z^2}$, we have

$$\frac{\partial r}{\partial x} = \frac{x}{r},\ \frac{\partial r}{\partial y} = \frac{y}{r},\ \text{and}\ \frac{\partial r}{\partial z} = \frac{z}{r}.$$

Therefore,

(i) $\nabla(r) = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(r) = \hat{i}\frac{\partial r}{\partial x} + \hat{j}\frac{\partial r}{\partial y} + \hat{k}\frac{\partial r}{\partial z}$

$$= \hat{i}\frac{x}{r} + \hat{j}\frac{y}{r} + \hat{k}\frac{z}{r} = \frac{x\hat{i} + y\hat{j} + z\hat{k}}{r} = \frac{\vec{r}}{r}.$$

(ii) $\nabla\left(\frac{1}{r}\right) = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)\left(\frac{1}{r}\right)$

$$= \hat{i}\frac{\partial}{\partial x}\left(\frac{1}{r}\right) + \hat{j}\frac{\partial}{\partial y}\left(\frac{1}{r}\right) + \hat{k}\frac{\partial}{\partial z}\left(\frac{1}{r}\right)$$

$$= \hat{i}\left(-\frac{1}{r^2}\frac{\partial r}{\partial x}\right) + \hat{j}\left(-\frac{1}{r^2}\frac{\partial r}{\partial y}\right) + \hat{k}\left(-\frac{1}{r^2}\frac{\partial r}{\partial z}\right)$$

$$= \hat{i}\left(-\frac{1}{r^2}.\frac{x}{r}\right) + \hat{j}\left(-\frac{1}{r^2}.\frac{y}{r}\right) + \hat{k}\left(-\frac{1}{r^2}.\frac{z}{r}\right)$$

$$= -\frac{1}{r^3}\left(x\hat{i} + y\hat{j} + z\hat{k}\right) = -\frac{\vec{r}}{r^3}.$$

(iii) $\nabla r^n = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(r^n)$

$$= \hat{i}\frac{\partial}{\partial x}(r^n) + \hat{j}\frac{\partial}{\partial y}(r^n) + \hat{k}\frac{\partial}{\partial z}(r^n)$$

$$= \hat{i}\left(nr^{n-1}\frac{\partial r}{\partial x}\right) + \hat{j}\left(nr^{n-1}\frac{\partial r}{\partial y}\right)$$

$$+ \hat{k}\left(nr^{n-1}\frac{\partial r}{\partial z}\right)$$

$$= \hat{i}\left(nr^{n-1}.\frac{x}{r}\right) + \hat{j}\left(nr^{n-1}.\frac{y}{r}\right) + \hat{k}\left(nr^{n-1}.\frac{z}{r}\right)$$

$$= nr^{n-2}\left(x\hat{i} + y\hat{j} + z\hat{k}\right) = nr^{n-2}\vec{r}.$$

(iv) $$\nabla\left(\frac{1}{r^2}\right) = \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)\left(\frac{1}{r^2}\right)$$

$$= \hat{i}\frac{\partial}{\partial x}\left(\frac{1}{r^2}\right)+\hat{j}\frac{\partial}{\partial y}\left(\frac{1}{r^2}\right)+\hat{k}\frac{\partial}{\partial z}\left(\frac{1}{r^2}\right)$$

$$= \hat{i}\left(-\frac{2}{r^3}\frac{\partial r}{\partial x}\right)+\hat{j}\left(-\frac{2}{r^3}\frac{\partial r}{\partial y}\right)$$

$$+\hat{k}\left(-\frac{2}{r^3}\frac{\partial r}{\partial z}\right)$$

$$= \hat{i}\left(-\frac{2}{r^3}.\frac{x}{r}\right)+\hat{j}\left(-\frac{2}{r^3}.\frac{y}{r}\right)$$

$$+\hat{k}\left(-\frac{2}{r^3}.\frac{z}{r}\right)$$

$$= -\frac{2}{r^4}\left(x\hat{i}+y\hat{j}+z\hat{k}\right) = -\frac{2\vec{r}}{r^4}.$$

EXAMPLE 7.16

Find the directional derivative of $f(x, y, z) = xy^2 + yz^3$ at the point (2, – 1,1) in the direction of the vector $\hat{i}+2\hat{j}+2\hat{k}$.

Solution. We have

$$\nabla f = \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)(xy^2+yz^3)$$

$$= y^2\hat{i}+(2xy+z^3)\hat{j}+(3yz^2)\hat{k}$$

$$= \hat{i}-3\hat{j}-3\hat{k} \text{ at the point } (2,-1,1).$$

The unit vector in the direction of the vector $\hat{i}+2\hat{j}+2\hat{k}$ is

$$\hat{a} = \frac{\hat{i}+2\hat{j}+2\hat{k}}{\sqrt{1+4+4}} = \frac{1}{3}(\hat{i}+2\hat{j}+2\,k).$$

Therefore, the directional derivative of f at (2, – 1,1) in the direction of $\hat{i}+2\hat{j}+2\hat{k}$ is

$$\nabla\,\phi.\hat{a} = \left(\hat{i}-3\hat{j}-3\hat{k}\right).\frac{1}{3}\left(\hat{i}+2\hat{j}+2\hat{k}\right)$$

$$= \frac{1}{3}(1-6-6) = -\frac{11}{3}.$$

EXAMPLE 7.17

Find the directional derivative of $\phi\,(x, y, z) = xy^2 + yz^3$ at the point (2, – 1,1) in the direction of the normal to the surface $x log\, z - y^2 + 4 = 0$ at (2, – 1,1).

Solution. We have $\phi = xy^2+yz^3$, $\frac{\partial\phi}{\partial x} = y^2$, $\frac{\partial\phi}{\partial y} = 2xy+z^3$, and $\frac{\partial\phi}{\partial z} = 3yz^2$. Therefore, as in Example 7.16,

$$\nabla\phi = \hat{i}-3\hat{j}-3\hat{k} \text{ at } (2,-1,1).$$

On the other hand,

$$\nabla(x\log z - y^2+4)$$

$$= \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)(x\log z - y^2+4)$$

$$= \hat{i}\frac{\partial}{\partial x}(x\log z - y^2+4)+\hat{j}\frac{\partial}{\partial y}(x\log z - y^2+4)$$

$$+\hat{k}\frac{\partial}{\partial z}(x\log z - y^2+4)$$

$$= \log z\hat{i} - 2y\hat{j}+\frac{x}{z}\hat{k} = -4\hat{j}-\hat{k} \text{ at } (-1,2,1).$$

But $\nabla(x \log z - y^2 + 4)$ is normal to the surface $x \log z - y^2 + 4$. Unit vector along $\nabla(x \log z - y^2 + 4)$ is

$$\hat{a} = \frac{-4\hat{j}-\hat{k}}{\sqrt{16+1}} = \frac{-4\hat{j}-\hat{k}}{\sqrt{17}}.$$

Therefore, the directional derivative of ϕ at (2, – 1,1) in the direction of the normal to the surface x log $z - y^2 + 4 = 0$ at (– 1,2,1) is

$$\nabla\,\phi.\hat{a} = \left(\hat{i}-3\hat{j}-3\hat{k}\right)\left(\frac{-4\hat{j}-\hat{k}}{\sqrt{17}}\right) = \frac{12+3}{\sqrt{17}} = \frac{15}{\sqrt{17}}.$$

EXAMPLE 7.18

Find the angle between the surfaces $x^2 + y^2 + z^2 = 9$ and $z = x^2 + y^2 - 3$ at the point (2, –1, 2).

Solution. Let $\phi(x, y, z) = x^2 + y^2 + z^2 - 9$ and $\psi\,(x, y, z) = x^2 + y^2 - 3 - z$. Then, the angle between the surfaces at the given point (2,–1,2) is the angle between the normal to the surfaces at that point. Also $\nabla\phi$ and $\nabla\psi$ are along the normal to ϕ and ψ, respectively. But,

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(x^2 + y^2 + z^2 - 9)$$
$$= 2x\hat{i} + 2y\hat{j} + 2z\hat{k}$$
$$= 4\hat{i} - 2\hat{j} + 4\hat{k} \text{ at the point}(2, -1, 2)$$

and

$$\nabla\psi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(x^2 + y^2 - 3 - z)$$
$$= 2x\hat{i} + 2y\hat{j} - \hat{k}$$
$$= 4\hat{i} - 2\hat{j} - \hat{k} \text{ at the point } (2, -1, 2).$$

If θ is the angle between $\nabla\phi$ and $\nabla\psi$, then

$$\cos\theta = \frac{\nabla\phi.\nabla\psi}{|\nabla\phi||\nabla\psi|} = \frac{\left(4\hat{i} - 2\hat{j} + 4\hat{k}\right).\left(4\hat{i} - 2\hat{j} - \hat{k}\right)}{\sqrt{16+4+16}\sqrt{16+4+1}}$$
$$= \frac{16+4-4}{6\sqrt{21}} = \frac{8}{3\sqrt{21}}.$$

Hence, the required angle is $\theta = \cos^{-1}\left(\frac{8}{3\sqrt{21}}\right)$.

EXAMPLE 7.19

In what direction from (3,1,–2) is the directional derivative of $\phi = x^2 y^2 z^4$ maximum and what is its magnitude?

Solution. The directional derivative at a given point of a given surface ϕ is maximum along the normal to the surface and grad ϕ acts along the normal. Therefore, the directional derivative is maximum along $\nabla\phi$. We have

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(x^2y^2z^4)$$
$$= (2xy^2z^4)\hat{i} + (2x^2yz^4)\hat{j} + (4x^2y^2z^3)\hat{k}$$
$$= 96\hat{i} + 288\hat{j} - 288\hat{k} \text{ at the point } (3, 1, -2).$$

Thus, the directional derivative is maximum in the direction of $96\hat{i} + 288\hat{j} - 288\hat{k}$. The magnitude of the maximum directional derivative is $|\nabla\phi| = 96\sqrt{1+9+9} = 96\sqrt{19}$.

EXAMPLE 7.20

Find the angles between the normal to the surface $xy = z^2$ at the points (4,1,2) and (3,3,–3).

Solution. Let $\phi(x, y, z) = xy - z^2$. Since $\nabla\phi$ is along the normal, it is sufficient to find angle between $\nabla\phi$ at (4, 1, 2) and $\nabla\phi$ at (3, 3, –3). Now,

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(xy - z^2) = y\hat{i} + x\hat{j} - 2z\hat{k}.$$

Therefore,

$$\nabla\phi \text{ at } (4, 1, 2) \text{ is } \hat{i} + 4\hat{j} - 4\hat{k} \text{ and}$$
$$\nabla\phi \text{ at } (3, 3, -3) \text{ is } 3\hat{i} + 3\hat{j} + 6\hat{k}.$$

Hence, the required angle θ is the angle between $\hat{i} + 4\hat{j} - 4\hat{k}$ and $3\hat{i} + 3\hat{j} + 6\hat{k}$. Therefore,

$$\cos\theta = \frac{\left(\hat{i} + 4\hat{j} - 4\hat{k}\right).\left(3\hat{i} + 3\hat{j} + 6\hat{k}\right)}{|\hat{i} + 4\hat{j} - 4\hat{k}|\ |3\hat{i} + 3\hat{j} + 6\hat{k}|} = \frac{-9}{\sqrt{33}\sqrt{54}}$$
$$= -\frac{1}{\sqrt{22}}.$$

Hence,

$$\theta = \cos^{-1}\left(-\frac{1}{\sqrt{22}}\right).$$

EXAMPLE 7.21

Find the constants a and b so that the surface $ax^2 - byz = (a + 2)x$ is orthogonal to the surface $4x^2 y + z^3 = 4$ at the point (1,–1,2).

Solution. The two given surfaces will be orthogonal if the angle between the normal to the surfaces at the point (1,–1,2) is $\frac{\pi}{2}$. Since $\nabla\phi$ acts along the normal, it is sufficient to find $\nabla\phi$ and $\nabla\psi$ at (1,–1,2), where $\phi = ax^2 - byz - (a + 2)x$ and $\psi = 4x^2 y + z^3 - 4$. We have

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(ax^2 - byz - (a+2)x)$$
$$= \hat{i}(2ax - a - 2) + \hat{j}(-bz) + \hat{k}(-by)$$
$$= (a-2)\hat{i} - 2b\hat{j} + b\hat{k} \text{ at } (1, -1, 2),$$

and

$$\nabla\psi = \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)\left(4x^2y+z^3-4\right)$$

$$= \hat{i}(8xy)+\hat{j}\left(4x^2\right)+\hat{k}\left(3z^2\right)$$

$$= -8\hat{i}+4\hat{j}+12\hat{k}\ at\ (1,-1,2).$$

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

$$\cos\frac{\pi}{2} = \frac{\left((a-2)\hat{i}-2b\hat{j}+b\hat{k}\right).\left(-8\hat{i}+4\hat{j}+12\hat{k}\right)}{\left|(a-2)\hat{i}-2b\hat{j}+b\hat{k}\right|\left|-8\hat{i}+4\hat{j}+12\hat{k}\right|}.$$

Hence,

$$\left((a-2)\hat{i}-2b\hat{j}+b\hat{k}\right).\left(-8\hat{i}+4\hat{j}+12\hat{k}\right)=0$$

or

$$(a-2)\,(-8)-8b+12b=0$$

or

$$-8a+4b=-16. \qquad (1)$$

Since the points (1, −1, 2) lie on both surfaces ϕ and ψ, we have from the surface ϕ,

$$a+2b=a+2 \text{ or } b=1. \qquad (2)$$

Putting the value of b from (2) in (1), we get

$$-8a=-16-4 \text{ or } a=\frac{20}{8}=\frac{5}{2}.$$

Hence, $a=\frac{5}{2}$ and $b=1$.

EXAMPLE 7.22

A paraboloid of revolution has the equation $2z = x^2 + y^2$. Find the equation of the normal and the tangent plane to the surface at the point (1, 3, 5).

Solution. Let $\phi = x^2 + y^2 - 2z$. Then, $\nabla\phi$ gives the vector normal to the surface. Thus, the normal vector to the surface is

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)\left(x^2+y^2-2z\right)$$

$$= 2x\hat{i}+2y\hat{j}-2\hat{k}$$

$$= 2\hat{i}+6\hat{j}-2\hat{k}, \text{ at the point } (1,3,5).$$

Therefore, the unit normal vector at the point (1,3,5) is

$$\hat{a} = \frac{2\hat{i}+6\hat{j}-2\hat{k}}{\sqrt{4+36+4}} = \frac{\hat{i}+3\hat{j}-\hat{k}}{\sqrt{11}}.$$

The equation of the line through the point (1,3,5) in the direction of this normal vector is

$$\frac{x-1}{1}=\frac{y-3}{3}=\frac{z-5}{-1}.$$

Therefore, the equation of the tangent plane to the surface at the point (1,3,5) is

$$1(x-1)+3(y-3)+(-1)\,(z-5)=0$$

or

$$x+3y-z=5.$$

EXAMPLE 7.23

Find the angle between the tangent planes to the surfaces $x \log z = y^2 - 1$ and $x^2\, y = 2 - z$ at the point (1, 1, 1).

Solution. The required angle will be the angle between the vectors normal to the given surfaces at the given point. The normal vectors to the surfaces $\phi = x \log z - y^2 + 1$ and $\psi = x^2\, y - 2 + z$ are given by

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)\left(x\log z-y^2+1\right)$$

$$= \log z\hat{i}-2y\hat{j}+\frac{x}{z}\hat{k}$$

$$= -2\hat{j}+\hat{k} \text{ at the point } (1,1,1)$$

and

$$\nabla\psi = \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)\left(x^2y-2+z\right)$$

$$= 2xy\hat{i}+x^2\hat{j}+\hat{k}$$

$$= 2\hat{i}+\hat{j}+\hat{k} \text{ at the point } (1,1,1).$$

Therefore, the required angle is given by

$$\cos\theta = \frac{\left(-2\hat{j}+\hat{k}\right).\left(2\hat{i}+\hat{j}+\hat{k}\right)}{\sqrt{5}.\sqrt{6}} = \frac{-2+1}{\sqrt{30}} = \frac{-1}{\sqrt{30}}.$$

Hence,

$$\theta = \cos^{-1}\left(\frac{-1}{\sqrt{30}}\right).$$

7.7 DIVERGENCE OF A VECTOR-POINT FUNCTION

If we want to consider the rate of change of a vector-point function $\vec{f}$, there are two ways of operating the vector operator ∇ to the vector $\vec{f}$. Thus, we have two cases to consider, namely,

$$\nabla \cdot \vec{f} \text{ and } \nabla \times \vec{f}.$$

These two cases lead us to the two concepts called *Divergence of a Vector Function* and *curl of a Vector Function*. If we consider a vector field as a fluid flow, then at every point in the flow, we need to measure the *rate of flow* of the fluid from that point and the *amount of spin* possessed by the particles of the fluid at that point. The above two concepts provide respectively, the two measures called divergence of $\vec{f}$ and curl of $\vec{f}$.

Let $\vec{f} = f_1\hat{i} + f_2\hat{j} + f_3\hat{k}$ be a vector function, where f_1, f_2, and f_3 are scalar-point functions, which is defined and differentiable at each point of the region of space. Then, the *divergence* of $\vec{f}$, denoted by $\nabla \cdot \vec{f}$ or div $\vec{f}$, is a scalar given by

$$\nabla \cdot \vec{f} = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right).\left(f_1\hat{i} + f_2\hat{j} + f_3\hat{k}\right)$$

$$= \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}.$$

The vector $\vec{f}$ is called *Solenoidal* if $\nabla \cdot \vec{f} = 0$.

7.8 PHYSICAL INTERPRETATION OF DIVERGENCE

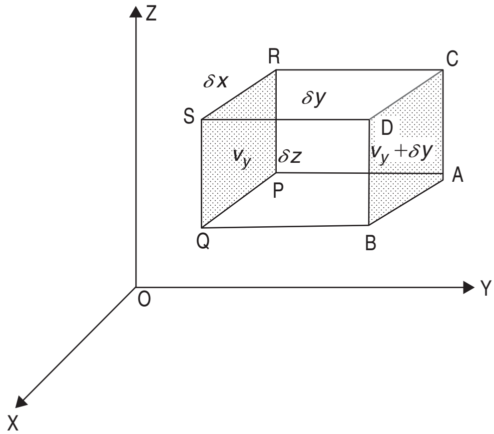
Consider the steady motion of the fluid having velocity $\vec{v} = v_x\hat{i} + v_y\hat{j} + v_z\hat{k}$ at a point $P(x, y, z)$ Consider a small parallelopiped with edges δx, δy, and δz parallel to the axes, with one of its corner at $P(x, y, z)$. The mass of the fluid entering through the face PQRS per unit time is $v_y\ \delta x\ \delta z$ and the mass of the fluid that flows out through the opposite face ABCD is $v_{y+\delta y}$ $\delta x\ \delta z$. Therefore, the change in the mass of fluid flowing across these two faces is equal to

$$v_{y+\delta y}\delta x\ \delta z - v_y\delta x\delta z = \left(v_y + \frac{\partial v_y}{\partial y}.\delta y\right)\delta x\ \delta z$$

$$- v_y\delta x\ \delta z = \frac{\partial v_y}{\partial y}\delta y\ \delta x\ \delta z.$$

Similarly, the changes in the mass of the fluid for the other two pairs of faces are

$$\frac{\partial v_x}{\partial x}\delta x\ \delta y\ \delta z \quad \text{and} \quad \frac{\partial v_z}{\partial z}\delta x\ \delta y\ \delta z.$$



Therefore, the total change in the mass of the fluid inside the parallelopiped per unit time is equal to

$$\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right)\delta x\ \delta y\ \delta z.$$

Hence, the rate of change of the mass of the fluid per unit time per unit volume is

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = \nabla \cdot \vec{v},$$

by the definition of divergence. Hence, div $\vec{v}$ gives the rate at which the fluid (the vector field) is flowing away at a point of the fluid.

EXAMPLE 7.24

Find div $\vec{v}$, where $\vec{v} = 3x^2y\hat{i} + z\hat{j} + x^2\hat{k}$.

Solution. We know that

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

Here, $v_1 = 3x^2\ y$, $v_2 = z$, and $v_3 = x^2$. Therefore,

$$\text{div } \vec{v} = 6xy.$$

EXAMPLE 7.25

Find the value of the constant λ such that the vector field defined by

$\vec{f} = (2x^2y^2 + z^2)\hat{i} + (3xy^3 - x^2z)\hat{j} + (\lambda xy^2z + xy)\hat{k}$

is solenoidal.

Solution. We have

$f_1 = 2x^2y^2 + z^2, f_2 = 3xy^3 - x^2z$, and $f_3 = \lambda xy^2z + xy$.

Therefore,

$$\operatorname{div}\vec{f} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} = 4xy^2 + 9xy^2 + \lambda xy^2.$$

The vector field shall be a solenoidal if $\operatorname{div}\vec{f} = 0$. So, we must have

$$4xy^2 + 9xy^2 + \lambda xy^2 = 0,$$

which yields $\lambda = -13$.

EXAMPLE 7.26

Find $\operatorname{div}\vec{f}$, where $\vec{f} = grad(x^3 + y^3 + z^3 - 3xyz)$.

Solution. We have

$$\vec{f} = \nabla(x^3 + y^3 + z^3 - 3xyz)$$
$$= \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(x^3 + y^3 + z^3 - 3xyz)$$
$$= (3x^2 - 3yz)\hat{i} + (3y^2 - 3xz)\hat{j} + (3z^2 - 3xy)\hat{k}$$
$$= f_1\hat{i} + f_2\hat{j} + f_3\hat{k}, \text{ say.}$$

Then,

$$div\vec{f} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}$$
$$= 6x + 6y + 6z = 6(x + y + z).$$

EXAMPLE 7.27

Find div $\left(3x^2\hat{i} + 5xy^2\hat{j} + xyz^3\hat{k}\right)$ at the point. (1, 2, 3).

Solution. Let $\vec{f} = 3x^2\hat{i} + 5xy^2\hat{j} + xyz^3\hat{k} = f_1\hat{i} + f_2\hat{j} + f_3\hat{k}$, say. Then,

$$div\,\vec{f} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} = 6x + 10xy + 3xyz^2$$
$$= 6 + 20 + 54 = 80 \text{ at } (1, 2, 3).$$

7.9 CURL OF A VECTOR-POINT FUNCTION

Let $\vec{f} = f_1\hat{i} + f_2\hat{j} + f_3\hat{k}$ be a vector-point function, where f_1, f_2, and f_3 are scalar-point functions. If $\vec{f}$ is defined and differentiable at each point (x, y, z) of the region of space, then the *curl* (or *rotation*) of $\vec{f}$, denoted, by curl $\vec{f}$, $\nabla \times \vec{f}$, or rot $\vec{f}$ is defined by

$$\text{Curl}\,\vec{f} = \nabla \times \vec{f}$$
$$= \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right) \times \left(f_1\hat{i} + f_2\hat{j} + f_3\hat{k}\right)$$
$$= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & f_3 \end{vmatrix}$$
$$= \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}\right)\hat{i} + \left(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}\right)\hat{j} + \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right)\hat{k}.$$

Obviously, *curl $\vec{f}$ is a vector-point function*.

7.10 PHYSICAL INTERPRETATION OF CURL

Consider a rigid body rotating about a fixed axis through the origin with angular velocity $\vec{\omega} = \omega_1\hat{i} + \omega_2\hat{j} + \omega_3\hat{k}$. Let $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$ be the position vector of any point $P(x,y,z)$ on the body. Then, the velocity $\vec{v}$ of P is given by

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

Therefore,

$$\text{Curl}\,\vec{v} = \nabla \times \vec{v}$$
$$= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \omega_2 z - \omega_3 y & \omega_3 x - \omega_1 z & \omega_1 y - \omega_2 x \end{vmatrix}$$
$$= 2\left(\omega_1\hat{i} + \omega_2\hat{j} + \omega_3\hat{k}\right), \text{ since } \omega_1, \omega_2, \text{ and } \omega_3 \text{ are constants}$$
$$= 2\vec{\omega}.$$

Hence, $\vec{\omega} = \frac{1}{2}$ curl $\vec{v}$. It follows, therefore, that the *angular velocity at any point is equal to half the curl of the linear velocity at that point of the body.*

Thus, curl is a *measure of rotation*. If curl $\vec{v} = 0$, then the vector $\vec{v}$ is called an *irrotational vector*.

7.11 THE LAPLACIAN OPERATOR ∇^2

If ϕ is a scalar-point function, then

$$\text{grad } \phi = \nabla\, \phi = \hat{i}\frac{\partial \phi}{\partial x} + \hat{j}\frac{\partial \phi}{\partial y} + \hat{k}\frac{\partial \phi}{\partial z}$$

(vector - point function)

and then,

$$\begin{aligned}\text{div } [\text{grad } \phi] &= \nabla \cdot \nabla\, \phi \\ &= \frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial \phi}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial \phi}{\partial z}\right) \\ &= \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} \\ &= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\phi \;= \nabla^2 \phi,\end{aligned}$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is called *Laplacian operator*.

A scalar-point function possessing second-order continuous partial derivatives and satisfying the *Laplacian equation* $\nabla^2 \phi = 0$ is called a *harmonic function.*

EXAMPLE 7.28

Find curl $\vec{F}$, where $\vec{F} = grad(x^3 + y^3 + z^3 - 3\,xyz)$.

Solution. We have

$$\begin{aligned}\vec{F} &= \text{grad}\left(x^3 + y^3 + z^3 - 3xyz\right) \\ &= \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)\left(x^3 + y^3 + z^3 - 3xyz\right) \\ &= \left(3x^2 - 3yz\right)\hat{i} + \left(3y^2 - 3xz\right)\hat{j} + \left(3z^2 - 3xy\right)\hat{k}.\end{aligned}$$

Therefore,

$$\begin{aligned}\text{Curl } \vec{F} &= \nabla \times \vec{F} = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right) \times \vec{F} \\ &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 3x^2 - 3yz & 3y^2 - 3xz & 3z^2 - 3xy \end{vmatrix} \\ &= \hat{i}(-3x + 3x) + \hat{j}(-3y + 3y) + \hat{k}(-3z + 3z) \\ &= \vec{0}.\end{aligned}$$

EXAMPLE 7.29

Show that the vector $\vec{v} = (yz)\hat{i} + (zx)\hat{j} + (xy)\hat{k}$ is irrotational.

Solution. It is sufficient to show that the curl $\vec{v} = \vec{0}$. We have

$$\begin{aligned}\text{Curl } \vec{v} &= \nabla \times \vec{v} = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right) \\ &\quad \times \left[(yz)\hat{i} + (zx)\hat{j} + (xy)\hat{k}\right] \\ &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ yz & zx & xy \end{vmatrix} \\ &= \hat{i}(x - x) + \hat{j}(y - y) + \hat{k}(z - z) \;= \vec{0}.\end{aligned}$$

EXAMPLE 7.30

If $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$, show that curl $\vec{r} = \vec{0}$.

Solution. We have

$$\begin{aligned}\text{curl } \vec{r} &= \nabla \times \vec{r} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x & y & z \end{vmatrix} \\ &= \hat{i}\left[\frac{\partial}{\partial y}(z) - \frac{\partial}{\partial z}(y)\right] + \hat{j}\left[\frac{\partial}{\partial z}(x) - \frac{\partial}{\partial x}(z)\right] \\ &\quad + \hat{k}\left[\frac{\partial}{\partial x}(y) - \frac{\partial}{\partial y}(x)\right] \\ &= \hat{i}\,(0) + \hat{j}(0) + \hat{k}(0) = \vec{0}.\end{aligned}$$

EXAMPLE 7.31

Show that curl curl $\vec{f} = \vec{0}$, where $\vec{f} = z\hat{i} + x\hat{j} + y\hat{k}$.

Solution. Let $\vec{f} = z\hat{i} + x\hat{j} + y\hat{k}$. Then

$$\begin{aligned}\text{curl } \vec{f} &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ z & x & y \end{vmatrix} \\ &= \left[\frac{\partial y}{\partial y} - \frac{\partial x}{\partial z}\right]\hat{i} + \left[\frac{\partial z}{\partial z} - \frac{\partial y}{\partial x}\right]\hat{j} + \left[\frac{\partial x}{\partial x} - \frac{\partial z}{\partial y}\right]\hat{k} \\ &= (1 - 0)\hat{i} + (1 - 0)\hat{j} + (1 - 0)\hat{k} = \hat{i} + \hat{j} + \hat{k}.\end{aligned}$$

Hence,

$$curl\ curl\ \vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 1 & 1 & 1 \end{vmatrix} = \left[\frac{\partial}{\partial y}(1) - \frac{\partial}{\partial z}(1)\right]\hat{i}$$

$$+ \left[\frac{\partial}{\partial z}(1) - \frac{\partial}{\partial x}(1)\right]$$

$$\hat{j} + \left[\frac{\partial}{\partial x}(1) - \frac{\partial}{\partial y}(1)\right]\hat{k}$$

$$= 0\hat{i} + 0\hat{j} + 0\hat{k} = \vec{0}.$$

EXAMPLE 7.32

If all second-order derivatives of ϕ and $\vec{v}$ are continuous, show that
(i) curl (grad ϕ) = $\vec{0}$, (ii) curl (curl $\vec{v}$) = grad div $\vec{v} - \nabla^2\vec{v}$, (iii) div (curl $\vec{v}$) = 0, and (iv) grad (div $\vec{v}$) = $\left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)\left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}\right)$.

Solution. (i) We have

$$\text{curl (grad } \phi) = \nabla \times \nabla\phi$$

$$= \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)$$

$$\times \left(\hat{i}\frac{\partial\phi}{\partial x} + \hat{j}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}\right)$$

$$= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial\phi}{\partial x} & \frac{\partial\phi}{\partial y} & \frac{\partial\phi}{\partial z} \end{vmatrix}$$

$$= \hat{i}\left(\frac{\partial^2\phi}{\partial y\partial z} - \frac{\partial^2\phi}{\partial z\partial y}\right) + \hat{j}\left(\frac{\partial^2\phi}{\partial z\partial x} - \frac{\partial^2\phi}{\partial x\partial z}\right)$$

$$+ \hat{k}\left(\frac{\partial^2\phi}{\partial x\partial y} - \frac{\partial^2\phi}{\partial y\partial x}\right)$$

$$= \hat{i}(0) + \hat{j}(0) + \hat{k}(0) = \vec{0}.$$

(ii) If $\vec{v} = v_1\hat{i} + v_2\hat{j} + v_3\hat{k}$, then

$$\text{curl}\,\vec{v} = \nabla \times \vec{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix} = \sum i\left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)$$

and so,

$$\text{curl}\,(\text{curl}\,\vec{v}) = \nabla \times (\nabla \times \vec{v})$$

$$= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z} & \frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x} & \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \end{vmatrix}$$

$$= \sum \hat{i}\left[\frac{\partial}{\partial y}\left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right) - \frac{\partial}{\partial z}\left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\right]$$

$$= \sum \hat{i}\left[\frac{\partial}{\partial x}\left(\frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}\right) - \left(\frac{\partial^2 v_1}{\partial y^2} + \frac{\partial^2 v_1}{\partial z^2}\right)\right]$$

$$= \sum \hat{i}\left[\frac{\partial}{\partial x}\left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}\right)\right.$$

$$\left. - \left(\frac{\partial^2 v_1}{\partial x^2} + \frac{\partial^2 v_1}{\partial y^2} + \frac{\partial^2 v_1}{\partial z^2}\right)\right]$$

$$= \sum \hat{i}\frac{\partial}{\partial x}(\text{div}\,\vec{v}) - \sum \nabla^2\hat{i}v_1$$

$$= \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(\text{div}\,\vec{v})$$

$$- \nabla^2\left(\hat{i}v_1 + \hat{j}v_2 + \hat{k}v_3\right) = \text{grad div}\,\vec{v} - \nabla^2\vec{v}.$$

(iii) As in (ii),

$$\text{curl}\ \vec{v} = \hat{i}\left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right) + \hat{j}\left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)$$

$$+ \hat{k}\left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right).$$

Therefore,

$$\text{div}(\text{curl}\,\vec{v}) = \nabla \cdot (\nabla \times \vec{v})$$

$$= \frac{\partial}{\partial x}\left[\left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\right]$$

$$+ \frac{\partial}{\partial z}\left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)$$

$$= \frac{\partial^2 v_3}{\partial x\partial y} - \frac{\partial^2 v_2}{\partial x\partial z} + \frac{\partial^2 v_1}{\partial y\partial z} - \frac{\partial^2 v_3}{\partial y\partial x}$$

$$+ \frac{\partial^2 v_2}{\partial z\partial x} - \frac{\partial^2 v_1}{\partial z\partial y}$$

$$= \frac{\partial^2 v_3}{\partial x\partial y} - \frac{\partial^2 v_2}{\partial x\partial z} + \frac{\partial^2 v_1}{\partial y\partial z} - \frac{\partial^2 v_3}{\partial x\partial y}$$

$$+ \frac{\partial^2 v_2}{\partial x\partial z} - \frac{\partial^2 v_1}{\partial y\partial z} \quad \text{since } \vec{v} \text{ is continuous}$$

$$= 0.$$

(iv) If $\vec{v} = v_1\hat{i} + v_2\hat{j} + v_3\hat{k}$, then

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

Therefore,

$$\text{grad}\,(\text{div}\,\vec{v}) = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)\left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}\right).$$

EXAMPLE 7.33

If $\vec{a}$ is a constant vector and $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$, show that
(i) $\text{div}(\vec{a} \times \vec{r}) = 0$ and (ii) $\text{curl}(\vec{a} \times \vec{r}) = 2\vec{a}$.

Solution. (i) We have $\vec{a} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k}$. Therefore,

$$\vec{a} \times \vec{r} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ a_1 & a_2 & a_3 \\ x & y & z \end{vmatrix}$$

$$= \hat{i}(a_2 z - a_3 y) - \hat{j}(a_1 z - a_3 x) + \hat{k}(a_1 y - a_2 x).$$

Hence,

$$div\ (\vec{v} \times \vec{r}) = \frac{\partial}{\partial x}(a_2\ z - a_3\ y) - \frac{\partial}{\partial y}(a_1\ z - a_3\ x) + \frac{\partial}{\partial z}(a_1\ y - a_2\ x) = 0 - 0 + 0 = 0$$

(ii) $\text{curl}(\vec{a} \times \vec{r}) = \nabla \times (\vec{a} \times \vec{r})$

$$= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_2\,z - a_3\,y & a_3\,x - a_1\,z & a_1\,y - a_2\,x \end{vmatrix}$$

$$= \hat{i}(a_1 + a_1) + \hat{j}(a_2 + a_2) + \hat{k}(a_3 + a_3)$$

$$= 2\left(a_1\hat{i} + a_2\hat{j} + a_3\hat{k}\right) = 2\vec{a}.$$

EXAMPLE 7.34

Determine curl curl $\vec{v}$ if $\vec{v} = x^2y\hat{i} + y^2z\hat{j} + z^2y\hat{k}$.

Solution. Let $\vec{v} = x^2y\hat{i} + y^2z\hat{j} + z^2y\hat{k}$. Then

$$\text{curl } \vec{v} = \nabla \times \vec{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2y & y^2z & z^2y \end{vmatrix}$$

$$= \hat{i}\left(z^2 - y^2\right) + \hat{j}(0) + \hat{k}\left(-x^2\right)$$

$$= \hat{i}\left(z^2 - y^2\right) + 0\hat{j} - x^2\hat{k}$$

Hence,

$$\text{curl curl } \vec{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ z^2 - y^2 & 0 & -x^2 \end{vmatrix}$$

$$= (2z + 2x)\hat{j} + 2y\hat{k}$$

EXAMPLE 7.35

Show that $r^n\vec{r}$ is irrotational.

Solution. It is sufficient to show that curl $r^n\vec{r} = \vec{0}$. We have

$$\text{curl } r^n\ \vec{r} = \nabla \times r^n\vec{r}$$

$$= \nabla \times \left[r^n\left(x\vec{i} + y\vec{j} + z\vec{k}\right)\right]$$

$$= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ r^n x & r^n y & r^n z \end{vmatrix}$$

$$= \left[\frac{\partial}{\partial x}(r^n z) - \frac{\partial}{\partial z}(r^n y)\right]\hat{i} + \left[\frac{\partial}{\partial z}(r^n x) - \frac{\partial}{\partial x}(r^n z)\right]\hat{j} + \left[\frac{\partial}{\partial x}(r^n y) - \frac{\partial}{\partial y}(r^n x)\right]\hat{k}$$

$$= \left[znr^{n-1}\frac{\partial r}{\partial y} - ynr^{n-1}\frac{\partial r}{\partial z}\right]\hat{i} + \left[xnr^{n-1}\frac{\partial r}{\partial z} - znr^{n-1}\frac{\partial r}{\partial x}\right]\hat{j} + \left[ynr^{n-1}\frac{\partial r}{\partial x} - xnr^{n-1}\frac{\partial r}{\partial y}\right]\hat{k}$$

$$= \left[znr^{n-1}\frac{y}{r} - ynr^{n-1}\frac{z}{r}\right]\hat{i} + \left[xnr^{n-1}\frac{z}{r} - znr^{n-1}\frac{x}{r}\right]\hat{j} + \left[ynr^{n-1}\frac{x}{r} - xnr^{n-1}\frac{y}{r}\right]\hat{k}$$

$$= 0\hat{i} + 0\hat{j} + 0\hat{k} = \vec{0}.$$

Hence, $r^n\ \vec{r}$ is irrotational.

7.12 PROPERTIES OF DIVERGENCE AND CURL

(A) **Properties of Divergence.** Let $\vec{f}$ and $\vec{g}$ be two vector-point functions and ϕ a scalar-point function. Then, the divergence has the following properties.

(i) $\text{div}\left(\vec{f} + \vec{g}\right) = \nabla \cdot \left(\vec{f} + \vec{g}\right) = \nabla \cdot \vec{f} + \nabla \cdot \vec{g}.$

(ii) $\text{div}\left(\phi\vec{f}\right) = \nabla\cdot\left(\phi\vec{f}\right) = (\nabla\phi)\cdot\vec{f} + \phi\left(\nabla\cdot\vec{f}\right)$
$= (grad\,\phi)\cdot\vec{f} + \phi\,\text{div}\vec{f}.$

(iii) $\text{div}\,\vec{f} = 0$ if $\vec{f}$ is a constant vector.

(iv) $\text{div}\left(\vec{f}\times\vec{g}\right) = \vec{g}.\text{curl}\,\vec{f} - \vec{f}.curl\,\vec{g}.$

Proof: Let $\vec{f} = f_1\hat{i} + f_2\hat{j} + f_3\hat{k}$ and $\vec{g} = g_1\hat{i} + g_2\hat{j} + g_3\hat{k}$.
(i) We have

$$\vec{f} + \vec{g} = (f_1 + g_1)\hat{i} + (f_2 + g_2)\hat{j} + (f_3 + g_3)\hat{k}.$$

Therefore,

$$\begin{aligned}
\nabla\cdot\left(\vec{f}+\vec{g}\right) &= \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right). \\
&\quad\left[(f_1+g_1)\hat{i}+(f_2+g_2)\hat{j}+(f_3+g_3)\hat{k}\right] \\
&= \frac{\partial}{\partial x}(f_1+g_1)+\frac{\partial}{\partial y}(f_2+g_2)+\frac{\partial}{\partial z}(f_3+g_3) \\
&= \left(\frac{\partial f_1}{\partial x}+\frac{\partial f_2}{\partial y}+\frac{\partial f_3}{\partial z}\right) \\
&\quad + \left(\frac{\partial g_1}{\partial x}+\frac{\partial g_2}{\partial y}+\frac{\partial g_3}{\partial z}\right) \\
&= \nabla\cdot\vec{f}+\nabla\cdot\vec{g}
\end{aligned}$$

(ii)
$$\begin{aligned}
\nabla\cdot\left(\phi\vec{f}\right) &= \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right).\left(\phi f_1\hat{i}+\phi f_2\hat{j}+\phi f_3\hat{k}\right) \\
&= \frac{\partial}{\partial x}(\phi f_1)+\frac{\partial}{\partial y}(\phi f_2)+\frac{\partial}{\partial z}(\phi f_3) \\
&= \left(\frac{\partial\phi}{\partial x}f_1+\phi\frac{\partial f_1}{\partial x}\right)+\left(\frac{\partial\phi}{\partial y}f_2+\phi\frac{\partial f_2}{\partial y}\right) \\
&\quad + \left(\frac{\partial\phi}{\partial z}f_3+\phi\frac{\partial f_3}{\partial z}\right) \\
&= \left(\frac{\partial\phi}{\partial x}f_1+\frac{\partial\phi}{\partial y}f_2+\frac{\partial\phi}{\partial z}f_3\right) \\
&\quad + \left(\phi\frac{\partial f_1}{\partial x}+\phi\frac{\partial f_2}{\partial y}+\phi\frac{\partial f_3}{\partial z}\right) \\
&= \left(\frac{\partial\phi}{\partial x}\hat{i}+\frac{\partial\phi}{\partial y}\hat{j}+\frac{\partial\phi}{\partial z}\hat{k}\right)\cdot\left(f_1\hat{i}+f_2\hat{j}+f_3\hat{k}\right) \\
&\quad + \phi\left(\frac{\partial f_1}{\partial x}+\frac{\partial f_2}{\partial y}+\frac{\partial f_3}{\partial z}\right) \\
&= (\nabla\phi)\cdot\vec{f}+\phi\left(\nabla\cdot\vec{f}\right) \\
&= (\text{grad}\,\phi).\vec{f}+\phi\,\text{div}\vec{f}.
\end{aligned}$$

(iii) Let $\vec{f} = f_1\hat{i} + f_2\hat{j} + f_3\hat{k}$. Then

$$\nabla\cdot\vec{f} = \frac{\partial f_1}{\partial x}+\frac{\partial f_2}{\partial y}+\frac{\partial f_3}{\partial z} = 0, \text{ since } \vec{f} \text{ is constant.}$$

(iv)
$$\begin{aligned}
\text{div}\left(\vec{f}\times\vec{g}\right) &= \nabla\cdot\left(\vec{f}\times\vec{g}\right) \\
&= \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right).\left(\vec{f}\times\vec{g}\right) \\
&= \sum\hat{i}.\frac{\partial}{\partial x}\left(\vec{f}\times\vec{g}\right) \\
&= \sum\hat{i}.\left(\frac{\partial\vec{f}}{\partial x}\times\vec{g}+\vec{f}\times\frac{\partial\vec{g}}{\partial x}\right) \\
&= \sum\hat{i}.\left(\frac{\partial\vec{f}}{\partial x}\times\vec{g}\right) \\
&\quad + \sum\hat{i}.\left(\vec{f}\times\frac{\partial\vec{g}}{\partial x}\right) \\
&= \sum\left(\hat{i}\times\frac{\partial\vec{f}}{\partial x}\right).\vec{g} \\
&\quad - \sum\hat{i}.\left(\frac{\partial\vec{g}}{\partial x}\times\vec{f}\right) \\
&= \sum\left(\hat{i}\times\frac{\partial\vec{f}}{\partial x}\right).\vec{g} \\
&\quad - \sum\left(\hat{i}\times\frac{\partial\vec{g}}{\partial x}\right).\vec{f} \\
&= \left(\nabla\times\vec{f}\right).\vec{g} - (\nabla\times\vec{g}).\vec{f} \\
&= \vec{g}.\left(\nabla\times\vec{f}\right) - \vec{f}\,(\nabla\times\vec{g})
\end{aligned}$$

(commutativity of dot product)

$$= \vec{g}\;curl\,\vec{f} - \vec{f}.curl\,\vec{g}.$$

(B) **Properties of Curl.** Let $\vec{f}$ and $\vec{g}$ be two vector-point functions and ϕ a scalar-point function, all having continuous second-order partial derivatives. Then,

(i) $\text{curl}\left(\vec{f}+\vec{g}\right) = \text{curl}\,\vec{f} + \text{curl}\,\vec{g}.$

(ii) $\text{curl}\left(\phi\vec{f}\right) = (\text{grad}\,\phi)\times\vec{f} + \phi\,\text{curl}\,\vec{f}.$

(iii) $\text{curl}\,\vec{f} = \vec{0}$, if $\vec{f}$ is a constant vector.

(iv) $\nabla\times\left(\vec{f}\times\vec{g}\right) = (\vec{g}.\nabla)\vec{f} - \left(\vec{f}.\vec{\nabla}\right)\vec{g}$
$+\vec{f}(\nabla\cdot\vec{g}) - \left(\nabla\cdot\vec{f}\right)\vec{g}.$

Proof: If $\vec{f}=f_1\hat{i}+f_2\hat{j}+f_3\hat{k}$ and $\vec{g}=g_1\hat{i}+g_2\hat{j}+g_3\hat{k}$, then

(i) $\vec{f}+\vec{g}=(f_1+g_1)\hat{i}+(f_2+g_2)\hat{j}+(f_3+g_3)\hat{k}$ and so,

$$curl\left(\vec{f}+\vec{g}\right)=\nabla\times\left(\vec{f}+\vec{g}\right)$$

$$=\begin{vmatrix}\hat{i} & \hat{j} & \hat{k}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ f_1+g_1 & f_2+g_2 & f_3+g_3\end{vmatrix}$$

$$=\left[\frac{\partial}{\partial y}(f_3+g_3)-\frac{\partial}{\partial z}(f_2+g_2)\right]\hat{i}$$

$$+\left[\frac{\partial}{\partial z}(f_1+g_1)-\frac{\partial}{\partial x}(f_3+g_3)\right]\hat{j}$$

$$+\left[\frac{\partial}{\partial x}(f_2+g_2)-\frac{\partial}{\partial y}(f_1+g_1)\right]\hat{k}$$

$$=\left[\left(\frac{\partial f_3}{\partial y}-\frac{\partial f_2}{\partial z}\right)\hat{i}+\left(\frac{\partial f_1}{\partial z}-\frac{\partial f_3}{\partial x}\right)\hat{j}\right.$$

$$\left.+\left(\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right)\hat{k}\right]$$

$$+\left[\left(\frac{\partial g_3}{\partial y}-\frac{\partial g_2}{\partial z}\right)\hat{i}+\left(\frac{\partial g_1}{\partial z}-\frac{\partial g_3}{\partial x}\right)\hat{j}\right.$$

$$\left.+\left(\frac{\partial g_2}{\partial x}-\frac{\partial g_1}{\partial y}\right)\hat{k}\right]$$

$$=\nabla\times\vec{f}+\nabla\times\vec{g}=\text{curl }\vec{f}+\text{curl }\vec{g}.$$

(ii) $curl\left(\phi\vec{f}\right)=\nabla\times\left(\phi\vec{f}\right)$

$$=\nabla\times\left(\phi f_1\hat{i}+\phi f_2\hat{j}+\phi f_3\hat{k}\right)$$

$$=\begin{vmatrix}\hat{i} & \hat{j} & \hat{k}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ \phi f_1 & \phi f_2 & \phi f_3\end{vmatrix}$$

$$=\left[\frac{\partial}{\partial y}(\phi f_3)-\frac{\partial}{\partial z}(\phi f_2)\right]\hat{i}$$

$$+\left[\frac{\partial}{\partial z}(\phi f_1)-\frac{\partial}{\partial x}(\phi f_3)\right]\hat{j}$$

$$+\left[\frac{\partial}{\partial x}(\phi f_2)-\frac{\partial}{\partial y}(\phi f_1)\right]\hat{k}$$

$$=\left[\phi\frac{\partial f_3}{\partial y}+\frac{\partial\phi}{\partial y}f_3-\phi\frac{\partial f_2}{\partial z}-\frac{\partial\phi}{\partial z}f_2\right]\hat{i}$$

$$+\left[\phi\frac{\partial f_1}{\partial y}+\frac{\partial\phi}{\partial z}f_1-\phi\frac{\partial f_3}{\partial x}-\frac{\partial\phi}{\partial x}f_3\right]\hat{j}$$

$$+\left[\phi\frac{\partial f_2}{\partial z}+\frac{\partial\phi}{\partial x}f_2-\phi\frac{\partial f_1}{\partial y}-\frac{\partial\phi}{\partial y}f_1\right]\hat{k}$$

$$=\phi\left[\left(\frac{\partial f_3}{\partial y}-\frac{\partial f_2}{\partial z}\right)\hat{i}+\left(\frac{\partial f_1}{\partial z}-\frac{\partial f_3}{\partial x}\right)\hat{j}+\left(\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right)\hat{k}\right]$$

$$+\left[\left(\frac{\partial\phi}{\partial y}f_3-\frac{\partial\phi}{\partial z}f_2\right)\hat{i}+\left(\frac{\partial\phi}{\partial z}f_1-\frac{\partial\phi}{\partial x}f_3\right)\hat{j}+\left(\frac{\partial\phi}{\partial x}f_2-\frac{\partial\phi}{\partial y}f_1\right)\hat{k}\right]$$

$$=\phi\left(\nabla\times\vec{f}\right)+\begin{vmatrix}\hat{i} & \hat{j} & \hat{k}\\ \frac{\partial\phi}{\partial x} & \frac{\partial\phi}{\partial y} & \frac{\partial\phi}{\partial z}\\ f_1 & f_2 & f_3\end{vmatrix}$$

$$=\phi\left(\nabla\times\vec{f}\right)+(\nabla\phi)\times\vec{f}$$

$$=\phi\,\text{curl}\,\vec{f}+(\text{grad}\,\phi)\times\vec{f}.$$

(iii) Let $\vec{f}=f_1\hat{i}+f_2\hat{j}+f_3\hat{k}$ be a constant vector. Then,

$$curl\,\vec{f}=\nabla\times\vec{f}$$

$$=\left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)\times\left(f_1\hat{i}+f_2\hat{j}+f_3\hat{k}\right)$$

$$=\begin{vmatrix}\hat{i} & \hat{j} & \hat{k}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ f_1 & f_2 & f_3\end{vmatrix}$$

$$=\hat{i}\left(\frac{\partial f_3}{\partial y}-\frac{\partial f_2}{\partial z}\right)+\hat{j}\left(\frac{\partial f_1}{\partial z}-\frac{\partial f_3}{\partial x}\right)$$

$$+\hat{k}\left(\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right)$$

$$=\hat{i}(0)+\hat{j}(0)+\hat{k}(0)=\vec{0}.$$

(iv) If $\vec{f}=f_1\hat{i}+f_2\hat{j}+f_3\hat{k}$ and $\vec{g}=g_1\hat{i}+g_2\hat{j}+g_3\hat{k}$, then

$$\nabla\times\left(\vec{f}\times\vec{g}\right)=\left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\right)\times\left(\vec{f}\times\vec{g}\right)$$

$$=\hat{i}\times\frac{\partial}{\partial x}\left(\vec{f}\times\vec{g}\right)+\hat{j}\times\frac{\partial}{\partial y}\left(\vec{f}\times\vec{g}\right)$$

$$+\hat{k}\times\frac{\partial}{\partial z}\left(\vec{f}\times\vec{g}\right)$$

$$=\hat{i}\times\left[\frac{\partial\vec{f}}{\partial x}\times\vec{g}+\vec{f}\times\frac{\partial\vec{g}}{\partial x}\right]$$

$$+\hat{j}\times\left[\frac{\partial\vec{f}}{\partial y}\times\vec{g}+\vec{f}\times\frac{\partial\vec{g}}{\partial y}\right]$$

$$+\hat{k}\times\left[\frac{\partial\vec{f}}{\partial z}\times\vec{g}+\vec{f}\times\frac{\partial\vec{g}}{\partial z}\right]$$

$$=\sum\hat{i}\times\left(\frac{\partial\vec{f}}{\partial x}\times\vec{g}\right)$$

$$+\sum\hat{i}\times\left(\vec{f}\times\frac{\partial\vec{g}}{\partial x}\right)$$

$$=\sum\left[(\hat{i}.\vec{g})\frac{\partial\vec{f}}{\partial x}-\left(\hat{i}.\frac{\partial\vec{f}}{\partial x}\right)\vec{g}\right]$$

$$+\sum\left[\left(\hat{i}.\frac{\partial\vec{g}}{\partial x}\right)\vec{f}-\left(\hat{i}.\vec{f}\right)\frac{\partial\vec{g}}{\partial x}\right]$$

(by property of vector triple product)

$$=\sum g_1\frac{\partial\vec{f}}{\partial x}-\left(\sum\frac{\partial f_1}{\partial x}\right)\vec{g}$$

$$+\left(\sum\frac{\partial g_1}{\partial x}\right)\vec{f}-\sum f_1\frac{\partial\vec{g}}{\partial x}$$

$$=\left(g_1\frac{\partial}{\partial x}+g_2\frac{\partial}{\partial y}+g_3\frac{\partial}{\partial z}\right)\vec{f}-\left(\nabla\cdot\vec{f}\right)\vec{g}$$

$$+(\nabla\cdot\vec{g})\vec{f}-\left(f_1\frac{\partial}{\partial x}+f_2\frac{\partial}{\partial y}+f_3\frac{\partial}{\partial z}\right)\vec{g}$$

$$=(\vec{g}.\nabla)\vec{f}-\left(\nabla\cdot\vec{f}\right)\vec{g}+(\nabla\cdot\vec{g})\vec{f}-\left(\vec{f}.\nabla\right)\vec{g}.$$

EXAMPLE 7.36

Show that

$$\text{div}(\text{grad } r^n)=\nabla\cdot(\nabla r^n)=\nabla^2(r^n)$$
$$=n(n+1)r^{n-2}.$$

Deduce that $\nabla^2\left(\frac{1}{r}\right)=0$.

Solution. From the definition of the Laplacian operator ∇^2, we have

$$\text{div}(\text{grad } r^n)=\nabla\cdot(\nabla r^n)$$
$$=\nabla^2(r^n).=\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)(r^n).\quad(1)$$

But,

$$\frac{\partial^2}{\partial x^2}(r^n)=\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}r^n\right)=\frac{\partial}{\partial x}\left(nr^{n-1}\frac{\partial r}{\partial x}\right)$$
$$=\frac{\partial}{\partial x}\left(nr^{n-1}.\frac{x}{r}\right),\text{since }\frac{\partial r}{\partial x}=\frac{x}{r}$$
$$=\frac{\partial}{\partial x}(nr^{n-2}x)=n\left[(n-2)r^{n-3}\frac{\partial r}{\partial x}.x+r^{n-2}\right]$$
$$=n(n-2)r^{n-3}.\frac{x}{r}.x+nr^{n-2}$$
$$=n(n-2)r^{n-4}x^2+nr^{n-2}.$$

Similarly,

$$\frac{\partial^2}{\partial y^2}(r^n)=n(n-2)r^{n-4}y^2+n\,r^{n-2}\text{ and}$$
$$\frac{\partial^2}{\partial z^2}(r^n)=n(n-2)r^{n-4}z^2+n\,r^{n-2}.$$

Hence, (1) reduces to

$$\nabla^2(r^n)=n[(n-2)r^{n-4}\left(x^2+y^2+z^2\right)+3r^{n-2}]$$
$$=n[(n-2)r^{n-2}+3\,r^{n-2}\theta]=n(n+1)r^{n-2}.$$

Putting $n=-1$, we get $\nabla^2\left(\frac{1}{r}\right)=0$.

Second Method: From Example 7.15, we have $\nabla r^n=nr^{n-2}\vec{r}$. Therefore,

$$\nabla^2(r^n)=\nabla\cdot(\nabla r^n)$$
$$=\nabla\cdot\left(nr^{n-2}\vec{r}\right)$$
$$=n\,\nabla\cdot\left(r^{n-2}\vec{r}\right).\qquad(2)$$

But, $\text{div}(\phi\vec{r})=\nabla\cdot(\phi\vec{r})=(\nabla\phi).\vec{r}+\phi(\nabla\cdot\vec{r})$. Therefore, (2) becomes

$$\nabla^2(r^n)=n\left[\left(\nabla r^{n-2}\right).\vec{r}+r^{n-2}(\nabla\cdot\vec{r})\right]$$
$$=n\left[(n-2)r^{n-4}\vec{r}.\vec{r}+3r^{n-2}\right]\text{ since }\nabla\cdot\vec{r}=3$$
$$=n[(n-2)r^{n-4}r^2+3r^{n-2}],\text{ since }\vec{r}.\vec{r}=r^2$$
$$=n(n+1)r^{n-2}.$$

EXAMPLE 7.37

Show that $\text{curl}(\text{grad}\,\vec{r}) = \nabla \times \nabla r^n = \vec{0}$.

Solution. We have seen that

$$\nabla r^n = nr^{n-2}\,\vec{r} = nr^{n-2}\left(x\hat{i} + y\hat{j} + z\hat{k}\right).$$

Therefore,

$$\nabla\times\nabla r^n = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ nr^{n-2}x & nr^{n-2}y & nr^{n-2}z \end{vmatrix}$$

$$= \sum\left[\frac{\partial}{\partial y}(nr^{n-2}z) - \frac{\partial}{\partial z}(nr^{n-2}y)\right]\hat{i}$$

$$= \sum\left[nz(n-2)r^{n-3}\frac{\partial r}{\partial y} - ny(n-2)r^{n-3}\frac{\partial r}{\partial z}\right]\hat{i}$$

$$= \sum\left[nz(n-2)r^{n-3}\frac{y}{r} - ny(n-2)r^{n-3}\frac{z}{r}\right]\hat{i}$$

$$= \sum\left[n(n-2)r^{n-4}yz - n(n-2)r^{n-4}yz\right]\hat{i}$$

$$= 0\hat{i} + 0\hat{j} + 0\hat{k} = \vec{0}.$$

EXAMPLE 7.38

If $\vec{f}$ and $\vec{g}$ are irrotational, show that $\vec{f}\times\vec{g}$ is solenoidal.

Solution. Since $\vec{f}$ and $\vec{g}$ are irrotational, we have

$$\nabla\times\vec{f} = \vec{0} \text{ and } \nabla\times\vec{g} = \vec{0}.$$

Now,

$$\text{div}\left(\vec{f}\times\vec{g}\right) = \vec{g}.\left(\nabla\times\vec{f}\right) - \vec{f}.(\nabla\times\vec{g})$$

$$= \vec{g}.\vec{0} - \vec{f}.\vec{0} = \vec{0}.$$

Hence, $\vec{f}\times\vec{g}$ is solenoidal.

EXAMPLE 7.39

Show that

$$\nabla^2 f(r) = f''(r) + \frac{2}{r}f'(r),$$

where $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$.

Solution. We have $r^2 = x^2 + y^2 + z^2$ and so, $\frac{\partial r}{\partial x} = \frac{x}{r}$, $\frac{\partial r}{\partial y} = \frac{y}{r}$, and $\frac{\partial r}{\partial z} = \frac{z}{r}$. Then,

$$\text{grad}\,r = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)r = \hat{i}\frac{\partial r}{\partial r} + \hat{j}\frac{\partial r}{\partial y} + \hat{k}\frac{\partial r}{\partial z}$$

$$= \hat{i}\,\frac{x}{r} + \hat{j}\,\frac{y}{r} + \hat{k}\,\frac{z}{r} = \frac{1}{r}\left(x\hat{i} + y\hat{j} + z\hat{k}\right) = \frac{\vec{r}}{r}. \quad (1)$$

Also, by Example 7.13,

$$\text{grad}\,f(r) = f'(r)\nabla\,r = f'(r)\text{grad}\,r. \qquad (2)$$

Therefore,

$$\nabla^2 f(r) = \nabla\cdot(\nabla f(r)) = \text{div}(\text{grad}\,f(r)) = \text{div}(f'(r)\nabla r)$$

$$= \text{div}\left[\frac{1}{r}f'(r)\vec{r}\right], \text{using}(1)$$

$$= \frac{1}{r}f'(r)\,\text{div}\vec{r} + \vec{r}.\text{grad}\left(\frac{1}{r}f'(r)\right),$$

by divergent property

$$= \frac{3}{r}f'(r) + \vec{r}.\left[\frac{d}{dr}\left(\frac{1}{r}f'(r)\right)\text{grad}\,r\right],$$

using (2) and $\text{div}\vec{r} = 3$.

$$= \frac{3}{r}f'(r) + \vec{r}.\left[-\frac{1}{r^2}f'(r) + \frac{1}{r}f''(r)\right]\frac{\vec{r}}{r},$$

using(1)

$$= \frac{3}{r}f'(r) + \frac{\vec{r}.\vec{r}}{r}\left[-\frac{1}{r^2}f'(r) + \frac{1}{r}f''(r)\right]$$

$$= \frac{3}{r}f'(r) + \frac{r^2}{r}\left[-\frac{1}{r^2}f'(r) + \frac{1}{r}f''(r)\right]$$

$$= \frac{3}{r}f'(r) + f''(r) - \frac{1}{r}f'(r) = \frac{2}{r}f'(r) + f''(r).$$

EXAMPLE 7.40

Show that $\text{div}(r^n\vec{r}) = (n+3)r^n$.

Solution. We know that

$$\text{div}\left(\phi\vec{f}\right) = (\text{grad}\,\phi).\vec{f} + \phi\,\text{div}\,\vec{f}.$$

Therefore,

$$\text{div}(r^n\vec{r}) = (\text{grad}\,r^n).\vec{r} + r^n\text{div}\vec{r}.$$

But, $\text{grad}\,r^n = nr^{n-2}\,\vec{r}$ and $\text{div}\,\vec{r} = 3$. Therefore,

$$\text{div}(r^n\,\vec{r}) = nr^{n-2}\,\vec{r}.\vec{r} + 3r^n = nr^{n-2}\,r^2 + 3r^n$$

$$= (n+3)r^n.$$

EXAMPLE 7.41

Show that $\text{curl}(\phi\,\text{grad}\,\phi) = \vec{0}$.

Solution. We have

$$\text{grad } \phi = \nabla\phi = \frac{\partial \phi}{\partial x}\hat{i} + \frac{\partial \phi}{\partial y}\hat{j} + \frac{\partial \phi}{\partial z}\hat{k}$$

and so,

$$\phi \text{ grad } \phi = \phi\frac{\partial \phi}{\partial x}\hat{i} + \phi\frac{\partial \phi}{\partial y}\hat{j} + \phi\frac{\partial \phi}{\partial z}\hat{k}.$$

Hence,

$$\text{curl}(\phi \text{grad}\phi) = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \phi\frac{\partial\phi}{\partial x} & \phi\frac{\partial\phi}{\partial y} & \phi\frac{\partial\phi}{\partial z} \end{vmatrix}$$

$$= \sum \left[\frac{\partial}{\partial y}\left(\phi\frac{\partial \phi}{\partial z}\right) - \frac{\partial}{\partial z}\left(\phi\frac{\partial \phi}{\partial y}\right)\right]\hat{i}$$

$$= \sum \left[\frac{\partial \phi}{\partial y}\cdot\frac{\partial \phi}{\partial z} + \phi\frac{\partial^2 \phi}{\partial y \partial z} - \frac{\partial \phi}{\partial z}\cdot\frac{\partial \phi}{\partial y} - \phi\frac{\partial^2 \phi}{\partial z \partial y}\right]\hat{i}$$

$$= \sum 0\hat{i} = 0\hat{i} + 0\hat{j} + 0\hat{k} = \vec{0}.$$

Also, it follows that ϕ grad ϕ is irrotational.

EXAMPLE 7.42

Show that the vector $f(r)\,\vec{r}$ is irrotational.

Solution. A vector $\vec{f}$ is irrotational if $curl\,\vec{f} = \vec{0}$. Also, we know that

$$\vec{r} \times \vec{r} = 0, \ \text{curl}\,\vec{r} = 0 \text{ and}$$

$$grad\, f(r) = f'(r) grad\ r = f'(r)\frac{\vec{r}}{r}.$$

Therefore,

$$\text{curl}[f(r)\vec{r}] = [\text{grad} f(r)] \times \vec{r} + f(r)\text{curl}\,\vec{r}$$

$$= [f'(r)\text{grad } r] \times \vec{r} + \vec{0}$$

$$= f'(r)\frac{\vec{r}}{r} \times \vec{r} + \vec{0}$$

$$= \frac{1}{r}f'(r)(\vec{r} \times \vec{r}) = \vec{0}.$$

Hence, $f(r)\,\vec{r}$ is irrotational.

7.13 INTEGRATION OF VECTOR FUNCTIONS

If $\vec{f}$ and $\vec{F}$ are two vector functions of the scalar variable t such that $\frac{d\vec{F}}{dt} = \vec{f}$, then $\vec{F}$ is called the *indefinite integral* of $\vec{f}$ with respect to t. Thus,

$$\int \vec{f}\,dt = \vec{F}.$$

If $\vec{f} = f_1\,\hat{i} + f_2\,\hat{j} + f_3\,\hat{k}$, where f_1, f_2, and f_3 are scalar functions of the scalar t, then

$$\int \vec{f}\ dt = \hat{i}\int f_1 dt + \hat{j}\int f_2 dt + \hat{k}\int f_3 dt.$$

Hence, to integrate a vector function, we integrate its components.

EXAMPLE 7.43

The acceleration of a particle at any time $t \geq 0$ is given by

$$\vec{a} = \frac{d\vec{v}}{dt} = 12\cos 2t\hat{i} - 8\sin 2t\hat{j} + 16t\hat{k}.$$

If the velocity $\vec{v}$ and displacement $\vec{r}$ are zero at $t = 0$, find $\vec{v}$ and $\vec{r}$ at any time t.

Solution. We have

$$\vec{a} = \frac{d\vec{v}}{dt} = 12\cos 2t\,\hat{i} - 8\sin 2t\,\hat{j} + 16t\,\hat{k}.$$

Integration with respect to t yields

$$\vec{v} = 6\ \sin\ 2t\,\hat{i} + 4\ \cos\ 2t\,\hat{j} + 8t^2\,\hat{k} + \vec{c},$$

where $\vec{c}$ is a constant of integration. But $\vec{v} = \vec{0}$ when $t = 0$. Therefore,

$$\vec{0} = 0\,\hat{i} + 4\,\hat{j} + 0\,\hat{k} + \vec{c}$$

and so, $\vec{c} = -4\,\hat{j}$. Therefore,

$$\vec{v} = 6\sin 2t\hat{i} + (4\cos\ 2t - 4)\hat{j} + 8t^2\,\hat{k}$$

or

$$\frac{d\vec{r}}{dt} = 6\sin\ 2t\,\hat{i} + (4\cos 2t - 4)\,\hat{j} + 8t^2\,\hat{k}.$$

Integrating again with respect to t, we get

$$\vec{r} = -3\cos\ 2t\,\hat{i} + (2\sin\ 2t - 4t)\,\hat{j} + \frac{8}{3}t^3\,\hat{k} + \vec{p},$$

where $\vec{p}$ is a constant of integration. But $\vec{r} = 0$ when $t = 0$. Therefore,

$$\vec{0} = -3\,\hat{i} + 0\,\hat{j} + 0\,\hat{k} + \vec{p}$$

and so, $\vec{p} = 3\,\hat{i}$. Hence,

$$\vec{r} = (3 - 3\cos 2t)\,\hat{i} + (2\sin\ 2t - 4t)\,\hat{j} + \frac{8}{3}t^3\,\hat{k}.$$

EXAMPLE 7.44

If $\vec{r} = \vec{0}$ when $t = 0$ and $\frac{d\vec{r}}{dt} = \vec{u}$ when $t = 0$, find the value of $\vec{r}$ satisfying the equation $\frac{d^2\vec{r}}{dt^2} = \vec{a}$, where $\vec{a}$ is a constant vector.

Solution. Integrating $\frac{d^2\vec{r}}{dt^2} = \vec{a}$ with respect to t, we get

$$\frac{d\vec{r}}{dt} = \vec{a}t + \vec{c},$$

where $\vec{c}$ is a constant vector of integration. When $t = 0$, $\frac{d\vec{r}}{dt} = \vec{u}$. Therefore,

$$\vec{u} = \vec{a}(0) + \vec{c} \text{ and so } \vec{c} = \vec{u}.$$

Therefore,

$$\frac{d\vec{r}}{dt} = \vec{a}t + \vec{u}.$$

Integrating again with respect to t, we get

$$\vec{r} = \frac{1}{2}\vec{a}\,t^2 + \vec{u}\,t + \vec{p},$$

where $\vec{p}$ is the constant vector of integration. When $t = 0$, $\vec{r} = \vec{0}$. Therefore, $\vec{0} = \vec{p}$. Hence,

$$\vec{r} = \vec{u}\,t + \frac{1}{2}\vec{a}\,t^2.$$

7.14 LINE INTEGRAL

An integral which is evaluated along a curve is called a *line integral*. Note, however, that *a line integral is not represented by the area under the curve.*

Consider any arc of the curve C enclosed between two points A and B. Let a and b be the values of the parameter t for A and B, respectively. Partition the arc between A and B into n parts as given in the following equation:

$$A = P_0, \quad P_1, \ldots, P_n = B.$$

Let $\vec{r}_0, \vec{r}_1, \ldots, \vec{r}_n$ be the position vectors of the points $P_0, P_1, \ldots, P_n$, respectively. Let $_i$ be any point on the subarc $P_{1-1}\ P_i$ and let $\delta\vec{r}_i = \vec{r}_i - \vec{r}_{i-1}$. Let $\vec{f}(\vec{r})$ be a continuous vector-point function. Then,

$$\lim_{\substack{n\to\infty \\ |\delta\,\vec{r}_i|\to 0}} \sum_{i=1}^{n} \vec{f}(\xi_i)\,.\,\delta\,\vec{r}_i, \qquad (1)$$

if exists, is called a line integral of $\vec{f}$ along C and is denoted by $\int_C \vec{f}.d\vec{r}$ or $\int_C \vec{f}.\frac{d\vec{r}}{dt}dt$. Thus, the line integral is a scalar and is also called the *tangential line integral* of $\vec{f}$ along the curve C.

If $\vec{f} = f_1\,\hat{i} + f_2\,\hat{j} + f_3\,\hat{k}$ and $\vec{r} = x\,\hat{i} + y\,\hat{j} + z\,\hat{k}$, then $d\,\vec{r} = \hat{i}dx + \hat{j}dy + \hat{k}dz$ and so,

$$\int_C \vec{f}.d\vec{r} = \int_C (f_1dx + f_2dy + f_3dz)$$

$$= \int_a^b \left(f_1\frac{dx}{dt} + f_2\frac{dy}{dt} + f_3\frac{dz}{dt}\right)dt,$$

where a and b are, respectively, values of the parameter t at the points A and B.

If we replace the dot product in (1) by a vector product, then the *vector line integral* is defined as $\int_C \vec{f} \times d\vec{r}$, which is a vector.

If C is a simple closed curve, then the tangent line integral of the vector function $\vec{f}$ around C is called the *circulation* of $\vec{f}$ around C and denoted by $\oint_C \vec{f}.d\vec{r}$.

The vector function $\vec{f}$ is said to be *irrotational* in a region R if the circulation of $\vec{f}$ around any closed curve in R is zero.

EXAMPLE 7.45

If $\vec{f} = (3x^2 + 6y)\hat{i} - 14yz\hat{j} + 20xz^2\,\hat{k}$, evaluate $\int_C \vec{f}.d\vec{r}$, where C is given by $x = t$, $y = t^2$, and $z = t^3$, and t varies from 0 to 1.

Solution. The parametric equation of C is

$$x = t,\ y = t^2, \text{and}\ z = t^3,$$

where t varies from 0 to 1.

Now, $\vec{r} = x\,\hat{i} + y\,\hat{j} + z\,\hat{k} = t\,\hat{i} + t^2\,\hat{j} + t^3\,\hat{k}$. Therefore,

$$\frac{d\vec{r}}{dt} = \hat{i} + 2t\hat{j} + 3t^2\,\hat{k}.$$

Further,

$$\begin{aligned}\vec{f} &= (3x^2 + 6y)\hat{i} - 14yz\hat{j} + 20xz^2\,\hat{k} \\ &= (3t^2 + 6t^2)\hat{i} - 14t^2.t^3\,\hat{j} + 20t.t^6\,\hat{k} \\ &= 9t^2\hat{i} - 14t^5\,\hat{j} + 20t^7\,\hat{k}.\end{aligned}$$

Therefore,

$$\int_C \vec{f}.d\vec{r} = \int_C \left(\vec{f}.\frac{d\vec{r}}{dt}\right)dt = \int_0^1 \Big[\left(9t^2\hat{i} - 14t^5\hat{j} + 20t^7\hat{k}\right) \cdot \left(\hat{i} + 2t\hat{j} + 3t^2\hat{j} + 3t^2\hat{k}\right)\Big]dt$$

$$= \int_0^1 \left(9t^2 - 28t^6 + 60t^9\right)dt$$

$$= \left[\frac{9t^3}{3} - \frac{28t^7}{7} + \frac{60t^{10}}{10}\right]_0^1 = 5.$$

EXAMPLE 7.46

If $\vec{f} = (\sin y)\hat{i} + x(1 + \cos y)\hat{j}$, evaluate the line integral $\int_C \vec{f}.d\vec{r}$ along the circular path C given by $x^2 + y^2 = a^2$ and $z = 0$.

Solution. The parametric equation of the circular path C are $x = a \cos t$, $y = a \sin t$, and $z = 0$, where t varies from 0 to 2π. Now,

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k} = (a\cos t)\hat{i} + (a\sin t)\hat{j}.$$

Therefore,

$$\frac{d\vec{r}}{dt} = (-a\sin t)\hat{i} + (a\cos t)\hat{j}.$$

Also, $\vec{f}$, in terms of parameter t, is given by

$$\vec{f} = \sin(a\sin t)\hat{i} + (a\cos t)\,(1 + \cos(a\sin t))\hat{j}.$$

Therefore,

$$\int_C \vec{f}.d\vec{r} = \int_C \left(\vec{f}.\frac{d\vec{r}}{dt}\right)dt$$

$$= \int_0^{2\pi} \left[\sin(a\sin t)\hat{i} + \left(a\cos t(1 + \cos(a\sin t))\hat{j}\right)\right] . \left[(-a\sin t)\hat{i} + (a\cos t)\hat{j}\right]dt$$

$$= \int_0^{2\pi} [-a\sin t\sin(a\sin t) + a^2\cos^2 t(1 + \cos(a\sin t))]dt$$

$$= \int_0^{2\pi} [-a\sin t\sin(a\sin t) + a^2\cos^2 t + a^2\cos^2 t\cos(a\sin t)]dt$$

$$= \int_0^{2\pi} \{d[a\cos t\sin(a\sin t)] + a^2\cos^2 t\}dt$$

$$= [a\cos t\sin(a\sin t)]_0^{2\pi} + \int_0^{2\pi} a^2\cos^2 t\,dt$$

$$= a^2\int_0^{2\pi} \frac{1 + \cos 2t}{2}dt = \frac{a^2}{2}\left[t + \frac{\sin 2t}{2}\right]_0^{2\pi}$$

$$= \frac{a^2}{2}(2\pi) = \pi a^2.$$

EXAMPLE 7.47

Calculate $\int_C \vec{f}.d\vec{r}$, where C is the part of the spiral $\vec{r} = (a\cos\theta,\ a\sin\theta,\ a\theta)$ corresponding to $0 \le \theta \le \frac{\pi}{2}$ and $\vec{f} = r^2\,\hat{i}$.

Solution. We have

$$\vec{r} = a\cos\theta\,\hat{i} + a\sin\theta\,\hat{j} + a\theta\,\hat{k} \text{ so that}$$

$$\frac{d\vec{r}}{d\theta} = -a\sin\theta\,\hat{i} + a\cos\theta\,\hat{j} + a\,\hat{k}.$$

Also,

$$\vec{f} = r^2\,\hat{i} = \left(a^2\cos^2\theta + a^2\sin^2\theta + a^2\theta^2\right)\hat{i} = [a^2\left(1 + \theta^2\right)\,\hat{i}.$$

Therefore,

$$\int_C \vec{f}.d\vec{r} = \int_C \left(\vec{f}.\frac{d\vec{r}}{d\theta}\right)d\theta = \int_0^{\frac{\pi}{2}} \{[a^2(1 + \theta^2)]\hat{i} . [(-a\sin\theta)\hat{i} + (a\cos\theta)\hat{j} + a\hat{k}]\}d\theta$$

$$= -\int_0^{\frac{\pi}{2}} a^3\left(1 + \theta^2\right)\sin\theta\,d\theta$$

$$= -a^3\int_0^{\frac{\pi}{2}} \left(\sin\theta + \theta^2\sin\theta\right)d\theta$$

$$= -a^3[\cos\theta + 2\theta\sin\theta - \theta^2\cos]_0^{\frac{\pi}{2}} = -a^3(\pi - 1).$$

EXAMPLE 7.48

If $\vec{f} = xy\,\hat{i} - z\,\hat{j} + x^2\,\hat{k}$ and C is the curve $x = t^2$, $y = 2t$, and $z = t^3$ from $t = 0$ to $t = 1$, find the *vector line integral* $\int_C \vec{f} \times d\,\vec{r}$.

Solution. We have

$$\vec{r} = x\,\hat{i} + y\hat{j}z + \hat{k} = t^2\,\hat{i} + 2t\,\hat{j} + t^3\,\hat{k},$$

$$\frac{d\,\vec{r}}{dt} = 2t\,\hat{i} + 2\,\hat{j} + 3t^2\,\hat{k},$$

and

$$\vec{f} = xy\,\hat{i} - z\,\hat{j} + x^2\,\hat{k} = 2t^3\,\hat{i} - t^3\,\hat{j} + t^4\,\hat{k}.$$

Therefore,

$$\vec{f} \times \frac{d\,\vec{r}}{dt} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 2t^3 & -t^3 & t^4 \\ 2t & 2 & 3t^2 \end{vmatrix}$$

$$= \left(-3t^5 - 2t^4\right)\hat{i} + \left(-4t^5\right)\hat{j} + \left(4t^3 + 2t^4\right)\hat{k}.$$

Hence,

$$\int_C \vec{f} \times d\vec{r} = \int_0^1 \left(\vec{f} \times \frac{d\vec{r}}{dt}\right) dt$$

$$= \int_0^1 \left[\left(-3t^5 - 2t^4\right)\hat{i} + \left(-4t^5\right)\hat{j} + \left(4t^3 + 2t^4\right)\hat{k}\right] dt$$

$$= \left[-\frac{3t^6}{6} - \frac{2t^5}{5}\right]_0^1 \hat{i} + \left[\left(-4\frac{t^6}{6}\right)\right]_0^1 \hat{j}$$

$$+ \left[4\frac{t^4}{4} + \frac{2t^5}{5}\right]_0^1 \hat{k}$$

$$= -\frac{9}{10}\hat{i} - \frac{2}{3}\hat{j} + \frac{7}{5}\hat{k}.$$

EXAMPLE 7.49

If $\vec{f} = (5xy - 6x^2)\hat{i} + (2y - 4x)\hat{j}$, evaluate $\int_C \vec{f}.d\vec{r}$ along the curve C in the xy plane $y = x^3$ from the point (1, 1) to (2, 8).

Solution. Substituting $x = t$, we get $y = t^3$. When $x = 1$, $t = 1$and when $x = 2$, $t = 2$. Then,

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k} = t\hat{i} + t^3\hat{j}, \quad \frac{d\vec{r}}{dt} = \hat{i} + 3t^2\hat{j} \quad \text{and}$$

$$\vec{f} = \left(5xy - 6x^2\right)\hat{i} + (2y - 4x)\hat{j} = \left(5t^4 - 6t^2\right)\hat{i} + \left(2t^3 - 4t\right)\hat{j}.$$

Hence,

$$\int_c \vec{f}.d\,\vec{r} = \int_1^2 \left(\vec{f}.\frac{d\,\vec{r}}{dt}\right) dt = \int_1^2 \{[(5t^4 - 6t^2)\hat{i}$$

$$+ (2t^3 - 4t)\hat{j}] \cdot (\hat{i} + 3t^2\hat{j})\}dt$$

$$= \int_1^2 [(5t^4 - 6t^2) + 3t^2(2t^3 - 4t)]dt$$

$$= \int_1^2 [6t^5 + 5t^4 - 12t^3 - 6t^2]\,dt$$

$$= \left[6\frac{t^6}{6} + 5\frac{t^5}{5} - 12\frac{t^4}{4} - \frac{6t^3}{3}\right]_1^2$$

$$= [t^6 + t^5 - 3t^4 - 2t^3]_1^2 = 35.$$

EXAMPLE 7.50

Evaluate $\int_A^B [2xydx + (x^2 - y^2)dy]$ along the arc of the circle $x^2 + y^2 = 1$ in the first quadrant from A (1, 0) to B(0, 1).

Solution. On the circle, $y = \sqrt{1 - x^2}$ so that $\frac{dy}{dx} = -x\,(1 - x^2)^{-\frac{1}{2}}$ or $dy = -2x(1 - x^2)^{-\frac{1}{2}}dx$. Therefore,

$$\int_A^B \left[2xydx + \left(x^2 - y^2\right)dy\right]$$

$$= \int_1^0 \left[2x\left(1 - x^2\right)^{\frac{1}{2}}dx - \left(2x^2 - 1\right)x\left(1 - x^2\right)^{-\frac{1}{2}}dx\right]$$

$$= \left[-\frac{2}{3}\left(1 - x^2\right)^{\frac{3}{2}}\right]_1^0 - \int_1^0 \left(x^2 - \frac{1}{2}\right)2x\left(1 - x^2\right)^{-\frac{1}{2}}dx$$

$$= -\frac{2}{3} + \left[2\left(x^2 - \frac{1}{2}\right)\left(1 - x^2\right)^{\frac{1}{2}}\right]_1^0$$

$$-2\int_1^0 2x\left(1 - x^2\right)^{\frac{1}{2}}dx$$

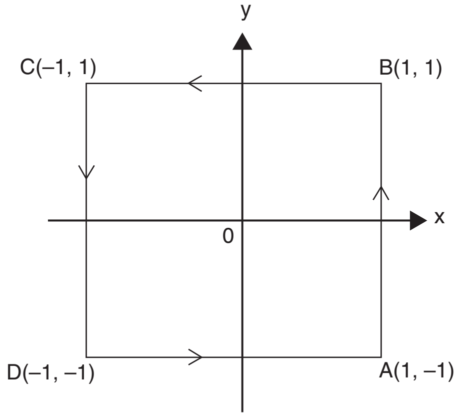
$$= -\frac{2}{3} - 1 + 2\left[\frac{\left(1 - x^2\right)^{\frac{3}{2}}}{\frac{3}{2}}\right]_1^0$$

$$= -\frac{2}{3} - 1 + \frac{4}{3} = -\frac{1}{3}.$$

EXAMPLE 7.51

Evaluate $\int_C [(x^2 + xy)dx + (x^2 + y^2)dy]$, where C is the square formed by the lines $y = \pm 1$ and $x = \pm 1$.

Solution. The curve C is as shown in the following figure:



We note that

$$\int_C [(x^2+xy)dx+(x^2+y^2)dy] = \int_{AB} + \int_{BC} + \int_{CD} + \int_{DA}.$$

Along AB, we have $x = 1$ and so, $dx = 0$. Also along AB, y varies from -1 to 1. Thus,

$$\int_{AB} [(x^2 + xy)dx + (x^2 + y^2)dy]$$

$$= \int_{-1}^{1} (1 + y^2)dy = \left[y + \frac{y^3}{3}\right]_{-1}^{1} = \frac{8}{3}.$$

Along BC, we have $y = 1$ so that $dy = 0$. Also along BC, x varies from 1 to -1. Thus,

$$\int_{BC} [(x^2 + xy)dx + (x^2 + y^2)\ dy]$$

$$= \int_{-1}^{1} (x^2 + x)dx = \left[\frac{x^3}{3} + \frac{x^2}{2}\right]_{1}^{-1} = -\frac{2}{3}.$$

Similarly,

$$\int_{CD} [(x^2 + xy)dx + (x^2 + y^2)\ dy]$$

$$= \int_{1}^{-1} (1 + y^2)dy = -\frac{8}{3} \text{ and}$$

$$\int_{DA} [(x^2 + xy)dx + (x^2 + y^2)\ dy]$$

$$= \int_{-1}^{1} (x^2 + x)dx = \frac{2}{3}.$$

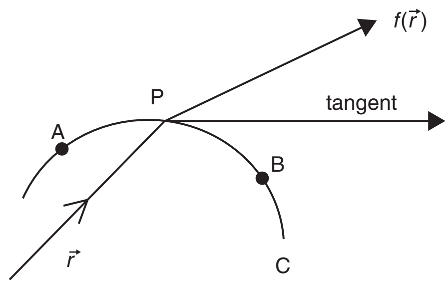
Hence,

$$\int_c [(x^2 + xy)dx + (x^2 + y^2)\ dy] = \frac{8}{3} - \frac{2}{3} - \frac{8}{3} + \frac{2}{3} = 0.$$

7.15 WORK DONE BY A FORCE

The work done as the point of application of a force $\vec{f}$ moves along a given path C can be expressed as a line integral. In fact, the work done, when the point of application moves from $P(\vec{r})$ to Q $(\vec{r} + \delta\vec{r})$, where $\vec{PQ} = \delta\vec{r}$, is

$$\delta W = |\ \delta\ \vec{r}|\ \left|\vec{f}\right| \cos\theta = \vec{f}.\delta\ \vec{r}.$$



Therefore, the total work done as P moves from A to B is

$$W = \int_A^B \vec{f}.d\ \vec{r}.$$

Now, suppose that the force $\vec{f}$ is conservative. Then, there exists a scalar function ϕ such that $\vec{f} = -\text{grad}\,\phi$, that is, $\vec{f} = -\nabla\phi = -\left[\hat{i}\frac{\partial\phi}{\partial x} + \hat{j}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}\right]$.

Therefore, the work done in this case is given by

$$W = \int_A^B -\left[\hat{i}\frac{\partial\phi}{\partial x} + \hat{j}\frac{\partial\phi}{\partial y} + \hat{j}\frac{\partial\phi}{\partial z}\right].\left(\hat{i}\,dx + \hat{j}\,dy + \hat{k}\,dz\right)$$

$$= \int_B^A \left[\frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz\right]$$

$$= \int_B^A d\phi = [\phi]_B^A = \phi_A - \phi_B.$$

Hence, in a conservative field, the work done depends on A and B and is the same for all paths joining A and B. Thus, in the case of conservative force, $\vec{f}(\vec{r}).d\,\vec{r}$ is an *exact differential* $-d\phi$. In such a case, ϕ is called the *potential energy*. The forces which do not have this property are said to be *dissipative* or *nonconservative*.

EXAMPLE 7.52

Find the total work done in moving a particle in a force field, given by $\vec{f} = 3xy\,\hat{i} - 5z\,\hat{j} + 10x\,\hat{k}$, along the curve $x = t^2 + 1$, $y = 2t^2$, and $z = t^3$, from $t = 1$ to $t = 2$.

Solution. The parametric equation of the curve is

$$x = t^2 + 1,\ y = 2t^2,\ \text{and } z = t^3,\ 1 \le t \le 2.$$

We have

$$\vec{f} = 3xy\,\hat{i} - 5z\,\hat{j} + 10x\,\hat{k}$$
$$= 3\left(t^2 + 1\right)\ \left(2t^2\right)\hat{i} - 5t^3\hat{j} + 10\left(t^2 + 1\right)\hat{k}$$
$$= 6\left(t^4 + t^2\right)\hat{i} - 5t^3\hat{j} + 10\left(t^2 + 1\right)\hat{k}$$

and

$$\vec{r} = x\,\hat{i} + y\,\hat{j} + z\,\hat{k} = \left(t^2 + 1\right)\hat{i} + 2t^2\hat{j} + t^3\hat{k}.$$

Therefore, $\frac{d\,\vec{r}}{dt} = 2t\,\hat{i} + 4t\,\hat{j} + 3t^2\,\hat{k}$ and so, the total work done is given by

$$W = \int_C \vec{f}.d\,\vec{r} = \int_1^2 \vec{f}.\left(\frac{d\,\vec{r}}{dt}\right)dt$$

$$= \int_1^2 \left[6\left(t^4 + t^2\right)\hat{i} - 5t^3\hat{j} + 10\left(t^2 + 1\right)\hat{k}\right] \cdot \left[2t\,\hat{i} + 4t\,\hat{j} + 3t^2\hat{k}\right]dt$$

$$= \int_1^2 \left[12\left(t^5 + t^3\right) - 20t^4 + 30\left(t^4 + t^2\right)\right]dt$$

$$= \int_1^2 \left(12t^2 + 10t^4 + 12t^3 + 30t^2\right)dt$$

$$= \left[12\frac{t^6}{6} + 10\frac{t^5}{5} - 12\frac{t^4}{4} + 30\frac{t^3}{3}\right]_1^2$$

$$= 320 - 17 = 303.$$

EXAMPLE 7.53

Find the work done by the force

$$\vec{f} = (2y + 3)\hat{i} + xz\,\hat{j} + (yz - x)\hat{k},$$

when it moves a particle from the point (0, 0, 0) to the point (2, 1, 1) along the curve $x = 2t^2$, $y = t$, and $z = t^3$.

Solution. The parametric equations of the curve are $x = 2t^2$, $y = t$, and $z = t^3$. Further,

$$\vec{f} = (2y + 3)\hat{i} + xz\,\hat{j} + (yz - x)\hat{k}$$
$$= (2t + 3)\hat{i} + 2t^5\hat{j} + \left(t^4 - 2t^2\right)\hat{k}$$

and

$$\vec{r} = x\,\hat{i} + y\,\hat{j} + z\,\hat{k} = 2t^2\,\hat{i} + t\,\hat{j} + t^3\,\hat{k}.$$

Therefore,

$$\frac{d\,\vec{r}}{dt} = 4t\,\hat{i} + \hat{j} + 3t^2\,\hat{k}.$$

The given points (0, 0, 0) and (2, 1, 1) correspond to $t = 0$ and $t = 1$. Therefore, the work done by the force is given by

$$W = \int_c \vec{f}.d\,\vec{r}$$

$$= \int_0^1 \left[(2t+3)\hat{i} + 2t^5\hat{j} + \left(t^4 - 2t^2\right)\hat{k}\right] \cdot \left[4t\,\hat{i} + \hat{j} + 3t^2\hat{k}\right] dt$$

$$= \int_0^1 \left[(2t+3)4t + 2t^5 + 3\left(t^4 - 2t^2\right)\, t^2\right] dt$$

$$= \int_0^1 \left[8t^2 + 12t + 2t^5 + 3t^6 - 6t^4\right] dt$$

$$= \int_0^1 \left[3t^6 + 2t^5 - 6t^4 + 8t^2 + 12t\right] dt$$

$$= \left[3\frac{t^7}{7} + 2\frac{t^6}{6} - 6\frac{t^5}{5} + 8\frac{t^3}{3} + 12\frac{t^2}{2}\right]_0^1$$

$$= \frac{3}{7} + \frac{1}{3} - \frac{6}{5} + \frac{8}{3} + 6 = \frac{288}{35}.$$

EXAMPLE 7.54

Find the work done in moving a particle in the force field $\vec{f} = 3x^2\,\hat{i} + (2xz - y)\hat{j} + z\,\hat{k}$ along

(i) the straight line from (0, 0, 0,) to (2, 1, 3) and

(ii) the curve defined by $x^2 = 4y$ and $3x^2 = 8z$ from $x = 0$ to $x = 2$.

Solution. (i) The curve C is the line joining (0, 0, 0) to (2, 1, 3) whose equation is

$$\frac{x-0}{2-0} = \frac{y-0}{1-0} = \frac{z-0}{3-0} \text{ or } \frac{x}{2} = \frac{y}{1} = \frac{z}{3} = t, \text{ say.}$$

Thus, $x = 2t$, $y = t$, and $z = 3t$ are the parametric equations of the line. The point (0, 0, 0) corresponds to $t = 0$ and the point (2, 1, 3) corresponds to $t = 1$. Also,

$$\vec{f} = 3x^2\,\hat{i} + (2xz - y)\hat{j} + z\,\hat{k}$$
$$= 12t^2\,\hat{i} + \left(12t^2 - t\right)\hat{j} + 3t\,\hat{k}$$

and

$$\vec{r} = x\,\hat{i} + y\,\hat{j} + z\,\hat{k} = 2t\,\hat{i} + t\,\hat{j} + 3t\,\hat{k}.$$

Therefore, $\frac{d\vec{r}}{dt} = 2\,\hat{i} + \hat{j} + 3\,\hat{k}$. Hence,

$$W = \int_C \vec{f}.d\,\vec{r} = \int_0^1 \left(\vec{f}.\frac{d\,\vec{r}}{dr}\right) dt$$

$$= \int_0^1 \left\{\left[12t^2\,\hat{i} + \left(12t^2 - t\right)\hat{j} + 3t\,\hat{k}\right] \cdot \left[2\,\hat{i} + \hat{j} + 3\,\hat{k}\right]\right\} dt$$

$$= \int_0^1 \left[24t^2 + \left(12t^2 - t\right) + 9t\right] dt$$

$$= \int_0^1 \left(36t^2 + 8t\right) dt = \left[12t^3 + 4t^2\right]_0^1 = 16.$$

(ii) Putting $x = t$ in the given curve, we get $y = \frac{t^2}{4}$ and $z = \frac{3t^3}{3}$, where $0 \le t \le 2$. Then,

$$\vec{f} = 3x^2\,\hat{i} + (2xz - y)\hat{j} + z\,\hat{k}$$
$$= 3t^2\,\hat{i} + \left(\frac{3}{4}t^4 - \frac{t^2}{4}\right)\hat{j} + \frac{3}{8}t^3\hat{k}$$

and

$$\vec{r} = x\,\hat{i} + y\,\hat{j} + z\,\hat{k} = t\,\hat{i} + \frac{t^2}{4}\hat{j} + \frac{3}{8}t^3\,\hat{k}, \quad 0 \le t \le 2.$$

Therefore,

$$\frac{d\,\vec{r}}{dt} = \hat{i} + \frac{1}{2}t\,\hat{j} + \frac{9}{8}t^2\,\hat{k}.$$

Hence,

$$W = \int_C \vec{f}.d\,\vec{r} = \int_0^2 \left(\vec{f}.\frac{d\,\vec{r}}{dt}\right) dt$$

$$= \int_0^2 \left(3t^2 + \frac{3}{8}t^5 - \frac{1}{8}t^3 + \frac{27}{64}t^5\right) dt$$

$$= \left[t^3 + \frac{t^6}{16} - \frac{t^4}{32} + \frac{9}{128}t^6\right]_0^2 = 16.$$

7.16 SURFACE INTEGRAL

An integral evaluated over a surface is called a *surface integral*. Two types of surface integral exist:

(i) $\iint_S f(x,y,z)dS$

and

(ii) $\iint_S \vec{f}(\vec{r}) \cdot \hat{n}\, dS = \iint_S \vec{f}(\vec{r}).d\,\vec{S}$.

In case (i), we have a scalar field f, whereas in case (ii), we have a vector field $\vec{f}(\vec{r})$, vector element of area $d\,\vec{S} = \hat{n}\, dS$, and $\hat{n}$ the outward-drawn unit normal vector to the element dS.

(i) Let $f(x,y,z)$ be a scalar-point function defined over a surface S of finite area. Partition the area S into n subareas $\delta S_1, \delta S_2, \ldots, \delta S_n$. In each area δS_i, choose an arbitrary point $P_i(x_i, y_i, z_i)$. Define $f(P_i) = f(x_i, y_i, z_i)$ and form the sum $\sum_{i=1}^{n} f(x_1, y_i, z_i)\delta S_i$. Then, the limit of this sum as $n \to \infty$ in such a way that the largest of the subarea δS_i approaches zero is called the *surface integral* of $f(x,y,z)$ over S and is denoted by $\iint_S f(x,y,z)dS$.

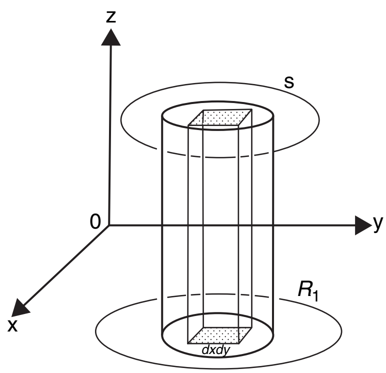
(ii) Now, let $\vec{f}$ be a vector-point function defined and continuous over a surface S. Let P be any point on the surface S and let $\hat{n}$ be the unit vector at P in the direction of the outward-drawn normal to the surface S at P. Then, $\vec{f}.\,\hat{n}$ is the normal component of $\vec{f}$ at P. The integral of $\vec{f}.\,\hat{n}$ over S is called the *normal surface integral* of $\vec{f}$ over S and is denoted by $\iint_S \vec{f}.\hat{n}\, dS$. This integral is also known as *flux of $\vec{f}$ over S*. If we associate with the differential of surface area dS, a vector $d\vec{S}$, with magnitude dS, and whose direction is that of $\hat{n}$, then $d\vec{S} = \hat{n}dS$ and hence,

$$\iint_S \vec{f}.\hat{n}\, dS = \iint \vec{f}.d\,\vec{S}.$$

The surface integrals are easily evaluated by expressing them as double integrals, taken over an orthogonal projection of the surface S on any of the coordinate planes. But, the condition for this is that any line perpendicular to the coordinate plane chosen meets the surface S in not more than one point. However, if S does not satisfy this condition, then S can be subdivided into surfaces satisfying this condition.

Let S be the surface such that any line perpendicular to the xy-plane meets S in not more than one point. Then, the equation of the surface S can be written as $z = h(x,y)$. Let R_1 be the projection of S on the xy-plane. Then, the projection of dS on the xy-plane is $dS \cos\gamma$, where γ is the acute angle which the normal $\hat{n}$ at P to the surface S makes with z-axis. Therefore,

$$dS\cos\gamma = dx\ dy.$$



But $\cos\gamma = |\hat{n}.\,\hat{k}|$, where $\vec{k}$ is, as usual, a unit vector along the z-axis. Thus,

$$dS = \frac{dxdy}{|\hat{n}\ .\ \hat{k}|}.$$

Hence,

$$\iint_S \vec{f}.\hat{n}\, dS = \iint_{R_1} \vec{f}.\hat{n}\frac{dxdy}{|\hat{n}\ .\ \hat{k}|}.$$

Similarly, if R_2 and R_3 are projections of S on the yz, and zx-plane, respectively, then

$$\iint_S \vec{f}.\ \hat{n}\ dS = \iint_{R_2} \vec{f}.\ \hat{n}\ \frac{dxdy}{|\hat{n}\ .\hat{i}|},$$

and

$$\iint_S \vec{f}.\ \hat{n}\ dS = \iint_{R_3} \vec{f}.\ \hat{n}\ \frac{dxdy}{|\hat{n}\ .\hat{j}|}.$$

EXAMPLE 7.55

Evaluate $\iint_S \vec{f}.\,\hat{n}\,dS$, where $\vec{f} = yz\,\hat{i} + zx\,\hat{j} + zy\,\hat{k}$ and S is that part of the surface of the sphere $x^2 + y^2 + z^2 = 1$, which lies in the first octant.

Solution. A vector normal to the surface of the given sphere is

$$\nabla(x^2 + y^2 + z^2 - 1) = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right) \times (x^2 + y^2 + z^2 - 1) = 2x\,\hat{i} + 2y\,\hat{j} + 2z\,\hat{k}.$$

Therefore, the unit normal to any point (x,y,z) of the surface is

$$\hat{n} = \frac{2x\,\hat{i} + 2y\,\hat{j} + 2z\,\hat{k}}{|2x\,\hat{i} + 2y\,\hat{j} + 2z\,\hat{k}|} = \frac{2x\,\hat{i} + 2y\,\hat{j} + 2z\,\hat{k}}{\sqrt{4x^2 + 4y^2 + 4z^2}} = \frac{2x\,\hat{i} + 2y\,\hat{j} + 2z\,\hat{k}}{2\sqrt{x^2 + y^2 + z^2}} = x\,\hat{i} + y\,\hat{j} + z\,\hat{k},$$

since $x^2 + y^2 + z^2 = 1$ on S. Now,

$$\vec{f}.\,\hat{n} = \left(yz\,\hat{i} + zx\,\hat{j} + xy\,\hat{k}\right).\left(x\,\hat{i} + y\,\hat{j} + z\,\hat{k}\right) = xyz + xyz + xyz = 3xyz$$

and

$$\vec{n}\,.\,\hat{k} = .\left(x\,\hat{i} + y\,\hat{j} + z\,\hat{k}\right).\,\hat{k} = z,$$

which gives $|\hat{n}\,.\,\hat{k}| = z$. Hence, in the first quadrant,

$$\iint_S \vec{f}.\,\hat{n}\,dS = \iint_S \vec{f}.\,\hat{n}\frac{dxdy}{|\hat{n}\,.\,\hat{k}|} = \iint_R \frac{3xyz}{z}dxdy = 3\iint_R xy\,dx\,dy = 3\int_0^1\left[\int_0^{\sqrt{1-x^2}} xy\,dy\right]dx = 3\int_0^1 x\left[\frac{y^2}{2}\right]_0^{\sqrt{1-x^2}} dx = \frac{3}{2}\int_0^1 x(1 - x^2)dx = \frac{3}{2}\left[\frac{x^2}{2} - \frac{x^4}{4}\right]_0^1 = \frac{3}{8}.$$

EXAMPLE 7.56

Evaluate $\iint_S \vec{f}.\,\hat{n}\,dS$, where $\vec{f} = 18z\,\hat{i} - 12\,\hat{j} + 3y\,\hat{k}$ and S is the surface $2x + 3y + 6z = 12$ in the first octant.

Solution. A vector normal to the surface S is

$$\nabla(2x + 3y + 6z = 12) = 2z\,\hat{i} + 3\,\hat{j} + 6\,\hat{k}.$$

Therefore, the unit normal vector to the surface S is

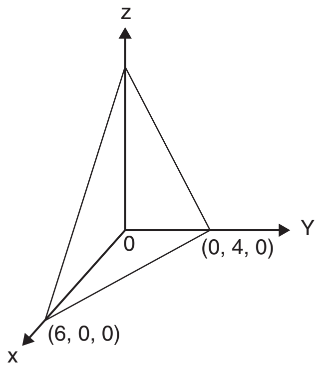
$$\hat{n} = \frac{2z\,\hat{i} + 3\hat{j} + 6\,\hat{k}}{\sqrt{4 + 9 + 36}} = \frac{2z\,\hat{i} + 3\hat{j} + 6\,\hat{k}}{7}$$

and so,

$$\hat{n}\,.\,\hat{k} = \left(\frac{2\hat{i} + 3\hat{j} + 6\hat{k}}{7}\right).\hat{k} = \frac{6}{7}.$$

Also,

$$\vec{f}.\,\hat{n} = \left(18z\,\hat{i} - 12\hat{j} + 3y\,\hat{k}\right)\cdot\left(\frac{2\,\hat{i} + 3\hat{j} + 6\,\hat{k}}{7}\right) = \frac{36z}{7} - \frac{36}{7} + \frac{18}{7}y = \frac{36}{7}\left[\frac{12 - 2x - 3y}{6}\right] - \frac{36}{7} + \frac{18}{7}y = \frac{36}{7} - \frac{12x}{7}.$$



Hence,

$$\iint_S \vec{f}.\,\hat{n}\,dS = \iint_R \vec{f}.\,\hat{n}\,\frac{dxdy}{|\hat{n}\,.\,\hat{k}|} = \frac{7}{6}\iint_R\left(\frac{36}{7} - \frac{12x}{7}\right)dx\,dy = \iint_R (6 - 2x)dx\,dy.$$

But, R is the region of projection of S (triangle) on the xy plane. Thus, the projection is a triangle bounded by x-axis, y-axis, and the line $2x+3y=12$ and $z=0$. Hence, the limits of x are from 0 to 6 and that of y are from 0 to $\frac{12-2x}{3}$. Therefore,

$$\iint_S \vec{f}.\,\hat{n}\;dS = \int_0^6 \left[\int_0^{\frac{12-2x}{3}} (6-2x)dy\right]dx$$

$$= \int_0^6 [6y-2xy]_0^{\frac{12-2x}{3}}\,dx$$

$$= \frac{1}{3}\int_0^6 \left(72-36x+4x^2\right)dx$$

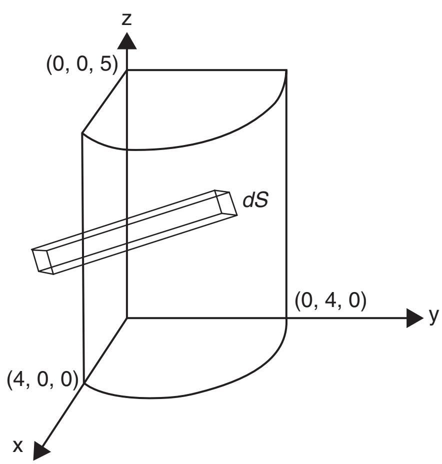
$$= \frac{1}{3}\left[72x-36\frac{x^2}{2}+\frac{4x^3}{3}\right]_0^6$$

$$= 144-216+96 = 24.$$

EXAMPLE 7.57

Evaluate $\iint_S \vec{f}.d\,\vec{S}$, where $\vec{f} = z\,\hat{i}+x\,\hat{j}-3y^2z\,\hat{k}$ and S is the surface of the cylinder $x^2+y^2=16$ in the first octant between $z=0$ and $z=5$.

Solution. The surface S is shown in the following figure:



A vector normal to the surface S is given by

$$\nabla\left(x^2+y^2-16\right) = 2x\,\hat{i}+2y\,\hat{j},$$

so that the unit normal vector $\hat{n}$ at any point of S is

$$\hat{n} = \frac{2x\,\hat{i}+2y\,\hat{j}}{\sqrt{4x^2+4y^2}} = \frac{2x\,\hat{i}+2y\,\hat{j}}{2\sqrt{x^2+y^2}}$$

$$= \frac{2x\,\hat{i}+2y\,\hat{j}}{8} = \frac{x\,\hat{i}+y\,\hat{j}}{4}.$$

Also,

$$\vec{f}.\hat{n} = \left(z\,\hat{i}+x\hat{j}-3y^2z\,\hat{k}\right).\left(\frac{x\,\hat{i}+y\hat{j}}{4}\right) = \frac{1}{4}(xz+xy).$$

Let R be the projection of the surface S on xz-plane. Then

$$\hat{n}\;.\,\hat{j} = \frac{1}{4}\left(x\,\hat{i}+y\,\hat{j}\right).\,(\hat{j}) = \frac{y}{4}.$$

Hence

$$\iint_S \vec{f}.\,dS = \iint_S \vec{f}.\,\hat{n}\,dS = \iint_R \frac{xz+xy}{4}\cdot\frac{dxdz}{|\hat{n}\;.\,\hat{j}|}$$

$$= \iint_R \frac{xz+xy}{4}\cdot\frac{dxdz}{\frac{y}{4}} = \iint_R \frac{xz+xy}{y}.dx\,dz,$$

where R is the rectangular region in the xz-plane bounded by $0\le x\le 4$, $0\le z\le 5$. Since the integrand is still evaluated on the surface, we have $y=\sqrt{16-x^2}$ and so,

$$\iint_S \vec{f}.d\vec{S} = \int_0^4 \left[\int_0^5 \left(x+\frac{xz}{\sqrt{16-x^2}}\right)dz\right]dx$$

$$= \int_0^4 \left[xz+\frac{xz^2}{2\sqrt{16-x^2}}\right]_0^5 dx$$

$$= \int_0^4 \left(5x+\frac{25x}{2\sqrt{16-x^2}}\right)dx$$

$$= \left[5\frac{x^2}{2}-\frac{25}{2}\sqrt{16-x^2}\right]_0^4 = 90.$$

EXAMPLE 7.58

Evaluate $\iint_S \vec{f}.\,\hat{n}\;dS$, where $\vec{f} = 4x\,\hat{i}-2y^2\hat{j}+z^2\hat{k}$ and S is the surface bounding the region $x^2+y^3=4$, $z=0$, and $z=3$.

Solution. The region is bounded by the cylinder $x^2 + y^3 = 4$, $z = 0$, and $z = 3$. Therefore, the surface S consists of three parts:

(i) S_1, the circular base of the cylinder in the plane $z = 0$,

(ii) S_2, the circular top in the plane $z = 3$, and

(iii) S_3, the curved surface of the cylinder given by $x^2 + y^2 = 4$.

Now, for the subsurface S_1, we have $z = 0$, $\hat{n} = -\hat{k}$, and $\vec{f} = 4x\,\hat{i} - 2y^2\hat{j}$.
Therefore,

$$\vec{f}\,.\,\hat{n} = \left(4x\,\hat{i} - 2y^2\hat{j}\right)\,.\,\left(-\,\hat{k}\right) = 0.$$

Hence, $\iint\limits_{S_1} \vec{f}.\hat{n}\; dS = 0.$

On S_2, we have $z = 3$, $\hat{n} = \hat{k}$, and $\vec{f} = 4x\,\hat{i} - 2y^2\hat{j} + 9\,\hat{k}$. Therefore,

$$\vec{f}\,.\,\hat{n} = \left(4x\,\hat{i} - 2y^2\hat{j} + 9\,\hat{k}\right).\left(\hat{k}\right) = 9.$$

Hence,

$$\iint\limits_{S_2} \vec{f}\,.\,\hat{n}\; dS = \iint\limits_{S_2} 9dxdy \;= 9(\pi.4) = 36\pi.$$

For the surface S_3, which is the curved surface of the cylinder and is given by $x^2 + y^2 = 4$, the vector normal to the surface is

$$\nabla\left(x^2 + y^2 - 4\right) = 2x\,\hat{i} + 2y\,\hat{j}.$$

Therefore, the unit normal vector to the surface S_3 is given by

$$\hat{n} = \frac{2x\,\hat{i} + 2y\,\hat{j}}{\sqrt{4x^2 + 4y^2}} = \frac{2x\,\hat{i} + 2y\,\hat{j}}{2\sqrt{x^2 + y^2}}$$

$$= \frac{2x\,\hat{i} + 2y\,\hat{j}}{2.2}$$

$$= \frac{x\,\hat{i} + y\,\hat{j}}{2}.$$

Therefore,

$$\vec{f}\,.\,\hat{n} = \left(4x\,\hat{i} - 2y^2\hat{j} + z^2\hat{k}\right).\left(\frac{x\,\hat{i} + y\,\hat{j}}{2}\right) = 2x^2 - y^3.$$

Now, on S_3, $x = 2\cos\theta$, $y = 2\sin\theta$, and $dS = 2d\theta\, dz$. For this surface, z varies from 0 to 3 and θ varies from 0 to 2π. Therefore,

$$\iint\limits_{S_3} \vec{f}\,.\,\hat{n}\; dS = \int_0^{2\pi}\int_0^{3} \left[2(2\cos\theta)^2 - (2\sin\theta)^3\right] 2d\theta\; dz$$

$$= \int_0^{2\pi} 16\left(\cos^2\theta - \sin^3\theta\right)[z]_0^3 d\theta$$

$$= 48\int_0^{2\pi} \left(\cos^2\theta - \sin^3\theta\right) d\theta$$

$$= 48\int_0^{2\pi} \cos^2\theta\; d\theta - 48\int_0^{2\pi} \sin^3\theta\; d\theta$$

$$= (48)(4)\int_0^{\frac{\pi}{2}} \cos^2\theta\; d\theta - 0$$

$$= 192\;.\frac{1}{2}.\frac{\pi}{2}$$

$$= 48\pi.$$

Hence,

$$\iint\limits_{S} \vec{f}\,.\,\hat{n}\; dS = \iint\limits_{S_1} \vec{f}\,.\,\hat{n}\; dS + \iint\limits_{S_2} \vec{f}\,.\,\hat{n}\; dS + \iint\limits_{S_3} \vec{f}\,.\,\hat{n}\; dS$$

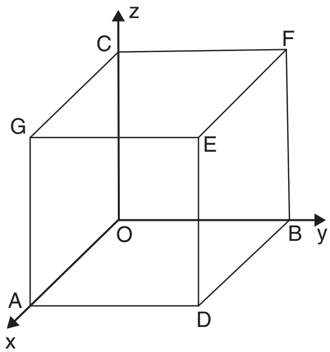
$$= 0 + 36\pi + 48\pi$$

$$= 84\pi.$$

EXAMPLE 7.59

If $\vec{f} = 4xz\,\hat{i} - y^2\hat{j} + yz\,\hat{k}$, evaluate $\int\limits_S \vec{f}\,.\,\hat{n}\; dS$, where S is the surface of the cube bounded by $x = 0, x = 1$, $y = 0, y = 1, z = 0$, and $z = 1$.

Solution. The surface of the cube is bounded by $x = 0$, $x = 1$, $y = 0$, $y = 1$, $z = 0$, and $z = 1$ and so, the surface can be subdivided into six parts in the following manner:



(i) S_1 is the surface formed by the face OADB, where

$$z = 0,\ \hat{n} = -\hat{k},\ \text{and}\ \vec{f} = -y^2\hat{j},\ \text{so that}$$

$$\vec{f}\,.\,\hat{n} = \left(-y^2\hat{j}\right)\,.\,\left(-\hat{k}\right) = 0$$

and

$$\iint_{S_1} \vec{f}\,.\,\hat{n}\ dS = \int_0^1\int_0^1 0 dxdy = 0.$$

(ii) S_2 is the surface formed by the face GEFC, where $z = 1$ $\hat{n} = \hat{k}$, and $dz = 0$. On this face, we have

$$\vec{f} = 4xz\ \hat{i} - y^2\hat{j} + yz\ \hat{k}\ = 4x\ \hat{i} - y^2\hat{j} + y\ \hat{k}$$

and so,

$$\vec{f}\,.\,\hat{n}\ = \left(4x\ \hat{i} - y^2\hat{j} + y\ \hat{k}\right)\,.\,\left(\hat{k}\right) = y.$$

Hence,

$$\iint_{S_2} \vec{f}\,.\,\hat{n}\ dS = \int_0^1\int_0^1 ydxdy = \int_0^1 \left[\frac{y^2}{2}\right]_0^1 dx$$

$$= \frac{1}{2}\int_0^1 dx = \frac{1}{2}.$$

(iii) S_3 is the surface formed by the face ADEG, where $\hat{n} = \hat{i}$, $x = 1$, and $dx = 0$. On this face,

$$\vec{f} = 4xz\,\hat{i} - y^2\hat{j} + yz\,\hat{k}\ = 4z\,\hat{i} - y^2\hat{j} + yz\,\hat{k}\ \text{and}$$

$$\vec{f}\,.\,\hat{n} = \left(4z\,\hat{i} - y^2\hat{j} + yz\,\hat{k}\right)\,.\,\hat{i} = 4z.$$

Hence,

$$\iint_{S_3} \vec{f}\,.\,\hat{n}\ ds = \int_0^1\int_0^1 4z\ dy\,dx$$

$$= 4\int_0^1 \left[\frac{z^2}{2}\right]_0^1 dy = 2\int_0^1 dy = 2.$$

(iv) S_4 is the surface formed by OBFC, where $\hat{n} = -\hat{i}$, $x = 0$, and $dx = 0$. On this face,

$$\vec{f} = 4xz\ \hat{i} - y^2\hat{j} + yz\ \hat{k}\ = -y^2\hat{j} + yz\ \hat{k}\ \text{and}$$

$$\vec{f}\,.\,\hat{n} = \left(-y^2\hat{j} + yz\ \hat{k}\right)\cdot\left(-\hat{i}\right) = 0.$$

Hence,

$$\iint_{S_4} \vec{f}.\hat{n}\ dS = \int_0^1\int_0^1 0\,dx\,dy = 0.$$

(v) S_5 is the surface formed by the face OCGA, where $\hat{n} = -\hat{j}$, $y = 0$, and $dy = 0$. On this face, $\vec{f} = 4xz\ \hat{i} - y^2\hat{j} + yz\ \hat{k} = 4xz\ \hat{i}$ and so,

$$\vec{f}\,.\,\hat{n} = \left(4xz\ \hat{i}\right)\,.\,\left(-\hat{j}\right) = 0.\ \text{Hence,}$$

$$\iint_{S_5} \vec{f}\,.\,\hat{n}\ dS = 0.$$

(vi) S_6 is the surface formed by the face DBFE, where $\hat{n} = \hat{j}$, $y = 1$, and $dy = 0$. On this face,

$$\vec{f} = 4xz\ \hat{i} - y^2\hat{j} + yz\ \hat{k} = 4xz\ \hat{i} - \hat{j} + z\ \hat{k}\ \text{and}$$

$$\vec{f}.\hat{n} = \left(4xz\ \hat{i} - \hat{j} + z\ \hat{k}\right)\cdot\left(\hat{j}\right) = -1.$$

Therefore,

$$\iint_{S_6} \vec{f}.\hat{n}\ dS = \int_0^1\int_0^1 (-1)dz\,dx = -\int_0^1 [z]_0^1 dx$$

$$= -1\int_0^1 dx = -1.$$

Hence,

$$\iint_S \vec{f}.\hat{n}\ ds = \iint_{S_1} + \iint_{S_2} + \iint_{S_3} + \iint_{S_4}$$

$$+ \iint_{S_5} + \iint_{S_6}$$

$$= 0 + \frac{1}{2} + 2 + 0 + 0 + -1 = \frac{3}{2}.$$

7.17 VOLUME INTEGRAL

Let ϕ be a scalar-point function defined throughout a given region of volume V. Partition the given region into n subregions of volumes $\delta V_1, \delta V_2, \ldots, \delta V_n$. Let $P(x_i, y_i, z_i)$ be any point inside or on the boundary of the subregion of volume δV_i. Then the limit

$$\lim_{\substack{n\to\infty \\ \delta V_i \to 0}} \sum_{i=1}^{n} \phi(P_i)\delta V_i,$$

if exists, for all mode of subdivision (partition), is called the *volume integral* of ϕ over the volume V, and this integral is denoted by $\iiint_V \phi \, dV$.

If we partition the region of volume V into small cuboids, by drawing lines parallel to the coordinate axes, then $dV = dx\ dy\ dz$ and so,

$$\iiint_V \phi \, dV = \iiint_V \phi \, dx\ dy\ dz.$$

Similarly, if $\vec{f} = f_1\ \hat{i} + f_2\ \hat{j} + f_3\ \hat{k}$ is a vector-point function, then

$$\iiint_V \vec{f} dV = \hat{i} \iiint_V f_1(x,y,z) dx\ dy\ dz + \hat{j} \iiint_V f_2(x,y,z) dx\ dy\ dz + \hat{k} \iiint_V f_3(x,y,z) dx\ dy\ dz.$$

EXAMPLE 7.60

If $\vec{f} = (2x^2 - 3z)\hat{i} - 2xy\ \hat{j} - 4x\ \hat{k}$, evaluate $\iiint_V \nabla \times \vec{f} dV$, where V is the region bounded by the coordinate planes and the plane $2x + 2y + z = 4$.

Solution. We have

$$\nabla \times \vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 2x^2 - 3z & -2xy & -4x \end{vmatrix}$$

$$= \hat{i}\left[\frac{\partial}{\partial y}(-4x) - \frac{\partial}{\partial z}(-2xy)\right] + \hat{j}\left[\frac{\partial}{\partial z}(2x^2 - 3z) - \frac{\partial}{\partial x}(-4x)\right] + \hat{k}\left[\frac{\partial}{\partial x}(-2xy) - \frac{\partial}{\partial y}(2x^2 - 3z)\right]$$

$$= \hat{j} - 2y\hat{k}.$$

The region V is bounded by the planes $x = 0, y = 0, z = 0$ and the plane $2x + 2y + z = 4$. Therefore, the limits of integration are:

z varies from 0 to $4 - 2x - 2y$,
y varies from 0 to $2 - x$, and
x varies from 0 to 2.

Hence,

$$\iiint_V \nabla \times \vec{f}\ dV$$

$$= \int_0^2 \int_0^{2-x} \int_0^{4-2x-2y} \left(\hat{j} - 2y\ \hat{k}\right) dx\ dy\ dz$$

$$= \int_0^2 \int_0^{2-x} \left(\hat{j} - 2y\ \hat{k}\right) [z]_0^{4-2x-2y} dx\ dy$$

$$= \int_0^2 \int_0^{2-x} \left[(4 - 2x - 2y)\hat{j} - \left(8y - 4xy + 4y^2\right)\hat{k}\right] dx\, dy$$

$$= \int_0^2 \left[\left(4y - 2xy - 2\frac{y^2}{2}\right)\hat{j} - \left(8\frac{y^2}{2} - 4x\frac{y^2}{2} + 4\frac{y^3}{3}\right)\hat{k}\right]_0^{2-x}$$

$$= \int_0^2 \left[(2 - x)^2\hat{j} - \frac{2}{3}(2 - x)^3\hat{k}\right] dx$$

$$= \left[\frac{(2 - x)^3}{-3}\hat{j} - \frac{2}{3}\frac{(2 - x)^4}{-4}\hat{k}\right]_0^2$$

$$= \frac{8}{3}\hat{j} - \frac{8}{3}\hat{k} = \frac{8}{3}\left(\hat{j} - \hat{k}\right).$$

EXAMPLE 7.61

If $\vec{f} = (2x^2 - 3z)\hat{i} - 2xy\ \hat{j} - 4x\ \hat{k}$, then evaluate $\iiint_V \nabla \cdot \vec{f}\ dV$, where V is bounded by the coordinate planes and the plane $2x + 2y + z = 4$.

Solution. We have

$$\nabla \cdot \vec{f} = \left(\frac{\partial}{\partial x}\hat{i} + \frac{\partial}{\partial y}\hat{j} + \frac{\partial}{\partial z}\hat{k}\right) \cdot \left[(2x^2 - 3z)\hat{i} - 2xy\ \hat{j} - 4x\ \hat{k}\right]$$

$$= 4x - 2x = 2x.$$

The limits of integration are as mentioned in Example 7.60. Therefore,

$$\iiint_V \nabla\cdot\vec{f}\,dV = \int_0^2\int_0^{2-x}\int_0^{4-2x-2y} dx\,dy\,dz$$

$$= 2\int_0^2\int_0^{2-x} x(4-2x-2y)dx\,dy$$

$$= 4\int_0^2\int_0^{2-x} x(2-x-y)dx\,dy$$

$$= \frac{4}{2}\int_0^2 x(2-x)^2dx$$

$$= 2\left[\frac{x^2}{4}-\frac{4x^3}{3}+\frac{4x^2}{2}\right]_0^2 = \frac{8}{3}.$$

EXAMPLE 7.62

Evaluate $\iiint_V \nabla\cdot\vec{f}\,dV$ if $\vec{f} = 4xy\,\hat{i} + yz\,\hat{j} - xy\,\hat{k}$ and V is bounded by $x = 0, x = 2, y = 0, y = 2, z = 0$ and $z = 2$.

Solution. We have

$$\iiint_V \nabla\cdot\vec{f}$$

$$= \iiint_V \left(\frac{\partial}{\partial x}(4xy) + \frac{\partial}{\partial y}(yz) + \frac{\partial}{\partial z}(-xy)\right)dV$$

$$= \iiint_V (4y+z)dV = \int_0^2\int_0^2\int_0^2 (4y+z)dz\,dy\,dx$$

$$= \int_0^2\int_0^2 \left[4yz+\frac{z^2}{2}\right]_0^2 dy\,dx = 2\int_0^2\int_0^2 (4y+1)dy\,dx$$

$$= 2\int_0^2 \left[\frac{4y^2}{2}+y\right]_0^2 dx = 4\int_0^2 5dx = 20[x]_0^2 = 40.$$

7.18 GAUSS'S DIVERGENCE THEOREM

The following theorem of Gauss is useful in evaluating the surface integral over a closed surface by reducing it to a volume integral (triple integral) and vice versa.

Theorem 7.10. (Gauss's Divergence Theorem). Let $\vec{f}$ be a vector-point function possessing continuous first-order partial derivatives at each point of a three-dimensional region V enclosed in a closed surface S. Then,

$$\iint_S \vec{f}.\,\hat{n}\,dS = \iiint_V \operatorname{div}\vec{f}\,dV = \iiint_V \nabla\cdot\vec{f}\,dV,$$

where $\hat{n}$ is the outward-drawn unit normal vector to the surface S.

The divergence theorem can be expressed in the form of Cartesian coordinates as follows:

Let $\vec{f} = f_1\,\hat{i} + f_2\,\hat{j} + f_3\,\hat{k}$. Then

$$\operatorname{div}\vec{f} = \nabla\cdot\vec{f} = \frac{\partial f_1}{\partial x}+\frac{\partial f_2}{\partial y}+\frac{\partial f_3}{\partial z}.$$

Let the outward-drawn unit normal vector $\hat{n}$ makes angles α, β, and γ, respectively, with positive directions of x-, y-, and z-axis. Thus, cos α, cos β, and cos γ are the direction cosines of $\hat{n}$ and so,

$$\hat{n} = \cos\alpha\,\hat{i} + \cos\beta\hat{j} + \cos\gamma\,\hat{k}$$

and then,

$$\vec{f}.\,\hat{n} = \left(f_1\,\hat{i}+f_2\,\hat{j}+f_3\,\hat{k}\right).\left(\cos\alpha\,\hat{i}+\cos\beta\hat{j}+\cos\gamma\hat{k}\right)$$
$$= f_1\cos\alpha + f_2\cos\beta + f_3\cos\gamma.$$

Hence, the Gauss's Divergence Theorem takes the form

$$\iiint_V \left(\frac{\partial f_1}{\partial x}+\frac{\partial f_2}{\partial y}+\frac{\partial f_3}{\partial z}\right)dx\,dy\,dz$$

$$= \iint_S (f_1dy\,dz + f_2dz\,dx + f_3dx\,dy),$$

since cos α $dS = dy\,dz$, cos β $dS = dz\,dx$, and cos γ $dS = dxdy$. This form of Gauss's Divergence Theorem is also known as *Green's Theorem in Space.*

Proof: Consider a closed surface S, which is such that it is possible to introduce a rectangular coordinate system, such that any line parallel to any coordinate axis cuts S in, at the most, two points. Let R be the projection of the surface S on the xy-plane. Then, in accordance to our assumption, a line through a point (x, y, 0) of R meets the boundary of S in two points. Suppose that the z coordinates of these points are $z = \phi_1(x, y)$ and $z = \phi_2(x, y)$, where

$\phi_2(x, y) \geq \phi_1(x, y)$. Then,

$$\iiint_V \frac{\partial f_3}{\partial z} dV = \iiint_V \frac{\partial f_3}{\partial z} dz\, dy\, dx$$

$$= \iint_R \left[\int_{\phi_1(x,y)}^{\phi_2(x,y)} \frac{\partial f_3}{\partial z} dz \right] dy\, dx$$

$$= \iint_R [f_3(x,y,z)]_{\phi_1(x,y)}^{\phi_2(x,y)} dy\, dx$$

$$= \iint_R [f_3(x,y,\phi_2) - f_3(x,y,\phi_1)]\, dy\, dx. \quad (1)$$

Let S_1 and S_2 be the portion of the surface S corresponding to $z = \phi_1\,(x, y)$ and $z = \phi_2\,(x, y)$, respectively. Let $\hat{n}_2$ be the outward-drawn unit normal vector to S_2, making an acute angle γ_2 with the positive direction $\left(\hat{k}\right)$ of z-axis. If dS_2 is projected on the xy-plane, then this projection $dy\, dx$ of dS_2 is

$$dy\, dx = \cos\gamma_2 dS_2 = \hat{k}.\, \hat{k}_2\; dS_2.$$

In the same fashion, let $\hat{n}_1$ be the outward-drawn unit normal vector to S_1, making an obtuse angle γ_1 with $\hat{k}$. Then,

$$dy\, dx = \cos(\pi - \gamma_1) dS_1 = -\cos\gamma_1 dS_1 = -\hat{k}.\, \hat{n}\; dS_1.$$

Therefore,

$$\iint_R f_3(x, y, \phi_2) dy\, dx = \iint_{S_2} f_3\; \hat{k}.\hat{n}_2 dS_2$$

and

$$\iint_R f_3(x, y, \phi_1) dy\, dx = -\iint_{S_1} f_3\; \hat{k}.\; \hat{n}_1\; dS_1.$$

Hence, (1) reduces to

$$\iiint_V \frac{\partial f_3}{\partial z} dV = \iint_{S_2} f_3\; \hat{k}.\hat{n}_2\; dS_2 + \iint_{S_1} f_3\; \hat{k}.\; \hat{n}_1\; dS_1$$

$$= \iint_S f_3\; \hat{k}.\; \hat{n}\; dS. \quad (2)$$

Similarly, projecting S on the remaining two coordinate planes, we have

$$\iiint_V \frac{\partial f_1}{\partial x} dV = \iint_S f_1\; \hat{i}.\; \hat{n}\; dS \quad \text{and} \quad (3)$$

$$\iiint_V \frac{\partial f_2}{\partial y} dV = \iint_S f_2\; \hat{j}.\; \hat{n}\; dS. \quad (4)$$

Adding (2), (3), and (4), we obtain

$$\iiint_V \left(\frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} \right) dV$$

$$= \iint_S \left(f_1\; \hat{i} + f_2\; \hat{j} + f_3.\; \hat{k} \right).\hat{n}\; dS$$

$$= \iint_S \vec{f}.\; \hat{n}\; dS.$$

This proves the theorem. The Gauss's Divergence Theorem can be extended to the surfaces which are such that lines parallel to the coordinate axes meet them in more than two points. For this, the region enclosed by S is partitioned into subregions whose surfaces satisfy the condition assumed in the above proof. Applying the theorem to each subregion and adding will yield the required result.

Deductions:

(i) If $\hat{n}$ is the outward-drawn unit normal vector to S, then

$$\iiint_V \vec{f}.\nabla\phi dV = \iint_S \phi\vec{f}.\; \hat{n}\; dS - \iiint_V \phi\; \text{div}\,\vec{f}\; dV.$$

Proof: By Gauss's Divergence Theorem,

$$\iint_S \vec{F}.\; \hat{n}\; dS = \iiint_V \text{div}\; \vec{F}\; dV.$$

Putting $\vec{F} = \phi\vec{f}$, we have

$$\iint_S \phi\,\vec{f}.\,\hat{n}\,dS = \iiint_V \operatorname{div}\left(\phi\,\vec{f}\right)dV$$

$$= \iiint_V \left(\phi\,\operatorname{div}\vec{f} + \vec{f}\,.\nabla\phi\right)dV$$

$$= \iiint_V \phi\,\operatorname{div}\vec{f}\,dV + \iiint_V \vec{f}\,.\nabla\phi dV$$

and so,

$$\iiint_V \vec{f}.\nabla\phi dV = \iint_S \phi\vec{f}.\hat{n}\,dS - \iiint_V \phi\,\operatorname{div}\vec{f}\,dV.$$

(ii) $\iint_S \vec{f} \times \hat{n}\,dS = -\iiint_V \operatorname{curl}\vec{f}\,dV.$

Proof: By Gauss's Divergence Theorem,

$$\iint_S \vec{F}.\,\hat{n}\,dS = \iiint_V \operatorname{div}\vec{F}\,dV.$$

Putting $\vec{F} = \vec{a} \times \vec{f}$, where $\vec{a}$ is an arbitrary constant vector, we have

$$\iint_S \left(\vec{a} \times \vec{f}\right).\,\hat{n}\,dS = \iiint_V \nabla\cdot\left(\vec{a} \times \vec{f}\right)dV$$

or

$$\iint_S \vec{a}.\,\left(\vec{f} \times \hat{n}\right)dS = -\iiint_S \vec{a}.\left(\nabla \times \vec{f}\right)dV$$

or

$$\vec{a}.\iint_S \left(\vec{f} \times \hat{n}\right)dS = -\vec{a}.\iiint_V \left(\nabla \times \vec{f}\right)dV$$

or

$$\vec{a}.\left[\iint_S \left(\vec{f} \times \hat{n}\right)dS + \iiint_V \left(\nabla \times \vec{f}\right)dV\right] = 0$$

or

$$\iint_S \left(\vec{f} \times \hat{n}\right). + \iiint_V \left(\nabla \times \vec{f}\right)dV = 0,$$

that is,

$$\iint_S \left(\vec{f} \times \hat{n}\right)dS = -\iiint_V \operatorname{curl}\vec{f}\,dV.$$

(iii) $\iint_S \phi\,\hat{n}\,dS = \iiint_V \operatorname{grad}\phi\,dV.$

Proof: By Gauss's Divergence Theorem

$$\iint_S \vec{f}.\hat{n}\,dS = \iiint_V \operatorname{div}\vec{f}\,dV.$$

Putting $\vec{f} = \vec{a}\,\phi$, where $\vec{a}$ is on arbitrary constant vector, we get

$$\iint_S (\vec{a}\,\phi)\,.\,\hat{n}\,dS = \iiint_V \operatorname{div}(\vec{a}\,\phi)dV$$

$$= \iiint_V \nabla\cdot(\vec{a}\,\phi)dV.$$

Thus,

$$\vec{a}.\iint_S (\phi\,\hat{n})dS = \vec{a}.\iiint_V (\nabla\phi)dV$$

or

$$\vec{a}.\left[\iint_S (\phi\,\hat{n})dS - \iiint_V (\nabla\phi)dV\right] = 0$$

or

$$\iint_S (\phi\,\hat{n})dS - \iiint_V (\nabla\phi)dV = 0$$

or

$$\iint_S \phi\,\hat{n}\,dS = \iiint_V \nabla\phi\,dV.$$

EXAMPLE 7.63

If S is a closed surface, $\hat{n}$ is the outward-drawn normal to S and V is the volume enclosed by S, show that

(i) $\iint_S \vec{r}.\,\hat{n}\,dS = 3V,$

(ii) $\iiint_V \operatorname{div}\hat{n}\,dV = S,$

(iii) $\iint_S \hat{n}\,dS = \vec{0},$ and

(iv) $\iint_S \vec{f}.\hat{n}\,dS = 6V,$

where $\vec{f} = x\,\hat{i} + 2y\,\hat{j} + 3z\,\hat{k}.$

Solution. (i) By the divergence theorem,

$$\iint_S \vec{r}.\,\hat{n}\,ds = \iiint_V \text{div}\,\vec{r}\,dV = \iiint_V \nabla\cdot\,\vec{r}\,dV$$
$$= \iiint_V \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)\cdot\left(x\,\hat{i}+y\,\hat{j}+z\,\hat{k}\right)dV$$
$$= 3\iiint_V dV = 3V.$$

(ii) By the divergence theorem,

$$\iiint_V \text{div}\,\hat{n}\,dV = \iint_S \hat{n}\,.\,\hat{n}\,dS$$
$$= \iint_S dS = S.$$

(iii) If $\hat{a}$ is any constant vector, then

$$\vec{a}.\iint_S \hat{n}\,dS = \iint_S \vec{a}.\,\hat{n}\,dS$$
$$= \iiint_V \text{div}\,\vec{a}\,dV,$$

by divergence theorem

$$= \vec{0},\ \text{because div}\,\vec{a} = \vec{0}.$$

(iv) By the Gauss's Divergence Theorem

$$\iint_S \vec{f}.\,\hat{n}\,dS = \iiint_V \nabla\cdot\vec{f}\,dV$$
$$= \iiint_V \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)\cdot\left(x\,\hat{i}+2y\,\hat{j}+3z\,\hat{k}\right)dV$$
$$= \iiint_V (1+2+3)dV$$
$$= 6\iiint_V dV = 6V.$$

EXAMPLE 7.64

Verify the divergence theorem for $\vec{f} = 4x\,\hat{i} - 2y^2\hat{j} + z^2\hat{k}$, taken over the region bounded by the cylinder $x^2 + y^2 = 4$, $z = 0$ and $z = 3$.

Solution. In Example 7.58, we have shown that

$$\iint_S \vec{f}.\,\hat{n}\,dS = 84\pi.$$

On the other hand,

$$\iiint_V \text{div}\vec{f}\,dV = \iiint_V \nabla\cdot\vec{f}\,dV$$
$$= \iiint_V \left[\frac{\partial}{\partial x}(4x)+\frac{\partial}{\partial y}(-2y^2)+\frac{\partial}{\partial z}(z^2)\right]dV$$
$$= \iiint_V (4-4y+3z)dx\,dy\,dz.$$

Since z varies from 0 to 3, y varies from $-\sqrt{4-x^2}$ to $\sqrt{4-x^2}$, and x varies from -2 to 2, we have

$$\iiint_V \text{div}\,\vec{f}\,dV$$
$$= \int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\left[\int_0^3 (4-4y+2z)dz\right]dy\,dx$$
$$= \int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\left[4z-4yz+2\frac{z^2}{2}\right]_0^3 dy\,dx$$
$$= \int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}[12-12y+9]\,dy\,dx$$
$$= 21\int_{-2}^{2}\left[\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} dy\right]dx,$$

since 12y is an odd function of y

$$= 42\int_{-2}^{2}\left[\int_0^{\sqrt{4-x^2}} dy\right]dx = 42\int_{-2}^{2}\sqrt{4-x^2}dx$$
$$= 84\int_0^2 \sqrt{4-x^2}\,dx,\ \text{since integrand is even}$$
$$= 84\left[\frac{x\sqrt{4-x^2}}{2}+\frac{4}{2}\sin^{-1}\frac{x}{2}\right]_0^2$$
$$= 84\left[0+2\sin^{-1}1\right] = 84\left[\frac{2\pi}{2}\right] = 84\pi.$$

Therefore,

$$\iint_S \vec{f}.\,\hat{n}\,dS = \iiint_V \text{div}\,\vec{f}\,dV$$

and thus, Gauss's Divergence Theorem is verified for the given function.

EXAMPLE 7.65

Verify Gauss's Divergence Theorem for the function $\vec{f} = 4xz\,\hat{i} - y^2\hat{j} + yz\,\hat{k}$ over the surface S of the cube bounded by $x = 0$, $x = 1$, $y = 0$, $y = 1$, $z = 0$, and $z = 1$.

Solution. In Example 7.59, we have shown that

$$\iint_S \vec{f}.\hat{n}\,dS = \frac{3}{2}.$$

On the other hand,

$$\iiint_V \text{div}\vec{f}\,dv$$
$$= \iiint_V \nabla\cdot\vec{f}\,dV$$
$$= \iiint_V \left[\frac{\partial}{\partial x}(4xz)+\frac{\partial}{\partial y}(-y^2)+\frac{\partial}{\partial z}(yz)\right]dV$$
$$= \iiint_V (4z-2y+y)dV = \iiint_V (4z-y)dV$$
$$= \int_0^1\int_0^1\left[\int_0^1 (4z-y)dz\right]dy\,dx = \int_0^1\int_0^1\left[\frac{4z^2}{2}-yz\right]_0^1 dy\,dx$$
$$= \int_0^1\left[\int_0^1 (2-y)dy\right]dx = \int_0^1\left[2y-\frac{y^2}{2}\right]_0^1 dx$$
$$= \int_0^1\left(2-\frac{1}{2}\right)dx = \frac{3}{2}\int_0^1 dx = \frac{3}{2}.$$

Therefore,

$$\iint_S \vec{f}.\,\hat{n}\,dS = \iiint_V \text{div}\,\vec{f}\,dV$$

and thus, the Gauss's Divergence Theorem is verified.

EXAMPLE 7.66

Evaluate $\iint_S \vec{f}.\,\hat{n}\,dS$, where $\vec{f} = ax\,\hat{i} + by\,\hat{j} + cz\,\hat{k}$ and S is the surface of the sphere $x^2 + y^2 + z^2 = 1$.

Solution. Since S is closed, by the divergence theorem, we have

$$\iint_S \vec{f}.\,\hat{n}\,dS = \iiint_V \text{div}\,\vec{f}\,dV$$
$$= \iiint_V \nabla\cdot\vec{f}\,dV$$
$$= \iiint_V \left[\frac{\partial}{\partial x}(ax) + \frac{\partial}{\partial y}(by) + \frac{\partial}{\partial z}(cz)\right]dV$$
$$= \iiint_V (a+b+c)dV$$
$$= (a+b+c)\iiint_V dV$$
$$= (a+b+c)\left(\frac{4}{3}\pi\right),$$

since $\iiint_V dV$ = volume of the sphere $x^2 + y^2 + z^2 = 1$, which is $\frac{4}{3}\pi.\,1^3$.

EXAMPLE 7.67

Verify Gauss's Divergence Theorem for $\vec{f} = (x^3 - yz)\hat{i} - 2x^2y\vec{j} + 2\hat{k}$, taken over the cube bounded by the planes $x = 0$, $x = a$, $y = 0$, $y = a$, $z = 0$, and $z = a$.

Solution. The surface F is a cube with six faces as shown in the following figure:

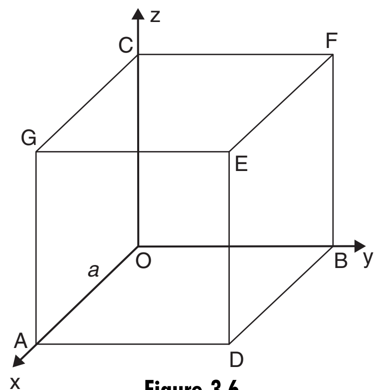


Figure 3.6

To calculate $\iint_S \vec{f}.\,\hat{n}\,dS$, we evaluate $\iint_S \vec{f}.\,\hat{n}\,dS$ over the six faces and then add those values.

(i) For the face OADB, we have $\hat{n} = -\hat{k}$ and $z = 0$. Therefore,

$$\iint_{OADB} \vec{f}.\hat{n}\,dS$$

$$= \iint_{OADB} \left[(x^3 - yz)\hat{i} - 2x^2y\hat{j} + 2\hat{k}\right].\left(-\hat{k}\right)dS$$

$$= \int_0^a\int_0^a (-2)dxdy = -2\int_0^a [y]_0^a dx$$

$$= -2a\int_0^a dx = -2a[x]_0^a$$

$$= -2a^2. \quad (1)$$

(ii) For the face CGEF, we have $\hat{n} = \hat{k}$ and $z = a$. Therefore,

$$\iint_{CGEF} \vec{f}.\,\hat{n}\,dS = \int_0^a\int_0^a 2dx\,dy$$

$$= 2a\int_0^a dx$$

$$= 2a^2. \quad (2)$$

(iii) For the face ADEG, we have $\hat{n} = \hat{i}$, $x = a$, and $dx = 0$. Therefore,

$$\iint_{ADEG} \vec{f}.\,\hat{n}\,dS = \int_0^a\int_0^a (a^3 - yz)dydz$$

$$= \int_0^a \left[a^3z - y\frac{z^2}{2}\right]_0^a dy$$

$$= \int_0^1 \left(a^4 - \frac{a^2y}{2}\right)dy$$

$$= \left[a^4y - \frac{a^2y^2}{4}\right]_0^a$$

$$= a^5 - \frac{a^4}{4}. \quad (3)$$

(iv) For the face OBFC, we have $\hat{n} = -\hat{i}$, $x = 0$, and $dx = 0$. Therefore,

$$\iint_{OBFC} \vec{f}.\,\hat{n}\,dS = \int_0^a\int_0^a (yz)dydz$$

$$= \int_0^a \left[y\frac{z^2}{2}\right]_0^a$$

$$= \frac{1}{2}\int_0^a a^2y\,dy$$

$$= \frac{a^2}{2}\left[\frac{y^2}{2}\right]_0^a$$

$$= \frac{a^2}{4}. \quad (4)$$

(v) For the face OAGC, we have $\hat{n} = -\hat{j}$, $y = 0$, and $dy = 0$. Therefore,

$$\iint_{\mathrm{OAGC}} \vec{f}.\hat{n}\,dS = \int_0^a\int_0^a 0\,dx\,dz = 0. \quad (5)$$

(vi) For the face DBFE, we have $\hat{n} = \hat{j}$, $y = a$, and $dy = 0$. Therefore,

$$\iint_{DBFE} \vec{f}.\hat{n}\,dS$$

$$= \int_0^a\int_0^a -2x^2a\,dxdz = -2a\int_0^a [x^2z]_0^a dx$$

$$= -2a^2\int_0^a x^2dx = -2a^2\left[\frac{x^3}{3}\right]_0^a$$

$$= -\frac{2a^5}{3}. \quad (6)$$

Adding (1)–(6), we get

$$\iint_S \vec{f}.\hat{n}\,dS = -2a^2 + 2a^2 + a^5$$

$$-\frac{a^4}{4} + \frac{a^2}{4} + 0 - \frac{2a^5}{3} = \frac{a^5}{3}.$$

On the other hand,

$$\iiint_V \operatorname{div}\vec{f}\, dV = \iiint_V \nabla\cdot\vec{f}\, dV$$

$$= \iiint_V \left[\frac{\partial}{\partial x}(x^3 - yz) + \frac{\partial}{\partial y}(-2x^2y) + \frac{\partial}{\partial z}(2)\right] dV$$

$$= \iiint_V (3x^2 - 2x^2)dV = \int_0^a\int_0^a\int_0^a x^2 dx\, dy\, dz$$

$$= \int_0^a\int_0^a [x^2z]_0^a\, dy\, dx = \int_a^a\left[\int_0^a x^2 a\, dy\right] dx$$

$$= a\int_0^a [x^2y]_0^a\, dx = a^2\int_0^a x^2 dx = a^2\left[\frac{x^3}{3}\right]_0^a = \frac{a^5}{3}.$$

Thus,

$$\iint_S \vec{f}.\hat{n}\, dS = \iiint_V \operatorname{div}\vec{f} dV$$

and thereby, Gauss's Divergence Theorem is verified.

EXAMPLE 7.68

Evaluate $\iint_S (x^3 - yz)\, dy\, dz - 2x^2y\, dz\, dx + z\, dx\, dy$ over the surface bounded by the coordinate planes and the planes $x = y = z = a$.

Solution. By giving divergence theorem in Cartesian form, we have

$$\iint_S (f_1 dy\, dz + f_2 dz\, dx + f_3 dx\, dy)$$

$$= \iiint_V \left(\frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}\right) dV$$

$$= \iiint_V \left[\frac{\partial}{\partial x}(x^3 - yz) + \frac{\partial}{\partial y}(-2x^2y) + \frac{\partial}{\partial z}(z)\right] dV$$

$$= \iiint_V (x^2 + 1)dV = \int_0^a\int_0^a\int_0^a (x^2 + 1)dx\, dy\, dz$$

$$= \int_0^a\int_0^a [x^2z + z]_0^a\, dydz = \int_0^a\left[\int_0^a a(x^2+1)dy\right] dx$$

$$= a\int_0^a [x^2y + y]_0^a\, dx = a^2\int_0^a (x^2+1)dx$$

$$= a^2\left[\frac{x^3}{3} + x\right]_0^a = \frac{a^5}{3} + a^3.$$

EXAMPLE 7.69

Using Green's Theorem in space, evaluate $\iint_S (4xzdy\, dz - y^2 dz\, dx + yzdx\, dy)$, where S is the surface of a cube bounded by the planes $x = 0$, $y = 0$, $z = 0$, $x = 1$, $y = 1$, and $z = 1$.

Solution. Let

$$f_1 = 4xz,\quad f_2 = -y^2,\ \text{and}\ f_3 = yz.$$

Then, by Green's Theorem in space, we have

$$\iint_S (f_1 dy\, dz + f_2 dz\, dx + f_3 dx\, dy)$$

$$= \iiint_V \left(\frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}\right) dx\, dy\, dz$$

$$= \iiint_V \left(\frac{\partial}{\partial x}(4xz) + \frac{\partial}{\partial y}(-y^2) + \frac{\partial}{\partial z}(yz)\right) dx\, dy\, dz$$

$$= \iiint_V (4z - y)dx\, dy\, dz$$

$$= \int_0^1\int_0^1 \left[4\frac{z^2}{2} - yz\right]_0^1 dy\, dx = \int_0^1\left[\int_0^1 (2-y)dy\right] dx$$

$$= \int_0^1 \left[2y - \frac{y^2}{2}\right]_0^1 = \frac{3}{2}\int_0^1 dx = \frac{3}{2}.$$

7.19 GREEN'S THEOREM IN A PLANE

A domain D is said to be a *quadratic* with respect to y-axis, if it is bounded by the curves of the form

$$y = \phi(x),\ y = \psi(x):\ x = a,\ x = b,$$

where ϕ and ψ are continuous functions and $\phi(x) \geq \psi(x)$ for all $x \in [a, b]$. Thus, a domain which is quadratic with respect to y-axis is such that a line parallel to y-axis and lying between $x = a$ and $x = b$ meets the boundary of D in just two points. Similarly, we can define domains which are quadratic with respect to x-axis.

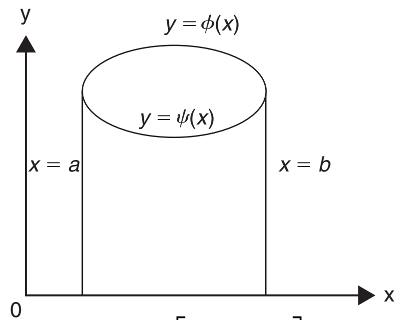
The Green's Theorem is useful in changing a line integral around a *closed curve C* into a double integral over the region R enclosed by C.

Theorem 7.11. (Green's Theorem). Let f, g, $\frac{\partial f}{\partial y}$, and $\frac{\partial f}{\partial y}$ are continuous in a region R, which can be split up in finite number of regions quadratic with respect to either axis. Then,

$$\oint_C [f(x,y)dx + g(x,y)dy] = \iint_R \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx\,dy,$$

where the integral on the left is a line integral around the boundary C of the region, taken in such a way that the interior of the region remains on the left as the boundary is described.

Proof: Consider the region R bounded by the curves $x = a$, $x = b$, $y = \phi(x)$, and $y = \psi(x)$, such that $\phi(x) \geq \psi(x)$ for all $x \in [a, b]$. Let f be a real-valued continuous function defined in R, and let $\frac{\partial f}{\partial y}$ exists and is continuous in R. Then,



$$\iint_R \frac{\partial f}{\partial y} dxdy = \int_a^b \left[\int_{\psi(x)}^{\phi(x)} \frac{\partial f}{\partial y} dy\right] dx$$

$$= \int_a^b f(x, .\phi(x))dx - \int_a^b f(x, \psi(x))dx$$

$$= -\int_b^a f(x, \phi(x))dx - \int_a^b f(x, \psi(x))dx$$

$$= -\left[\int_a^b f(x, \psi(x))dx + \int_b^a f(x, \phi(x))dx\right]$$

$$= -\oint_C f(x,y)dx.$$

Therefore,

$$\oint_C f(x,y)dx = -\iint_R \frac{\partial f}{\partial y} dx\,dy. \qquad (1)$$

Similarly, it can be shown that

$$\oint_C g(x,y)dy = \iint_R \frac{\partial g}{\partial x} dx\,dy. \qquad (2)$$

Adding (1) and (2), we obtain

$$\oint_C [f(x,y)dx + g(x,y)dy] = \iint_R \left[\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right] dx\,dy.$$

Deductions:

(i) If $f(x, y) = -y$ and $g(x, y) = x$, then by Green's Theorem, we have

$$\oint_C (xdy - ydx) = \iint_R (1+1)dxdy$$

$$= 2\iint_R dxdy = 2A,$$

where A denotes the area of the region R. Thus,

$$A = \frac{1}{2}\oint_C [xdy - ydx].$$

(ii) Putting $f(x, y) = -y$ and $g(x, y) = 0$, the Green's Theorem implies

$$-\oint ydx = \iint_R dxdy = \text{ Area of the region R.}$$

(iii) Putting $g(x, y) = x$ and $f(x, y) = 0$, we get

$$\oint_C xdy = \iint_R dx\,dy = \text{ Area of the region R.}$$

Hence, the area of a closed region R is given by any of the three formulae

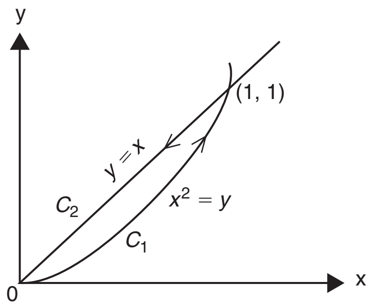
$$\oint_C xdy, \ -\oint_C ydx, \ \text{or } \frac{1}{2}\oint_C (xdy - ydx),$$

where C denotes the boundary of the closed region R described in the positive sense.

EXAMPLE 7.70

Verify Green's theorem in the plane for $\oint_C [(xy + y^2)dx + x^2dy]$, where C is the closed curve of the region bounded by $y = x$ and $y = x^2$.

Solution. The region is bounded by the straight line $y = x$ and the parabola $y = x^2$. The point of intersection of $y = x$ and $y = x^2$ are (0, 0) and (1, 1). We note that



$$\oint_C [(xy + y^2)dx + x^2dy]$$

$$= \int_{C_1} [(xy + y^2)dx + x^2dy]$$

$$+ \int_{C_2} [(xy + y^2)dx + x^2dy] \qquad (1)$$

For the line integral along C_1, we have $y = x^2$ and so, $dy = 2xdx$ and x varies from 0 to 1. Thus,

$$\int_{C_1} [(xy + y^2)dx + x^2dy]$$

$$= \int_0^1 [(x^3 + x^4)dx + x^2(2x)dx]$$

$$= \int_0^1 (x^4 + 3x^3)dx = \left[\frac{x^5}{5} + \frac{3x^4}{4}\right]_0^1 = \frac{19}{20}.$$

For the line integral along (2), we have $y = x$ and so, $dy = dx$, and x varies from 1 to 0. Therefore,

$$\int_{C_2} [(xy+y^2)dx+x^2dy] = \int_1^0 [(x^2+x^2)dx+x^2dx]$$

$$= \int_1^0 3x^2dx = \left[\frac{3x^3}{3}\right]_1^0 = -1.$$

Hence, (1) yields

$$\int_C [(xy + y^2)dx + x^2dy] = \frac{19}{20} - 1 = -\frac{1}{20}.$$

On the other hand,

$$\iint_S \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) dx\,dy$$

$$= \iint_S \left[\frac{\partial}{\partial x}(x^2) - \frac{\partial}{\partial y}(xy+y^2)\right] dx\,dy$$

$$= \int_0^1 \int_{y=x^2}^{y=x} [2x-(x+2y)]\,dy\,dx$$

$$= \int_0^1 \left[\int_{x^2}^{x} (x-2y)dy\right] dx = \int_0^1 [xy-y^2]_{x^2}^{x} dx$$

$$= \int_0^1 (x^4-x^3)dx = \left[\frac{x^5}{5} - \frac{x^4}{4}\right]_0^1 = \frac{1}{5} - \frac{1}{4} = -\frac{1}{20}.$$

Hence,

$$\oint_C [f_1(x,y)dx + f_2(x,y)dy] = \iint_S \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) dx\,dy,$$

and thus, Green's theorem is verified.

EXAMPLE 7.71

Apply Green's Theorem to show that the area bounded by a simple closed curve C is given by $\frac{1}{2}\oint_C (xdy - ydx)$. Hence, find the area of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

Solution. In Deduction (i) of Green's Theorem, we have shown that the area A bounded by a simple closed curve C is equal to $\frac{1}{2}\oint_C (xdy - ydx)$.

For the second part, we know that the parametric equations of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ are $x = a\cos\theta$ and $y = b\sin\theta$. Thus, $dx = -a\sin\theta\, d\theta$ and $dy = b\cos\theta\, d\theta$. Therefore, the area A of the ellipse is given by

$$A = \frac{1}{2}\oint_C (xdy - ydx)$$

$$= \frac{1}{2}\int_0^{2\pi} [(a\cos\theta)(b\cos\theta)d\theta$$

$$- (b\sin\theta)(-a\sin\theta)d\theta]$$

$$= \frac{1}{2}\int_0^{2\pi} [ab\cos^2\theta + ab\sin^2\theta]d\theta$$

$$= \frac{ab}{2}\int_0^{2\pi} d\theta = \frac{ab}{2}[\theta]_\theta^{2\pi} = \pi ab.$$

EXAMPLE 7.72

Verify Green's theorem in the plane for $\oint_C [(3x^2 - 8y^2)dx + (4y - 6xy)dy]$, where C is the boundary of the region bounded by $x = 0$, $y = 0$, and $x + y = 1$.

Solution. We have

$$f_1(x, y) = 3x^2 - 8y^2 \text{ and } f_2(x, y) = 4y - 6xy.$$

Therefore,

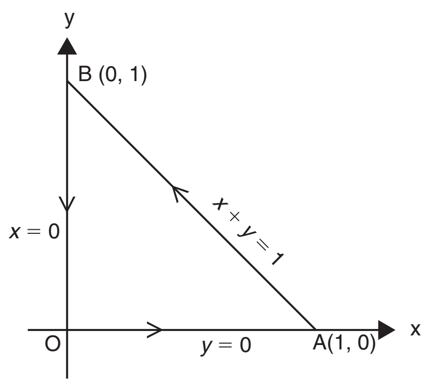
$$\iint_S \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) dxdy$$

$$= \int_0^1 \left[\int_0^{1-x} (-6y + 16y)dy\right] dx = 10 \int_0^1 \left[\frac{y^2}{2}\right]_0^{1-x} dx$$

$$= 5 \int_0^1 (1 - x)^2 dx = -\frac{5}{3}[(1 - x)^3]_0^1 = \frac{5}{3}. \qquad (1)$$

Further, the line integral splits into three parts:

$$\oint_C [f_2 dx + f_1 dy] = \int_{OA} + \int_{AB} + \int_{BO} .$$



Along OA, we have $y = 0$ so that $dy = 0$ and x varies from 0 to 1. Hence,

$$\int_{OA} = \int_0^1 3x^2 dx = 3\left[\frac{x^3}{3}\right]_0^1 = 1.$$

Along AB, we have $y = 1 - x$ and so $dy = -dx$ and x varies from 1 to 0. Therefore,

$$\int_{AB} = \int_1^0 [3x^2 - 8(1 - x)]^2 dx$$

$$+ [4(1 - x) - 6x(1 - x)](-1)dx$$

$$= \int_1^0 (-11x^2 + 26x - 12)dx$$

$$= \left[-11\frac{x^3}{3} + 26\frac{x^2}{2} - 12x\right]_1^0$$

$$= \frac{11}{3} - 13 + 12 = \frac{8}{3}.$$

Along BO, we have $x = 0$ so that $dx = 0$ and y varies from 1 to 0. Therefore,

$$\int_{BO} = \int_1^0 4ydy = 4\left[\frac{y^2}{2}\right]_1^0 = -2.$$

Hence,

$$\oint_C [f_2 dx + f_1 dy] = 1 + \frac{8}{3} - 2 = \frac{5}{3}. \qquad (2)$$

From (1) and (2), it follows that

$$\oint_C [f_2 dx + f_1 dy] = \iint_S \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) dx\, dy,$$

and thus, Green's theorem is verified.

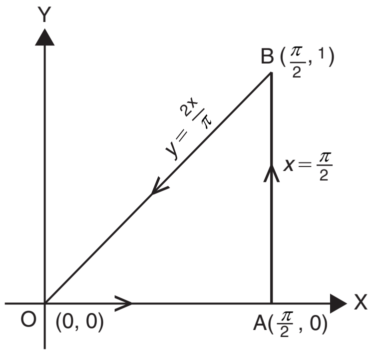
EXAMPLE 7.73

Using Green's theorem in a plane, evaluate $\oint_C [(y - \sin x)dx + \cos x\, dy]$, where C is the triangle with vertices (0,0), $\left(\frac{\pi}{2}, 0\right)$, and $\left(\frac{\pi}{2}, 1\right)$.

Solution. We have

$$f_1(x, y) = y - \sin x \text{ and } f_2(x, y) = \cos x.$$

The closed curve C is the triangle with vertices (0,0), $\left(\frac{\pi}{2}, 0\right)$, and $\left(\frac{\pi}{2}, 1\right)$ as shown in the following figure. The equation of the line OB is $\frac{y-0}{x-0} = \frac{1-0}{\frac{\pi}{2}-0}$, that is, $y = \frac{2x}{\pi}$.



By Green's Theorem, we have

$$\oint_C [f_1 dx + f_2 dy] = \iint_S \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) dx\, dy$$

$$= \iint_S \left[\frac{\partial}{\partial x}(\cos x) - \frac{\partial}{\partial y}(y - \sin x)\right] dx\, dy$$

$$= \iint -(\sin x + 1) dx\, dy$$

$$= -\int_0^{\frac{\pi}{2}} \left[\int_0^{\frac{2x}{\pi}} (1 + \sin x) dy\right] dx$$

$$= -\int_0^{\frac{\pi}{2}} [y + y \sin x]_0^{\frac{2x}{\pi}} dx$$

$$= -\frac{2}{\pi} \int_0^{\frac{\pi}{2}} x(1 + \sin x) dx$$

$$= -\frac{2}{\pi} \left\{ [x(x - \cos x)]_0^{\frac{\pi}{2}} - \int_0^{\frac{\pi}{2}} (x - \cos x) dx \right\}$$

$$= -\frac{\pi}{2} + \frac{2}{\pi} \int_0^{\frac{\pi}{2}} (x - \cos x) dx$$

$$= -\frac{\pi}{2} + \frac{2}{\pi} \left[\frac{x^2}{2} - \sin x\right]_0^{\frac{\pi}{2}}$$

$$= -\frac{\pi}{2} + \frac{2}{\pi} \left[\frac{\pi^2}{8} - 1\right]$$

$$= -\frac{\pi}{2} + \frac{\pi}{4} - \frac{2}{\pi} = -\frac{\pi}{4} - \frac{2}{\pi}.$$

EXAMPLE 7.74

Evaluate, by Green's Theorem, $\oint_C [(3x - y)dx + (2x + y)dy]$, where C is the curve $x^2 + y^2 = a^2$.

Solution. We have

$$f_1(x, y) = 3x - y \text{ and } f_2(x, y) = 2x + y.$$

By Green's Theorem,

$$\int_C (f_1 dx + f_2 dy)$$

$$= \iint_R \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) dx\, dy = \iint_R (2 - 1) dx\, dy$$

$$= \int_{-a}^{a} \int_{-\sqrt{a^2 - x^2}}^{\sqrt{a^2 - x^2}} dx\, dy = 4 \int_0^a \int_0^{\sqrt{a^2 - x^2}} dx\, dy$$

$$= 4 \int_0^a \sqrt{a^2 - x^2} dx$$

$$= 4a^2 \int_0^{\frac{\pi}{2}} \cos^2\theta\, d\theta, \; x = a \sin\theta$$

$$= 4a^2 . \frac{1}{2} . \frac{\pi}{2} = \pi a^2.$$

EXAMPLE 7.75

Compute the area of the loop of Descartes's Folium, $x^3 + y^3 = 3axy$.

Solution. Putting $y = tx$, we get the parametric equations of the contour of the folium as

$$x = \frac{3at}{1 + t^3} \text{ and } y = \frac{3at^2}{1 + t^3}.$$

The loop is described as t varies from 0 to ∞, since $t = \frac{y}{x} = \tan\theta$, where θ varies from 0 to $\frac{\pi}{2}$. Thus, $dx = 3a \frac{1 - 2t^3}{(1 + t^3)^2} dt$ and $dy = 3a \frac{2t - t^4}{(1 + t^3)^2} dt$. Hence, by Green's Theorem,

$$\text{Area } = \frac{1}{2} \oint_C (x dy - y dx) = \frac{9a^2}{2} \int_0^{\infty} \frac{t^2 dt}{(1 + t^3)^2} = \frac{3}{2} a^2.$$

7.20 STOKE'S THEOREM

The Stoke's Theorem provides a relation between a surface integral taken over a surface to a line integral along the boundary curve of the surface.

Theorem 7.12. (Stoke's Theorem). Let $\vec{f}$ be a vector-point function possessing continuous first-order partial derivatives and S be a surface bounded by a closed curve C. Then,

$$\oint_C \vec{f}.d\vec{r} = \iint_S curl\,\vec{f}.\hat{n}dS,$$

where $\hat{n}$ is a unit normal vector at any point of S, drawn in the sense in which a right-handed screw would move when rotated in the sense of description of the curve C.

Proof: Let the unit normal vector $\hat{n}$ makes angles α, β and γ with the positive directions of coordinate axes x, y, and z, respectively. Then, $\hat{n}=\cos\alpha\hat{i}+\cos\beta\hat{j}+\cos\gamma\hat{k}$. Since $\vec{r}=x\hat{i}+y\hat{j}+z\hat{k}$, we have $d\vec{r}=\hat{i}dx+\hat{j}dy+\hat{k}dz$. Let $\vec{f}=f_1\hat{i}+f_2\hat{j}+f_3\hat{k}$. Then,

$$curl\,\vec{f}=\nabla\times\vec{f}=\begin{vmatrix}\hat{i} & \hat{j} & \hat{k}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ f_1 & f_2 & f_3\end{vmatrix}$$

$$=\hat{i}\left(\frac{\partial f_3}{\partial y}-\frac{\partial f_2}{\partial z}\right)+\hat{j}\left(\frac{\partial f_1}{\partial z}-\frac{\partial f_3}{\partial x}\right)+\hat{k}\left(\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right)$$

and so,

$$curl\,\vec{f}.\hat{n} = \left(\frac{\partial f_3}{\partial y}-\frac{\partial f_2}{\partial z}\right)\cos\alpha + \left(\frac{\partial f_1}{\partial z}-\frac{\partial f_3}{\partial x}\right)\cos\beta + \left(\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right)\cos\gamma.$$

On the other hand,

$$\vec{f}.d\vec{r} = \left(f_1\hat{i}+f_2\hat{j}+f_3\hat{k}\right)\,.\,\left(\hat{i}dx+\hat{j}dy+\hat{k}dz\right)$$
$$=f_1dx+f_2dy+f_3dz.$$

Therefore, Stoke's Theorem takes the form

$$\oint_C (f_1dx+f_2dy+f_3dz)$$
$$=\iint_S\left[\left(\frac{\partial f_3}{\partial y}-\frac{\partial f_2}{\partial z}\right)\cos\alpha+\left(\frac{\partial f_1}{\partial z}-\frac{\partial f_3}{\partial x}\right)\cos\beta+\left(\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right)\cos\gamma\right]dS.$$

We now prove the theorem in this form. Suppose $z=\phi(x, y)$ be the equation of the surface S and R be the projection of S on the xy-plane. Then, the projection of the curve C on the xy-plane shall be the curve C_1, which enclose the region R. Therefore,

$$\oint_C f_1(x,y,z)dx = \oint_{C_1} f_1(x,y,\phi(x,y))dx$$
$$= \int_{C_1}[f_1(x,y,\phi)dx - 0dy]$$
$$=\iint_R\left[\frac{\partial}{\partial x}(0)-\frac{\partial}{\partial y}f_1(x,y,\phi)\right]dx\,dy$$

(by Green's theorem in plane)

$$=-\iint_R\frac{\partial}{\partial y}f_1(x,y,\phi)dx\,dy$$
$$=-\iint_R\left[\frac{\partial f_1}{\partial y}+\frac{\partial f_1}{\partial z}\frac{\partial\phi}{\partial y}\right]dx\,dy. \quad (1)$$

Since the direction ratios of the normal $\hat{n}$ to the surface S are $\frac{\partial\phi}{\partial x}$, $\frac{\partial\phi}{\partial y}$, and -1, we have

$$\frac{\cos\alpha}{\frac{\partial\phi}{\partial x}}=\frac{\cos\beta}{\frac{\partial\phi}{\partial y}}=\frac{\cos\gamma}{-1}\text{ and so }\frac{\partial\phi}{\partial y}=-\frac{\cos\beta}{\cos\gamma}.$$

Moreover, $dxdy$ being the projection of dS on the xy-plane, we have

$$dxdy = \cos\gamma dS.$$

Hence, (1) reduces to

$$\oint_C f_1(x,y,z)dx$$
$$=-\iint_S\left[\frac{\partial f_1}{\partial y}+\frac{\partial f_1}{\partial z}\left(-\frac{\cos\beta}{\cos\gamma}\right)\right]\cos\gamma\,dS$$
$$=\iint_S\left[\frac{\partial f_1}{\partial z}\cos\beta-\frac{\partial f_1}{\partial y}\cos\gamma\right]dS. \quad (2)$$

Similarly, it can be established that

$$\oint_C f_2(x,y,z)dx=\iint_S\left[\frac{\partial f_2}{\partial x}\cos\gamma-\frac{\partial f_2}{\partial z}\cos\alpha\right]dS \quad (3)$$

and

$$\oint_C f_3(x,y,z)dz=-\iint_S\left[\frac{\partial f_3}{\partial y}\cos\alpha-\frac{\partial f_3}{\partial x}\cos\beta\right]dS. \quad (4)$$

Adding (2), (3), and (4), we get

$$\oint_C (f_1 dx + f_2 dy + f_3 dz)$$

$$= \iint_S \left[\left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z} \right) \cos\alpha + \left(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x} \right) \cos\beta + \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right) \cos\gamma \right] dS.$$

This completes the proof of the theorem.

Remark 7.1. The equivalent statement of Stoke's Theorem is that
The line integral of the tangential component of a vector-point function $\vec{f}$ taken around a simple closed curve C is equal to the surface integral of the normal component of the curl of $\vec{f}$ taken over any surface S having C as its boundary.

EXAMPLE 7.76

Verify Stoke's Theorem for the function $\vec{f} = x^2 \hat{i} + xy\hat{j}$, integrated around the square in the plane z = 0, whose sides are along the lines $x = 0$, $x = a$, $y = 0$, and $y = a$.

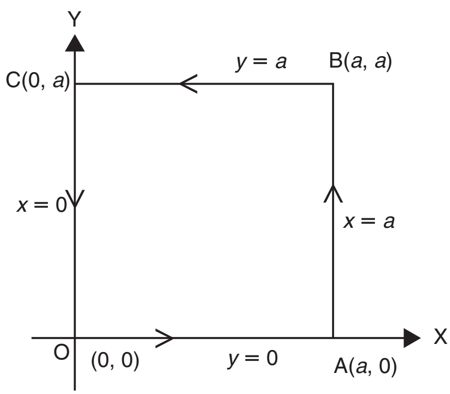
Solution. Since $\vec{f} = x^2\, \hat{i} + xy\, \hat{j}$, we have

$$\vec{f}.d\,\vec{r} = (x^2\,\hat{i} + xy\,\hat{j})\,.\,(\hat{i}\,dx + \hat{j}\,dy) = x^2 dx + xydy.$$

Therefore,

$$\oint_C \vec{f}.d\,\vec{r} = \oint_C (x^2 dx + xydy),$$

where C is the square shown in the figure.



Thus,

$$\oint_C \vec{f}.d\,\vec{r} = \int_{OA} + \int_{AB} + \int_{BC} + \int_{CA}. \quad (1)$$

Along OA, we have y = 0 and so, dy = 0. Thus,

$$\int_{AB} \vec{f}.d\,\vec{r} = \int_0^a x^2 dx = \left[\frac{x^3}{3}\right]_0^a = \frac{a^3}{3}.$$

Along AB, x = a and so, dx = 0. Thus,

$$\int_{AB} \vec{f}.\,d\,\vec{r} = \int_0^a aydy = \frac{a^2}{2}.$$

Along BC, we have y = a and so, dy = 0. Thus,

$$\int_{BC} \vec{f}.d\,\vec{r} = \int_a^0 x^2 dx = -\frac{a^3}{3}.$$

Along CO, we have x = 0 and so, dx = 0. Thus,

$$\int_{CO} \vec{f}.d\,\vec{r} = \int_a^0 0dy = 0.$$

Hence, (1) yields

$$\oint_C \vec{f}.d\,\vec{r} = \frac{a^3}{3} + \frac{a^2}{2} - \frac{a^3}{3} = \frac{a^2}{2}.$$

On the other hand,

$$curl\,\vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2 & xy & 0 \end{vmatrix} = y\,\hat{k}.$$

Since the square (surface) lies in the xy-plane, $\hat{n} = \hat{k}$. Therefore,

$$curl\,\vec{f}.\hat{n} = y\,\hat{k}.\hat{k} = y$$

and so,

$$\iint_S curl\,\vec{f}.\hat{n}\,dS = \int_0^a \int_0^a ydx\,dy = \int_0^a \left[\frac{y^2}{2}\right]_0^a dx$$

$$= \frac{a^2}{2} \int_0^a dx = \frac{a^3}{2}.$$

Hence,

$$\oint_C \vec{f}.\,d\,\vec{r} = \iint_S curl\,\vec{f}.\,\hat{n}\,dS.$$

EXAMPLE 7.77

Verifies Stoke's Theorem for $\vec{f} = y\,\hat{i} + z\,\hat{j} + x\,\hat{k}$, where S is the *upper-half* surface of the surface $x^2 + y^2 + z^2 = 1$ and C is its boundary.

Solution. Here, C is the boundary of the upper-half surface of $x^2 + y^2 + z^2 = 1$, that is, C is the boundary of the circle $x^2 + y^2 = 1$ in the xy-plane. Thus, the parametric equations of C are $x = \cos t$, $y = \sin t$, $z = 0$, and $0 \le t \le 2\pi$. Therefore,

$$\oint_C \vec{f}\,.\,d\,\vec{r} = \oint_C (f_1 dx + f_2 dy + f_3 dz)$$

$$= \oint_C [ydx + zdy + xdz]$$

$$= \int_0^{2\pi} [\sin t(-\sin t)]dt$$

$$= -\int_0^{2\pi} \sin^2 t\; dt$$

$$= -4\int_0^{\frac{\pi}{2}} \sin^2 t dt = -4.\frac{1}{2}.\frac{\pi}{2} = -\pi.$$

On the other hand,

$$curl\,\vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & z & x \end{vmatrix}$$

$$= \hat{i}\left(\frac{\partial x}{\partial y} - \frac{\partial z}{\partial z}\right) + \hat{j}\left(\frac{\partial y}{\partial z} - \frac{\partial x}{\partial x}\right) + \hat{k}\left(\frac{\partial z}{\partial x} - \frac{\partial y}{\partial y}\right)$$

$$= -\left(\hat{i} + \hat{j} + \hat{k}\right)$$

and so,

$$curl\,\vec{f}\,.\,\hat{n} = -\left(\hat{i} + \hat{j} + \hat{k}\right)\,.\,\hat{k} = -1.$$

Therefore,

$$\iint_S curl\,\vec{f}\,.\,\hat{n}\; dS = -\iint dx\; dy = -\pi(1)^2 = -\pi.$$

Hence,

$$\oint_C \vec{f}\,.\,d\,\vec{r} = \iint_S curl\,\vec{f}\,.\,\hat{n}\; dS$$

and thus, Stoke's theorem is verified.

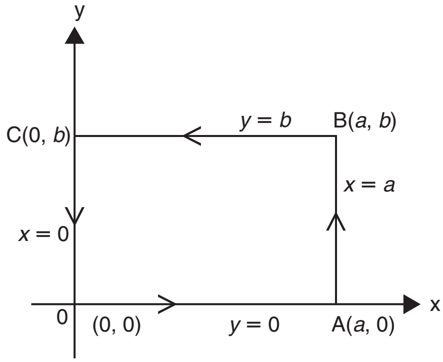
EXAMPLE 7.78

Verify Stoke's theorem for the vector field $\vec{f} = (x^2 - y^2)\hat{i} + 2xy\,\hat{j}$, integrated around the rectangle z = 0, and bounded by the lines $x = 0$, $y = 0$, $x = a$, and $y = b$.

Solution. Since $\vec{f} = (x^2 - y^2)\hat{i} + 2xy\,\hat{j}$, we have

$$\vec{f}\,.\,d\,\vec{r} = \left[(x^2 - y^2)\hat{i} + 2xy\,\hat{j}\right]\,.\,\left(\hat{i}\,dx + \hat{j}\,dy\right)$$

$$= \left(x^2 - y^2\right)dx + 2xydy.$$



Therefore,

$$\oint_C \vec{f}\,.\,d\,\vec{r} = \oint_C \left[\left(x^2 - y^2\right)dx + 2xydy\right]$$

$$= \int_{OA} + \int_{AB} + \int_{BC} + \int_{CO}.$$

Along OA, we have y = 0 and dy = 0. Therefore,

$$\int_{OA} \vec{f}\,.\,d\,\vec{r} = \int_0^a x^2 dx = \left[\frac{x^3}{3}\right]_0^a = \frac{a^3}{3}.$$

Along AB, we have x = a and dx = 0. Therefore,

$$\int_{AB} \vec{f}\,.\,d\,\vec{r} = \int_0^b 2aydy = ab^2.$$

Along BC, we have y = b and dy = 0. Therefore,

$$\int_{BC} \vec{f}\,.\,d\,\vec{r} = \int_a^0 (x^2 - b^2)dx = \left[\frac{x^3}{3} - b^2 x\right]_a^0$$

$$= -\frac{a^3}{3} + ab^2.$$

Along CO, we have x = 0 and dx = 0. Therefore,

$$\int_{CO} \vec{f}\,.\,d\,\vec{r} = 0.$$

Hence,

$$\oint_C \vec{f}.\, d\,\vec{r} = \frac{a^3}{3} + ab^2 - \frac{a^3}{3} + ab^2 = 2ab^2.$$

On the other hand,

$$curl\,\vec{f} = \nabla \times \vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2 - y^2 & 2xy & 0 \end{vmatrix} = 4y\,\hat{k}.$$

Therefore,

$$curl\,\vec{f}.\hat{n} = \left(4y\,\hat{k}\right).\,\left(\hat{k}\right) = 4y$$

and so,

$$\iint_S curl\,\vec{f}.\,\hat{n}\,dS = 4\int_0^a \left(\int_0^b y dy\right) dx = 4\int_0^a \left[\frac{y^2}{2}\right]_0^b dx$$

$$= 2b^2 \int_0^a dx = 2ab^2.$$

Hence,

$$\oint_C \vec{f}.\,d\,\vec{r} = \iint_S curl\,\vec{f}.\,\hat{n}\,dS$$

and thus, Stoke's theorem is verified.

EXAMPLE 7.79

Apply Stoke's Theorem to evaluate

$$\int_C [(x+y)dx + (2x-z)dy + (y+z)dz],$$

where C is the boundary of the triangle with vertices (2,0,0), (0,3,0), and (0,0,6).

Solution. Taking projection on three planes, we note that the surface S consists of three triangles, OAB in xy-plane, OBC in yz-plane, and OAC in xz-plane. Using two-point formula (or intercept form), the equations of the lines AB, BC, and CA are respectively $3x + 2y = 6$, $2y + z = 6$ and $3x + z = 6$. We have

$$\vec{f} = (x+y)\hat{i} + (2x-z)\hat{j} + (y+z)\hat{k}.$$

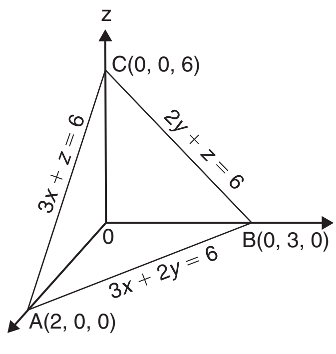
Therefore,

$$curl\,\vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x+y & 2x-z & y+z \end{vmatrix}$$

$$= \hat{i}\left[\frac{\partial}{\partial y}(y+z) - \frac{\partial}{\partial z}(2x-z)\right]$$

$$+\hat{j}\left[\frac{\partial}{\partial z}(x+y) - \frac{\partial}{\partial x}(y+z)\right]$$

$$+\hat{k}\left[\frac{\partial}{\partial x}(2x-z) - \frac{\partial}{\partial y}(x+y)\right] = 2\hat{i} + \hat{k}.$$



Now, by Stoke's Theorem,

$$\oint_C \vec{f}.\,d\,\vec{r} = \iint_S curl\,\vec{f}.\,\hat{n}\,ds$$

$$= \iint_{OAB} + \iint_{OBC} + \iint_{OAC}$$

$$= \iint_{OAB} \left(2\,\hat{i} + \hat{k}\right).\left(\hat{k}\right)ds$$

$$+ \iint_{OBC} \left(2\,\hat{i} + \hat{k}\right).\,(\hat{j})ds$$

$$+ \iint_{OAC} \left(2\,\hat{i} + \hat{k}\right).\,(\hat{i})ds$$

$$= \int_0^3 \left[\int_0^{\frac{6-2y}{3}} dx\right] dy + 0 + 2\int_0^6 \left[\int_0^{\frac{6-z}{2}} dy\right] dz$$

$$= \int_0^3 \frac{6-2y}{3}dy + 2\int_0^6 \frac{6-z}{2}dz$$

$$= \frac{1}{3}\left[6y - 2\frac{y^2}{2}\right]_0^3 + \frac{2}{2}\left[6z - \frac{z^2}{2}\right]_0^6$$

$$= 3 + 18 = 21.$$

EXAMPLE 7.80

Evaluate $\oint_C \vec{f}.\, d\,\vec{r}$ by Stoke's theorem, where $\vec{f} = y^2\ \hat{i} + x^2\hat{j} - (x+z)\hat{k}$ and C is the boundary of the triangle with vertices at (0, 0, 0), (1, 0, 0), and (1, 1, 0).

Solution. We have

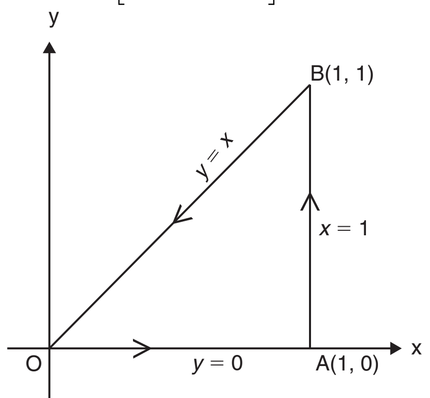
$$curl\,\vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y^2 & x^2 & -(x+z) \end{vmatrix} = \hat{j} + 2(x-y)\ \hat{k}.$$

Therefore,

$$curl\,\vec{f}.\ \hat{n} = \left[\hat{j} + 2(x-y)\hat{k}\right].\ \hat{n}\ .$$

We note that the z-coordinate of each vertex of the triangle is zero. Therefore, the triangle lies in the xy-plane. Hence, $\hat{n} = \hat{k}$. Thus,

$$curl\,\vec{f}.\ \hat{n} = \left[\hat{j} + 2(x-y)\hat{k}\right].\ \hat{k} = 2(x-y).$$



The equation of OB is y = x. Therefore, by Stoke's theorem, we have

$$\oint_C \vec{f}.\ d\,\vec{r} = \iint_S curl\,\vec{f}.\ \hat{n}\ dS = \int_0^1\int_0^x 2(x-y)dy\ dx$$

$$= 2\int_0^1 \left[xy - \frac{y^2}{2}\right]_0^x dx = 2\int_0^1 \left(x^2 - \frac{x^2}{2}\right)dx$$

$$= \int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}.$$

7.21 MISCELLANEOUS EXAMPLES

EXAMPLE 7.81

Evaluate $\int_C (x^2 + yz)dz$, where C is the curve defined by $x = t$, $y = t^2$, $z = 3t$ for t lying in the interval $1 \le t \le\ 2$.

Solution. The parametric equation of the curve C are $x = t$, $y = t^2$ and $z = 3t$. Therefore

$$\int_C (x^2 + yz)dz = 3\int_1^2 (t^2 + 3t^3)dt$$

$$= 3\left[\frac{t^3}{3} + 3\frac{t^4}{4}\right]_1^2$$

$$= \frac{163}{4}.$$

EXAMPLE 7.82

Find the work done when a force $\bar{F} = (x^2 - y^2 + x)i - (2xy + y)j$ moves a particle from the origin to the point (1, 1) along $y = x^2$.

Solution. We put $x = t$ and get $y = t^2$. Then $x = 0 \Rightarrow t = 0$ and $x = 1 \Rightarrow t = 1$. Thus

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k} = t\hat{i} + t^2\hat{j} \ \text{ implies } \ \frac{d\vec{r}}{dt}$$

$$= \hat{i} + 2t\hat{j} \ \text{ and } \ \vec{F} = (t^2 - t^4 + t)\hat{i} - (2t^3 + t^2)\hat{j}.$$

Therefore

$$W = \int_c \vec{F}\cdot d\vec{r} = \int_0^1 \left(\vec{F}\cdot\frac{d\vec{r}}{dt}\right)dt$$

$$= \int_0^1 \left\{\left[(t^2 - t^4 + t)\hat{i} - (2t^3 + t^2)\hat{j}\right]\cdot\left(\hat{i} + 2t\hat{j}\right)\right\}dt$$

$$= \int_0^1 [(t^2 - t^4 + t) - 2t(2t^3 + t^2)]dt$$

$$= \int_0^1 [t^2 - t^4 + t - 4t^4 - 2t^3]dt$$

$$= \int_0^1 (-5t^4 - 2t^3 + t^2 + t)dt$$
$$= \left[\frac{-5t^5}{5} - \frac{2t^4}{4} + \frac{t^3}{3} + \frac{t^2}{2}\right]_0^1$$
$$= -1 - \frac{1}{2} + \frac{1}{3} + \frac{1}{2} = -\frac{2}{3}.$$

EXAMPLE 7.83

If $\vec{f} = (x+3y)\hat{i} + (y-2z)\hat{j} + (x+\lambda z)\hat{k}$ is Solenoidal, find λ

Solution. As in Example 7.25, we have

$$f_1 = x + 3y,$$
$$f_2 = y - 2z,$$
$$f_3 = x + \lambda z.$$

Then

$$div\,\vec{f} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} = 1 + 1 + \lambda = 2 + \lambda.$$

The vector will be solenoidal if $div\,\vec{f} = 0$, that is, if $2 + \lambda = 0$ or if $\lambda = -2$.

EXAMPLE 7.84

(a) Show that the vector field $\vec{F} = 2x(y^2 + z^3)\hat{i} + 2x^2y\hat{j} + 3x^2z^2\hat{k}$ is conservative.

(b) Prove that $(y^2 - z^2 + 3yz - 2x)\hat{i} + (3xz + 2xy)\hat{j} + (3xy - 2xz + 2z)\hat{k}$ is both solenoidal and irrotational.

Solution. (a) Since every irrotational field is conservative, it is sufficient to show that $curl\,\vec{F} = 0$. We note that

$$curl\,\vec{F} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 2x(y^2+z^3) & 2x^2y & 3x^2z^2 \end{vmatrix}$$
$$= \hat{i}\left[\frac{\partial}{\partial y}(3x^2z^2) - \frac{\partial}{\partial z}(2x^3y)\right]$$
$$-\hat{j}\left[\frac{\partial}{\partial x}(3x^2z^2) - \frac{\partial}{\partial z}(2x(y^2+z^3))\right]$$
$$+\hat{k}\left[\frac{\partial}{\partial x}(2x^2y) - \frac{\partial}{\partial y}(2x(y^2+z^3))\right]$$
$$= 0 - (6xz^2 - 6xz^2)\hat{j} + \hat{k}\,(4xy - 4xy) = 0.$$

Hence the force $\vec{F}$ is conservative.

(b) $\vec{f} = (y^2 - z^2 + 3yz - 2x)\hat{i} + (3xz + 2xy)\hat{j} + (3xy - 2xz + 2z)\hat{k}$

Then

$$\nabla \cdot \vec{f} = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right).\vec{f}$$
$$= -2 + 2x + 2 - 2x = 0.$$

Hence $\vec{f}$ is solenoidal. Further,

$$curl\,\vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y^2 - z^2 + 3yz - 2x & 3xz + 2xy & 3xy - 2xz + 2z \end{vmatrix}.$$
$$= \hat{i}[3x - 3x] - \hat{j}[3y - 2z + 2z - 3y]$$
$$+ \hat{k}[3z + 2y - 2y - 3z] = \vec{0}$$

Hence $\vec{f}$ is irrotational.

EXAMPLE 7.85

Find the gradient of the sealer field $f(x,y) = y^2 - 4xy$ at $(l, 2)$.

Solution.

$$\nabla f = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(y^2 - 4xy)$$
$$= \hat{i}\frac{\partial}{\partial x}(y^2 - 4xy) + \hat{j}\frac{\partial}{\partial y}(y^2 - 4xy)$$
$$+ \hat{k}\frac{\partial}{\partial z}(y^2 - 4xy)$$
$$= -4y\hat{i} + (2y - 4x)\hat{j} + 0$$
$$= -8\hat{i} + 0\hat{j} = -8\hat{i}\text{ at }(1, 2).$$

EXAMPLE 7.86

A particle moves on the curve $x = 2t^2$, $y = t^2 - 4t$, $z = 3t - 5$, where t is time. Find the components of velocity and acceleration at time $t = 1$ in the direction of $i - 3j + 2k$.

Solution. Proceeding as in Example 7.8, we have

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k} = 2t^2\hat{i} + (t^2 - 4t)\hat{j} + (3t - 5)\hat{k}.$$

Therefore

$$\vec{v} = \frac{d\vec{r}}{dt} = 4t\hat{i} + (2t-4)\hat{j} + 3\hat{k}$$
$$= 4\hat{i} - 2\hat{j} + 3\hat{k} \quad \text{at} \quad t = 1,$$
$$\vec{a} = \frac{d^2\vec{r}}{dt^2} = 4\hat{i} + 2\hat{j}$$

The unit vector in the direction of $\hat{i} - 3\hat{j} + 2\hat{k}$ is

$$\hat{n} = \frac{\hat{i} - 3\hat{j} + 2\hat{k}}{|\hat{i} - 3\hat{j} + 2\hat{k}|} = \frac{\hat{i} - 3\hat{j} + \hat{k}}{\sqrt{14}}.$$

Therefore the components of velocity and acceleration in the direction of $\hat{i} - 3\hat{j} + 2\hat{k}$ are

$$\vec{v}.\hat{n} = (4\hat{i} - 2\hat{j} + 3\hat{k}) \cdot \frac{\hat{i} - 3\hat{j} + 2\hat{k}}{\sqrt{14}} = \frac{4+6+6}{\sqrt{14}}$$
$$= \frac{16}{\sqrt{14}}$$

or

$$\vec{a}.\hat{n} = (4\hat{i} + 2\hat{j}) \cdot \frac{\hat{i} - 3\hat{j} + 2\hat{k}}{\sqrt{14}} = -\frac{2}{\sqrt{14}}.$$

EXAMPLE 7.87

Find the values of a and b so that the surfaces $ax^3 - by^2z = (a+3)x^2$ and $4x^2y - z^3 = 11$ may cut orthogonally at $(2, -1, -3)$.

Solution. Following Example 7.21, we have

$$\phi = ax^3 - by^2z - (a+3)x^2$$
$$\psi = 4x^2y - z^3 - 11.$$

Then

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(ax^3 - by^2z - a + z)x^2$$
$$= \hat{i}\,[3ax^2 - 2\,(a+3)x] + \hat{j}\,[-2byz] + \hat{k}[-by^2]$$
$$= \hat{i}\,(8a-12) - 6b\hat{j} - b\hat{k} \quad at \quad (2, -1, -3)$$
$$\nabla\psi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(4x^3y - z^3 - 11)$$
$$= \hat{i}\,(8xy) + \hat{j}\,(4x^2) + \hat{k}\,(-3z^2)$$
$$= 16\hat{i} + 16\hat{j} - 27\hat{k} \quad at \quad (2, -1, -3).$$

Then $\nabla\phi \cdot \nabla\psi = 0$ implies

$$[(8a-12)\hat{i} - 6b\hat{j} - b\hat{k}].[-16\hat{i} + 16\hat{j} - 27\hat{k}] = 0$$
$$\Rightarrow 128a + 69b = 192. \qquad (1)$$

Also $(2, -1, -3)$ lies on ϕ. Therefore

$$8a + 3b - 4a - 12 = 0$$

or

$$4a + 3b = 12 \qquad (2)$$

Solving (1) and (2), we get $a = -\frac{7}{3}$, $b = \frac{64}{9}$.

EXAMPLE 7.88

(a) Find the directional derivative of $\varphi = x^2\,yz + 4xz^2$ at the point $(1, -2, -1)$ in the direction of the vector $2\hat{i} - \hat{j} - 2\hat{k}$.

(b) Find a unit normal vector $\hat{n}$ of the cone of revolution $z^2 = 4(x^2 + y^2)$ at the point $P(1, 0, 2)$.

(c) Find the directional derivative of $f(x, y, z) = 2x^2 + 3y^2 + z^2$ at the point $P(2, 1, 3)$ in the direction of the vector $\hat{a} = \hat{i} - 2\hat{k}$.

Solution. (a) We have

$$\nabla f = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(x^2yz + 4xz^2)$$
$$= \hat{i}(2yz + 4z^2) + \hat{j}(x^2z) + \hat{k}(x^2y + 8xz)$$
$$= 8\hat{i} - \hat{j} - 10\hat{k} \text{ at } (1, -2, -1).$$

The unit vector in the direction of the given vector $2\hat{i} - \hat{j} - 2\hat{k}$ is

$$\hat{n} = \frac{2\hat{i} - \hat{j} - 2\hat{k}}{\sqrt{4+4+1}} = \frac{2\hat{i} - \hat{j} - 2\hat{k}}{3}.$$

Therefore the directional derivative of f at $(1, -2, -1)$ in the direction of $2\hat{i} - \hat{j} - 2\hat{k}$ is

$$\nabla f.\hat{n} = (8\hat{i} - \hat{j} - 10\hat{k}).\left(\frac{2\hat{i} - \hat{j} - 2\hat{k}}{3}\right) = \frac{37}{3}.$$

(b) Similar to Exercise 16 of Chapter 7.

Let $\phi = z^2 - 4x^2 - 4y^2$. Then $\nabla\phi$ is along the normal vector.

But

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(z^2 - 4x^2 - 4y^2)$$
$$= -8x\hat{i} - 8y\hat{j} + 2z\hat{k}$$
$$= -8\hat{i} + 4\hat{k} \text{ at the point } (1, 0, 2).$$

Therefore unit normal vector $\hat{n}$ to the given cone at (1,0,2) is

$$\hat{n} = \frac{-8\hat{i}+4\hat{k}}{\sqrt{64+16}} = \frac{-8\hat{i}+4\hat{k}}{\sqrt{80}} = \frac{-2\hat{i}+\hat{k}}{\sqrt{5}}$$

(c) Similar to Example 7.16.
We have

$$\nabla f = \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)(2x^2+3y^2+z^2)$$

$$= 4x\hat{i}+6y\hat{j}+2z\hat{k}$$

$$= 8\hat{i}+6\hat{j}+6\hat{k} \quad \text{at the point } (2,13).$$

Now unit vector in the direction of $\hat{i}-2\hat{k}$ is

$$\hat{u} = \frac{\hat{i}-2\hat{k}}{\sqrt{1+4}} = \frac{1}{\sqrt{5}}\left(\hat{i}-2\hat{k}\right).$$

Therefore, the directional derivative at (2, 1, 3) in the direction of $\hat{i}-2\hat{k}$ is

$$\nabla f.\hat{u} = \frac{1}{\sqrt{5}}\left(8\hat{i}+6\hat{j}+6\hat{k}\right)\left(\hat{i}-2\hat{k}\right)$$

$$= \frac{1}{\sqrt{5}}[8+0-12] = -\frac{4}{\sqrt{5}}$$

EXAMPLE 7.89

If $r = |\vec{r}|$ and $\vec{a}$ is a constant vector, prove that

$$\nabla\times\left[\frac{\vec{a}\times\vec{r}}{r^n}\right] = \frac{2-n}{r^n}\vec{a} + \frac{n}{r^{n+2}}(\vec{a}\cdot\vec{r})\,\vec{r}\,.$$

Solution. If $r = |\vec{r}|$ and $\vec{a}$ is a constant vector, then using 7.12 (B) (ii), we have

$$\nabla\times\left[\frac{\vec{a}\times\vec{r}}{r^n}\right] = \nabla\times[r^{-n}(\vec{a}\times\vec{r})]$$

$$= (\nabla r^{-n})\times(\vec{a}\times\vec{r}) + r^{-n}[\nabla\times(\vec{a}\times\vec{r})$$

$$= \left[-\frac{n\vec{r}}{r^{n+2}}\right]\times(\vec{a}\times\vec{r}) + r^{-n}[\nabla\times(\vec{a}\times\vec{r})]$$

$$= -\frac{n}{r^{n+2}}[\vec{r}\times(\vec{a}\times\vec{r})] + r^{-n}[\nabla\times(\vec{a}\times\vec{r})]$$

$$= -\frac{n}{r^{n+2}}[(\vec{r}.\vec{r})\vec{a}-(\vec{r}.\vec{a})\vec{r}] + r^{-n}[\nabla\times(\vec{a}\times\vec{r})]$$

$$= -\frac{n}{r^{n+2}}[r^2\vec{a}-(\vec{a}.\vec{r})\vec{r}] + r^{-n}[\nabla\times(\vec{a}\times\vec{r})]$$

$$= -\frac{n}{r^n}\vec{a} + \frac{n}{r^{n+2}}(\vec{a}.\vec{r})\vec{r}] + r^{-n}[\nabla\times(\vec{a}\times\vec{r})].$$

Also by Example 7.33, (ii)

$$\nabla\times(\vec{a}\times\vec{r}) = curl(\vec{a}\times\vec{r}) = 2\vec{a}.$$

Therefore

$$\nabla\times\left[\frac{\vec{a}\times\vec{r}}{r^n}\right] = \frac{-n}{r^n}\vec{a} + \frac{n}{r^{n+2}}(\vec{a}.\vec{r})\vec{r} + \frac{2\vec{a}}{r^n}$$

$$= \frac{2-n}{r^n}\vec{a} + \frac{n}{r^{n+2}}(\vec{a}.\vec{r})\vec{r},$$

which is the required result.

EXAMPLE 7.90

Show that

$$grad(\vec{f}.\vec{g}) = \vec{f}\times curl\,\vec{g} + \vec{g}\times curl\vec{f} + (\vec{f}.\nabla)\vec{g} + (\vec{g}.\nabla)\vec{f}\,.$$

Solution. We have

$$grad(\vec{f}\cdot\vec{g}) = \sum\hat{i}\frac{\partial}{\partial x}\left(\vec{f}\cdot\vec{g}\right)$$

$$= \sum\hat{i}\frac{\partial}{\partial x}\left(\vec{f}\cdot\frac{\partial\vec{g}}{\partial x}+\frac{\partial\vec{f}}{\partial x}.\vec{g}\right)$$

$$= \sum\hat{i}\left(\vec{f}\cdot\frac{\partial\vec{g}}{\partial x}\right) + \sum\hat{i}\left(\frac{\partial\vec{f}}{\partial x}.\vec{g}\right). \quad (1)$$

On the other hand

$$\vec{f}\times(\hat{i}\times\frac{\partial\vec{g}}{\partial x}) = (\vec{f}.\frac{\partial\vec{g}}{\partial x})\hat{i} - (\vec{f}.\hat{i})\frac{\partial\vec{g}}{\partial x}$$

or

$$\left(\vec{f}\cdot\frac{\partial\vec{g}}{\partial x}\right)\hat{i} = \vec{f}\times\left(\hat{i}\times\frac{\partial\vec{g}}{\partial x}\right) + \left(\vec{f}\cdot\hat{i}\right)\frac{\partial\vec{g}}{\partial x}.$$

Therefore

$$\sum\left(\vec{f}.\frac{\partial\vec{g}}{\partial x}\right)\hat{i} = \vec{f}\times\left(\sum\hat{i}\times\frac{\partial\vec{g}}{\partial x}\right) + \sum\left(\vec{f}.\hat{i}\right)\frac{\partial\vec{g}}{\partial x}$$

$$= \vec{f}\times curl\,\vec{g} + \left(\vec{f}\cdot\nabla\right)\vec{g}. \quad (2)$$

Interchanging $\vec{f}$ and $\vec{g}$ in (2), we get

$$\sum\left(\vec{g}.\frac{\partial\vec{f}}{\partial x}\right)\hat{i} = \vec{g}\times curl\vec{f} + (\vec{g}.\nabla)\vec{f}. \quad (3)$$

From (1), (2) and (3), it follows that

$$grad\left(\vec{f}\cdot\vec{g}\right)=\vec{f}\times curl\,\vec{g}+\vec{g}\times curl\,\vec{f}+\left(\vec{f}\cdot\nabla\right)\vec{g}+\left(\vec{g}\cdot\nabla\right)\vec{f}.$$

EXAMPLE 7.91

Verify divergence theorem for

$$\vec{F}=(x^2-yz)\hat{i}+(y^2-zx)\hat{j}+(z^2-xy)\hat{k},$$

taken over the rectangular parallelopiped $0\le x\le a$, $0\le y\le b,\ 0\le z\le c$.

Solution. To verify Gauss divergence theorem, we have to show that

$$\iiint_v div\ \vec{F}\ \ dv=\iint_s \vec{F}.\hat{n}\ \ ds.$$

Firstly,

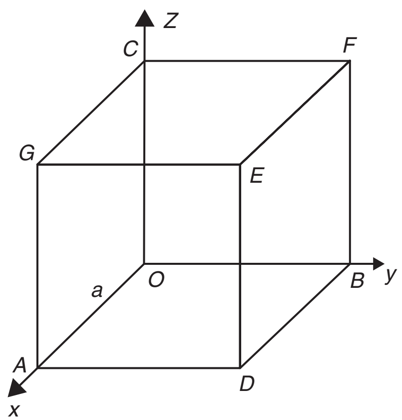
$$\iiint_v div\ \vec{F}\ \ dv$$

$$=\int_0^c\int_0^b\int_0^a\left[\frac{\partial}{\partial x}(x^2-yz)+\frac{\partial}{\partial y}(y^2-zx)+\frac{\partial}{\partial z}(z^2-xy)\right]dxdy$$

$$=\int_0^c\int_0^b\int_0^a 2(x+y+z)dxdydz$$

$$=a^2bc+ab^2c+abc^2=abc(a+b+c).$$

Now to calculate $\iint_s \vec{F}\cdot\hat{n}\,ds$, we divide the surface s of the parallelepiped $0\le x\le a,\ 0\le y\le b,$ $0\le z\le c$ into six parts.



(i) For the face $OADB$, we have $\hat{n}=-\hat{k},\ z=0.$ Therefore

$$\iint_{OADB}\vec{F}.\hat{n}.\ ds=\iint_{OADB}(x^2\hat{i}+y^2\hat{j}-xy\hat{k}).(-\hat{k})ds$$

$$=\int_0^b\int_0^a xy\ dx\ dy=\frac{a^2b^2}{4}.$$

(ii) For the face $CGEF$, we have $z=c,\ \hat{n}=\hat{k}.$ Therefore

$$\iint_{CGEF}\vec{F}.\hat{n}.\ ds$$

$$=\iint_{CGEF}[(x^2-cy)\hat{i}+(y^2-cx)\hat{j}+(c^2-xy)\hat{k})]\cdot\hat{k}\ ds$$

$$=\int_0^b\int_0^a(c^2-xy)dx\ dy=abc^2-\frac{a^2b^2}{4}.$$

(iii) For the face $ADEG$, we have $\hat{n}=\hat{i},x=a$ and $dx=0$. Therefore

$$\iint_{ADEG}\vec{F}.\hat{n}.\ ds=\int_0^c\int_0^b(a^2-yz)dy\ dz.$$

$$=a^2bc-\frac{b^2c^2}{4}$$

(iv) For the face $OBFC$, we have $\hat{n}=-\hat{i},x=0,$ $dx=0$. Therefore

$$\iint_{OBFC}\vec{F}.\hat{n}.\ ds=\int_0^a\int_0^b yz\ dy\ dz=\frac{b^2c^2}{4}.$$

(v) For the face $OAGC$, we have $\hat{n}=-\hat{j},\ y=0,$ $dy=0$ Therefore

$$\iint_{OAGC}\vec{F}.\hat{n}.\ ds=\int_0^a\int_0^c zx\ dz\ dx=\frac{a^2c^2}{4}.$$

(vi) For the face $DBFE$, we have $\hat{n}=\hat{j},\ y=b,$ $dy\ =\ 0$ Therefore

$$\iint_{DBFE}\vec{F}.\hat{n}.\ ds=\int_0^a\int_0^c(b^2-zx)dz\,dx=ab^2c-\frac{a^2c^2}{4}.$$

Hence adding the values of the above integrals, we get

$$\iint_S \vec{F}.\hat{n}.\ ds = abc(a+b+c).$$

Hence

$$\iiint_V div\ \vec{F}\ dv = \iint_S \vec{F}.\hat{n}\ ds,$$

which verifies the Gauss's divergence theorem.

EXAMPLE 7.92

Evaluate using divergence theorem $\iint_S (\vec{v}\cdot\hat{n})dA$, where $\vec{v} = x^2z\hat{i} + y\hat{j} - xz^2\hat{k}$ and S is the boundary of the region bounded by the paraboloid $z = x^2 + y^2$ and the plane $z = 4y$.

Solution. We have

$$\vec{v} = x^2z\hat{i} + y\hat{j} - xz^2\hat{k}.$$

Therefore

$$div\ \vec{v} = \nabla\cdot\vec{v} = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)$$

$$\times\left(x^2z\hat{i} + y\hat{j} - xz^2\hat{k}\right) = 2xz + 1 - 2xz = 1.$$

Using Divergence Theorem, we have

$$\iint (\vec{v}\cdot\hat{n})dA = \iiint_v div\,\vec{v}\,dv$$

$$= \iiint_v dv, \text{ since } div\,\vec{v} = 1.$$

$$= \int_0^4 \int_{-\sqrt{4y-y^2}}^{\sqrt{4y-y^2}} \int_{x^2+y^2}^{4y} dz\,dx\,dy$$

$$= \int_0^4 \int_{-\sqrt{4y-y^2}}^{\sqrt{4y-y^2}} \left[4y - x^2 - y^2\right] dx\,dy$$

$$= 2\int_0^4 \int_0^{\sqrt{4y-y^2}} \left[4y - x^2 - y^2\right] dx\,dy$$

(even integrand is x)

$$= 2\int_0^4 \left[(4y-y^2)x - \frac{x^3}{3}\right]_0^{\sqrt{4y-y^2}} dy$$

$$= \frac{4}{3}\int_0^4 (4y-y^2)^{\frac{3}{2}}\,dy = \frac{4}{3}\int_0^4 \left[4-(y-2)^2\right]^{\frac{3}{2}} dy$$

Substituting $y - 2 = 2\sin t$, we have $dy = 2\cos t\,dt$ and so

$$\iiint_v div\,\vec{v}\,dv = \frac{8}{3}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (4 - \sin^2 t)\cos t\ dt$$

$$= \frac{64}{3}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^3 t\,\cos t\,dt$$

$$= \frac{128}{3}\int_0^{\frac{\pi}{2}} \cos^4 t\ dt = \frac{128}{3}\cdot\frac{3}{8}\cdot\frac{\pi}{2} = 8\pi.$$

EXAMPLE 7.93

(a) Using Green's theorem in the plane evaluate $\int_C [(2x^2 - y^2)dx + (x^2 + y^2)dy]$ where C is the boundary of the region bounded by $x = 0$, $y = 0$, $x + y = 1$.

(b) Using Green's Theorem find the area of the region in the first quadrant bounded by the curves

$$y = x,\ y = \frac{1}{x},\ y = \frac{x}{4}.$$

Solution. (a) We have

$$\iint_S \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) dx\,dy$$

$$= \int_0^1 \int_0^{1-x} 2(x-y)dy\,dx = 2\int_0^1 \left[xy - \frac{y^2}{2}\right]_0^{1-x} dx$$

$$= \int_0^1 \left[2xy - y^2\right]_0^{1-x} dx$$

$$= \int_0^1 [2x(1-x) - (1-x)^2]dx$$

$$= \int_0^1 [2x - 2x^2 - 1 - x^2 + 2x]dx$$

$$= \int_0^1 (4x - 3x^2 - 1)dx$$

$$= \left[\frac{4x^2}{2} - \frac{3x^3}{3} - x\right]_0^1 = 2 - 1 - 1 = 0.$$

(b) Using Green's Theorem,

$$A=\frac{1}{2}\oint_{c}(x\,dy-ydx),$$

$$=\frac{1}{2}\left[\oint_{C_1}+\oint_{C_2}+\oint_{c_3}\right],$$

$$=\frac{1}{2}\left[\int_{c_1}(xdy-ydx)+\int_{c_2}(xdy-ydx)\right.$$

$$\left.+\int_{c_3}(xdy-ydx)\right],$$

where c_1 is $y=\frac{x}{4}$, c_2 is $y=\frac{1}{x}$ and c_3 is $y=x$.

Along c_1, we have $y=\frac{x}{4}$ so that $y=\frac{1}{x}dx$ and x varies from 0 to 2. Therefore

$$\oint_{c_1}(xdy-ydx)=\int_0^2\left(\frac{x}{4}dx-\frac{x}{4}dx\right)=0.$$

Along c_2 we have $y=\frac{1}{x}$ so that $dy=-\frac{1}{x^2}dx$ and x varies from 2 to 1. Therefore

$$\oint_{c_2}(xdy-ydx)=\int_2^1\left(\frac{-1}{x}dx-\frac{1}{x}dx\right)$$

$$=-2\int_0^1\frac{1}{x}dx=2\log 2.$$

Along c_3, we have $y=x$ so that $dy=dx$ and x varies from 1 to 0. Therefore

$$\oint_{c_3}(xdy-ydx)=\oint_{c_3}(xdx-xdx)=0.$$

Hence

$$\oint_{c}(xdy-ydx)=\frac{1}{2}[0+2\log 2+0]=\log 2.$$

EXAMPLE 7.94

Verify Gauss divergence theorem for the function $\bar{F}=yi+xj+z^2k$ over the cylindrical region bounded by $x^2+y^2=9, z=0$ and $z=2$.

Solution. We have

$$\iiint_V div\vec{f}\ dv$$

$$=\int_{-3}^{3}\int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}}\left[\int_0^2 2zdz\right]dy\,dx$$

$$=4\left[\int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}}dy\right]dx=8\int_{-3}^{3}\left[\int_0^{\sqrt{9-x^2}}dy\right]dx$$

$$=16\int_0^3\sqrt{9-x^2dx}=16\left[\frac{x\sqrt{9-x^2}}{2}+\frac{9}{2}\sin^{-1}\frac{x}{3}\right]_0^3$$

$$=16\left[0+\frac{9}{2}\sin^{-1}\frac{3}{3}\right]=\frac{16\times 9\times\pi}{2\times 2}=36\pi.$$

Similarly (Proceeding as in Example7.58), we have

$$\iint_S\vec{F}\cdot\hat{n}\,ds=36\pi.$$

Hence the theorem is verified.

EXAMPLE.7.95

Evaluate $\iint_S\vec{f}.d\vec{S}$ if $\vec{f}=yz\,\hat{i}+2y^2\,\hat{j}+xz^2\,\hat{k}$ and S is the surface of the cylinder $x^2+y^2=9$ contained in the *first octant between* the planes $z=0$ and $z=2$.

Solution. By Gauss's Divergence Theorem

$$\iint_S\vec{f}\cdot\hat{n}\,ds=\iiint_V div\vec{f}\ dv$$

$$=\iiint_V\left[\frac{\partial}{\partial x}(yz)+\frac{\partial}{\partial y}(2y^2)+\frac{\partial}{\partial z}(xz^2)\right]dv$$

$$=\iiint[4y+2xz]dz\,dy\,dx$$

$$=\int_0^3\int_0^{\sqrt{9-x^2}}\int_0^2(4y+2xz)dz\,dy\,dx$$

$$=\int_0^3\int_0^{\sqrt{9-x^2}}(8y+4x)dy\,dx$$

$$=\int_0^3[4(9-x^2)+4x(9-x^2)^{\frac{1}{2}}]dx=108.$$

EXAMPLE 7.96

Verify Stoke's theorem for $\vec{F} = xy^2\hat{i} + y\hat{j} + z^2x\hat{k}$ for the surface of a rectangular lamina bounded by $x = 0$, $y = 0$, $x = 1$, $y = 2$, $z = 0$.

Solution. Similar to Example 7.78.
We have

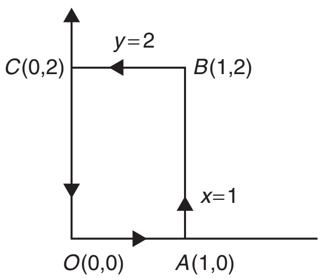
$$\vec{f} = xy^2\hat{i} + y\hat{j} + z^2x\hat{k}.$$

Therefore

$$\vec{f}\cdot d\vec{r} = (xy^2\hat{i} + y\hat{j} + z^2x\hat{k}).(\hat{i}\,dx + \hat{j}\,dy + \hat{k}\,dz)$$
$$= xy^2dx + ydy + z^2xdz.$$

Therefore

$$\oint_C \vec{f}\cdot d\vec{r} = \oint_C (xy^2dx + ydy + z^2xdz)$$
$$= \int_{OA} + \int_{AB} + \int_{BC} + \int_{CO}$$



Along OA, we have $y = 0$ and $dy = 0$. Therefore

$$\int_{OA} \vec{f}.d\vec{r} = \int_0^1 0\,dx = 0.$$

Along AB, we have $x = 1$ and $dx = 0$. Therefore

$$\int_{AB} \vec{f}\cdot d\vec{r} = \int_0^2 ydy = \left[\frac{y^2}{2}\right]_0^2 = 2.$$

Along BC, we have $y = 2$ and $dy = 0$. Therefore

$$\int_{BC} \vec{f}.d\vec{r} = \int_1^0 4x\,dx = 4\left[\frac{x^2}{2}\right]_1^0 = -2.$$

Along CO, we have $x = 0$ and $dx = 0$. Therefore

$$\int_{CO} \vec{f}\cdot d\vec{r} = \int_2^0 ydy = \left[\frac{y^2}{2}\right]_2^0 = -2.$$

Hence

$$\oint_C \vec{f}\cdot d\vec{r} = 0 + 2 - 2 - 2 = -2.$$

On the other hand,

$$curl\,\vec{f} = \nabla\times\vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xy^2 & y & z^2x \end{vmatrix}$$
$$= \hat{i}[0] + \hat{j}[0 - z^2] + \hat{k}[-2xy]$$
$$= -z^2\hat{j} + 2xy\hat{k}.$$

Therefore

$$curl\,\vec{f}\cdot\hat{n}\cdot ds = (-z^2\hat{j} + 2xy\hat{k})\,\hat{k} = -2xy$$

and so

$$\iint_S curl\,\vec{f}\cdot\hat{n}\cdot ds = -2\int_0^1\left(\int_0^2 xy\,dy\right)dx$$
$$= -2\int_0^1\left[x\frac{y^2}{2}\right]_0^2 dx$$
$$= -2\int_0^1 2x\,dx = -4\left[\frac{x^2}{2}\right]_0^1 = -2.$$

Hence

$$\oint_C \vec{f}\cdot d\vec{r} = \iint_S curl\,\vec{f}\cdot\hat{n}\,ds$$

and Stoke's Theorem is verified.

EXERCISES

Differentiation of Vectors

1. If $\hat{r}$ is a unit vector in the direction of $\vec{r}$, show that $\hat{r}\times\frac{d\hat{r}}{dt} = \frac{1}{r^2}\,\vec{r}\times\frac{d\vec{r}}{dt}$, where $|\vec{r}| = r$.

2. If $\vec{a}=t^2\hat{i}-t\hat{j}+(2t+1)\hat{k}$ and $\vec{b}=(2t-3)\hat{i}+\hat{j}-t\hat{k}$, find (*i*) $\frac{d}{dt}\left(\vec{a}.\vec{b}\right)$ and (*ii*) $\frac{d}{dt}\left(\vec{a}\times\vec{b}\right)$, when $t=1$.

   Ans. (*i*) -6, (*ii*) $7\hat{j}+3\,\hat{k}$.

3. If the vector $\vec{a}$ has a constant magnitude, show that $\vec{a}$ and $\frac{d\vec{a}}{dt}$ are perpendicular, provided $\left|\frac{d\vec{a}}{dt}\right|\neq 0$.
   Hint: $\vec{a}.\vec{a}=|\vec{a}|^2$ = constant implies $\frac{d}{dt}(\vec{a}.\vec{a})=0$ or $\vec{a}.\frac{d\vec{a}}{dt}+\vec{a}\frac{d\vec{a}}{dt}=0$ or $2\vec{a}.\frac{d\vec{a}}{dt}=0$ and so, $\vec{a}$ is orthogonal to $\frac{d\vec{a}}{dt}$, if $\left|\frac{d\vec{a}}{dt}\right|\neq 0$.

4. If $\vec{a},\vec{b}$, and $\vec{c}$ are constant vectors, show that the vector $\vec{r}=\vec{a}\,t^2+\vec{b}\,t+\vec{c}$ is the position vector of a point moving with a constant acceleration.
   Hint: $\frac{d^2\vec{r}}{dt}=2\,\vec{a}$ (constant).

5. A particle moves along the curve $x=t^3+1$, $y=t^2$, and $z=2t+5$, where t is the time. Find the component of its velocity and acceleration at time $t=1$ in the direction $\hat{i}+\hat{j}+3\hat{k}$.

   Ans. $\sqrt{11}$, $\frac{8}{\sqrt{11}}$.

6. A particle moves so that its position vector is given by $\vec{r}=\cos\omega t\,\hat{i}+\sin\omega t\,\hat{j}$, where ω is constant. Show that (i) the velocity $\vec{v}$ of the particle is perpendicular to $\vec{r}$, (ii) the acceleration is directed toward the origin and has a magnitude proportional to the distance from the origin, and (iii) $\vec{r}\times\vec{v}$ is a constant vector.
   Hint: $\vec{r}.v=0$ and $\vec{a}=\frac{d\,\vec{v}}{dt}=-\omega^2\cos t\,\hat{i}-\omega^2\sin\omega t\,\hat{j}=-\omega^2.\,\vec{r}$.

7. Find the unit tangent vector at any point on the curve $x=t^2+2$, $y=4t-5$, and $z=2t^2-6t$, where t is any variable.
   Hint: $\hat{T}=\frac{\frac{d\vec{r}}{dt}}{\left|\frac{d\vec{r}}{dt}\right|}$, where $\vec{r}=x\,\hat{i}+y\,\hat{j}+z\,\hat{k}$.

   Ans. $\frac{2\,\hat{i}+2\,\hat{j}+\hat{k}}{3}$.

8. Find the angle between the tangents to the curve $x=t$, $y=t^2$, and $z=t^3$ at $t=\pm 1$.
   Hint: $\vec{r}=x\,\hat{i}+y\,\hat{j}+z\,\hat{k}=t\,\hat{i}+t^2\,\hat{j}+t^3\,\hat{k}$. Find $\frac{d\vec{r}}{dt}$ and put $t=1$ and $t=-1$ to get $\vec{T}_1$ and $\vec{T}_2$. Then, the angle between $\vec{T}_1$ and $\vec{T}_2$ is given by $\cos\theta=\frac{\vec{T}_1\,\vec{T}_2}{|\vec{T}_1\,\vec{T}_2|}$.

   Ans. $\cos^{-1}\frac{3}{7}$.

9. If $\vec{a}=x^2yz\,\hat{i}-2xz^3\,\hat{j}+xz^2\,\hat{k}$, find the value of $\frac{\partial^2}{\partial x\partial y}\left(\vec{a}\times\vec{b}\right)$ at the point (1,0,−2).

   Ans. $-4\,\hat{i}-8\hat{j}$.

10. The position vector of a point at any time t is given by $\vec{r}=e^t\left(\cos t\,\hat{i}+\sin t\,\hat{j}\right)$. Show that (i) $\vec{v}$, where $\vec{a}$ and $\vec{v}$ are respectively acceleration and velocity of the particle and (ii) the angle between the radius vector and the acceleration is constant.

    Hint $\vec{v}=\dfrac{d\vec{r}}{dt}=e^t(\cos t-\sin t)\hat{i}+e^t(\sin t+\cos t)\hat{j}$,
    $\vec{a}=\sinh(t-1)\hat{i}+2e^t\cos t\hat{j}$.

    Clearly $\vec{a}=2(\vec{v}-\vec{r})$.

11. The position vector of a particle at time t is $\vec{r}=\cos(t-1)\hat{i}+\sinh(t-1)\hat{j}+Kt^3\hat{k}$. Find the value of K such that at time $t=1$, the acceleration is normal to the position vector $\vec{r}$.
    Hint: $\frac{d^2\,\vec{r}}{dt^2}$ at $t=1$ is $-\,\hat{i}+6K\,\hat{k}$ and $\vec{r}$ at $t=1$ is $\hat{i}+K\hat{k}$. Therefore, $\cos\theta=\frac{\left(-\hat{i}+6K\hat{k}\right).\left(\hat{i}+K\hat{k}\right)}{\sqrt{7}(1+k^2)^{\frac{1}{2}}}=\frac{6K^2-1}{\sqrt{7}(1+K^2)^{\frac{1}{2}}}$. For normality, $6K^2-1=0$ and so, $K=\pm\frac{1}{\sqrt{6}}$.

12. A particle moves along the curve $x=t^3+1$, $y=t^2$, and $z=2t+5$, where t represents the time. Find the component of its velocity and acceleration at time $t=1$ in the direction of $\hat{i}+\hat{j}+\hat{k}$.
    Hint: $\vec{r}=(t^3+1)\hat{i}+t^2\hat{j}+(2t+5)\hat{k}$ and $\frac{d\vec{r}}{dt}=3t^2\hat{i}+2t\hat{j}+2\hat{k}$. At $t=1,\vec{v}=3\hat{i}+2\hat{j}+2\hat{k}$. The unit vector in the direction of $\hat{i}+\hat{j}+3\hat{k}$ is $\frac{\hat{i}+\hat{j}+3\hat{k}}{\sqrt{11}}$. Therefore, the component of $\vec{v}$ along $\hat{i}+\hat{j}+3\hat{k}$ is $\left(3\hat{i}+2\hat{j}+2\hat{k}\right)\left(\frac{\hat{i}+\hat{j}+3\hat{k}}{\sqrt{11}}\right)=\sqrt{11}$. Similarly, proceed for acceleration, which will be $\frac{8}{\sqrt{11}}$.

13. If $\vec{F}=xyz\,\hat{i}+xz^2\,\hat{j}-y^3\,\hat{k}$ and $\vec{g}=x^3\,\hat{i}-xyz\,\hat{j}+x^2z\,\hat{k}$, then determine $\frac{\partial^2\vec{f}}{\partial y^2}\times\frac{\partial^2\vec{g}}{\partial x^2}$ at the point (1, 1, 0).

    Ans. $-36\hat{j}$.

Gradient- and Fractional Derivatives

14. If $\vec{r}$ is the position vector of a point and $\vec{a}$ is any vector, show that grad $\left[\vec{r}\;\vec{a}\;\vec{b}\right]=\vec{a}\times\vec{b}$.
    Hint: $\vec{r}=x\hat{i}+y\hat{j}+z\hat{k}$ and $\vec{a}=a_1\hat{i}+a_2\hat{j}+a_3\hat{k}$. Then, $\vec{a}\,.\vec{r}=a_1x+a_2y+a_3z$ and $\nabla(\vec{a}\,.\vec{r})=\vec{a}$. Therefore,
    $\nabla\left(\vec{r}\vec{a}\vec{b}\right)=\nabla\left[\vec{r}.\left(\vec{a}\times\vec{b}\right)\right]=\nabla\left[\left(\vec{a}\times\vec{b}\right).\vec{r}\right]=\vec{a}\times\vec{b}.$

15. If $\phi(x, y, z) = 3xy^2 - y^3 z^2$, find $\nabla \phi$ at the point $(-1, 2, -1)$.

Ans. $12\,\hat{i} - 24\,\hat{j} + 16\,\hat{k}$.

16. Find a unit normal to the surface $x^2 y + 2xz = 4$ at the point $(2, -2, 3)$. Hence, find the equation of the normal to the surface at $(2, -2, 3)$.

Hint: Let $\phi = x^2y + 2xz - 4$. Then, $\nabla\phi = 2(xy+z)\hat{i} + x^2\hat{j} + 2x\,\hat{k}$ and $\nabla\phi$ at $(2,-2,3) = -2\hat{i}+4\hat{j}+4\hat{k}$. Thus, the unit normal vector to the surface at $(2, -2, 3)$ is $\frac{-2\hat{i}+4\hat{j}+4\hat{k}}{\sqrt{36}} = \frac{1}{3}\left(-\hat{i}+2\hat{j}+2\hat{k}\right)$. Thus, the equation of normal is $\frac{x-2}{-1} = \frac{y+2}{2} = \frac{z-3}{2}$.

17. Find a unit normal vector to the surface $x^3 + y^3 + 3xyz = 3$ at the point $(1, 2, -1)$.

Ans. $\frac{-\hat{i}+3\hat{j}+2\hat{k}}{\sqrt{14}}$.

18. Find the directional derivative of $\phi(x, y, z) = x^2yz + 4xz^2$ at the point $(1, -2, -1)$ in the direction of $2\,\hat{i} - \hat{j} - 2\,\hat{k}$.

Hint: Proceed as in Example 7.16.

Ans. $\frac{37}{3}$.

19. Find the directional derivative of the function $\phi(x, y, z) = x^2 - y^2 + 2z^2$ at $P(1, 2, 3)$, in the direction of the line PQ, where Q is the point (5, 0,4). In what direction the directional derivative will be maximum?

Hint: $\nabla\phi = 2x\,\hat{i} - 2y\,\hat{j} + 4z\,\hat{k}$. Therefore, $\nabla\phi$ at $(1, 2, 3)$ is $2\,\hat{i} - 4\,\hat{j} + 12\,\hat{k}$. Also, $\vec{PQ} = \vec{OQ} - \vec{OP} = (5\hat{i}+4\hat{j}) - \left(\hat{i}+2\hat{j}+3\hat{k}\right) = 4\hat{i}-2\hat{j}+\hat{k}$. Unit vector $\hat{a}$ in the direction of $\vec{PQ}$ is $\frac{4\hat{i}-2\hat{j}+\hat{k}}{\sqrt{21}}$. Then, the required directional derivative $=\nabla\phi(at(1,2,3)).\,\hat{a} = \frac{4}{3}\sqrt{21}$. It will be maximum in the direction of the normal to ϕ, that is, in the direction of $\nabla\phi$, which is equal to $2\hat{i} - 4\hat{j} + 12\hat{k}$. Its maximum value is $|\nabla\phi| = \sqrt{4+16+144} = \sqrt{164} = 2\sqrt{41}$.

20. If $\phi(x, y, z) = 2xy + z^2$, find the directional derivative of ϕ in the direction of $\hat{i} + 2\,\hat{j} + 2\,\hat{k}$.

Hint: $\nabla\phi = 2y\hat{i} + 2x\hat{j} + 2z\hat{k} = -2\hat{i} + 2\hat{j} + 6\hat{k}$ at $(1, -1, 3)$. Unit vector $\hat{a}$ in the direction of $\hat{i}+2\hat{j}+2\hat{k}$ is $\frac{\hat{i}+2\hat{j}+2\hat{k}}{3}$. Therefore, the required directional derivative is $\nabla\phi.\hat{a} = \left(-2\hat{i}+2\hat{j}+6\hat{k}\right).\left(\frac{\hat{i}+2\hat{j}+2\hat{k}}{3}\right) = \frac{14}{3}$.

21. Find the greatest rate of increase of $u = x^2 + yz^2$ at the point $(1, -1, 3)$.

Ans. $|\nabla u| = \sqrt{121}$.

22. Find the equation of the tangent plane to the surface $z = x^2 + y^2$ at the point $(2, -1, 5)$.

Hint: $\nabla\phi$ at $(2, -1, 5)$ is $4\,\hat{i} - 2\,\hat{j} - \hat{k}$. The unit normal vector at $(2, -1, 5)$ is $\hat{a} = \frac{4\,\hat{i}-2\,\hat{j}-\hat{k}}{\sqrt{21}}$.

The equation of the line through $(2, -1, 5)$ in the direction of normal vector $\hat{a}$ is $\frac{x-2}{4} = \frac{y+1}{-2} = \frac{z-5}{-1}$. Therefore, the equation of tangent plane to the surface at $(2, -1, 5)$ is $4(x-2) - 2(y+1) - (z-5) = 0$ or $4x - 2y - z = 5$. We may also find a tangent plane using $(\vec{r} - \vec{a}).\nabla\phi = 0$. Therefore, in the present case, we have

$$\left[\left(x\,\hat{i}+y\,\hat{j}+z\,\hat{k}\right) - \left(2\,\hat{i}-\hat{j}+5\,\hat{k}\right)\right].\left(4\,\hat{i}-2\,\hat{j}-\hat{k}\right) = 0 \text{ or } 4(x-2)$$
$$-2(y+1)-(z-5) = 0 \text{ or } 4x-2y-z = 5.$$

Divergence and Curl of Vector-Point Function

23. Show that the vector $(-x^2 + yz)\hat{i} + (4y - z^2x)\hat{j} + (2xz - 4z)\hat{k}$ is solenoidal.
Hint: Show that $\nabla\cdot\vec{f} = 0$.

24. If $f = (x^2 + y^2 + z^2)^{-n}$, find div grad f and also n, so that div grad $f = 0$.

Ans. $\frac{2n(2n-1)}{(x^2+y^2+z^2)^{n+1}}$ and $n = \frac{1}{2}$.

25. Show that $\operatorname{div}\left(\frac{\vec{r}}{r^3}\right) = 0$, where $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$.

Hint: Use $\operatorname{div}\left(\phi\vec{f}\right) = \phi\operatorname{div}\vec{f} + grad\phi.\vec{f}$. We get

$$\operatorname{div}\left(\frac{\vec{r}}{r^3}\right) = \operatorname{div}\left(r^{-3}\vec{r}\right) = r^{-3}\operatorname{div}\vec{r} + \vec{r}.grad\,r^{-3} =$$

$$3r^{-3} + \vec{r}.\left(-3r^{-4}grad\,r\right) = 3r^{-3} + \vec{r}.\left[-3r^{-4}\frac{\vec{r}}{r}\right] =$$

$$3r^{-3} - 3r^{-5}(\vec{r}.\vec{r}) = 3r^{-3} - 3r^{-5}\left(r^2\right) = 0.$$

Thus, it also follows that $\frac{\vec{r}}{r^3}$ is solenoidal.

26. Show that the function $\frac{1}{r}$, where $r = |\vec{r}| = \sqrt{x^2 + y^2 + z^2}$, is a harmonic function, if $r \neq 0$.

Hint: Show that $\nabla^2\left(\frac{1}{r}\right) = 0$ (see Example 7.36).

27. If $\vec{f} = \frac{1}{u}\nabla v$, where u and v are scalar fields and $\vec{f}$ is a vector field, show that $\vec{f}$ curl $\vec{f} = a$.

Hint: $curl\vec{f}=\nabla\times\left(\frac{1}{u}\nabla u\right).\nabla\frac{1}{u}\times\nabla v+\frac{1}{u}\nabla\times\nabla v=\nabla\frac{1}{u}\times\nabla v+0$. Hence, $\vec{f}$. Curl $\vec{f}=\frac{1}{u}\nabla v.(\nabla\frac{1}{u}\times\nabla v)=0$.

28. Show that the vector $\nabla\phi\times\nabla\psi$ is solenoidal.
29. Find the value of a so that

$$\vec{f}=\left(ax^2y+yz\right)\hat{i}+\left(xy^2-xz^2\right)\hat{j}+\left(2xyz-2x^2y^2\right)\hat{k}$$

is solenoidal. Also find the curl of this solenoidal vector.

Hint: div $\vec{f}=2(a+2)xy$. Now, $\vec{f}$ will be solenoidal if div $\vec{f}=0$, which yields $a=-2$. $curl\,\vec{f}$ can be found.

30. If $\vec{r}=x\hat{i}+y\hat{j}+z\hat{k}$ and $r=|\vec{r}|=\sqrt{x^2+y^2+z^2}$, show that div$(\hat{r})=\frac{2}{r}$.

Hint: $\hat{r}=\dfrac{x\hat{i}+y\hat{j}+z\hat{k}}{\sqrt{x^2+y^2+z^2}}=\dfrac{\vec{r}}{r}$ and

$$\begin{aligned}\text{div}\left(\frac{\vec{r}}{r}\right)&=\frac{1}{r}div\vec{r}+grad\frac{1}{r}.\vec{r}\\&=\frac{3}{r}+\left[(-1)r^{-2}grad\,r\right].\vec{r}\\&=\frac{3}{r}-\frac{1}{r^2}\left(\frac{\vec{r}}{r}\right).\vec{r}=\frac{3}{r}-\frac{1}{r^3}(\vec{r}.\vec{r})\\&=\frac{3}{r}-\frac{r^2}{r^3}=\frac{3}{r}-\frac{1}{r}=\frac{2}{r}.\end{aligned}$$

31. Show that $\nabla^2(r\,\vec{r})=\left(\frac{4}{r}\right)\vec{r}$.

Hint:

$$\begin{aligned}\nabla^2(r\vec{r})&=\sum\frac{\partial^2}{\partial x^2}(r\vec{r})\\&=\sum\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}(r\vec{r})\right)\\&=\sum\frac{\partial}{\partial x}\left[\frac{\partial r}{\partial x}\vec{r}+r\frac{\partial\vec{r}}{\partial x}\right]\\&=\sum\frac{\partial}{\partial x}\left[\frac{x}{r}\vec{r}+ri\right]\\&=\sum\left(\left\{\frac{\vec{r}}{r}+\frac{x}{r}(i)-\frac{x}{r^2}\left(\frac{x}{r}\right)\vec{r}\right\}+i\frac{x}{r}\right)\\&=\left(\frac{4}{r}\right)\vec{r}.\end{aligned}$$

32. Show that the vector field $\vec{v}=(\sin y+z)\hat{i}+(x\cos y-z)\hat{j}+(x-y)\hat{k}$ is irrotational.
33. If $\vec{r}=x\hat{i}+y\hat{j}+z\hat{k}$, determine $\nabla\cdot\left(\frac{\vec{r}}{r}\right)$. (See Exercise 25).

Vector Integration and Line Integrals

34. If $\vec{r}(t)=2\hat{i}-\hat{j}+2\hat{k}$ for $t=2$ and $\vec{r}(t)=4\hat{i}-2\hat{j}+3\hat{k}$ for $t=3$, show that $\int_2^3\left[\vec{r}.\frac{d\vec{r}}{dt}\right]dt=10$.

Hint: $\frac{d}{dt}\left[(\vec{r})^2\right]=2\vec{r}\frac{d\vec{r}}{dt}$ implies $\vec{r}\frac{d\vec{r}}{dt}=\frac{1}{2}\frac{d}{dt}\left[(\vec{r})^2\right]$.

Therefore,

$$\int_2^3\left[\vec{r}\frac{d\vec{r}}{dt}\right]dt=\frac{1}{2}\left[(\vec{r})^2\right]_2^3=\frac{1}{2}[29-9]=10,\text{ using}$$

$\vec{r}(t)=4\hat{i}-2\hat{j}+3\hat{k}$ for $t=3$ and $2\hat{i}-\hat{j}+2\hat{k}$ for $t=2$.

35. Evaluate $\int_C\vec{f}.d\vec{r}$, where $\vec{f}=(x^2+y^2)\hat{i}-2xy\hat{j}$ and the curve C is the rectangle bounded by $y=0$, $x=a$, $y=b$, and $x=0$.

Ans. $-2ab^2$.

36. If $\vec{f}=2y\hat{i}-z\hat{j}+x\hat{k}$, find the vector line integral $\int_C\vec{f}\times d\vec{r}$ along the curve $x=\cos t$, $y=\sin t$, and $z=2\cos t$ and from $t=0$ to $t=\frac{\pi}{2}$.

Ans. $\left(2-\frac{\pi}{4}\right)\hat{i}+\left(\pi-\frac{1}{2}\right)\hat{j}$.

37. Evaluate $\int_C\vec{f}.d\vec{r}$, where $\vec{f}=yz\hat{i}+zx\hat{j}+xy\hat{k}$ and C is the portion of the curve $\vec{r}=a\cos t\hat{i}+b\sin t\hat{j}+ct\hat{k}$ from $t=0$ to $t=\frac{\pi}{2}$.

Hint: Parametric equations of the curve are $x=a\cos t$, $y=b\sin t$, and $z=ct$.
Also $\frac{d\vec{r}}{dt}=-a\sin t\hat{i}+b\cos t\hat{j}+c\hat{k}$. Putting the values of x, y, z (in terms of t) in $\vec{f}$, we see that $\int_C\vec{f}.\,d\vec{r}=\int_C\vec{f}.\left(\frac{d\vec{r}}{dt}\right)dt=abc(0)=0$.

38. Evaluate $\int_C[y^2dx-x^2dy]$ along the triangle whose vertices are (1, 0), (0, 1) and (−1, 0).

Hint: Find the equations of three sides by a two-point formula and evaluate the integral over those sides.

Ans. $-\frac{2}{3}$.

39. If $\vec{f}=(2x+y)\hat{i}+(3y-x)\hat{j}$, evaluate $\int_C\vec{f}.\,d\vec{r}$, where C is the curve in the xy-plane consisting of straight lines from (0, 0) to (2, 0) and (2, 0) to (3, 2).

Hint: If O (0, 0), A (2, 0), and B (3, 2) are the points, then C consists of two lines OA and AB. On OA, we have $y=0$ so that $dy=0$ and x varies from 0 to 2. The equation of AB is $y=2x-4$ so that $dy=2dx$ and on this line, x

varies from 2 to 3. Therefore, $\int_C \vec{f}.\, d\vec{r} = \int_{OA} + \int_{AB}$, which will come out to be 4 + 7 = 11.

40. Find the circulation of $\vec{f}$ around the curve C, where $\vec{f} = y\hat{i} + z\hat{j} + x\hat{k}$ and C is the circle $x^2 + y^2 = 1$ and $z = 0$.
Hint: Parametric equations of C are $x = \cos t$, $y = \sin t$, and $z = 0$, where t varies from 0 to 2π. Then, $\oint_C \vec{f}.\, d\vec{r} = -\int_0^{2\pi} \frac{1-\cos 2t}{2} dt = -\pi$.

41. Find the work done when a force $\vec{f} = (x^2 - y^2 + x)\hat{i} - (2xy + y)\hat{j}$ moves a particle in xy-plane from (0, 0) to (1, 1) along the parabola $y^2 = x$.
Hint: Proceed as in Example 7.49.
Ans. $-\frac{2}{3}$.

42. Compute the work done by a force $\vec{f} = x\hat{i} - z\hat{j} + 2y\hat{k}$ to displace a particle along a closed path C consisting of the segments C_1, C_2, and C_3, such that

$$0 \le x \le 1,\ y = x,\ z = 0 \ on\ C_1,$$
$$0 \le z \le 1,\ x = 1,\ y = 1 \ on\ C_2,\ \text{and}$$
$$0 \le x \le 1,\ y = z = x \ on\ C_3.$$

Ans. $\frac{3}{2}$.

43. Find the work done in moving a particle once around a circle C in the xy-plane, if the circle has its center at the origin with a radius 3, and if the force field is given by $\vec{f} = (2x - y + z)\hat{i} + (x + y - z^2)\hat{j} + (3x - 2y + 4z)\hat{k}$.
Hint: Parametric equations of C are $x = 3 \cos t$, $y = 3 \sin t$, and $0 \le t \le 2\pi$.
Ans. 18π.

Surface Integrals

44. Evaluate $\iint_S \vec{f}.\, \hat{n}\, ds$, where $\vec{f} = 12x^2y\hat{i} - 3yz\hat{j} + 2z\hat{k}$ and S is the portion of the plane $x + y + z = 1$, included in the first octant.
Hint: $\hat{n} = \frac{\hat{i}+\hat{j}+\hat{k}}{\sqrt{3}}$ and $\vec{f}.\hat{n} = \frac{1}{\sqrt{3}}(12x^2y - 3yz + 2z)$

$$= \frac{1}{\sqrt{3}}[12x^2y - 3y(1-x-y) + 2(1-x-y).$$

$\hat{n}.\hat{k} = \frac{1}{\sqrt{3}}$. Evaluate $\iint \hat{f}.\, \hat{n}\, ds$.
Ans. $\frac{49}{120}$.

45. Evaluate $\iint_S \vec{f}.\, \hat{n}\, dS$, where $\vec{f} = (x + y^2)\hat{i} - 2x\hat{j} + 2yz\hat{k}$ and S is the surface of the plane $2x + y + 2z = 6$, in the first octant.
Hint: Proceed as in Example 7.56. **Ans.** 81.

46. Evaluate $\iint_S \vec{f}.\, \hat{n} dS$, where $\hat{f} = 4xy\,\hat{i} + yz\,\hat{j} - xy\,\hat{k}$ and S is the surface bounded by the planes $x = 0$, $x = 2$, $y = 0$, $y = 2$, $z = 0$, and $z = 2$.
Hint: Proceed as in Example 7.59. **Ans.** 40.

47. Evaluate $\iint_S \phi\, \hat{n}\, dS$, where $\phi = \frac{3}{8}xyz$ and S is the surface of the cylinder $x^2 + y^2 = 16$, included in the first octant between $z = 0$ and $z = 5$.
Hint: $\nabla(x^2 + y^2 - 16) = 2x\hat{i} + 2y\hat{j}$, $\hat{n} = \frac{2x\hat{i}+2y\hat{j}}{\sqrt{4x^2+4y^2}}$, and

$$\hat{n}.\hat{j} = \frac{(x\hat{i}+y\hat{j})}{4}.\hat{j} = \frac{y}{4} \text{ Therefore,}$$

$$\iint_S \phi\hat{n}\, ds = \iint_S \phi\hat{n}\frac{dxdz}{|\hat{n}.\hat{j}|}$$

$$= \iint_R \frac{3}{8}xyz\frac{(x\hat{i}+y\hat{j})}{4} . \frac{dxdz}{\frac{y}{4}}$$

$$= \frac{3}{8}\iint_R xz(x\hat{i}+y\hat{j})dxdz$$

$$= \frac{3}{8}\int_0^5\int_0^4 \left(x^2z\hat{i} + xz\sqrt{16-x^2}\hat{j}\right)dxdz = 100(\hat{i}+\hat{j}).$$

48. Evaluate $\iint_S \vec{f}\, .\, \hat{n}\, dS$, where $\vec{f} = y\hat{i} + 2x\,\hat{j} - z\hat{k}$ and S is the surface of the plane $2x + y = 6$, in the first octant cut off by the plane $z = 4$.
Ans. 108.

Volume Integral

49. Evaluate $\iint_S \phi dV$, where $\phi = 45x^2 y$ and V is the region bounded by the planes $4x + 2y + z = 8$, $x = 0$, $y = 0$, and $z = 0$.
Ans. 128.

50. Evaluate $\iiint_V (2x + y)dV$, where V is the closed region bounded by the cylinder $z = 4 - x^2$ and the planes $x = 0$, $y = 0$, $y = 2$, and $z = 0$.
Hint: The limits of integration are $x = 0$ to $x = 2$, $y = 0$ to $y = 2$, and $z = 0$ to $z = 4 - x^2$.
Ans. $\frac{80}{3}$.

51. Evaluate $\iiint_V \operatorname{div} \vec{f}\, dV$, where $\vec{f} = (x^2 - yz)\hat{i} - 2x^2y\hat{j} + 2\hat{k}$ and the region V is enclosed by the planes $x = 0, x = a, y = 0, y = a.$ and $z = 0$, and $z = a$.
Hint: See Example 7.67. **Ans.** $\frac{a^5}{3}$.

52. If $\vec{f} = 2xz\hat{i} - x\hat{j} + y^2\hat{k}$, evaluate $\iiint_V \vec{f}\, dV$, where V is the region bounded by the surfaces $x = 0$, $x = 2$, $y = 0$, $y = 6$, $z = x^2$ and $z = 4$.
Ans. $128\hat{i} - 24\hat{j} + 384\hat{k}$.

Gauss's Divergence Theorem

53. If $\vec{f} = x\hat{i} + 2y\hat{j} + 7z\hat{k}$, evaluate $\iint_S \vec{f}\,.\,\hat{n}dS$, where S is the surface enclosing volume V.
Hint: By Divergence Theorem,

$$\iint_S \vec{f}.\hat{n}dS = \iiint_V \left(\nabla.\vec{f}\right)dV$$
$$= \iiint_V \left[\frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(2y) + \frac{\partial}{\partial z}(7z)\right]dV$$
$$= \iiint_V (1+2+7)dV = 10V.$$

54. Verify divergence theorem for $\vec{f} = (x^2 - yz)\hat{i} + (y^2 - zx)\hat{j} + (z^2 - xy)\hat{k}$, taken over the rectangular parallelopiped $0 \le x \le a$, $0 \le y \le b$, and $0 \le z \le c$.
Hint: $\operatorname{div}\vec{f} = 2(x + y + z)$. Therefore,

$$\iiint_V \operatorname{div}\vec{f}\, dV = \int_0^a \int_0^b \int_0^c 2(x+y+z)dxdydz$$
$$= abc(a + b + c).$$

Evaluate $\iint \vec{f}\,.\,\hat{n}\, dS$ on all the six faces and add. We shall get $\iint_S \vec{f}.\,\hat{n}\, dS = abc(a + b + c)$. Thus, the theorem is verified.

55. Evaluate $\iint_S \vec{f}.\hat{n}dS$, where $\vec{f} = (2x + 3z)\hat{i} - (xz + y)\hat{j} + (y^2 + 2z)\hat{k}$ and S is the surface of the sphere having a radius 3.
Hint: By divergence theorem,

$$\iint_S \vec{f}.\hat{n}\, dS = \iiint_V \operatorname{div}\vec{f}\, dV$$
$$= \iiint_V \left[\frac{\partial}{\partial x}(2x+3z) + \frac{\partial}{\partial y}(-xz+y) + \frac{\partial}{\partial z}(y^2+2z)\right]dV$$
$$= \iiint_V 3dV$$
$$= 3V = 3\left(\frac{4}{3}\pi(3)^3\right) = 108\pi.$$

56. Verify Divergence Theorem for the function $\vec{f} = y\hat{i} + x\hat{j} + z^2\hat{k}$ over the cylindrical region bounded by $x^2 + y^2 = 9$, $z = 0$, and $z = 2$.
Hint: Proceed as in Example 7.64.

57. Verify Divergence Theorem for $\vec{f} = y\hat{i} + x\hat{j} + z^2\hat{k}$ over the cylindrical region bounded by $x^2 + y^2 = a^2$, $z = 0$, and $z = h$. (Similar to Exercise 56.)

58. Evaluate $\iint_S \vec{f}.\hat{n}dS$, where $\vec{f} = x^3\hat{i} + y^3\hat{j} + z^3\hat{k}$ and S is the surface of the sphere $x^2 + y^2 + z^2 = a^2$.
Hint: By divergence theorem,

$$\iint_S \vec{f}.\hat{n}\, dS = \iiint_V \operatorname{div}\vec{f}\, dV$$
$$= \iiint_V \left(3x^2 + 3y^2 + 3z^2\right)dV$$
$$= 3a^2 \iiint_V dV = 3a^2 V$$
$$= 3a^2.\frac{4}{3}\pi a^3 = 4\pi a^5.$$

59. Evaluate $\iint_S \vec{f}.\hat{n}\, dS$ for $\vec{f} = x\hat{i} - y\hat{j} + 2z\hat{k}$ over the sphere $x^2 + y^2 + (z - 1)^2 = 1$.
Ans. $\frac{8}{3}\pi$.

Green's Theorem

60. Verify Green's theorem in the xy-plane for $\oint_C [(xy^2 - 2xy)dx + (x^2y + 3)dy]$ around the boundary C of the region enclosed by $y^2 = 8x$ and $x = 2$.
Ans. $\oint_C (f_1dx + f_2dy) = \iint_S \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right)dxdy = \frac{128}{5}$.

61. Evaluate by Green's Theorem $\oint_C e^{-x}(\sin y dx + \cos y dy)$, where C is the rectangle with vertices (0,0), $(\pi, 0)$, $\left(\pi, \frac{\pi}{2}\right)$, and $\left(0, \frac{\pi}{2}\right)$.
Ans. $2e^{-x} - 2$.

62. Verify Green's theorem in the plane for $\oint_C [(2xy-x^2)dx+(x^2+y^2)dy]$, where C is the boundary of the region enclosed by $y=x^2$ and $y^2=x$.
Hint: The two parabolas intersect at (0,0) and (1,1).
$$\iint_S \left(\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right)dxdy=\int_0^1\int_{y=x^2}^{\sqrt{x}}(2x-2x)dxdy=0.$$
Along the lower portion C_1, we have $x^2=y$, so that $2xdx=dy$ and x varies from 0 to 1.
$$\int_{c_1}(f_1dx+f_2dy)=\int_0^1 [2x^3-x^2+(x^2+x^4)2x\,dx$$
$$=\int_0^1 (4x^3+2x^5-x^2)dx$$
$$=\left[\frac{4x^4}{5}+\frac{2x^6}{6}-\frac{x^3}{3}\right]_0^1$$
$$=1+\frac{1}{3}-\frac{1}{3}=1.$$
Along the upper portion, we have $y^2=x$ so that $2\,ydy=dx$ and y varies from 1 to 0. Thus,
$$\int_{c_2}(f_1dx+f_2dy)$$
$$=\int_1^0 [(2y^3-y^4)2y+(y^4+y^2)]\,dy$$
$$=\int_1^0 [4y^4-2y^5+y^4+y^2]\,dy$$
$$=\int_1^0 (5y^4-2y^5+y^2)dy=\left[\frac{5y^5}{2}-\frac{2y^6}{6}+\frac{y^3}{3}\right]_1^0$$
$$=-1-\frac{1}{3}+\frac{1}{3}=-1.$$
So, $\oint_C (f_1dx+f_2dy)=1-1=0.$

63. Using Green's theorem in a plane, evaluate $\oint_C [(2x^2-y^2)dx+(x^2+y^2)]\,dy$, where C is the boundary in the xy-plane of the area enclosed by the x-axis and the semi-circle $x^2+y^2=1$ in the upper half of the xy-plane.
Hint: $\oint_C (f_1dx+f_2dy)=\iint_S\left(\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right)dxdy=\int_{-1}^{1}\int_0^{\sqrt{1-x^2}} 2(x+y)dydx=\frac{4}{3}.$

64. Verify Green's theorem in the plane for $\int_C (3x^2-8y^2)dx+(4y-6xy)dy$, where C is the boundary of the region bounded by the parabolas $y=\sqrt{x}$ and $y=x^2$.
Ans. $\oint_C (f_1dx+f_2dy)=\iint_S\left(\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right)dxdy=\frac{3}{2}.$

Stoke's Theorem

65. Verify Stoke's Theorem for the function $\vec{f}=(x^2+y^2)\hat{i}-2xy\,\hat{j}$ taken around the rectangle bounded by $x=\pm a$, $y=0$, and $y=b$.
Hint: $curl\vec{f}=\begin{vmatrix}\hat{i}&\hat{j}&\hat{k}\\ \frac{\partial}{\partial x}&\frac{\partial}{\partial y}&\frac{\partial}{\partial z}\\ x^2+y^2&-2xy&0\end{vmatrix}=-4y\hat{k}.$
For the given surface, $\hat{n}=\hat{k}$. Therefore, $\iint_S curl\vec{f}.\hat{n}dS=\int_0^b\int_{-a}^{a}-4yxdy=-4ab^2$. It can be seen that the line integral $\oint_C \vec{f}.\,d\vec{r}=-4ab^2$.

66. Evaluate by Stoke's Theorem, the integral $\oint_C (e^x dx+2ydy-dz)$, where C is the curve $x^2+y^2=4$ and $z=2$.
Hint: $curl\vec{f}=0$ and so, $curl\vec{f}.\,\hat{n}=0$. Hence, $\iint_S curl\vec{f}.\hat{n}ds=0$.

67. Verify Stoke's Theorem for the function $\vec{f}=(2x-y)\hat{i}-yz^2\hat{j}-y^2z\hat{k}$, where S is the upper-half surface of the sphere $x^2+y^2+z^2=1$, bounded by its projection on the xy-plane.
Hint: Parametric equations of C are $x=\cos t$, $y=\sin t$, $z=0$, and $0\le t\le 2\pi$. Therefore,
$$\oint \vec{f}.\,d\vec{r}=\int(f_1dx+f_2dy+f_3dz)$$
$$=\int_0^{2\pi}(-2\sin t\cos t+\sin^2 t)dt=\pi.$$
Further, $curl\vec{f}=\hat{k}$. Therefore, $\iint curl\vec{f}.\hat{n}dS=\iint_R \hat{n}.\hat{k}\frac{dxdy}{|\hat{n}.\hat{k}|}$, where R is the projection of S on

xy-plane. Then, $\iint_R \hat{n}.\hat{k}\frac{dxdy}{|\hat{n}.\hat{k}|} = \iint_R dxdy =$ area of $R = \pi(1)^2 = \pi$.

68. Transform the integral $\iint_S curl\vec{f}.\,\hat{n}dS$ into a line integral, if S is a part of the surface of the paraboloid $z = 1 - x^2 - y^2$ for which, $z \geq 0$ and $\vec{f} = y\hat{i} + z\hat{j} + x\hat{k}$.

Hint: Surface S is $x^2 + y^2 = 1$ and $z = 0$ with parametric equations $x = \cos\theta, y = \sin\theta, z = 0$, and $0 \leq \theta < 2\pi$. Use Stoke's Theorem to transform the given integral into a line integral. The value of the line integral will come out to be $-\pi$.

EXAMINATION PAPERS WITH SOLUTIONS

B.TECH
(SEM I) ODD SEMESTER THEORY EXAMINATION
2009–10
MATHEMATICS–I

Time: 3 Hours Total marks: 100

Note : *Attempt **all** questions.*

1. Attempt any **two** parts of the following:

(a) Reduce the matrix:

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

to column echelon form and find its rank.

(b) Verify the Cayley-Hamilton theorem for the matrix:

$$A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix}$$

And hence find A^{-1}.

(c) Find the eigenvalues and the corresponding eigen vectors of the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$$

2. Attempt any **two** parts of the following:

(a) If $y = (x^2 - 1)^n$, prove that

$$(x^2 - 1)y_{n+2} + 2xy_{n+1} - n(n+1)y_n = 0$$

(b) If $u = x^3 + y^3 + z^3 + 3xyz$, show that

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = 3u$$

(c) If $u = f(r)$ where $r^2 = x^2 + y^2$, prove that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f''(r) + \frac{1}{r}f'(r)$$

3. Attempt any **two** parts of the following:

(a) If $x = r\sin\theta\cos\phi,\ y = r\sin\theta\sin\phi,\ z = r\cos\theta$, show that

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = r^2\sin\theta$$

(b) Determine the points where the function

$$f(x, y) = x^3 + y^3 - 3xy$$

has a maximum or minimum.

(c) A rectangular box open at the top is to have a given capacity. Find the dimensions of the box requiring least material for its construction.

4. Attempt any **two** parts of the following:

(a) Evaluate

$$\iint_A xy\ dx\ xy$$

where A is the domain bounded by *x-axis*, ordinate $x = 2a$ and the curve $x^2 = 4ay$.

(b) Find the volume bounded by the cylinder $x^2 + y^2 = 4$ and the planes $y + z = 4$ and $z = 0$.

(c) Evaluate:

$$\iiint_R (x + y + z)dx\ dy\ dz, \quad \text{where}$$

$R : 0 \le x \le 1,\ 1 \le y \le 2,\ 2 \le z \le 3$

5. Attempt any **two** parts of the following:

(a) Find a unit normal vector $\hat{n}$ of the cone of revolution

$$z^2 = 4(x^2 + y^2) \text{ at the point } P : (1, 0, 2).$$

(b) Using Green's theorem evaluate

$$\int_C (x^2 + xy)dx + (x^2 + y^2)dy,$$

where C is the square formed by the lines $y = \pm 1, x = \pm 1$

(c) Verify Stock's theorem for $\vec{F} = xy^2\hat{\imath} + y\hat{\jmath} + z^2x\hat{k}$ for the surface of a rectangular lamina bounded by $x = 0,\ y = 0,\ x = 1,\ y = 2,\ z = 0$

SOLUTIONS

1. (a) Please see Example 4.74(c). Matrix is said to be in column echelon form if

   (i) The first non-zero entry in each non-zero column is 1.

   (ii) The column containing only zeros occurs next to all non-zero columns.

   (iii) The number of zeros above the first non-zero entry in each column is less than the number of such zeros in the next column.

   The given matrix is

$$A = \begin{bmatrix} 1 & 1 & -1 & 1 \\ -1 & 1 & -3 & -3 \\ 1 & 0 & 1 & 2 \\ 1 & -1 & 3 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & -4 & -2 \\ 1 & -1 & 2 & 1 \\ 1 & -2 & 4 & 2 \end{bmatrix} \begin{matrix} C_2 \to C_2 - C_1 \\ C_3 \to C_3 + C_1 \\ C_4 \to C_4 - C_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & -4 & 0 \\ 1 & -1 & 2 & 0 \\ 1 & -2 & 4 & 0 \end{bmatrix} C_4 \to C_4 + C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -2 & 0 & 0 \end{bmatrix} C_3 \to C_3 + 2C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 1 & -\frac{1}{2} & 0 & 0 \\ 1 & -1 & 0 & 0 \end{bmatrix} R_2 \to \frac{1}{2} R_2,$$

   which is column echelon form. The number of non-zero column is two and therefore $\rho(A) = 2$.

(b) Cayley-Hamilton theorem states that ''every matrix satisfies its characteristic equation''. We are given that

$$A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix}.$$

Therefore

$$A^2 = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 17 & -16 & -16 \\ -16 & 41 & 32 \\ -16 & 32 & 41 \end{bmatrix}$$

and

$$A^3 = A^2.A = \begin{bmatrix} 81 & -144 & -180 \\ -144 & 333 & 324 \\ -180 & 324 & 315 \end{bmatrix}.$$

On the other hand

$$|A-\lambda I| = \begin{vmatrix} 1-\lambda & 0 & -4 \\ 0 & 5-\lambda & 4 \\ -4 & 4 & 3-\lambda \end{vmatrix}$$
$$= (1-\lambda)[(5-\lambda)(3-\lambda)-16]-4[4(5-\lambda)]$$
$$= (1-\lambda)(\lambda^2-8\lambda-1)-80+16\lambda$$
$$= -\lambda^3+9\lambda^2+9\lambda-81.$$

Therefore, the characteristic equation of the matrix A is

$$\lambda^3-9\lambda^2-9\lambda+81=0. \quad (1)$$

To verify Cayley Hamilton Theorem, we have to show that

$$A^3-9A^2-9A+81I=0. \quad (2)$$

We note that

$$\begin{bmatrix} 81 & -144 & -180 \\ -144 & 333 & 324 \\ -180 & 324 & 315 \end{bmatrix} - 9\begin{bmatrix} 17 & -16 & -16 \\ -16 & 41 & 32 \\ -16 & 32 & 41 \end{bmatrix} - 9\begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix}$$
$$+81\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence A satisfies its characteristic equation.

(c) We have

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}.$$

The characteristic equation of A is

$$|A-\lambda I| = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & 2-\lambda & 1 \\ 2 & 0 & 3-\lambda \end{vmatrix} = 0$$

or

$$\lambda^3-6\lambda^2+11\lambda-6=0,$$

which yields $\lambda = 1,\ 2,\ 3$. Hence the characteristic roots are 1, 2 and 3.

The eigenvector corresponding to $\lambda = 1$ is given by $(A-I)X=0$, that is, by

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 2 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Thus, we have

$$x_2+x_3=0,$$
$$2x_1+2x_3=0.$$

Hence $x_1 = x_2 = -x_3$. Taking $x_3 = -1$, we get the vector

$$X_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

The eigenvector corresponding to the eigenvalue 2 is given by $(A - 2I)X = 0$, that is, by

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This equation yields

$$X_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ as one of the vector.}$$

Similarly, the eigenvector corresponding to $\lambda = 3$ is given by $(A - 3I)X = 0$ or by

$$\begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 1 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which yield $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ as the of the solution. Hence $X_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

2. (a) We have

$$y = (x^2 - 1)^n.$$

Therefore

$$y_1 = 2nx(x^2 - 1)^{n-1}$$
$$y_2 = 2^2x^2n(n - 1)(x^2 - 1)^{n-2} + 2n(x^2 - 1)^{n-1}.$$

Thus

$$\begin{aligned}(x^2 - 1)y_2 &= 2(n - 1)x[2nx(x^2 - 1)^{n-1}] + 2n(x^2 - 1)^n \\ &= 2(n - 1)xy_1 + 2ny. \end{aligned} \tag{1}$$

Differentiating (1) n times by Leibnitz's Theorem, we get

$$(x^2 - 1)y_{n+2} + {}^nc_1y_{n+1}(2x) + {}^nc_2y_n(2) - 2(n - 1)[y_{n+1}(x) + {}^nc_1y_n] - 2ny_n = 0$$

or

$$(x^2 - 1)y_{n+2} + 2nxy_{n+1} + n(n - 1)y_n - 2(n - 1)xy_{n+1} - 2n(n - 1)y_n - 2ny_n = 0$$

or

$$(x^2 - 1)y_{n+2} + 2xy_{n+1} - n(n + 1)y_n = 0.$$

(b) We have

$$u = x^3 + y^3 + z^3 + 3xyz.$$

Replacing x by tx, y by ty and z by tz, we get

$$\begin{aligned}u(tx, ty, tz) &= t^3x^3 + t^3y^3 + t^3z^3 + 3txtytz \\ &= t^3(x^3 + y^3 + z^3 + 3xyz) = t^3u(x, y, z).\end{aligned}$$

Hence $u(x,y,z)$ is a homogeneous function of degree 3 in $u(x,y,z)$. Therefore, by Euler's Theorem,

$$x\frac{\partial u}{\partial z}+y\frac{\partial u}{\partial y}+z\frac{\partial u}{\partial z}=3u.$$

(c) Example 3.9

3. (a) Example 3.72
   (b) Example 3.61 (putting $a=1$)
   (c) Example 3.63 (replacing 32 by V)

4. (a) Example 6.9
   (b) Example 6.47
   (c) We have

$$\int_2^3\int_1^2\int_0^1 (x+y+z)dx\,dy\,dz = \int_2^3\int_1^2\left[\frac{x^2}{2}+xy+xz\right]_0^1 dy\,dz$$

$$= \int_2^3\int_1^2\left[\frac{1}{2}+y+z\right]dy\,dz$$

$$= \int_2^3\left[\frac{1}{2}y+\frac{y^2}{2}+yz\right]^2 dz$$

$$= \int_2^3\left[1+2+2z-\left(\frac{1}{2}+\frac{1}{2}+z\right)\right]dz$$

$$= \int_2^3 [z+2]dz = \left[\frac{z^2}{2}+2z\right]_2^3$$

$$= \left(\frac{9}{2}+6\right)-(2+4)=\frac{9}{2}.$$

5. (a) Similar to Exercise 16 of Chapter 7.
   Let $\phi = z^2-4x^2-4y^2$. Then $\nabla\phi$ is along the normal vector.
   But

$$\nabla\phi = \left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)(z^2-4x^2-4y^2)$$

$$= -8x\,\hat{i}-8y\,\hat{j}+2z\,\hat{k}$$

$$= -8\,\hat{i}+4\,\hat{k} \text{ at the point } (1,0,2).$$

Therefore unit normal vector $\hat{n}$ to the given cone at $(1,0,2)$ is

$$\hat{n}=\frac{-8\,\hat{i}+4\,\hat{k}}{\sqrt{64+16}}=\frac{-8\,\hat{i}+4\,\hat{k}}{\sqrt{80}}=\frac{-2\,\hat{i}+\hat{k}}{\sqrt{5}}$$

(b) Let

$$f_1(x,y) = x^2 + xy \quad \text{and} \quad f_2(x,y) = x^2 + y^2.$$

Then, by Green's Theorem,

$$\int_C (f_1 dx + f_2 dy) = \iint_R \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) dx\, dy$$

$$= \int_{-1}^{1} \int_{-1}^{1} (2x - x) dx\, dy$$

$$= \int_{-1}^{1} \int_{-1}^{1} x dx\, dy$$

$$= \int_{-1}^{1} \left[\frac{x^2}{2}\right]_{-1}^{1} dy = 0$$

(c) Please see Example 7.96.
We have

$$\vec{f} = xy^2 \hat{i} + y \hat{j} + z^2 x \hat{k}.$$

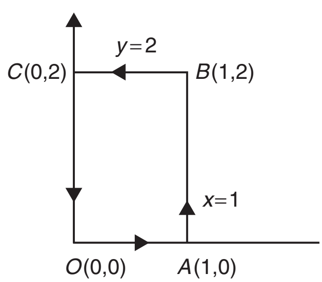
Therefore

$$\vec{f} \cdot d\vec{r} = (xy^2 \hat{i} + y \hat{j} + z^2 x \hat{k}) \cdot (\hat{i}\, dx + \hat{j}\, dy + \hat{k}\, dz)$$
$$= xy^2 dx + y dy + z^2 x dz.$$

Therefore

$$\oint_C \vec{f} \cdot d\vec{r} = \oint_C (xy^2 dx + y dy + z^2 x dz)$$

$$= \int_{OA} + \int_{AB} + \int_{BC} + \int_{CO}$$



Along OA, we have $y = 0$ and $dy = 0$. Therefore

$$\int_{OA} \vec{f} \cdot d\vec{r} = \int_0^1 0\,dx = 0.$$

Along AB, we have $x = 1$ and $dx = 0$. Therefore

$$\int_{AB} \vec{f} \cdot d\vec{r} = \int_0^2 y dy = \left[\frac{y^2}{2}\right]_0^2 = 2.$$

Along BC, we have $y = 2$ and $dy = 0$. Therefore

$$\int_{BC} \vec{f} \cdot d\vec{r} = \int_1^0 4x\,dx = 4\left[\frac{x^2}{2}\right]_1^0 = -2.$$

Along CO, we have $x = 0$ and $dx = 0$. Therefore

$$\int_{CO} \vec{f} \cdot d\vec{r} = \int_2^0 y dy = \left[\frac{y^2}{2}\right]_2^0 = -2.$$

Hence

$$\oint_C \vec{f} \cdot d\vec{r} = 0 + 2 - 2 - 2 = -2.$$

On the other hand,

$$\begin{aligned} curl\,\vec{f} &= \nabla \times \vec{f} \\ &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xy^2 & y & z^2 x \end{vmatrix} \\ &= \hat{i}[0] + \hat{j}[0 - z^2] + \hat{k}[-2xy] \\ &= -z^2\,\hat{j} + 2xy\,\hat{k}. \end{aligned}$$

Therefore

$$curl\,\vec{f} \cdot \hat{n}\,.ds = (-z^2\,\hat{j} + 2xy\,\hat{k})\,\hat{k} = -2xy$$

and so

$$\iint_S curl\,\vec{f}\cdot\hat{n}\cdot ds = -2\int_0^1\left(\int_0^2 xy\,dy\right)dx$$

$$= -2\int_0^1\left[x\frac{y^2}{2}\right]_0^2 dx$$

$$= -2\int_0^1 2x\,dx = -4\left[\frac{x^2}{2}\right]_0^1 = -2.$$

Hence

$$\oint_C \vec{f}\cdot d\vec{r} = \iint_S curl\,\vec{f}\,.\hat{n}\,ds$$

and Stoke's Theorem is verified.

U.P. TECHNICAL UNIVERSITY, LUCKNOW

B.TECH. (C.O.)
FIRST SEMESTER EXAMINATION, 2008–2009
MATHEMATICS–I
(PAPER ID: 9916)

Time : 3 Hours Total Marks : 100

Note: *Attempt **all** Questions.*

1. Attempt any **two** parts of the following:

   (a) Find all values of μ for which rank of the matrix.

$$A = \begin{bmatrix} \mu & -1 & 0 & 0 \\ 0 & \mu & -1 & 0 \\ 0 & 0 & \mu & -1 \\ -6 & 11 & -6 & 1 \end{bmatrix} \text{ is equal to 3.}$$

   (b) If $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, then show that $A^n = A^{n-2} + A^2 - I$ for $n \geq 3$.

   (c) Show that the matrix $A = \begin{bmatrix} 3 & 1 & -1 \\ -2 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$ is diagonalizable. Hence, find P such that $P^{-1}\,AP$ is a diagonal matrix.

2. Attempt any **two** parts of the following:

   (a) Find $(y_n)_0$ when $y = \sin(\alpha \sin^{-1} x)$.

   (b) If $u = e^{xyz}$, show that $\frac{\partial^2 u}{\partial x \partial y \partial z} = (1 + 3xyz + x^2y^2z^2)e^{xyz}$.

   (c) Trace the curve $y^2(a+x) = x^2(3a-x)$.

3. Attempt any **two** parts of the following:

   (a) Show that the functions $u = x^2 + y^2 + z^2,\ v = x + y + z,\ w = yz + zx + xy$ are not independent of one another.

(b) The height h and the semi-vertical angle α of a cone are measured, and from them A, the total surface area of the cone, including the base, is calculated. If h and α are in error by small quantities δh and $\delta\alpha$ respectively, find corresponding error in the area. Show further that, $\alpha = \frac{\pi}{6}$, an error of $+1$ percent in h will be approximately compensated by an error of $-0^\circ.33$ in α.

(c) Determine the points where the function $x^2 + y^3 - 3axy$ has a maximum or minimum.

(d) Find the point upon the plane $ax + by + cz = p$ at which the function $f = x^2 + y^2 + z^2$ has a mnimum value and find this minimum f.

4. Attempt any **two** parts of the following:

(a) Evaluate: $\iint_R xy\,dx\,dy$
where R is the quadrant of the circle $x^2 + y^2 = a^2$ where $x \geq 0$ and $y \geq 0$.

(b) Find the volume common to the cylinders $x^2 + y^2 = a^2$ and $x^2 + z^2 = a^2$

(c) Evaluate: $\iiint_R (x - 2y + z)dxdydz$ where R: $0 \leq x \leq 1, 0 \leq y \leq x^2, 0 \leq z \leq x + y$.

5. Attempt any **two** parts of the following:

(a) Find the directional derivative of $f(x, y, z) = 2x^2 + 3y^2 + z^2$ at the point $P(2, 1, 3)$ in the direction of the vector $\hat{a} = \hat{i} - 2\hat{k}$.

(b) Show that $\iint_S \vec{F} \cdot \hat{n} dS = \frac{3}{2}$, where $\vec{F} = 4xz\hat{i} - y^2\hat{i} + yz\hat{k}$ and S is the surface of the cube bounded by the planes $x = 0,\ x = 1,\ y = 0,\ y = 1,\ z = 0,\ z = 1$.

(c) Use the Stoke's theorem to evaluate $\int_C [(x + 2y)\,dx + (x - z)\,dy + (y - z)dz]$ where C is the boundary of the triangle with vertices $(2, 0, 0), (0, 3, 0)$ and $(0, 0, 6)$ oriented in the anti-clockwise direction.

SOLUTIONS

1. (a) Similar to Remark 4.5
We are given that

$$A = \begin{bmatrix} \mu & -1 & 0 & 0 \\ 0 & \mu & -1 & 0 \\ 0 & 0 & \mu & -1 \\ -6 & 11 & -6 & 1 \end{bmatrix}$$

Therefore

$$|A| = \mu \begin{vmatrix} \mu & -1 & 0 \\ 0 & \mu & -1 \\ -6 & 11 & -6 \end{vmatrix} + 1 \begin{vmatrix} 0 & -1 & 0 \\ 0 & \mu & -1 \\ -6 & -6 & 1 \end{vmatrix} = \mu^3 - 6\mu^2 + 11\mu - 6 = 0 \text{ if } \mu = 1,\ 2,\ 3.$$

For $\mu = 3$, we have the singular matrix

$$\begin{bmatrix} 3 & -1 & 0 & 0 \\ 0 & 3 & -1 & 0 \\ 0 & 0 & 3 & -1 \\ -6 & 11 & -6 & 1 \end{bmatrix},$$

which has non-singular sub-matrix

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

Thus for $\mu = 3$, the rank of the matrix A is 3. Similarly, the rank is 3 for $\mu = 2$ and $\mu = 1$. For other values of μ, we have $|A| \neq 0$ and so $\rho(A) = 4$ for other values of μ.

(b) We have

$$A^2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Then

$$A + A^2 - I = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Also

$$A^3 = A^2A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Hence for $n = 3$, the relation

$$A^n = A^{n-2} + A^2 - I \tag{1}$$

holds. We want to show that it holds for $n \geq 3$ We prove the result using mathematical induction. We have

$$\begin{aligned} A^{n+1} &= A^n.A = [A^{n-2} + A^2 - I]A \\ &= A^{(n+1)-2} + A^3 - AI \\ &= A^{(n+1)-2} + [A + A^2 - I] - A \\ &= A^{(n+1)-2} + A^2 - I. \end{aligned}$$

Hence, by mathematical induction, the result holds for all $n \geq 3$.

(c) The characteristic matrix of the given matrix A is

$$|A - \lambda I| = \begin{vmatrix} 3-\lambda & 1 & -1 \\ -2 & 1-\lambda & 2 \\ 0 & 1 & 2-\lambda \end{vmatrix} = 0$$

or

$$(3-\lambda)[(1-\lambda)(2-\lambda) - 2] - 1[-4 + 2\lambda] - 1(-2) = 0$$

or

$$(3-\lambda)(1-\lambda)(2-\lambda) - 6 + 2\lambda + 4 - 2\lambda + 2 = 0$$

or

$$(3-\lambda)(1-\lambda)(2-\lambda) = 0.$$

Hence the given matrix A has distinct characteristic roots $\lambda = 1, 2, 3$. Consequently it is diagonalizable. Now the eigenvector corresponding to $\lambda = 1$ is given by $(A - I)X = 0$, that is, by

$$\begin{bmatrix} 2 & 1 & -1 \\ -2 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus

$$\begin{aligned} &2x_1 + x_2 - x_3 = 0 \\ &-2x_1 + 0x_2 + 2x_3 = 0 \\ &0x_1 + x_2 + x_3 = 0 \end{aligned}$$

and so $x_1 = x_3 = -x_2$. Taking $x_2 = -1$, we get an eigenvector corresponding to $\lambda = 1$ as

$$X_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Now eigenvector corresponding to $\lambda = 2$ is given by $(A - 2I)X = 0$, that is, by

$$\begin{bmatrix} 1 & 1 & -1 \\ -2 & -1 & 2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus

$$\begin{aligned} &x_1 + x_2 - x_3 = 0 \\ &-2x_1 - x_2 + 2x_3 = 0 \\ &x_2 = 0 \end{aligned}$$

For this system $x_1 = 1,\ x_2 = 0,\ x_3 = 1$ is a solution. Therefore

$$X_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

An eigenvector corresponding to $\lambda = 3$ is given by $(A - 3I)X = 0$, that is, by

$$\begin{bmatrix} 0 & 1 & -1 \\ -2 & -2 & 2 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus

$$\begin{aligned} &x_2 - x_3 = 0 \\ &-2x_1 - 2x_2 + 2x_3 = 0 \\ &x_2 - x_3 = 0, \end{aligned}$$

which yields $x_1 = 0,\ x_2 = 1,\ x_3 = 1$ as one of the solution. Thus

$$X_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Therefore the transforming matrix is

$$P = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

and so the diagonal matrix is

$$P^{-1}AP = \begin{bmatrix} -1 & -1 & 1 \\ 2 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & -1 \\ -2 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

2. (a) We are given that

$$y = \sin(a \sin^{-1} x)$$

Therefore

$$y_1 = \cos(a \sin^{-1} x) \frac{a}{\sqrt{1 - x^2}} \tag{1}$$

or

$$\sqrt{1 - x^2}\, y_1 = a \cos(a\ \sin^{-1} x)$$

or

$$(1 - x^2) y_1^2 = a^2 \cos^2(a \sin^{-1} x) = a^2[1 - \sin^2(a \sin^{-1} x)]$$

or

$$(1 - x^2) y_1^2 = a^2(1 - y^2).$$

Differentiating again with respect to x, we get

$$(1 - x^2) 2y_1 y_2 - 2xy_1^2 = -2a^2 yy_1$$

or

$$(1 - x^2) y_2 - xy_1 + a^2 y = 0 \tag{2}$$

Differentiating (2), n times by Leibnitz-theorem, we have

$$[y_{n+2}(1 - x^2) + {}^n c_1 y_{n+1}(-2x) + {}^n c_2 y_n(-2)]$$
$$- [y_{n+1}(x) + {}^n c_1 y_n(1)] + a^2 y_n = 0$$

or

$$(1 - x^2) y_{n+2} - (2n + 1) xy_{n+1} - (n^2 - a^2) y_n = 0 \tag{3}$$

Putting $x = 0$ in (1), (2) and (3) we get

$$y_1(0) = a(1 - a^2), \quad y_2(0) = 0$$
$$y_n(0) = a(1^2 - a^2)(3^2 - a^2) \ldots [(n - 2)^2 - a^2] \text{ for odd } n$$

and

$$y_n(0) = 0 \text{ for even } n.$$

(b) We have

$$u = e^{xyz}.$$

Therefore

$$\frac{\partial u}{\partial z} = (xy)e^{xyz},$$
$$\frac{\partial^2 u}{\partial y\,\partial z} = (xy)(xz)e^{xyz} + xe^{xyz} = e^{xyz}[x + x^2yz]$$

and then

$$\begin{aligned}\frac{\partial^3 u}{\partial x\,\partial y\,\partial z} &= e^{xyz}(yz)[x + x^2yz] + e^{xyz}[1 + 2xyz]\\ &= e^{xyz}[xyz + x^2y^2z^2 + 1 + 2xyz]\\ &= (1 + 3xyz + x^2y^2z^2)e^{xyz}.\end{aligned}$$

(c) Similar to Example 2.24. The only difference is that the curve intersect the x-axis at $x = 0$ and $x = 3a$, that is, at the points $(0, 0)$ and $(3a, 0)$. Tangent at $(3a, 0)$ is parallel to y-axis. Also, $y = \pm\sqrt{3}x$ are two real and distinct tangents at the origin and so origin is a node. Also $x = -a$ is asymptote parallel to the x-axis. There is no oblique asymptote.

3. (a) We have

$$u = x^2 + y^2 + z^2, \quad v = x + y + z, \quad w = yz + zx + xy.$$

Then

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix} = \begin{vmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \\ z + y & z + x & x + y \end{vmatrix}$$

$$\begin{aligned}&= 2x[(x + y) - (z + x) - 2y[(x + y) - (z + y)] + 2z[(z + x) - (z + y)]\\ &= 0.\end{aligned}$$

Since Jacobian $J(u, v, w) = 0$, there exists a functional relation connecting some or all of the variables x, y and z. Hence u, v, w are not independent.

(b) Radius of the base $= r = h\tan\alpha$. Further, slant height $= l = h\sec\alpha$. Therefore

$$\begin{aligned}\text{Total area } &= \pi r^2 + \pi r l = \pi r(r + l)\\ &= \pi h \tan\alpha(h \tan\alpha + h\sec\alpha)\\ &= \pi h^2(\tan^2\alpha + \sec\alpha\tan\alpha).\end{aligned}$$

Then the error in A is given by

$$\begin{aligned}\delta A &= \frac{\partial A}{\partial h}\delta h + \frac{\partial A}{\partial \alpha}\delta\alpha\\ &= 2\pi h(\tan^2\alpha + \sec\alpha\tan\alpha)\delta h + \pi h^2(2\tan\alpha\sec^2\alpha + \sec^3\alpha + \sec\alpha\tan^2)\delta\alpha\end{aligned}$$

For the second part of the question,

$$\alpha = \frac{\pi}{6}, \quad \delta h = \frac{h}{100}.$$

Therefore

$$\delta A = 2\pi h\left[\frac{1}{3}+\frac{2}{3}\right]\frac{h}{100}+\pi h^2\left(\frac{2}{\sqrt{3}}\left(\frac{4}{3}\right)+\frac{8}{3\sqrt{3}}+\frac{2}{3\sqrt{3}}\right)\delta\alpha$$

$$= \frac{\pi h^2}{50} + 2\sqrt{3}\pi h^2 \delta\alpha. \qquad (1)$$

But after compensation $\delta A = 0$. Therefore (1) implies

$$\delta\alpha = -\frac{1}{100\sqrt{3}} \text{ radians } = -\frac{57.3^\circ}{173.2} = -0.33^\circ.$$

4. (a) Example 6.8
   (b) Example 6.50
   (c) We want to evaluate $I = \iiint_R (x-2y+z)dx\,dy\,dz$,
   where

$$R : 0 \le x \le 1,\ \ 0 \le y \le x^2,\ \ 0 \le z \le x+y.$$

We have

$$I = \int_0^1 \int_0^{x^2} \left[\int_0^{x+y} (x-2y+z)dz\right] dy\,dx$$

$$= \int_0^1 \int_0^{x^2} \left[xz - 2yz + \frac{z^2}{2}\right]_0^{x+y} dy\,dx$$

$$= \int_0^1 \int_0^{x^2} \left[x^2 - xy - 2y^2 + \frac{x^2}{2} + xy + \frac{y^2}{2}\right] dy\,dx$$

$$= \frac{3}{2}\int_0^1 \int_0^{x^2} (x^2 - y^2) dy\,dx$$

$$= \frac{3}{2}\int_0^1 \left[x^2 y - \frac{y^3}{3}\right]_0^{x^2} dx$$

$$= \frac{3}{2}\int \left(x^4 - \frac{x^6}{3}\right) dx$$

$$= \frac{3}{2}\left[\frac{x^5}{5} - \frac{x^7}{21}\right]_0^1 = \frac{8}{35}.$$

5. (a) Similar to Example 7.16.
   We have

$$\nabla f = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(2x^2 + 3y^2 + z^2)$$

$$= 4x\,\hat{i} + 6y\,\hat{j} + 2z\,\hat{k}$$

$$= 8\hat{i} + 6\hat{j} + 6\hat{k} \text{ at the point } (2,1,3).$$

Now unit vector in the direction of $\hat{i} - 2\hat{k}$ is

$$\hat{u} = \frac{\hat{i} - 2\hat{k}}{\sqrt{1+4}} = \frac{1}{\sqrt{5}}\left(\hat{i} - 2\hat{k}\right).$$

Therefore, the directional derivative at (2, 1, 3) in the direction of $\hat{i} - 2\hat{k}$ is

$$\nabla f \,.\, \hat{u} = \frac{1}{\sqrt{5}}\left(8\hat{i} + 6\hat{j} + 6\hat{k}\right)\left(\hat{i} - 2\hat{k}\right) = \frac{1}{\sqrt{5}}[8 + 0 - 12] = -\frac{4}{\sqrt{5}}$$

(b) Example 7.59

(c) Similar to Example to 7.79. Here

$$\begin{aligned}
curl\,\vec{f} &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x+2y & x-z & y-z \end{vmatrix} \\
&= \hat{i}\left[\frac{\partial}{\partial y}(y-z) - \frac{\partial}{\partial x}(x-z)\right] \\
&\quad + \hat{j}\left[\frac{\partial}{\partial z}(x+2y) - \frac{\partial}{\partial x}(y-z)\right] + \hat{k}\left[\frac{\partial}{\partial x}(x-z) - \frac{\partial}{\partial y}(x+2y)\right] \\
&= 2\hat{i} - 2\hat{k}.
\end{aligned}$$

Now, by Stoke's Theorem,

$$\begin{aligned}
\oint \vec{f}\cdot d\vec{r} &= \iint_S curl\,\vec{f}\cdot\hat{n}\,ds \\
&= \iint_{OAB} + \iint_{OBC} + \iint_{OAC} \quad \text{(see figure of Example 7.79)} \\
&= \iint_{OAB} \left(2\hat{i} - 2\hat{k}\right).\left(\hat{k}\right) ds \\
&\quad + \iint_{OAB} \left(2\hat{i} - 2\hat{k}\right).\left(\hat{j}\right) ds \\
&\quad + \iint_{OAC} \left(2\hat{i} - 2\hat{k}\right).\left(\hat{i}\right) ds \\
&= -2\int_0^3 \left[\int_0^{\frac{6-2y}{3}} dx\right] dy + 0 + 2\int_0^6 \left[\int_0^{\frac{6-z}{2}} dy\right] dz \\
&= -6 + 18 = 12
\end{aligned}$$

U.P. TECHNICAL UNIVERSITY, LUCKNOW

B.TECH
FIRST SEMESTER EXAMINATION, 2008–2009
MATHEMATICS–I
(PAPER ID: 9601)

Time: 3 Hours | Total marks: 100

SECTION A

All parts of the question are compulsory: **(2 × 10 = 20)**

1. (a) For which value of 'b' the rank of the martix

$$A = \begin{bmatrix} 1 & 5 & 4 \\ 0 & 3 & 2 \\ b & 13 & 10 \end{bmatrix} \text{is } 2,\ b = \ldots$$

(b) Determine the constants a and b such that the curl of vector $\bar{A} = (2xy + 3yz)\hat{i} + (x^2 + axz - 4z^2)\hat{j} + (3xy + 2byz)\ \hat{k}$ is zero, $a = \ldots,\ b = \ldots.$

(c) The n^{th} derivative (y_n) of the function $y = x^2 \sin x$ at $x = 0$ is $\ldots.$

(d) With usual notations, match the items on right hand side with those on left had side for properties of Max^m and minimum:

(i) Max^m	(p) $rt - s^2 = 0$
(ii) Min^m	(q) $rt - s^2 < 0$
(iii) Saddle point	(r) $rt - s^2 > 0,\ r > 0$
(iv) Failure case	(s) $rt - s^2 > 0$ and $r < 0$

(e) Match the items on the right hand side with those on left hand side for the following special functions: (Full marks is awarded if all matching are correct)

(i) $\beta(p, q)$	(p) $\Gamma(1/2)$
(ii) $\frac{\Gamma p \Gamma q}{\Gamma(p+q)}$	(q) $\int_0^\infty \frac{y^{p-1}}{(1+y)^{p+q}} dy$
(iii) $\sqrt{\pi}$	(r) $\beta(p, q)$
(iv) $\frac{\pi}{\sin p\pi}$	(s) $\Gamma p \Gamma(1 - p)$

Indicate True of False for the following statements:

(f) (i) If $|A| = 0$, then at least on eigen value is zero. (True/False)

(ii) A^{-1} exists if 0 is an eigen value of A (True/False)

(iii) If $|A| \neq 0$, then A is known as singular matrix (True/False)

(iv) Two vectors X and Y is said to be orthogonal $Y, X^T Y = Y^T X \neq 0$. (True/False)

(g) (i) The curve $y^2 = 4ax$ is symmetric about x-axis. (True/False)

(ii) The curve $x^3 + y^3 = 3axy$ is symmetric about the line $y = -x$ (True/False)

(iii) The curve $x^2 + y^2 = a^2$ is symmetric about both the axis x and y. (True/False)

(iv) The curve $x^3 - y^3 = 3axy$ is symmetric about the line $y = x$. (True/False)

Pick the correct answer of the choice given bolow:

(h) If $\bar{r} = x\hat{i} + y\hat{j} + z\hat{k}$ is position vector, then value of $\nabla(\log r)$ is

(i) $\frac{\bar{r}}{r}$ (ii) $\frac{\bar{r}}{r^2}$

(iii) $-\frac{\bar{r}}{r^3}$ (iv) None of the above

(i) The Jacobian $\frac{\partial(u,v)}{\partial(x,y)}$ for the function $u = e^x \sin y, v = (x + \log \sin y)$ is

(i) 1 (ii) $\sin x \sin y - xy \cos x \cos y$

(iii) 0 (iv) $\frac{e^x}{x}$.

(j) The volume of the solid under the surface $az = x^2 + y^2$ and whose base R is the circle $x^2 + y^2 = a^2$ is given as

(i) $\pi/2a$ (ii) $\pi a^3/2$

(iii) $\frac{4}{3}\pi a^3$ (iv) None of the above.

SECTION B

2. **Attempt any three parts of the following:** **(3 × 10 = 30)**

(a) If $y = (\sin^{-1} x)^2$, prove that $y_n(0) = 0$ for b odd and $y_n(0) = 2.2^2.4^2.6^2 \ldots (n-2)^2, n \neq 2$ for n is even.

(b) Find the dimension of rectangular box of maximum capacity whose surface area is given when (a) box is open at the top (b) box is closed.

(c) Find a matrix P which diagonalizes the matrix $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$, verify $P^{-1}AP = D$ where D is the diagonal matrix.

(d) Find the area and the mass contained in the first quadrant enclosed by the curve $\left(\frac{x}{a}\right)^\alpha + \left(\frac{y}{b}\right)^\beta = 1$ where $\alpha > 0,\ \beta > 0$ given that density at any point $\rho(x, y)$ is $k\sqrt{xy}$.

(e) Using the divergence theorem, evaluate the surface integral $\iint_S (yz\,dy\,dz + zx\,dz\,dx + xy\,dy\,dx)$ where S: $x^2 + y^2 + z^2 = 4$.

SECTION C

Attempt any two parts from each question. All questions are compulsory: (5 × 10 = 50)

3. (a) Trace the curve $r^2 = a^2 \cos 2\theta$

(b) If $u = \log\left(\frac{(x^2+y^2)}{(x+y)}\right)$, prove that $x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 1$

(c) If $V = f(2x - 3y, 3y - 4z, 4z - 2x)$, compute the value of $6V_x + 4V_y + 3V_z$.

4. (a) The temperature 'T' at any point (x, y, z) in space is $\mathrm{T}(x, y, z) = \mathrm{K}xyz^2$ where K is constant. Find the highest temperature on the surface of the sphere $x^2 + y^2 + z^2 = a^2$.

(b) Verify the chain rule for Jacobians if $x = u, y = u \tan v, z = w$.

(c) The time 'T' of a complete oscillation of a simple pendulum of length 'L' is governed by the equation $\mathrm{T} = 2\pi\sqrt{\frac{\mathrm{L}}{g}}$, g is constant, find the approximate error in the calculated value of T corresponding to an error of 2% in the value of L.

5. (a) Determine 'b' such that the system of homogeneous equation $2x + y + 2z = 0; x + y + 3z = 0;$ $4x + 3y + bz = 0$ has (i) Trivial solution (ii) Non-Trivial solution. Find the Non-Trivial solution using matrix method.

(b) Verify Cayley–Hamilton theorem for the matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$ and hence find A^{-1}.

(c) Find the eigen value and corresponding eigen vectors of the matrix

$$I = \begin{pmatrix} -5 & 2 \\ 2 & -2 \end{pmatrix}$$

6. (a) Find the directional derivative of $\nabla(\nabla f)$ at the point $(1, -2, 1)$ in the direction of the normal to the surface $xy^2z = 3x + z^2$ where $f = 2x^3y^2z^4$.

(b) Using Green's theorem, find the area of the region in the first quadrant bounded by the curves $y = x,\ y = \frac{1}{x},\ y = \frac{x}{4}$.

(c) Prove that $(y^2 - z^2 + 3yz - 2x)\hat{i} + (3xz + 2xy)\hat{j} + (3xy - 2xz + 2z)\hat{k}$ is both solenoidal and irrotational.

7. (a) Changing the order of integration of $\int_0^\infty \int_0^\infty e^{-xy} \sin nx\, dx\, dy$
Show that $\int_0^\infty \left(\frac{\sin nx}{x}\right) dx = \frac{\pi}{2}$

(b) Determine the area bounded by the curves $xy = 2, 4y = x^2$ and $y = 4$.

(c) For a β function, show that

$$\beta(p, q) = \beta(p + 1, q) + \beta(p, q + 1).$$

SOLUTIONS

1. (a) We have to show that $|A| = 0$. But

$$|A| = \begin{vmatrix} 1 & 5 & 4 \\ 0 & 3 & 2 \\ b & 13 & 10 \end{vmatrix}$$
$$= 1[30 - 26] - 5[-2b] + 4[-3b]$$
$$= 4 + 10b - 12b = 4 - 2b.$$

Thus $|A| = 0$ if $2b = 4$ or $b = 2$.

(b) We note that

$$curl\,\vec{A} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 2xy + 3yz & x^2 + axz - 4z^2 & 2xy + 2byz \end{vmatrix}$$
$$= \hat{i}\left[\frac{\partial}{\partial y}(3xy + 2byz) - \frac{\partial}{\partial z}(x^2 + axz - 4z^2\right]$$
$$- \hat{j}\left[\frac{\partial}{\partial x}(3xy + 2byz) - \frac{\partial}{\partial z}(2xy + 3yz)\right]$$
$$+ \hat{k}\left[\frac{\partial}{\partial x}(x^2 + axz - 4z^2) - \frac{\partial}{\partial y}(2xy + 3yz)\right]$$
$$= \hat{i}[(3x + 2bz) - (ax - 8z)]$$
$$+ \hat{j}[3y - 3y] + \hat{k}[(2x + az) - (2x + 3z)]$$
$$= \hat{i}[x(3 - a) + z(2b - 8)] + \hat{k}[(a - 3)z]$$
$$= 0 \text{ for } a = 3 \text{ and } b = 4.$$

(c) Let $y = x^2 \sin x$. Take

$$u = \sin x \quad \text{and} \quad v = x^2$$

Then

$$u_n = \sin\left(x + \frac{n\pi}{2}\right), \;\; v_1 = 2x, \;\; v_2 = 2, \;\; v_3 = 0.$$

Therefore, by Leibnitz-theorem, we have

$$y_n = \sin\left(x+\frac{n\pi}{2}\right)(x^2) + {}^nc_1 \sin\left(x+\frac{(n-1)\pi}{2}\right)(2x)$$
$$+ {}^nc_2 \sin\left(x+\frac{(n-2)\pi}{2}\right)(2)$$
$$= x^2\sin\left(x+\frac{n\pi}{2}\right) + 2nx\sin\left[x+\frac{(n-1)\pi}{2}\right]$$
$$+ n(n-1)\sin\left[x+\frac{(n-2)\pi}{2}\right].$$

Hence

$$(y_n)_0 = n(n-1)\sin\frac{(n-2)\pi}{2} = (n-n^2)\sin\frac{n\pi}{2}.$$

(d) Matching yields

i. $Max \rightarrow rt - s^2 > 0,\quad r < 0$
ii. $Min \rightarrow rt - s^2 > 0,\quad r < 0$
iii. $saddle\,point \rightarrow rt - s^2 < 0$
iv. $Failure\,case \rightarrow rt - s^2 = 0$

(e) The Matching yields

(i) $\beta(p,q) \rightarrow \int_0^\infty \frac{y^{p-1}}{(1+y)^{p+2}}\,dy$
(ii) $\frac{\Gamma(p)\,\Gamma(q)}{\Gamma\,(p+q)} \rightarrow \beta(p,q)$
(iii) $\sqrt{\pi} \rightarrow \Gamma\left(\frac{1}{2}\right)$
(iv) $\frac{\pi}{\sin p\,\pi} \rightarrow \Gamma(p)\,\Gamma(1-p)$

(f)

(i) True
(ii) False
(iii) False
(iv) The statement is wrong

(g)

(i) False
(ii) False
(iii) True
(iv) False

(h)

$$\nabla(\log r) = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(\log r)$$
$$= \hat{i}\left(\frac{1}{r}\frac{\partial r}{\partial x}\right) + \hat{j}\left(\frac{1}{r}\frac{\partial r}{\partial y}\right) + \hat{k}\left(\frac{1}{r}\frac{\partial r}{\partial z}\right)$$

But $r = x\hat{i} + y\hat{j} + z\hat{k}$. Therefore

$$r = \left|\vec{r}\right| = \sqrt{x^2 + y^2 + z^2}$$

or

$$r^2 = x^2 + y^2 + z^2.$$

Thus

$$2r\frac{\partial r}{\partial x} = 2x$$

or

$$\frac{\partial r}{\partial x} = \frac{x}{r}.$$

Similarly,

$$\frac{\partial r}{\partial y} = \frac{y}{r} \quad \text{and} \quad \frac{\partial r}{\partial z} = \frac{z}{r}.$$

Therefore

$$\begin{aligned}\nabla(\log r) &= \hat{i}\left[\frac{1}{r}\left(\frac{x}{r}\right)\right] + \hat{j}\left[\frac{1}{r}\left(\frac{y}{r}\right)\right] + \hat{k}\left[\frac{1}{r}\left(\frac{z}{r}\right)\right] \\ &= \hat{i}\left(\frac{x}{r^2}\right) + \hat{j}\left(\frac{y}{r^2}\right) + \hat{k}\left(\frac{z}{r^2}\right) \\ &= \frac{1}{r^2}\left(x\hat{i} + y\hat{j} + 2\hat{k}\right) = \frac{\vec{r}}{r^2}.\end{aligned}$$

Hence choice (ii) is true

(i) We have

$$u = e^x \sin y, \qquad v = x + \log \sin y.$$

Therefore

$$\begin{aligned}\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} e^x \sin y & e^x \cos y \\ 1 & \cot y \end{vmatrix} \\ = e^x \cos y - e^x \cos y = 0\end{aligned}$$

Hence the choice (iii) is correct

(j) (iii) is correct.

2. (a) We have

$$y = (\sin^{-1} x)^2.$$

Then

$$y_1 = 2\sin^{-1} x.\frac{1}{\sqrt{1-x^2}}$$

or

$$(1-x^2)y_1^2 = 4(\sin^{-1} x)^2 = 4y \tag{1}$$

Differentiating (1) again, we get

$$(1-x^2)y_2 - xy_1 - 2 = 0 \tag{2}$$

Differentiating (2) n times by Leibnitz Theorem, we get

$$(1-x^2)y_{n+2} + {}^n c_1 y_{n+1}(-2x) + {}^n c_2 y_n(-2) - xy_{n+1} - {}^n c_1 y_n = 0$$

or

$$(1-x^2)y_{n+2} - (2n+1)xy_{n+1} - n^2 y_n = 0. \tag{3}$$

Therefore (1), (2), (3) imply

$$(y_1)_0 = 0, \quad (y_2)_0 = 2, \quad (y_{n+2})_0 = n^2(y_n)_0.$$

Taking $n = 1, 2, 3, \ldots$ in $(y_{n+2})_0 = n^2(y_n)_0$, we get

$$\begin{aligned}
(y_3)_0 &= (y_1)_0 = 0\\
(y_4)_0 &= 2^2(y_2)_0 = 2.2^2\\
(y_5)_0 &= 3^2(y_3)_0 = 0\\
(y_6)_0 &= 4^2(y_4)_0 = 2.2^2.4^2
\end{aligned}$$

and so on. Hence, in general

$$(y_n)_0 = 0 \text{ for odd } n$$

and

$$(y_n)_0 = 2.2^2.4^2.6^2.8^2 \ldots (n-2)^2.$$

(b) Please see Example 3.63

Replace 32 by V in this Example and get

$$x = y = (2V)^{\frac{1}{3}}, \quad z = \frac{(2V)^{\frac{1}{3}}}{2} \text{ (for open at the top)}$$

For the closed box

$$S = 2xy + 2yz + 2zx$$

Now proceed as in Example 3.63

(c) Please see Example 4.87(b). We have

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$$

The characteristic equation of A is

$$|A - \lambda I| = \begin{vmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{vmatrix} = 0$$

or

$$(4-\lambda)(3-\lambda) - 2 = 0$$

or

$$\lambda^2 - 7\lambda + 10 = 0$$

The characteristic roots are $\lambda = \frac{7\pm 3}{2} = 2, 5$. Since the eigenvalues are distinct, the matrix A is diagonalizable. The eigenvector corresponding to $\lambda = 2$ is given by $(A - 2I)X = 0$, that is, by

$$\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or by

$$2x_1 + x_2 = 0 \qquad \text{or} \qquad x_1 = -\frac{x_2}{2}.$$

Putting $x_2 = 2$, we get

$$X_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Similarly, eigenvector corresponding to $\lambda = 5$ is given by $(A - 5I)X = 0$ or by

$$\begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

or by

$$\begin{aligned} -x_1 + x_2 &= 0 \\ 2x_1 - 2x_2 &= 0 \end{aligned}$$

and so $x_1 = x_2$. Putting $x_2 = 1$, we get

$$X_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Thus the transforming matrix is

$$P = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}$$

and

$$P^{-1} = \begin{bmatrix} -\frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Then

$$P^{-1}AP = \begin{bmatrix} -\frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{2}{3} & \frac{2}{3} \\ \frac{10}{3} & \frac{5}{3} \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}.$$

(d) The equation of the curve is

$$\left(\frac{x}{a}\right)^{\alpha} + \left(\frac{y}{b}\right)^{\beta} = 1, \quad \alpha, \beta > 0.$$

The parametric form of the curve is

$$x = a\cos^{\frac{2}{\alpha}} t, \quad y = b\sin^{\frac{2}{\beta}} t.$$

Therefore, the required area is

$$A = \int_{\frac{\pi}{2}}^{0} y\,dx = \int_{\frac{\pi}{2}} y\frac{dx}{dt}dt$$

$$= \int_{\frac{\pi}{2}}^{0} (b\sin^{\frac{2}{\beta}} t)\left(-\frac{2a}{\alpha}\cos^{\left(\frac{2}{\alpha}-1\right)} t\right) \sin t\,dt$$

$$= \frac{2ab}{\alpha} \int_{0}^{\frac{\pi}{2}} \sin^{\left(\frac{2}{\beta}+1\right)} t \cos^{\left(\frac{2}{\alpha}-1\right)} t\,dt$$

$$= \frac{2ab}{\alpha}\left[\frac{\Gamma\left(\frac{\left(\frac{2}{\beta}+2\right)}{2}\right)\Gamma\left(\frac{\left(\frac{2}{\alpha}-1+1\right)}{2}\right)}{2\Gamma\left(\frac{\frac{2}{\beta}+1+\frac{2}{\alpha}-1+2}{2}\right)}\right]$$

$$= \frac{2ab}{2\alpha\beta}\left[\frac{\Gamma\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{1}{\beta}\right)}{\Gamma\left(\frac{1}{\alpha}+\frac{1}{\beta}+1\right)}\right]$$

$$= \frac{ab}{\alpha+\beta}\frac{\Gamma\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{1}{\beta}\right)}{F\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)}$$

(e) By Divergence Theorem,

$$\iint_S (yz\,dy\,dz + zx\,dz\,dx + xy\,dx\,dy)$$

$$= \iiint_V \left[\frac{\partial}{\partial x}(yz) + \frac{\partial}{\partial y}(zx) + \frac{\partial}{\partial z}(xy)\right] dV$$

$$= \iiint_V 0 dV = 0.$$

3. (a) Hints: The given curve is $r^2 = a^2 \cos 2\theta$. It is symmetrical about the initial line. It is symmetrical about the origin. Intersects the initial line at the points $\left(\pm\frac{\pi}{4}\right)$. The curve has no asymptote. At $(a, 0)$ and $(-a, 0)$ the tangents are perpendicular to the initial line. Tangents at the pole are $\theta = \pm\frac{\pi}{4}$. The curve passes through the pole. The figure of the curve is similar to the figure of Example 2.20.

(b) We have

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

Therefore

$$e^u = \frac{x^2 + y^2}{x + y},$$

which is homogeneous function of degree 1 in x and y. Therefore, by Euler's Theorem, we have

$$x\frac{\partial}{\partial x}(e^u) + y\frac{\partial}{\partial y}(e^u) = e^u$$

or

$$x\, e^u \frac{\partial u}{\partial x} + y\, e^u \frac{\partial u}{\partial y} = e^u$$

or

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 1.$$

(c) We have

$$V = f(2x - 3y,\ 3y - 4z,\ 4z - 2x).$$

Let

$$r = 2x - 3y, \quad s = 3y - 4z \quad \text{and} \quad t = 4z - 2x.$$

Then

$$V = f(r, s, t).$$

Further,

$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial x} + \frac{\partial V}{\partial s}\frac{\partial s}{\partial x} + \frac{\partial V}{\partial t}\cdot\frac{\partial t}{\partial x}$$

$$= 2\frac{\partial V}{\partial r} + 0 - 2\frac{\partial V}{\partial t} = 2\frac{\partial V}{\partial r} - 2\frac{\partial V}{\partial t}. \quad (1)$$

$$\frac{\partial V}{\partial y} = \frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial y} + \frac{\partial V}{\partial s}\frac{\partial s}{\partial y} + \frac{\partial V}{\partial t}\cdot\frac{\partial t}{\partial y}$$

$$= -3\frac{\partial V}{\partial r} + 3\frac{\partial V}{\partial s} + 0 = 3\frac{\partial V}{\partial r} + 3\frac{\partial V}{\partial s}. \quad (2)$$

and

$$\frac{\partial V}{\partial z} = \frac{\partial V}{\partial r}\cdot\frac{\partial r}{\partial z} + \frac{\partial V}{\partial s}\frac{\partial s}{\partial z} + \frac{\partial V}{\partial t}\cdot\frac{\partial t}{\partial z}$$

$$= 0 - 4\frac{\partial V}{\partial s} + 4\frac{\partial V}{\partial t}$$

$$= -4\frac{\partial V}{\partial s} + 4\frac{\partial V}{\partial t}.$$

The relations (1), (2) and (3) yields

$$6V_x + 4V_y + 3V_z = 6\left(2\frac{\partial V}{\partial r} - 2\frac{\partial V}{\partial t}\right) + 4\left(-3\frac{\partial V}{\partial r} + 3\frac{\partial V}{\partial s}\right) + 3\left(-4\frac{\partial V}{\partial s} + 4\frac{\partial V}{\partial t}\right) = 0.$$

4. (a) We have

$$T(x, y, z) = kxyz^2 \quad (1)$$

and

$$\phi(x, y, z) = x^2 + y^2 + z^2 - a^2 \quad (2)$$

Taking λ as the Lagrange's multiplier, we have

$$F(x, y, z) = T(x, y, z) + \lambda\phi(x, y, z)$$

$$= kxyz^2 + \lambda(x^2 + y^2 + z^2 - a^2).$$

For maxima or minima of $F(x, y, z)$, we should have

$$\frac{\partial F}{\partial x} = \frac{\partial T}{\partial x} + \lambda\frac{\partial \phi}{\partial x} = Kyz^2 + 2x\lambda = 0 \quad (3)$$

$$\frac{\partial F}{\partial y} = \frac{\partial T}{\partial y} + \lambda\frac{\partial \phi}{\partial y} = kxz^2 + 2\lambda y = 0 \quad (4)$$

$$\frac{\partial F}{\partial z} = \frac{\partial T}{\partial z} + \lambda\frac{\partial \phi}{\partial z} = 2kxyz + 2\lambda z = 0 \quad (5)$$

Multiplying (3) by x, (4) by y and (5) by z and adding we get

$$4kxyz^2 + 2\lambda(x^2 + y^2 + z^2) = 0$$

or

$$4kxyz^2 + 2\lambda a^2 = 0 \quad (\text{using } x^2 + y^2 + z^2 = a^2).$$

Therefore

$$\lambda = -\frac{2kxyz^2}{a^2}.$$

Substituting this value of λ in (3), we have

$$kyz^2 - \frac{4kx^2yz^2}{a^2} = 0$$

or

$$1 - \frac{4x^2}{a^2} = 0$$

Hence $x = \pm\frac{a}{2}$. Similarly $y = \pm\frac{a}{2}$. Substituting the value of x, y, and λ is (5), we get

$$1 - 2\frac{z^2}{a^2} = 0 \quad \text{and so} \quad z = \pm\frac{a}{\sqrt{2}}.$$

Then the highest temperature on the sphere is

$$T = kxyz^2 = k\left(\frac{a}{2}\right)\left(\frac{a}{2}\right)\left(\frac{a^2}{2}\right) = \frac{k}{8}a^4$$

(b) We are given thet

$$x = u,\ y = u\tan v,\ \ z = w. \tag{1}$$

Then

$$J = \frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ \tan v & u\sec^2 v & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$= u\sec^2 v.$$

Also, from (1), we have

$$u = x,\ v = \tan^{-1}\frac{y}{x} \text{ and } w = z.$$

Therefore

$$J' = \frac{\partial(u,v,w)}{\partial(x,y,z)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix} - \begin{vmatrix} 1 & 0 & 0 \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$= \frac{x}{x^2+y^2} = \frac{1}{x\left[1+\left(\frac{y}{x}\right)^2\right]}$$

$$= \frac{1}{u(1+\tan^2 v)}, \text{ since } \frac{y}{u} = \tan v$$

$$= \frac{1}{u \sec^2 v}.$$

Hence

$$J\,J' = 1, \text{ which proves the chain rule.}$$

(c) We have

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

Taking logarithm, we get

$$\log T = \log 2\pi + \frac{1}{2}\log\, l - \frac{1}{2}\log g. \tag{1}$$

Differentiating (1), we get

$$\frac{1}{T}\delta T = \frac{1}{2}\frac{\delta l}{l} - \frac{1}{2}\frac{\delta g}{g}$$

or

$$\frac{\delta T}{T} \times 100 = \frac{1}{2}\left[\frac{\delta l}{l} \times 100 - \frac{1}{2}\frac{\delta g}{g} \times 100\right]$$

$$= \frac{1}{2}[2-0] = 1.$$

Hence the approximate error is 1%.

5. (a) The give system of equation is

$$x + y + 3z = 0$$
$$2x + y + 2z = 0$$
$$4x + 3y + bz = 0.$$

The system in matrix form is

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 2 \\ 4 & 3 & b \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This homogenous system will have a non-trivial solution only if $|A| = 0$. Thus for non-trivial solution

$$\begin{vmatrix} 1 & 1 & 3 \\ 2 & 1 & 2 \\ 4 & 3 & b \end{vmatrix} = 0$$

or

$$1(b-6) - 1(2b-8) + 3(6-4) = 0$$

or

$$-b + 8 = 0, \text{ which yields } b = 8.$$

Thus for non-trivial solution $b = 8$. The coefficient matrix for non-trivial solution is

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 2 \\ 4 & 3 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 3 \\ 0 & -1 & -4 \\ 0 & -1 & -4 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 4R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 3 \\ 0 & -1 & -4 \\ 0 & 0 & 0 \end{bmatrix} R_3 \to R_3 - R_2$$

The last matrix is of rank 2. Thus the given system is equivalent to

$$x + y + 3z = 0$$
$$-y - 4z = 0.$$

Hence $y = -4z$ and then $x = z$. Taking $z = t$ the general solution is

$$x = t, \ y = -4t, \ \ z = t.$$

(b) We have

$$A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}.$$

The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 \\ 2 & -1-\lambda \end{vmatrix} = 0$$

or

$$\lambda^2 - 5 = 0. \tag{1}$$

We note that

$$A^2 = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}.$$

Then

$$A^2 - 5I = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} - \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence A satisfies its characteristic equation. Premultiplication by A^{-1} yields

$$A - 5A^{-1} = 0$$

or

$$A^{-1} = \frac{1}{5}A = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & -\frac{1}{5} \end{bmatrix}.$$

(c) The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix} = 0$$

or

$$(-5-\lambda)(-2-\lambda) - 4 = 0$$

or

$$\lambda^2 + 7\lambda + 6 = 0$$

or

$$\lambda = \frac{-7 \pm \sqrt{49-24}}{2} = -6,\ -1.$$

The characteristic vector corresponding to $\lambda = -2$ is given by $(A + I)X = 0$, that is, by

$$\begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or by

$$-4x_1 + 2x_2 = 0$$
$$2x_1 - x_2 = 0.$$

Thus

$$X_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

The characteristic vector corresponding to $\lambda = -6$ is given by

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or by

$$x_1 + 2x_2 = 0$$
$$2x_1 + 4x_2 = 0.$$

These equations imply

$$X_2 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

6 (a) We have

$$f = 2x^3y^2z^4$$

Therefore

$$\frac{\partial f}{\partial x} = 6x^2y^2z^4, \quad \frac{\partial f}{\partial y} = 4x^3yz^4 \quad \text{and} \quad \frac{\partial f}{\partial z} = 8x^3y^2z^3$$

and so

$$\nabla f = 6x^2y^2z^4 \hat{i} + 4x^3yz^4 \hat{j} + 8x^3y^2z^3 \hat{k}.$$

Then

$$\begin{aligned} \phi = \nabla(\nabla f) &= \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)\left(6x^2y^2z^4 \hat{i} + 4x^3yz^4 \hat{j} + 8x^2y^2z^3 \hat{k}\right) \\ &= 12xy^2z^4 + 4x^3z^4 + 24x^2y^2z^2. \end{aligned}$$

Now

$$\begin{aligned} \nabla\phi &= \frac{\partial\phi}{\partial x}\hat{i} + \frac{\partial\phi}{\partial y}\hat{j} + \frac{\partial\phi}{\partial z}\hat{k} \\ &= (12y^2z^4 + 12x^2z^4 + 48xy^2z^2)\hat{i} + (24zyz^4 + 48x^2yz^2)\hat{j} \\ &\quad + (48xy^2z^3 + 16x^3z^3 + 48x^2y^2z)\hat{k} \\ &= 56\hat{i} - 144\hat{j} + 16\hat{k}. \end{aligned}$$

Now normal to the surface $\psi = xy^2z - 3x - z^2$ is

$$\nabla\psi = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(xy^2z - 3x - z^2)$$
$$= (y^2z - 3)\hat{i} + (2xyz)\hat{j} + (xy^2 - 2z)\hat{k}.$$

The unit vector in the direction of $\nabla\psi$ is

$$\hat{a} = \frac{(y^2z - 3)\hat{i} + (2xyz)\hat{j} + (xy^2 - 2z)\hat{k}}{\sqrt{(y^2z - 3)^2 + (2xyz)^2 + (xy^2 - 2z)^2}}$$
$$= \frac{-\hat{i} - 4\hat{j} + 2\hat{k}}{\sqrt{1 + 16 + 4}} = \frac{1}{\sqrt{21}}\left(-\hat{i} - 4\hat{j} + 2\hat{k}\right)$$

Therefore directional derivative of $\nabla(\nabla f)$ at $(1, -2, 1)$ in the given direction is

$$\nabla\phi.\hat{a} = \left(56\hat{i} - 144\hat{j} + 16\hat{k}\right)\frac{1}{\sqrt{21}}\left(-\hat{i} - 4\hat{j} + 2\hat{k}\right)$$
$$= \frac{1}{\sqrt{21}}[-56 + 576 + 32] = \frac{542}{\sqrt{21}}.$$

(b) Using Green's Theorem,

$$A = \frac{1}{2}\oint_c (x\,dy - ydx),$$
$$= \frac{1}{2}\left[\oint_{C_1} + \oint_{C_2} + \oint_{c_3}\right],$$
$$= \frac{1}{2}\left[\int_{c_1}(xdy - ydx) + \int_{c_2}(xdy - ydx) + \int_{c_3}(xdy - ydx)\right],$$

where c_1 is $y = \frac{x}{4}$, c_2 is $y = \frac{1}{x}$ and c_3 is $y = x$.

Along c_1, we have $y = \frac{x}{4}$ so that $y = \frac{1}{x}dx$ and x varies from 0 to 2. Therefore

$$\oint_{c_1}(xdy - ydx) = \int_0^2\left(\frac{x}{4}dx - \frac{x}{4}dx\right) = 0.$$

Along c_2 we have $y = \frac{1}{x}$ so that $dy = -\frac{1}{x^2}dx$ and x varies from 2 to 1. Therefore

$$\oint_{c_2}(xdy - ydx) = \int_2^1\left(\frac{-1}{x}dx - \frac{1}{x}dx\right) = -2\int_0^1\frac{1}{x}dx = 2\log 2.$$

Along c_3, we have $y = x$ so that $dy = dx$ and x varies from 1 to 0. Therefore

$$\oint_{c_3} (xdy - ydx) = \oint_{c_3} (xdx - xdx) = 0.$$

Hence

$$\oint_{c} (xdy - ydx) = \frac{1}{2}[0 + 2\log 2 + 0] = \log 2.$$

(c)

$$\vec{f} = (y^2 - z^2 + 3yz - 2x)\hat{i} + (3xz + 2xy)\hat{j} + (3xy - 2xz + 2z)\hat{k}$$

Then

$$\nabla.\vec{f} = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right).\vec{f}$$
$$= -2 + 2x + 2 - 2x = 0.$$

Hence $\vec{f}$ is solenoidal. Further,

$$curl\,\vec{f} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y^2 - z^2 + 3yz - 2x & 3xz + 2xy & 3xy - 2xz + 2z \end{vmatrix}$$
$$= \hat{i}[3x - 3x] - \hat{j}[3y - 2z + 2z - 3y] + \hat{k}[3z + 2y - 2y - 3z] = \vec{0}$$

Hence $\vec{f}$ is irrotational.

7 (a).

$$I = \int_0^\infty \int_0^\infty e^{-xy} \sin nx\, dx\, dy$$
$$= \int_0^\infty \sin nx \left[\int_0^\infty e^{-xy} dy\right] dx$$
$$= \int_0^\infty \sin nx \left[\frac{e^{-xy}}{-x}\right]_0^\infty dx$$
$$= \int_0^\infty \frac{\sin nx}{x} dx. \quad (1)$$

On the other hand,

$$
\begin{aligned}
I &= \int_0^\infty \int_0^\infty e^{-xy} \sin nx \, dx \, dy \\
&= \int_0^\infty \left[\int_0^\infty e^{-xy} \sin nx \, dx \right] dy \\
&= \int_0^\infty \left[\frac{e^{-xy}}{n^2 + y^2} (n \cos nx + y \sin nx) \right]_0^\infty dy \\
&= \int_0^\infty \frac{n}{n^2 + y^2} dy = \left[\tan^{-1} \frac{y}{p} \right]_0^\infty = \frac{\pi}{2}. \qquad (2)
\end{aligned}
$$

From (1) and (2), it follows that

$$
\int_0^\infty \frac{\sin nx}{x} dx = \frac{\pi}{2}.
$$

(b)

$$
\begin{aligned}
A &= \int_1^4 \left[\int_{\frac{2}{y}}^{2\sqrt{y}} dx \right] dy \\
&= \int_1^4 [x]_{\frac{2}{y}}^{2\sqrt{y}} dy \\
&= \int_1^4 \left[2\sqrt{y} - \frac{2}{y} \right] dy \\
&= 2 \left[\frac{y^{\frac{3}{2}}}{\frac{3}{2}} - \log y \right]_1^4 \\
&= 2 \left[\left(\frac{16}{3} - 2 \log 2 \right) - \frac{2}{3} \right] \\
&= \frac{28}{3} - 4 \log 2.
\end{aligned}
$$

(c) Example 5.5

Index

Engineering Mathematics I

For UPTU

Semester I

Babu Ram

Engineering Mathematics I is designed as per the specific requirements of the first-semester paper offered in the BE/BTech syllabus of Uttar Pradesh Technical University (UPTU). With an emphasis on problem-solving techniques, engineering applications, as well as detailed explanations of the mathematical concepts, this book will give the students a complete grasp of the mathematical skills that are needed by engineers. The focus on practice rather than theory ensures complete mastery over the topics covered in the semester.

KEY FEATURES

- Mapped to the latest UPTU syllabus
- Exercises provided at the end of each chapter along with hints and answers
- 458 solved examples and 285 practice problems
- Reviewed by 15 eminent academicians
- All the examples and chapter-end exercises have been checked for accuracy
- Three UPTU solved question papers

CONTENTS

THE AUTHOR'S OTHER BOOKS WITH PEARSON

1. *Engineering Mathematics II (for UPTU)*
2. *Engineering Mathematics III (for UPTU)*
3. *Engineering Mathematics*
4. *Numerical Methods*
5. *Discrete Mathematics*

PEARSON

Online resources available at www.pearsoned.co.in/BabuRam

ISBN 978-81-317-3335-6

9 788131 733356

www.pearsoned.co.in