PREFACE

In a bid to standardise higher education in the country, the University Grants Commission (UGC) has introduced Choice Based Credit System (CBCS) based on five types of courses: core, generic discipline specific elective, and ability/ skill enhancement for graduate students of all programmes at Elective/ Honours level. This brings in the semester pattern, which finds efficacy in tandem with credit system, credit transfer, comprehensive and continuous assessments and a graded pattern of evaluation. The objective is to offer learners ample flexibility to choose from a wide gamut of courses, as also to provide them lateral mobility between various educational institutions in the country where they can carry acquired credits. I am happy to note that the University has been recently accredited by National Assessment and Accreditation Council of India (NAAC) with grade “A”.

UGC (Open and Distance Learning programmes and Online Programmes) Regulations, 2020 have mandated compliance with CBCS for all the HEIs in this mode. Welcoming this paradigm shift in higher education, Netaji Subhas Open University (NSOU) has resolved to adopt CBCS from the academic session 2021-22 at the Under Graduate Degree Programme level. The present syllabus, framed in the spirit of syllabi recommended by UGC, lays due stress on all aspects envisaged in the curricular framework of the apex body on higher education. It will be imparted to learners over the six semesters of the Programme.

Self Learning Materials (SLMs) are the mainstay of Student Support Services (SSS) of an Open University. From a logistic point of view, NSOU has embarked upon CBCS presently with SLMs in English. Eventually, these will be translated into Bengali too, for the benefit of learners. As always, we have requisitioned the services of the best academics in each domain for the preparation of new SLMs, and I am sure they will be of commendable academic support. We look forward to proactive feedback from all stake-holders who will participate in the teaching-learning of these study materials. It has been a very challenging task well executed, and 1 congratulate all concerned in the preparation of these SLMs.

I wish the venture a grand success.

Professor (Dr.) Subha Sankar Sarkar

Vice-Chancellor

Under Graduate Degree Programme

Choice Based Credit System (CBCS)

Subject: Honours in Physics (HPH)

Course :Mathematical Methods in Physics

Course Code : CC-PH-04

First Print — July, 2021

Printed in accordance with the regulations of the Distance Education
Bureau of the University Grants Commission.

Netaji Subhas Open University

Under Graduate Degree Programme

Choice Based Credit System (CBCS)

Subject: Honours in Physics (HPH)

Course :Mathematical Methods in Physics

Course Code : CC-PH-04

: Course Writing :

Dr. Biswajit Mukherjee
Retd. Associate Professor of Physics
Triveni Devi Bhalotia College

: Course Editing :

Dr. Prasanta Kr. De Sarkar
Retd. Principal
Asansol Girls' College

: Format Editor:

Dr. Gautam Kumar Mallik
Associate Professor of Physics
Netaji Subhas Open University

Netaji Subhas
Open University

UG : Physics-I
(HPH)

Mathematical Methods in Physics-I
Code : CC-PH-04

Unit 1 ❐ Calculus

Structure

1.1 ❐ Objectives

While you go through the pages of this chapter, you will learn

1. Continuity and differentiability of functions using intuitive ideas.
2. Method of approximation.
3. How to find out average and instantaneous values of functions defined in appropriate domain.

1.2 ❐ Introduction

Calculus is widely used to create mathematical models in order to arrive at an optimal solution. In physics calculus is used in a lot of its concepts, in the mathematical study of continuous change, in dynamics, astronomy, astrophysics and quantum mechanics. However in the pages to follow we will indulge in some introductory topics in the form of recapitulations of some basic ideas of calculus. Even in thermodynamics and statistical mechanics, differential are redefined to apply the rules of calculus.

Key words

Limits, continuity, Differentiability, Taylor's and Binomial series. Approximate solution, average value of functions.

1.3 ❒ Recapitulation

1.3.1 Limit of a function

We say that a function $f(x)$ has a limit L at a if and only if for every $\varepsilon > 0$ there exists a positive number δ depending on ε such that for any x in the domain of $f(x)$ with the property

$$0 < |x - a| < \delta \text{ we have } |f(x) - L| < \varepsilon .$$

In symbol we write

$$\lim_{x\to a} f(x) = L \,; \text{ or, } \quad f(x) \to L \,;\, as\ x \to a$$

A similar definition extends to functions in two variables. We say that L is the limit of a function $f(x, y)$ at the point (a, b), written,

$$\lim_{(x,y)\to(a,b)} f(x,y) = L$$

If $f(x, y)$ is as close to L as we please whenever the distance from the point (x, y) to the point (a, b) is sufficiently small, but not zero.

Using $\varepsilon - \delta$ definition we say that L is the limit of $f(x, y)$ as (x, y) approaches (a, b) if and only if for every given $\varepsilon > 0$ we can find a $\delta > 0$ such that for any point (x, y) where

$$0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta \text{, we have } |f(x, y) - L| < \varepsilon .$$

1.3.2. Continuity of a function

A function $f(x)$ is said to be continuous at $x = a$, if; $f(x)$ has a definite value at $x = a$; $\lim_{x\to a} f(x)$ exists and $\lim_{x\to a} f(x) = f(a)$.

In other words, $f(x)$ is said to be continuous at $x = a$, if

$$\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = f(a)$$

Using the $\varepsilon - \delta$ definition, the single valued function $f(x)$ is said to be continuous at $x = a$ provided $f(x)$ possess a definite finite value at $x = a$ and given any pre-assigned positive quantity ε, however small, we can determine another positive quantity δ (whose value depends on ε), such that, $|f(x)-f(a)|<\varepsilon$ for all x in $|x-a|<\delta$

A function $f(x, y)$ is continuous at the point (a, b) if the following two conditions are satisfied :

a) Both $f(a, b)$ and $\lim_{(x,y)\to(a,b)} f(x, y)$ exist.

b) $\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b)$

Example : Art (1.4.1) :

Example 1: Find $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^3-y^3}{x-y}$

Solution : $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^3-y^3}{x-y} = \lim\limits_{(x,y)\to(0,0)} x^2+xy+y^2 = 0\,, \qquad x-y\neq 0$

Example 2: Find $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2-xy}{\sqrt{x}-\sqrt{y}}$

Solution : $\lim\limits_{(x,y)\to(0,0)} \dfrac{\left(x^2-xy\right)\left(\sqrt{x}+\sqrt{y}\right)}{\left(\sqrt{x}-\sqrt{y}\right)\left(\sqrt{x}+\sqrt{y}\right)} = \dfrac{\lim\limits_{(x,y)\to(0,0)}\left(x(x-y)\left(\sqrt{x}+\sqrt{y}\right)\right)}{x-y}$

$= \lim\limits_{(x,y)\to(0,0)} x\left(\sqrt{x}+\sqrt{y}\right) = 0$

Exercise Art 1.4.1 :

1) Find the limit : $\lim\limits_{x\to 0} \dfrac{\sqrt{2+x}-\sqrt{2}}{x}\left[ans : \dfrac{1}{2\sqrt{2}}\right]$

2) Find $\lim\limits_{(x,y)\to(1,1)} \dfrac{(x^2y)}{x^4+y^2}\left[ans : \dfrac{1}{2}\right]$

1.4 ❒ Continuity and Differentiability

If a function $f(x)$ is differentiable at $x = a$ then $f(x)$ must be continuous at $x = a$. However the converse is not always true i.e. if a functions $f(x)$ is continuous $= a$, it is not necessarily differentiable at $x = a$.

Proof :

Now since $f(x)$ is differentiable at $x = a$

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \text{ is a finite quantity.}$$

Now $\lim_{h \to 0} \left[f(a+h) - f(a) \right] = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \times \lim_{h \to 0} h = f'(a) \times 0 = 0$

Therefore $\lim_{h \to 0} f(a+h) = f(a)$ i.e. $f(x)$ is continuous at $x = a$.

However from the definition of continuity we cannot always arrive at the differentiability of a function, as is discussed in the following examples.

Example of Art 1.5:

Example 1: A function $f(x)$ is defined as follows

$$f(x) = \begin{cases} x \text{ when } x \geq 0 \\ -x \text{ when } x < 0 \end{cases}$$

Examine the continuity and differentiability of $f(x)$ at $x = 0$.

Solution : We have $\lim_{x \to 0^+} f(x) = \lim_{x \to 0^+} x = 0$

Since $f(x) = x$ when $x > 0$

And $\lim_{x \to 0^-} f(x) = \lim_{x \to 0^-} (-x) = 0$

Since $f(x) = -x$ when $x < 0$

Again $f(0) = 0$

Therefore, $\lim_{x \to 0^+} f(x) = \lim_{x \to 0^-} f(x) = f(0)$

Therefore the function is continuous at $x = 0$

Now $f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$

$$\therefore \quad f'(0) = \lim_{h \to 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0} \frac{f(h)}{h}$$

Since $f(0) = 0$

Now $\lim_{h \to 0^+} \frac{f(h)}{h} = \lim_{h \to 0^+} \frac{h}{h} = 1$

Since $f(x) = x$ when $x > 0$ and $h \to 0^+$ means $h \neq 0$

And $\lim_{h \to 0^-} \frac{f(h)}{h} = \lim_{h \to 0^-} \frac{-h}{h} = -1$

Therefore right hand limit and left hand limit of $f'(0)$, though exist, are unequal. Therefore $f(x)$ is not differentiable at $x = 0$

Example 2 : Examine the continuity and differentiability of $f(x) = 2x^2 + 3$ at $x = 1$

Solution : $\lim_{x \to 1^+} f(x) = \lim_{x \to 1^+} 2x^2 + 3 = 5$

$$\lim_{x \to 1^-} f(x) = \lim_{x \to 1^-} 2x^2 + 3 = 5$$

And $f(1) = 5$

Therefore $\lim_{x \to 1^+} f(x) = \lim_{x \to 1^-} f(x) = f(1)$

And $f(x)$ is continuous at $x = 1$

Now $f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$

$\therefore \; f'(1) = \lim_{h \to 0} \frac{f(1+h) - f(1)}{h}$

Or, $f'(1) = \lim_{h \to 0} \frac{2(1+h)^2 + 3 - 5}{h}$ $\qquad \because \; f(1) = 5$

$$= \lim_{h \to 0} \frac{2h^2 + 4h}{h}$$

Now Right hand limit of $f'(1) = \lim\limits_{h\to 0^+} \dfrac{2h^2+4h}{h} = \lim\limits_{h\to 0^+} 2h+4=4 \quad [\because h \neq 0]$

And left hand limit of $f'(1) = \lim\limits_{h\to 0^-} \dfrac{2h^2+4h}{h} = \lim\limits_{h\to 0^-} 2h+4=4$

Therefore right hand limit of $f'(1)$ and left hand limit of $f'(1)$, both exist and are equal.

And so, $f'(1)$ exists and its value is 4 i.e. $f(x)$ is differentiable at $x= 1$.

Exercise of Arts 1.4 and 1.5 :

1) $$f(x) = \begin{cases} \dfrac{1}{x}\sin(x^2) & when\ x \neq 0 \\ 0 & when\ x = 0 \end{cases}$$

Discuss the continuity and differentiability of $f(x)$ at $x = 0$.

2) Find the co-efficient a and b such that the following function f is continuous and differentiable at $x = 1$

$$f(x) = \begin{cases} \dfrac{1}{|x|} & when\,|x|>1 \\ 2ax^2+b & when\,|x|<1 \end{cases}$$

Solution Exercise of Arts : 1.4 and 1.5 :

Solutions (1) : Differentiability at $x = 0$, we have by the definition

$$f'(0) = \lim_{h\to 0} \frac{f(0+h)-f(0)}{h},$$ therefore right hand limit,

$$= \lim_{h\to 0^+} \frac{f(h)}{h} = \lim_{h\to 0^+} \frac{\frac{1}{h}\sin h^2}{h} = \lim_{h\to 0^+} \frac{\sin(h^2)}{h^2} = \lim_{z\to 0^+} \frac{\sin z}{z} = 1;$$

where, $h^2 = z \quad \therefore\ z \to 0^+$ as $h \to 0^+$

Similarly, $\lim\limits_{h\to 0^-} \dfrac{f(h)}{h} = 1$

Therefore the function $f(x)$ is differentiable at $x = 0$. Since $f(x)$ is differentiable at = 0, it must be continuous at $x = 0$.

Solution 2 : Since $f(x)$ is continuous at $x = 1$, $\lim\limits_{x\to 1^-} f(x) = f(1) = 1$ or

$$\lim_{x\to 1^-}(2ax^2+b)=1 \text{ i.e. } 2a + b = 1$$

Again, $f(x)$ is differentiable at $x = 1$.

$f'(1) = \lim\limits_{h\to 0}\dfrac{f(x+h)-f(1)}{h} = \lim\limits_{h\to 0}\dfrac{f(1+h)-1}{h}$ and right hand $f'(1)$ = left hand $f'(1)$

$$\text{Or, } \lim_{h\to 0^+}\frac{f(1+h)-1}{h} = \lim_{h\to 0^-}\frac{f(1+h)-1}{h}$$

$$\text{Or, } \lim_{h\to 0^+}\frac{1}{h}\left[\frac{1}{1+h}-1\right] = \lim_{h\to 0^-}\frac{1}{h}\left[2a(1+h)^2+b-1\right]$$

$$\left[\because\ f(x)=\frac{1}{x} \text{ when } x>1 \text{ and } f(x)=2ax^2+b, \text{ when } x<1\right]$$

$$\text{Or, } \lim_{h\to 0^+}\frac{1}{h}\left[\frac{-h}{1+h}\right] = \lim_{h\to 0^-}\frac{1}{h}\left[2a+b-1+2a(2h+h^2)\right] \qquad (\because 2a+b=1)$$

$$\lim_{h\to 0^+}\left[\frac{-1}{1+h}\right] = \lim_{h\to 0^-}\frac{2ah(2+h)}{h}$$

$$\text{Or, } -1 = \lim_{h\to 0^-}2a(2+h) = 4a$$

$$\therefore\ a = -\frac{1}{4}$$

Therefore $b = 1-2a = 1+\dfrac{1}{2} = \dfrac{3}{2}$

1.5 ❒ Intuitive Ideas of Continuous and Differentiable Function

1) The function $f(x)=\dfrac{1}{x-1}$ is discontinuous at x = 1 because at = 1, f(1) is not defined (has 0 as denominator) and because $\lim\limits_{h\to 1} f(x)$ does not exist (equals ∞). The function is however continuous every where except at x = 1, where it is said to have an infinite discontinuity (Fig 1.6.1].

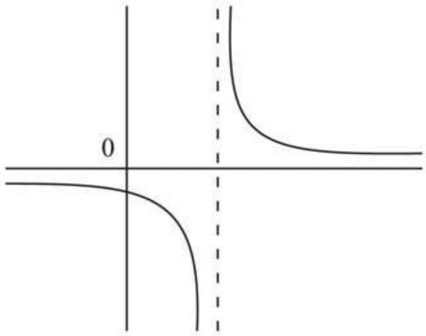


Fig. (1.6.1)

2) The function $f(x)=\dfrac{x^2-a^2}{x-a}$ is discontinuous at x = a because $f(a)$ is not defined (has zero, both numerator and denominator) and because $\lim\limits_{x\to a} f(x)=2a$. The discontinuity here is called removable since it may be removed by redefining the function as $f(x)=\dfrac{x^2-a^2}{x-a}$ for $x\neq a,\ f(a)=2a$.

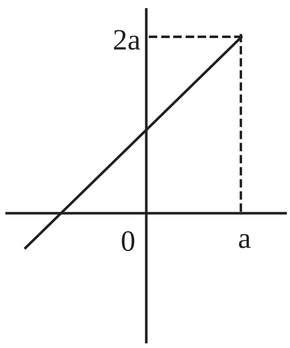


Fig. (1.6.2)

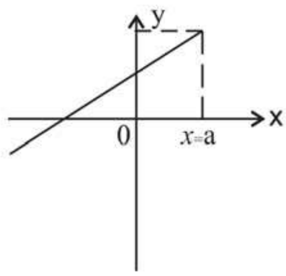


Fig. (1.6.3)

(Note that the discontinuity in example (1) cannot be so removed because the limit also does not exist). The graphs of $f(x)=\dfrac{x^2-a^2}{x-a}$ and $g(x) = x + a$ are identical except at $x = a$, where the former has a break [fig 1.6.2]. Removing the discontinuity consists simply of joining the break [fig 1.6.3].

A differentiable function of one real variable is a function whose derivative exists at each point in its domain. As a result, the graph of a differentiable function must have a tangent line (non-vertical) at each point in its domain, be relatively smooth and cannot contain any breaks, bends or cusps.

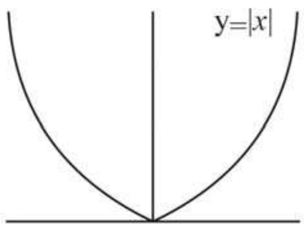


Fig. (1.6.4)

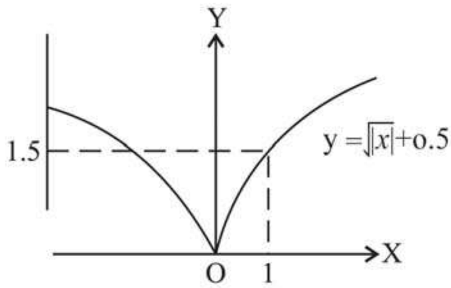


Fig. (1.6.5)

In fig (1.6.4) the absolute value function is continuous i.e. it has no gap. It is differentiable everywhere except at the point x= 0, where it makes a sharp turn as it crosses the y – axis.

A cusp on the graph of a continuous function fig (1.6.5) at x = 0. The function is continuous but not differentiable.

A function with a bend, cusp a vertical tangent may be continuous but fails to be differentiable at the location of the anomaly.

Below are graphs of functions that are not differentiable at x = 0 for various reasons.

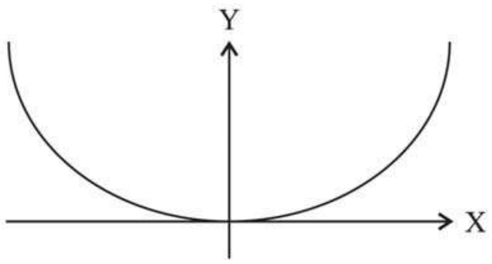


Fig. (1.6.6) : no tangent at x = 0

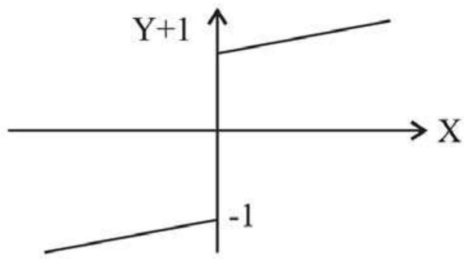


Fig. (1.6.7) : jump in the value of function at x = 0

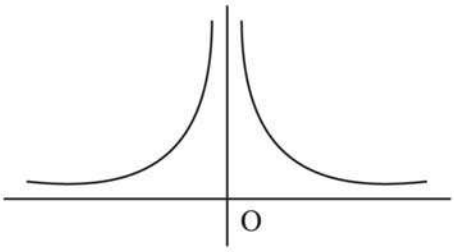


Fig. (1.6.8) : function increases indefinitely at x = 0

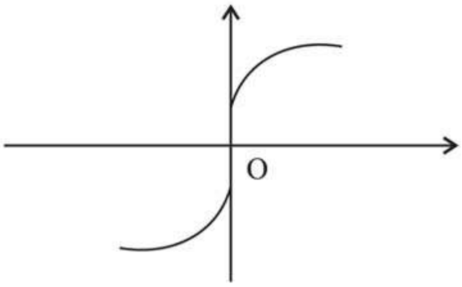


Fig. (1.6.9) : tangent at x = 0 is vertical.

Example : Art (1.6)

Find the $\lim\limits_{x\to 1} f(x)$, where $f(x) = x + x^2$ by intuitive ideas.

Solution : We tabulate the values of $f(x)$ near $x = 1$ in the following table

x	0.9	0.99	0.999	1.01	1.1	1.2
f(x)	1.11	1.9701	1.997001	2.0301	2.31	2.64

From this it is reasonable to say that, $\lim x \to 1^- f(x) = \lim x \to 1^+ f(x) = \lim x +1$ $f(x) = 2$

1.6 ❒ Average and Instantaneous values of functions :

Suppose a function $f(x)$ is continuous on the interval (a, b). We want to find the average value of $f(x)$ in the interval (a, b). We divide the interval by n numbers of intervals $x_1 = a, x_2$, ..., $x_n = b$ and find the n numbers of values of $f(x)$ [fig 1.7.1] e.g. at x_1, $f(x_1)$; at x_2, $f(x_2)$; and so on and we get the approximate average value of $f(x)$ on (a, b) as :

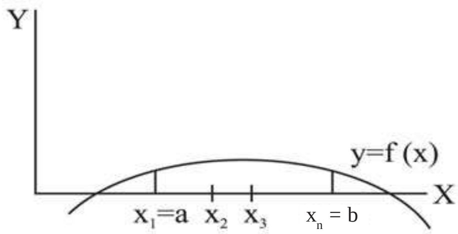


Fig. 1.7.1

$$\frac{f(x_1)+f(x_2)+...+f(x_n)}{n} \qquad (1.1)$$

Now let the points x_1, x_2, ... be Δx apart. And we multiple the numerator and the denominator of the approximate average by Δx , then average of $f(x)$ on (a, b) is approximately equal to

$$\frac{\left[f(x_1)+f(x_2)+\ldots+f(x_n)\right]\Delta x}{n\Delta x}$$

Where $n\Delta x = b - a$.

Now in the limit $\Delta n \to \infty, \Delta x = 0$ and we get the average value of $f(x)$ as

$$\langle f(x)\rangle = \frac{\int_a^b f(x)dx}{b-a} = \frac{\int_a^b f(x)dx}{\int_a^b dx} \tag{1.2}$$

We see that $f(x_1)$, $f(x_2)$, ... etc are the instantaneous values of $f(x)$ at x_1, x_2, ... etc.

In alternating current theory we take the instantaneous value of current as $i = i_0 \sin \omega t$ or $i_0 \cos \omega t$, i_0 is the amplitude. Now average value of i for the period $T = \frac{2\pi}{\omega}$ is given as,

$$< i > = \frac{\int_0^T i_0 \sin \omega t \, dt}{\int_0^T dt} = \frac{1}{T}\int_0^T i_0 \sin \omega t \, dt \, .$$

If we put $\omega t = \theta, \; \omega dt = d\theta, \; dt = \frac{d\theta}{\omega}$

Or, $< i > = \frac{1}{2\pi}\int_0^{2\pi} i_0 \sin\theta \; d\theta = 0, \;\; \omega T = 2\pi$

Thus we see that the average value of sin θ or cos θ for a complete period or any number of periods is zero. In such cases the average of the square of the function is taken to define a significant mean like root mean square current.

1.7 ❒ Approximation

There are many problems in physics which can be written as an infinite series and its solution lies in finding the sum of the infinite series. However it is often found that the results differ very little if we would have taken a finite number of terms at the beginning of the series rather than taking the entire infinite series. In this way we can find an approximate solution of the problems which cannot be solved exactly. The accuracy of the solution can be reached to the desired value, by taking as many terms of the series as required to reach the desired accuracy. Also many functions can be expanded in infinite power series (i.e. a series expanded in powers of x having infinite number of term).

Taylor's series :

We can write the Taylor series for a function $f(x)$ about $x = a$;

$$f(x) = f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + ... + \frac{1}{n!}(x-a)^n f^{(n)}(a) + ... \quad (1.3)$$

Where $f^{(n)}(x)$ represents n^{th} derivative of $f(x)$

Or, $$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \quad (1.4)$$

The Maclaurin series for $f(x)$ is the Taylor's series about the origin. Putting $a = 0$ in equation (1.4) we obtain the Maclaurin series for $f(x)$:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \quad (1.5)$$

The polynomial formed by taking some initial terms of the Taylor's series is called Taylor's polynomial.

A Taylor's series is a representation of function as an infinite sum of terms that are calculated from the values of the function's derivatives at a single point. The function can be approximated by using the Taylor polynomial of suitable number of terms. Taylor's theorem gives quantitative estimate on the error introduced by the use of such an approximation.

Example 1 : The sine function is closely approximated by its Taylor polynomial of degree 7 (dotted) for a full period centred at the origin.

The dotted curve is a polynomial of degree seven.

$$\sin x \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}$$

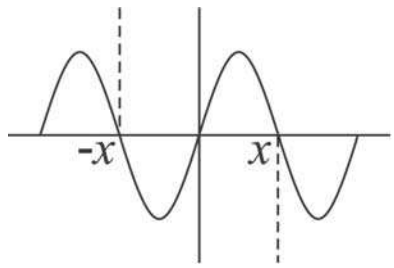


Fig. 1.8.1

The error in this approximation is no more than $\frac{|x|^9}{9}$. In particular for $-1 < x < 1$ the error is less than 0.000003 (fig. 1.8.1).

Example 2 : Using the quadratic Taylor polynomial for $f(x)=\frac{1}{x^2}$, approximate the value of $\frac{1}{4.41}$.

Ans. The quadratic Taylor polynomial is $f_2(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2$

Now $f'(x)=-\frac{2}{x^3}$; $f''(x)=\frac{6}{x^4}$; we write 4.41 = $(2 + 0.1)^2$, implies a = 2 & x = 2.1

$$\therefore\ f_2(x)-f(2)+\frac{f'(2)}{1!}(2.1-2)+\frac{f''(2)}{2!}(2.1-2)^2$$

$$=f(2)+\frac{\left(-\frac{2}{8}\right)}{1!}(2.1-2)+\frac{\left(\frac{6}{16}\right)}{2!}(2.1-2)^2$$

$$= 0.25 - 0.025 + 0.001875 = 0.226875$$

The actual value is $\frac{1}{4.41}$ = 0.226775

So the approximation deviates only about 0.05%.

Example 3 : What is the quadratic approximation of the function $f(x)=\sqrt{4+x}$ at $x = 0$

Solution : $f_2(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2=2+\frac{1}{4}x+\frac{\frac{1}{32}}{2!}x^2=2+\frac{1}{4}x-\frac{1}{14}x^2$

Binomial series : The binomial series can be written as :

$$(1+x)^p=\sum_{n=0}^{\infty}\binom{p}{n}x^n=1+px+\frac{p(p-1)}{2!}x^2+\frac{p(p-1)(p-2)}{3!}x^3+\ldots \quad (1.6)$$

P is any real number, positive or negative or fractional and the expansion is an infinite series, $\binom{p}{n}$ is called a binomial co-efficient.

The binomial co-efficients are

$$\binom{p}{0}=1;\ \binom{p}{1}=\frac{p}{1!};\quad \binom{p}{2}=\frac{p(p-1)}{2!};\quad \binom{p}{3}=\frac{p(p-1)(p-2)}{3!};$$

$$\binom{p}{n}=\frac{p(p-1)(p-2)...(p-n+1)}{n!}$$

To get an approximation we can consider a few terms from the expansion (1.6).

Example 1 : For small x, $1 + px$ is a reasonable approximation for $(1 + x)^p$. Notice that this correspond to picking the first two terms from (1.6). Now suppose $x = 0.0007$. Therefore

$$(1.0007)^9 = 1 + 0.0007 \times 9 \approx 1.0063$$

Now actual value of $(1.0007)^9 = 1.006317668842...$

Therefore our result is correct up to four decimal place.

1.8 ❐ Summary

Definitions of limit, continuity and differentiability are recapitulated. Application of series in finding the approximate solution of physical problems which cannot be solved exactly has been discussed. Average and instantaneous values of function is defined.

Unit 2 ❒ Second Order Differential Equation

Structure

2.1 ❒ Objectives

In going through the chapter you will learn :

1. To classify second order differential equation.
2. To find out solution of second order differential equation in terms of complementary functions and particular integral.
3. To find out the general solution and to define linearly dependent and independent solutions in terms of wronskian.
4. Statement of existence and uniqueness theorem for initial value problems and their applications.

2.2 ❒ Introduction

An universally accepted method for formulations and solutions of physical problems is to construct the relevant differential equation and then attempt to solve it. Thus differential equations are at the centre to many physical problems. In this chapter we shall limit our discussions to the techniques of solving second order linear homogeneous or inhomogeneous equations with constant co-efficients and also some relevant physical applications will be discussed.

An equation containing derivatives is called differentials equation, which may be classified as ordinary or partial.

Ordinary differential equation : Differential equation containing only one independent variable is called ordinary differential equation.

Example : $\frac{d^2x}{dt^2}+\omega^2 x=0$, where $x = x(t)$ i.e. functions of one independent variable t.

Partial differential equations : Differential equations containing partial derivatives of the dependent variable $y\,(x_1, x_2, x_3 \ldots x_n)$ with respect to more than one independent variables $x_1, x_2, \ldots x_n$ are called partial differential equations.

Example : $\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}+\lambda\psi=0$ where $\psi=\psi(x, y, z)$

Order of differential equation : It is defined as the order of the highest derivative on the equation.

Degree of differential equation : It is defined as the power of the highest derivative in the equation after fractional powers have been removed.

Examples :

i) $\frac{d^2y}{dx^2}+\left(\frac{dy}{dx}\right)^2+xy=0$ is a second order and 1st degree differential equation, while the equation.

ii) $\frac{d^2y}{dx^2}+\sqrt{\frac{dy}{dx}}+xy=0$ is a second order, 2nd degree differential equation, because when the square root is removed, $\left(\frac{dy}{dx}\right)^2$ appears in the equation as the highest power of highest order.

Linear differential equation : A linear differential equation satisfies the following properties :

i) Power of each derivative and of the dependent variable must be unity.

ii) The coefficient of all derivatives occurring in the equations may be constant or may be functions of the independent variables.

iii) The dependent variables and its derivatives is not multiplied together is the differential equation.

Differential equation obeying no such properties or property are termed non-linear differential equation.

Example : $\frac{d^2y}{dx^2} + y\frac{dy}{dx} + y = 0$; the presence of the term $y\frac{dy}{dx}$ makes it non-linear.

Homogeneous and in-homogeneous differential equation : When the right hand member of the differential equation is either zero or constant, the differential equation is called homogeneous, otherwise the differential equation is in-homogeneous, when the right hand side is function of independent variable.

Example : $P_0\frac{d^2y}{dx^2} + P_1\frac{dy}{dx} + P_2 y = Q(x)$ (2.1)

If $Q(x) = 0$ or constant, equation (2.1) is homogeneous, otherwise in-homogeneous.

Solution of Differential equation : A solution of differential equation is a function which, when substituted in the differential equation produces an identity.

Key Words

Homogeneous and inhomogeneous equations, wronskian, complementary function, particular integral. Existence and uniqueness theorem.

2.3 ❒ Linear second order differential equation :

A linear second order differential equation has the following general form

$$P_0(x)\frac{d^2y}{dx^2} + P_1(x)\frac{dy}{dx} + P_2(x)y = Q(x) \qquad (2.2)$$

Where $P_0(x)$, $P_1(x)$, and $P_2(x)$ are called co-efficient functions and $Q(x)$ is the force function. If the function $Q(x) = 0$, the equation is called homogeneous.

For the homogeneous equation, it is to be noted that the function $y(x) = 0$ always satisfy the equation, regardless what $P_0(x)$, $P_1(x)$ and $P_2(x)$ are. The solution $y(x) = 0$ is called the trivial solution of the homogeneous equation.

2.3.1 : Second order Linear homogeneous differential equations with constant co-efficient

When P_0, P_1 and P_2 are constants and $Q(x) = 0$, $P_0 \neq 0$, we get from equation (2.2)

$$P_0\frac{d^2y}{dx^2} + P_1\frac{dy}{dx} + P_2 y = 0 \qquad (2.3)$$

$$\Rightarrow \frac{d^2y}{dx^2} + \frac{P_1}{P_0}\frac{dy}{dx} + \frac{P_2}{P_0}y = 0$$

$$\Rightarrow \frac{d^2y}{dx^2} + \alpha\frac{dy}{dx} + \beta y = 0, \qquad (2.4)$$

Where $\alpha = \frac{P_1}{P_0}$ and $\beta = \frac{P_2}{P_0}$,

We seek solution equation (2.4) as $y = e^{\lambda x}$ (2.5)

Substituting the solution (2.5) in equation (2.4), we find that $e^{\lambda x}\left(\lambda^2 + \alpha\lambda + \beta\right) = 0$ (2.6)

Which shows that $e^{\lambda x}$ is a solution of (2.4) only when $\lambda^2 + \alpha\lambda + \beta = 0$ (2.7)

Equation (2.7) is called the characteristic equation. The characteristic roots are

$$\begin{aligned} \lambda_1 &= -\frac{\alpha}{2} + \frac{1}{2}\sqrt{\alpha^2 - 4\beta} \\ \lambda_1 &= -\frac{\alpha}{2} - \frac{1}{2}\sqrt{\alpha^2 - 4\beta} \end{aligned} \qquad (2.8)$$

Thus, solutions of equation (2.4) are given by : $y_1 = e^{\lambda_1 x},\ y_2 = e^{\lambda_2 x}$ (2.9)

Now the Wronskian of the solutions (2.9) is

$$W(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = \begin{vmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{vmatrix}$$

$$= (\lambda_2 - \lambda_1) e^{(\lambda_1 + \lambda_2)x} \tag{2.10}$$

Thus the solutions y_1, y_2 given by equation (2.10) will be linearly independent only when $\lambda_1 \neq \lambda_2$. In this case, the general solution is : $y(x) = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}$ (2.11)

Now we put $l = \alpha^2 - 4\beta$ (2.12)

If $l > 0$ the two roots λ_1, λ_2 are real.

In such a case, the solution (2.11) takes the form :

$$y(x) = e^{-\frac{\alpha x}{2}} \left\{ C_1 e^{\frac{\sqrt{l}x}{2}} + C_2 e^{-\frac{\sqrt{l}x}{2}} \right\} \tag{2.13}$$

or, $$y(x) = e^{-\frac{\alpha x}{2}} \left\{ C \sinh \frac{\sqrt{l}}{2} x + D \cosh \frac{\sqrt{l}}{2} x \right\} \tag{2.14}$$

If $l < 0$, the two roots λ_1, λ_2 are imaginary. Then the general solution (2.11) takes the form

$$y(x) = e^{-\frac{\alpha x}{2}} \left\{ C_1 e^{\frac{i\sqrt{l}x}{2}} + C_2 e^{\frac{-i\sqrt{l}x}{2}} \right\} \tag{2.15}$$

or, $$y(x) = e^{-\frac{\alpha x}{2}} \left\{ C \sin \frac{\sqrt{l}}{2} x + D \cos \frac{\sqrt{l}}{2} x \right\} \tag{2.16}$$

When $l = 0$, the two roots $\lambda_1 = \lambda_2 = -\frac{\alpha}{2}$ are equal and the wronskian of the solution (2.11) is $W(x) = 0$. Therefore, the two solutions given by equation (2.9) are not linearly independent and the solution (2.11) is not acceptable as general solution.

Since the solutions $y_1(x)$ and $y_2(x)$ are now dependent we try the second solution as,

$$y_2(x) = z(x)e^{\lambda x} \tag{2.17}$$

Substituting equation (2.17) in equation (2.4)

We get $$e^{\lambda x}\left[\frac{d^2z}{dx^2} + (2\lambda+\alpha)\frac{dz}{dx} + \left(\lambda^2+\alpha\lambda+\beta\right)z\right] = 0 \tag{2.18}$$

Now co-efficient of $z = 0$ by equation (2.7) and co-efficient of $\frac{dz}{dx} = 0$ since we assumed equal root of (2.7) which are $\lambda = -\frac{\alpha}{2}$. Thus from equation (2.18) we get since $e^{\lambda x} \neq 0$

$$\frac{d^2z}{dx^2} = 0$$ (2,19), which implies $z = x$.

Thus the second solution is,

$$y_2(x) = xe^{-\frac{\alpha x}{2}} \tag{2.20}$$

Thus, the general solution is

$$y(x) = C_1 e^{-\frac{\alpha x}{2}} + C_2 x e^{-\frac{\alpha x}{2}} \tag{2.21}$$

2.3.2 : Second Order Linear Homogeneous Differential Equation with Constant Co-Efficient : Working Rules for Solutions (Complementary Function)

We consider a linear homogeneous second order differential equation with constant co-efficient.

$$P_0\frac{d^2y}{dx^2} + P_1\frac{dy}{dx} + P_2 y = 0\,; \tag{2.22}$$

$P_0 \neq 0$; P_1 and P_2 are constants.

In terms of the linear operator $D \equiv \frac{d}{dx}$ where $Dy = \frac{dy}{dx}, D^2 y = \frac{d^2 y}{dx^2}$, equation (2.22) can be written as

$$\left(P_0 D^2 + P_1 D + P_2\right) y = 0 \tag{2.23}$$

Or, $\quad F(D)y=0 \quad$ (2.24)

Where $F(D)$ is a polynomial in the variable D. Now the polynomial $F(D)$ can be factored as :

$F(D) = P_0 D^2 + P_1 D + P_2 = P_0(D - m_1)(D - m_2)$ and equation (2.24) reduces to

$$P_0(D - m_1)(D - m_2)y = 0 \tag{2.25}$$

The equation $F(D) = P_0(D - m_1)(D - m_2) = 0$

Or, $\quad (D - m_1)(D - m_2) = 0$, since $P_0 \neq 0 \quad$ (2.26)

Is called the characteristic equation of (2.24) and the roots m_1, m_2 are called characteristic roots.

Now to solve equation (2.22) we first write it in the form of equation (2.25) and then write its characteristic equation. The characteristic roots $m_1 m_2$ are found out.

Now we are ready to write down the solution depending on the nature of the roots. The solutions are termed complementary function (y_c).

Rule I : If the roots are real and different i.e. if $m_1 \neq m_2$, then the solution is $y_C = C_1 e^{m_1 x} + C_2 e^{m_2 x}$; C_1, C_2 are arbitrary constant.

Rule II : If roots are equal i.e. $m_1 = m_2$, then the solution is : $y_C = C_1 e^{m_1 x} + x C_2 e^{m_1 x}$; C_1, C_2 are arbitrary constant.

Rule III : If the roots are imaginary i.e. $m_1 = \alpha + i\beta$ and $m_2 = \alpha - i\beta$, then the solution is $y_c = e^{\alpha x}\left[C_1 \cos \beta x + C_2 \sin \beta x\right]$; C_1, C_2 are arbitrary constant.

Example of Rule I : Find the complementary function of the equation :

$$\frac{d^2 x}{dt^2} - 5\frac{dx}{dt} + 6x = 0 \ \rightarrow (1)$$

Solution : Equation (1) can be rewritten as $(D^2 - 5D + 9)x = 0$, where $D \equiv \dfrac{d}{dt}$

Auxiliary equation is $(D^2 - 5D + 6) = 0$

Or, $(D - 3)(D - 2) = 0 \rightarrow (2)$

Roots of auxiliary equation (2) are $m_1 = 3$, $m_2 = 2$; both are real and different. Therefore solution of equation (1) is $x_c = C_1 e^{3t} + C_2 e^{2t}$; x_c is the complementary function.

Example of rule II : Find complementary function of the equation :

$$\frac{d^2y}{dt^2} + 6\frac{dy}{dt} + 9y = 0 \rightarrow (1)$$

Solution : Equation (1) can be re-written as $(D^2 + 6D + 9)y = 0 \rightarrow (2)$

Auxiliary equation is, $\left[D \equiv \dfrac{d}{dt}\right]$

$$(D^2 + 6D + 9) = 0$$

Or, $(D + 3)(D + 3) = 0 \rightarrow (3)$

Roots of auxiliary equation are $m_1 = -3$, $m_2 = -3$

Therefore, roots are real and equal.

Therefore solution $y_c = c_1 e^{-3x} + x c_2 e^{-3x}$

Example of Rule III : Find the complementary function of $\dfrac{d^2y}{dx^2} + 4y = 0 \rightarrow (1)$

Solution : Equation (1) can be re-written as $(D^2 + 4)y = 0$; $D \equiv \dfrac{d}{dx}$

Auxiliary equation is $D^2 + 4 = 0$

Or, $(D + 2i)(D - 2i) = 0$; $i = \sqrt{-1}$

Therefore, roots are imaginary i.e. $m_1 = +2i$, $m_2 = -2i$ and the complementary function is

$$y_C = C_1 \cos 2x + C_2 \sin 2x$$

2.4 ❐ The Existence and Uniqueness Theorem

Consider the initial value problem $\frac{d^2y}{dx^2}+\alpha(x)\frac{dy}{dx}+\beta(x)y=Q(x),\ y(x_0)=y_0,$ $y'(x_0)=y'_0$.

If the functions α, β and Q are continuous on the interval $I : p < x < q$ containing the point $x = x_0$; then there exists a unique solution $y=\phi(x)$ of the problem, and this solution exists throughout the interval *I*.

That is the theorem guarantees that the given initial value problem will always have (existence of) exactly one (uniqueness) twice differentiable solution, on any interval containing x_0 as long as all three functions $\alpha(x)$, $\beta(x)$ and $Q(x)$ are continuous on the same interval. Conversely neither existence nor uniqueness of a solution is guaranteed at a discontinuity of $\alpha(x)$, $\beta(x)$ or $Q(x)$. All the initial conditions in an initial value problem must be taken at the same point x_0.

The set of conditions where the values are taken at different points [e.g. : $x = x_0$; $x=x'_0,\ x=x''_0$ etc.] are known as boundary conditions. A boundary value problem does not have the existence and uniqueness guarantee.

Example 1 : Find the largest interval where : $\left(x^2-1\right)y''+3xy'+\cos xy=c^x;$ $y(0)$ = 4, $y'(0)=5$ is guaranteed to have a unique solution.

Solution : The given equation can be rewritten as : $y''+\frac{3x}{x^2-1}y'+\frac{\cos x}{x^2-1}y=\frac{e^x}{x^2-1}$

Comparing with the standard form, we get $\alpha(x)=\frac{3x}{x^2-1},\ \beta(x)=\frac{\cos x}{x^2-1}$ and $Q(x)=\frac{e^x}{x^2-1}$

But they are continue on – |< x < | containing the point $x = 0$

We see that $\alpha(x)$, $\beta(x)$ and $Q(x)$ all have discontinuities at $x = -1$ and $x = 1$. Thus the theorem tells us that there is a unique solution on the interval $I : -1 < x < 1$. Since $\alpha(x)$, $\beta(x)$ and $Q(x)$ are all continuous on $-1 < x < 1$ containing $x = 0$.

Now we investigate solutions to linear homogeneous differential equations :

$$y''+\alpha(x)y'+\beta(x)y=0 \tag{2.27}$$

where $y'' = \dfrac{d^2y}{dx^2}$ and $y' = \dfrac{dy}{dx}$. Now if $y_1(x)$ and $y_2(x)$ are solutions of equations (2.27), then $y(x) = C_1y_1(x) + C_2y_2(x)$ is also a solution. This is known as theorem of superposition principle.

Proof : since $y_1(x)$ is a solution of (2.27), $\dfrac{d^2y}{dx^2} + \alpha(x)y_1' + \beta(x)y_1 = 0$ (2.28)

similarly $y_2(x)$ is a solution of (2.27) $\dfrac{d^2y_2}{dx^2} + \alpha(x)y_2' + \beta(x)y_2 = 0$ (2.29)

multiplying equation (2.28) by C_1 and equation (2.29) by C_2 and adding the result since C_1 and C_2 are two arbitrary constants, we get

$$\frac{d^2\left(C_1y_1 + C_2y_2\right)}{dx^2} + \alpha(x)\frac{d\left(C_1y_1 + C_2y_2\right)}{dx} + \beta(x)\,(C_1y_1 + C_2y_2) = 0 \tag{2.30}$$

Equation (2.30) shows that $y(x) = C_1y_1(x) + C_2y_2(x)$ (2.31) also is a solution of equation (2.27).

Next we investigate the initial conditions. If we find a general solution to the homogeneous system, can we choose constants such that the solution satisfies the initial conditions ? That is can we find C_1 and C_2 such that

$$\begin{aligned} C_1y_1(x_0) + C_2y_2(x_0) &= y_0 \\ C_1y_1'(x_0) + C_2y_2'(x_0) &= y_0' \end{aligned} \tag{2.32}$$

In matrix form, equation (2.32) can be written as

$$\begin{pmatrix} y_1(x_0) & y_2(x_0) \\ y_1'(x_0) & y_2'(x_0) \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_0' \end{pmatrix} \tag{2.33}$$

Equation (2.32) has a unique solution if and only if the determinant of the matrix is not zero. This determinant $\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}$ is called wronskian.

Thus our discussion proves the following theorem.

Let $y'' + \alpha(x)y' + \beta(x)y = 0,\ y(x_0) = y_0,\ y'(x_0) = y_0'$ be a homogeneous linear second order differential equation and let y_1 and y_2 be two solutions (Any initial value), then if the wronskin.

$y_1 y_2' - y_2 y_2'$ is non-zero, there exists a solution to the any initial value problem of the form $y = C_1 y_1 + C_2 y_2$

Example 2 : Construct the wronskian of the solution of the differential equation : $y'' + 2y' - 8y = 0$ and show that any initial value problem will have a unique solution.

Solution : The general solution of the given equation : $y = C_1 e^{2x} + C_2 e^{-4x}$, now the wronkian of $y_1 = e^{2x}$ and $y_2 = e^{-4x}$

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = \begin{vmatrix} e^{2x} & e^{-4x} \\ 2e^{2x} & -4e^{-4x} \end{vmatrix} = -4e^{-2x} - 2e^{-2x} = -6e^{-2x} \neq 0$$

Thus W is never zero and we can conclude that any initial value problem will have a unique solution of the form $y = C_1 e^{2x} + C_2 e^{-4x}$.

Exercise 2.4.3 : For each IVP below, find the largest interval on which a unique solution is guaranteed to exist.

a) $(x+2)y'' + xy' + \cot xy = x^2 + 1, \;\; y(2) = 11, \;\; y'(2) = -2$

Solution : The standard form is

$$y'' + \frac{x}{x+2} y' + \frac{\cos x}{(x+2)\sin t} y = \frac{x^2+1}{x+2}$$ and $x_0 = 2$. The discontinuity α, β and Q are $x = -2$, 0, $\pm\pi, \pm 2\pi, \pm 3\pi$ ... ..., $x = -2$ respectively. The largest interval that contain $x_0 = 2$ but none of the discontinuities is, therefore $(0, \pi)$.

b) $\sqrt{16-x^2}\, y'' + \ln(x+1) y' + \cos xy = 0, \;\; y(0) = 2, \;\; y'(0) = 0$

Solution : The standard form is :

$$y'' + \frac{\ln(x+1)}{\sqrt{16-x^2}} y' + \frac{\cos x}{\sqrt{16-x^2}} y = 0$$

$\alpha(x)$ is only defined (and is continuous) on the interval $(-4,4)$ and similarly $\beta(x)$.

$Q(x)$ is continuous everywhere.

Combining these we see that α, β and Q are all continuous on the interval (– 4,4). Since this interval contains $x_0 = 0$, it must be the largest interval on which the solution is guaranteed to exist uniquely.

2.5 ❒ Linearly Dependent and Linearly Independent Solution of Differential Equation : Wronskian

Definition : If $y_1(x)$ and $y_2(x)$ are any two solutions of the differential equation :

$$y'' + \alpha(x)y' + \beta(x)y = 0 \tag{2.34}$$

On the interval $[a, b]$, then their wronskian, defined by :

$$W(y_1, y_2) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} = y_1(x)y_2'(x) - y_2(x)y_1'(x) \tag{2.35}$$

Is either identically zero or never zero on the interval [a, b]

When $W(y_1, y_2) = 0$, $y_1(x)$ & $y_2(x)$ are linearly dependent solutions (or function) of the equation (2.34) on [a, b]. In this case $\dfrac{y_2(x)}{y_1(x)} =$ constant.

When $W(y_1, y_2) \neq 0$, $y_1(x)$ & $y_2(x)$ are linearly independent solutions of equation (2.34) on the interval [a, b]. In this case $\dfrac{y_2(x)}{y_1(x)} \neq$ constant and the general solution of equation (2.34) can be written as : $y(x) = C_1y_1(x) + C_2y_2(x)$ (2.36)

Where C_1 and C_2 are two arbitrary constant.

Proof : We suppose $y_1(x)$ and $Y_2(x)$ are linearly dependent solutions of equation (2.34) on an interval [a, b]. Then we may assume $y_2(x) = Cy_1(x)$ where C is a constant. Therefore

$$W(y_1, y_2) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} = y_1(x)y_2'(x) - y_2(x)y_1'(x)$$

Again $y_2'(x) = Cy_2'(x)$

Therefore $W(y_1, y_2) = C\, y_1(x) y_1'(x) - C\, y_1(x)\, y_1'(x) = 0$ (2.37)

Again we assume that $y_1(x)$ and $y_2(x)$ are two linearly independent solutions of equation (2.34) on the interval [a, b]. Then we have

$$y_1''(x) + \alpha(x) y_1'(x) + \beta(x) y_1(x) = 0 \tag{2.38}$$

And $$y_2''(x) + \alpha(x) y_2'(x) + \beta(x) y_2(x) = 0 \tag{2.39}$$

Now multiplying equation (2.38) by $y_2(x)$ and (2.39) by $y_1(x)$, we get respectively :

$$y_1''(x) y_2(x) + \alpha(x) y_1'(x) y_2(x) + \beta(x) y_1(x) y_2(x) = 0 \tag{2.40}$$

And $$y_2''(x) y_1(x) + \alpha(x) y_2'(x) y_1(x) + \beta(x) y_2(x) y_1(x) = 0 \tag{2.41}$$

Subtracting equation (2.40) from (2.41),

$$\left[y_2''(x) y_1(x) - y_1''(x) y_2(x)\right] + \alpha(x)\left[y_2'(x) y_1(x) - y_1'(x) y_2(x)\right] = 0 \quad (2.42$$

Now wronskian $W\left(y_1(x), y_2(x)\right) = \left[y_2'(x) y_1(x) - y_1'(x) y_2(x)\right]$

Therefore $W'\left(y_1(x), y_2(x)\right) = \left[y_2''(x) y_1(x) - y_1''(x) y_2(x)\right]$

And equation (2.42) can be written as $W'\left(y_1(x), y_2(x)\right) + \alpha(x) W = 0$

Or $$\frac{dw}{dx} + \alpha(x) W = 0 \tag{2.43}$$

Solution of equation (2.43) is $W = Ce^{-\int \alpha(x)dx}$ (2.44)

Where C is a constant, depending on y_1 and y_2, but not on x. We assume that the function $\alpha(x)$ is continuous on the interval [a, b], then $\int \alpha(x)dx$ will also be continuous, on [a, b], so that $e^{-\int \alpha(x)dx} \neq 0$ in the assumed interval and therefore wronskian of two linearly independent solutions is never zero for all x in [a, b].

Now we have the following theorem :

Let $y_1(x)$ and $y_2(x)$ be differentiable on [a, b]. If wronskian $W(y_1, y_2)$ is non-zero for some x_0 in [a, b], then $y_1(x)$ and $y_2(x)$ are linearly independent on [a, b]. If $y_1(x)$ and $y_2(x)$ all linearly dependent then the wronskian is zero for all x in [a, b].

2.5.1 : Non-Homogeneous Linear Equation with Constant Co-Efficient

We consider the differentiable equation, $\frac{d^2y}{dx^2} + \alpha\frac{dy}{dx} + \beta y = Q(x)$ (2.45)

Where α and β are constants.

The general solution of equation (2.45) can be written as $y = y_c + y_p$ (2.46)

Where y_c is the complementary function (solution) corresponding to $Q(x) = 0$ in equation (2.45) and y_p is the particular integral (solution) of equation (2.45); corresponding to $Q(x) \neq 0$. In sec 2.4.2A we have discussed how to find y_c. Now we discuss how to find the particular integral y_p using the method of undetermined co-efficient.

Case I : If $Q(x)$ is a polynomial of degree n and zero is not a root of the characteristic equation, then y_p can be written as,

$$y_p = A_0 + A_1x + ... + A_nx^n \tag{2.47}$$

If however zero is a single root of the characteristic equation then

$$y_p = x(A_0 + A_1x + ... + A_nx^n) \tag{2.48}$$

Case II : If $Q(x) = ce^{\lambda x}$ and λ is not a solution of the characteristic equation, then

$$y_p = Ae^{\lambda x} \tag{2.49}$$

If λ is a single root of the characteristic equation, then $y_p = Axe^{\lambda x}$ (2.50)

If λ is a double root of the characteristic equation, then $y_p = Ax^2e^{\lambda x}$ (2.51)

Case III : If $Q(x) = \left.\begin{matrix} C\sin\lambda x \\ D\cos\lambda x \end{matrix}\right\}$ and $i\lambda$ is not a root of the characteristic equation, then

$$y_p = A\cos\lambda x + B\sin\lambda x \tag{2.52}$$

If however $i\lambda$ is a root of the characteristic equation, then

$$Y_P = x(A\cos\lambda x + B\sin\lambda x) \tag{2.53}$$

Case IV : Use of complex exponentials :

To find particular solution of $F(D)y = Q(x)$ where $Q(x)\begin{cases} C\sin\lambda x \\ C\cos\lambda x \end{cases}$ and where $F(D)$

= 0 is the characteristic equation, first solve $F(D)y = Ce^{i\lambda x}$ and then take the real and imaginary part.

Case V : When $Q(x)$ is an exponential times a Polynomial i.e. $Q(x) = e^{\lambda x}P_n(x)$

Where $P_n(x)$ is a Polynomial of degree n, a particular solution Y_p of

$$F(D)y = (D-a)(D-b)y = e^{\lambda x}P_n(X), \text{ is}$$

$$Y_p = \begin{cases} e^{\lambda x}Q_n(x) & \text{if } \lambda \text{ is not } a \text{ or } b \\ xe^{\lambda x}Q_n(x) & \text{if } \lambda \text{ equals } a \text{ or } b,\ a \neq b \\ x^2e^{\lambda x}Q_n(x) & \text{if } \lambda = a = b \end{cases} \tag{2.54}$$

Where $Q(x)$ is a polynomial of the same degree as $P_n(x)$ with undetermined co-efficient to be found to satisfy the given differential equation. Note that sines and cosines are included in $e^{\lambda x}$ by use of Complex exponentials as in Case IV.

Example of Case (I) : Find the general solution of $y'' + y' - 2y = x^2 - x$ (1)

Solution : The characteristic equation is $\lambda^2 + \lambda - 2 = 0$ ie $\lambda = \dfrac{-1+\sqrt{9}}{2}$

or, $\lambda_1 = \dfrac{-1+3}{2} = 1;\quad \lambda_2 = \dfrac{-1-3}{2} = -2$

Therefore the roots of the characteristic equation are (1, –2) and zero is not a root of the characteristic equation.

Therefore $\quad y_c = C_1e^x + C_2e^{-2x}$ (2)

And y_p may be taken in the form

$$y_p = A_0 + A_1x + A_2x^2 \tag{3}$$

Substituting equation (3) in equation (1), we get

$$Aa_2 + A_1 + 2A_2x - 2\left(A_0 + A_1x + A_2x^2\right) = x^2 - x$$

Or, $\quad 2A_2 - 2A_2x^2 + A_1 - 2A_1x + 2A_2x - 2A_0 = x^2 - x$

Or, $\quad -2A_2x^2 - 2A_1x + 2A_2x + 2A_2 + A_1 - 2A_0 = x^2 - x$

Comparing the co-efficient of various power of x, we get

$$x^0 \rightarrow 2A_2 + A_1 - 2A_0 = 0$$

$$x^1 \rightarrow 2A_2 - 2A_1 = -1, \quad A_1 = 0$$

$$x^1 \rightarrow -2A_2 = 1, \qquad A_2 = -\tfrac{1}{2}$$

$$A_0 = -\tfrac{1}{2}$$

Hence $y_p = -\frac{1}{2} - \frac{1}{2}x^2$

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

Therefore general solution y is given by

$$y = y_c + y_p = C_1 e^x + C_2 e^{-2x} - \frac{1}{2}\left(1 + x^2\right)$$

Example of Case (II) : Solve the equation $(D - 1)(D + 5)y = 7e^{2x}$ (1)

Solution : Characteristic equation is $(D - 1)(D + 5) = 0$ (2)

$\therefore$ Roots of characteristic equation

$$\lambda_1 = 1 \text{ and } \lambda_2 = -5$$

$$\therefore\ y_c = C_1 e^x + C_2 e^{-5x}$$

We see that $\lambda_2 = 2$ is not a root of the characteristic equation.

To find a particular solution we take $y_p = Ae^{\lambda x} = Ae^{2x}$ (3)

Now equation (1) can be re-written as

$$(D^2 + 4D - 5)y = 7e^{2x}$$

Or, $$\frac{d^2y}{dx^2} + 4\frac{dy}{dx} - 5y = 7e^{2x} \quad (4)$$

Substituting (3) in equation (4)

$$4Ae^{2x} + 8Ae^{2x} - 5Ae^{2x} = 7e^{2x}$$

$$7Ae^{2x} = 7e^{2x} \qquad \because e^{2x} \neq 0, A = 1$$

$$\therefore\ y_p = e^{2x} \quad (5)$$

Therefore general solution : $y = C_1 e^x + C_2 e^{-5x} + e^{2x}$

Example of Case (III) : Solve the initial value problem $\dfrac{d^2y}{dx^2} + 4y = 2\sin x$, given

$$y(0) = 1, \frac{dy}{dx}(0) = 0$$

Solution : Here $\lambda = 1$

Now characteristic equation $D^2 + 4 = 0$, $\therefore \lambda_1 = 2i, \lambda_2 = -2i$, are the roots of the characteristic equation. Hence $\lambda_1 = 1$ is not a root of the characteristic equation.

Now Complementary function $y_c = C_1 \cos 2x + C_2 \sin 2x$ (1)

We assume particular integral $y_p = A \sin x + B \cos x$ (2)

Substituting (1) in the original equation we get,

$$(-A \sin x - B \cos x) + 4(A \sin x + B \cos x) = 2 \sin x$$

Comparing co-efficient of sin x and cos x on both sides, we find $A = \dfrac{2}{3}, B = 0$

$$\therefore y_p = \frac{2}{3}\sin x$$

And the general solution $y = y_c + y_p = C_1 \sin 2x + C_2 \sin 2x + \dfrac{2}{3}\sin x$ (3)

Now $\dfrac{dy}{dx} = -2C_1 \sin 2x + 2C_2 \cos 2x + \dfrac{2}{3}\cos x$

Using the initial value $y(0) = C_1 = 1$ and $\dfrac{dy}{dx}(0) = 2C_2 + \dfrac{2}{3} = 0$

$$\therefore C_2 = -\frac{1}{3}$$

Hence the specific solution is $y = \cos 2x - \dfrac{1}{3}\sin 2x + \dfrac{2}{3}\sin x$

Example of Case (IV) : Find the complementary function of $y'' + y' - 2y = 4\sin 2x$ (1)

Solution : To find the particular solution y_p of equation (1), we find Y_p for the equation $Y'' + Y' - 2Y = 4e^{2ix}$ (2) and take its imaginary part.

We observe that $2i$ is not equal to a root of the auxiliary equation of (2).

Following the method written in case (II), we assume a solution.

$Y_p = Ae^{2ix}$ (3) and substitute it in equation (2) to get

$$(-4+2i-2)Ae^{2ix} = 4e^{2ix}$$

$$\therefore A = \frac{4}{2i-6} = \frac{4(2i+6)}{-40} = -\frac{1}{5}(i+3)$$

$$\therefore Y_p = -\frac{1}{5}(i+3)e^{2ix}$$

Taking the imaginary part of Y_p,

We get y_p of equation (1)

Or, $$y_p = I_m\left[-\frac{i}{5}e^{2ix} - \frac{3}{5}e^{2ix}\right]$$

$$\therefore y_p = I_m\left[-\frac{i}{5}(\cos 2x + i\sin 2x)\right] - I_m\left[\frac{3}{5}(\cos 2x + i\sin 2x)\right]$$

$$= -\frac{1}{5}\cos 2x - \frac{3}{5}\sin 2x$$

Where I_m means Imaginary Part.

Example of Case (V) : Find a particular solution of $y'' - 6y' + 9y = 12xe^{3x}$ (1)

Solution : Equation (1) is re-written as $(D-3)(D-3)y = 12xe^{3x}$ (2)

We observe that $\lambda = 3$ is equal to either of the roots of the auxiliary equation i.e. $\lambda_1 = a = b = \lambda_2$ also $P_n(x) = 12x = P_1(x)$ is a polynomial of degree 1. Then $Q(x)$ is also a polynomial of degree 1 namely $AX + B$. Since $\lambda = a = b$, we write $y_p = x^2e^{3x}(Ax + B) = e^{3x}(Ax^3 + Bx^2)$ (3)

We substitute equation (3) in equation (1) and find A and B so that we have an identity:

$$y_p'' - 6y_p' + 9y_p = 12xe^{3x}$$

We find A = 2 and B = 0,

$$\therefore y_p = e^{3x}(2x^3)$$

2.5.2 : Inhomogeneous Linear Equation with Constant Co-efficient : Working Rule for Finding the Particular Integral :

We have linear inhomogeneous differential equation with constant co-efficient $F(D)y = Q(x)$ (2.55)

As started earlier the solution of equation (2.55) consists of two parts

1) Solution for $F(D)Y = 0$, which is called complementary function (C.F) and

2) Any particular integral (P.I) of equation (2.55) given by

$$PI = \frac{1}{F(D)}Q(x) \tag{2.56}$$

Thus the solution of equation (2.55) is $y = CF + PI$

Rules for finding the particular integral (Pl) :

Rule I : When $Q(x) = e^{ax}$, a is a constant, $PI = \frac{1}{F(D)}Q(x) = \frac{1}{F(a)}e^{ax};\ F(a) \neq 0$

If $F(a) = 0$; $PI = x\frac{1}{F'(a)}e^{ax}; F'(a) \neq 0$

If $F'(a) = 0; PI = x^2\frac{1}{F''(a)}e^{ax}; F''(a) \neq 0$

And so on

Where $F'(D) = \frac{d}{dD}F(D);\ F'(a) = \frac{d}{dD}F(D)\Big|_{d=a}$ etc.

Rule II : When $Q(x) = e^{ex}V(x)$; When $V(x)$ is any function of x;

$$PI = \frac{1}{F(D)}e^{ax}V(x) = e^{ax}\frac{1}{F(D+a)}V(x)$$

Rule III : When $Q(x) = xV(x)$, When $V(x)$ is of the form $\sin(ax + b)$ or $\cos(ax + b)$;

Then, $PI = \frac{1}{F(D)}xV(x) = x\frac{1}{F(D)}V(x) - \frac{F'(D)}{\{F(D)\}^2}V(x)$

Rule IV : If Q(x) = x^m, m being a constant; $PI = \frac{1}{F(D)} x^m$.

Find $\frac{1}{F(D)}$ by actual division in ascending powers of D and retains term up to D^m.

Rule V : If Q(X) = sin (ax + b) or cos(ax+ b); where a & b are constant, then follow the method shown in worked out example.

Example of Rule I Find the particular integral of : $y'' - 3y' - 2y = e^{2x}$

Solution : $F(D) = D^2 - 3D + 2$

$$F(a) = F(2) = 2^2 - 3.2 + 2 = 0$$

Now $F'(D) = 2D - 3$

And $F'(2) = 2 \times 2 - 3 = 1 \neq 0$

Therefore, $PI = x \frac{1}{F'(2)} e^{2x} = xe^{2x}$

Example of Rule II : Find the particular integral of the equation :

$$y'' - 2y' + 4y = e^x \cos x$$

Solution : Here $a = 1$, $V(x) = \cos x$

$$PI = \frac{1}{F(D)} e^{ax} V(x) = e^{ax} \frac{1}{F(D+a)} V(x)$$

Now, $F(D) = D^2 - 2D + 4$

$$F(D + 1) = (D + 1)^2 - 2(D + 1) + 4$$

$$= D^2 + 2D + 1 - 2D - 2 + 4$$

$$= D^2 + 3$$

$\therefore$ $PI = e^x \frac{1}{D^2 + 3} \cos x = e^x \frac{1}{-1^2 + 3} \cos x$ (see example of Rule V)

$$= e^x \frac{1}{2} \cos x = \frac{1}{2} e^x \cos x$$

Example of rule III : Find the particular integral of the equation :

$(D^2 + 3D + 2)y = x \sin 2x$

Solution : Here $Q(x) = xV(x) = x \sin 2x$

$F(D) = D^2 + 3D + 2$

$F'(D) = 2D + 3$

$$\therefore PI = \frac{1}{D^2 + 3D + 2}(x \sin 2x)$$

$$= x\frac{1}{D^2 + 3D + 2}\sin 2x - \frac{2D+3}{\left(D^2 + 3D + 2\right)^2}\sin 2x$$

$$= x\frac{\sin 2x}{-2^2 + 3D + 2} - \frac{(2D+3)}{\left(-2^2 + 3D + 2\right)^2}\sin 2x$$

Where we have substituted $D^2 = (-a^2)$; here $a = 2$

$$= x\frac{\sin 2x}{3D-2} - \frac{2D+3}{(3D-2)^2}\sin 2x$$

Since the 1st term contains $3D - 2$ in the denominator we make it $9D^2 - 4$ by multiplying both numerator and denominator by $3D + 2$.

$$= x\frac{(3D+2)\sin 2x}{9D^2 - 4} - \frac{(2D+3)\sin 2x}{9D^2 - 12D + 4}$$

$$= x\frac{(3D+2)\sin 2x}{9(-2)^2 - 4} - \frac{(2D+3)\sin 2x}{9(-2)^2 - 12D + 4}$$

Where we have substituted $D^2 = -a^2 = (-2^2)$

$$= x\frac{(3D+2)\sin 2x}{-40} + \frac{(2D+3)\sin 2x}{(12D+32)}$$

In the 2nd term since there is no term contain D^2 in the denominator, we multiply both numerator and denominator of the 2nd term by $12D - 32$

$$= x\frac{(6\cos 2x + 2\sin 2x)}{-40} + \frac{(2D+3)\,(12D-32)}{(12D+32)\,(12D-32)}\sin 2x$$

$$= -x\frac{(6\cos 2x + 2\sin 2x)}{40} + \frac{(2D+3)\,(D-8)\sin 2x}{4(9D^2-64)}$$

or $$PI = -x\frac{3\cos 2x + \sin 2x}{20} + \frac{(2D+3)\,(6\cos 2x - 8\sin 2x)}{4(-100)}$$

$$= -\frac{x(3\cos 2x + \sin 2x)}{20} + \frac{48\sin 2x + 14\cos 2x}{400}$$

Also see example of Rule – V.

Example of Rule IV : Find the particular integral of the equation :

$$(2D^2 + 2D + 3)y = X^2 + 2x + 1$$

Solution : $Q(x) = X^2 + 2x + 1$

$$PI = \frac{1}{F(D)}\left(x^2 + 2x + 1\right)$$

Where $F(D) = 2D^2 + 2D + 3$

Now $\frac{1}{3+2D+2D^2}$ is found by actual division (not using any formula) and retaining up to the term containing D^2 in the quotient, since the degree of the polynomial $x^2 + 2x + 1$ is 2.

Therefore $PI = \dfrac{1}{F(D)}(x^2 + 2x + 1) = \dfrac{1}{3 + 2D + 2D^2}(x^2 + 2x + 1)$

$$= \left(\frac{1}{3} - \frac{2}{9}D + \frac{2}{27}D^2\right)(x^2 + 2x + 1)$$

$$= \frac{1}{3}x^2 - \frac{4}{9}x - \frac{4}{27} + \frac{2x}{3} - \frac{4}{9} + \frac{1}{3}$$

$$= \frac{1}{3}(x^2 + 2x - 1) - \frac{2}{9}(2x + 2) - \frac{2}{27}(2)$$

Example of Rule V : Find the particular integral of the equation :

$$(D^2 + 3D - 4)y = \sin 2x$$

Solution : Here $F(D) = D^2 + 3D - 4$; $a = 2$, $b = 0$

i.e. $\sin(ax + b) = \sin 2x$

Now $PI = \dfrac{1}{F(D)} Q(x) = \dfrac{1}{D^2 + 3D - 4} \sin 2x$

Now putting $D^2 = -a^2 = -2^2$, we get

$$PI = \frac{1}{3D - 8} \sin 2x = \frac{(3D + 8) \sin 2x}{(3D - 8)(3D + 8)}$$

$$= \frac{(3D + 8) \sin 2x}{9D^2 - 64}$$

Again putting $D^2 = -a^2 = -2^2$

$$PI = \frac{(3D + 8) \sin 2x}{-100}$$

$$= \frac{6 \cos 2x + 8 \sin 2x}{-100}$$

Case VI : When Q(x) is sum of several terms consisting of exponential, polynomial and trigonometric functions etc. ; then particular integral will be algebraic sum of individual particular integrals according to the principle of superposition. For example, if $Q(x) = (e^{ax}) + (4 \sin bx) + (ax^2 + bx)$;

Then the particular integral will be algebraic sum of the individual particular integrals corresponding to the respective function.

We take an imaginary differential equation :

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 2y = \left[e^{ax}\right] + [4 \sin bx] + \left[ax^2 + bx\right]$$

Then the particular integrals of the given differential equation will be given by :

$$y_p = y_{p_1} + y_{p_2} + y_{p_3}$$

When y_{p_1} is the particular integral for $e^{ax} y_{p_2}$ is the particular integral for 4 sin bx and y_{p_3} is the particular integral for $[ax^2 + bx]$.

Exercise (2.4.5) :

1) Obtain the general solution of the equation : $\frac{d^2y}{dx^2}-3\frac{dy}{dx}+2y=x^3$

2) Solve the equation $(D^2 + 4)y = \sin 2x$

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

Solution (1) : In D-operator notation, the equation become, $(D^2 - 2D + 2)y = x^3$

Auxiliary equation is : $D^2 - 3D + 2 = 0$

Roots of the auxiliary equation are $m_1 = 1$, $m_2 = 2$

Complementary function $C = y_c = Ae^x + Be^{2x}$

Particular integral is given by $y_p=\frac{1}{2-3D+D^2}x^3=\frac{1}{2}\left(x^3+\frac{9}{2}x^2+\frac{21}{2}x+\frac{45}{4}\right)$

$\therefore$ the complete solution is $y = y_c + y_p$

Solution (2) : The auxiliary equation is $D^2 + 4 = 0$

Roots of auxiliary equation is $m = \pm 2i$

Complementary function $y_c = A\cos 2x + B\sin 2x$

Where A and B are arbitrary constant.

Particular integral $y_p=\frac{1}{D^2+4}\sin 2x$

Now $D^2 = -(2)^2 = -4$ $\quad\therefore D^2 + 4 = 0$

$$\therefore\ y_p=x\frac{1}{2D}\sin 2x=\frac{x}{2}\frac{1}{D}\sin 2x=-\frac{x\cos 2x}{4}$$

$$\therefore\ y=y_c+y_p=A\cos 2x+B\sin 2x-\frac{x}{4}\cos 2x$$

Solution 3 : Complementary function $y_c = (A + Bx)e^x$

Particular integral,

$$y_p=\frac{1}{D^2-2D+1}e^x\,x\sin x=e^x\frac{1}{(D+1)^2-2(D+1)+1}x\sin x$$

$$= e^x \frac{1}{D^2} x \sin x = e^x \frac{1}{D} \int x \sin x dx = e^x \frac{1}{D} \left[x(-\cos x) - \int (-\cos x)\, dx \right]$$

$$= e^x \int (-x \cos x + \sin x)\, dx$$

$$= e^x \left[-\left\{ x \sin x - \int \sin x\, dx \right\} - \cos x \right]$$

$$= -\, e^x\, (x \sin x + 2 \cos x)$$

$\therefore$ Complete solution of $y = y_c + y_p$

2.6 ❐ Summary

i) Classification of second order differential equation explained.

ii) Different method of finding particular integral have been exemplified.

iii) Rules for finding complementary function and particular integral have also been included.

iv) Existence and uniqueness theorems for IVP have been illustrated with examples.

v) Use of wronskian to identify linear dependent and independent solutions have been discussed.

Unit 3 ❐ Calculus of Functions of More than one Variable

Structure

3.1 ❐ Objectives

1. To know what is partial and total derivatives and differentials.
2. To make an idea about exact and inexact differential.
3. To know how to convert inexact differential into exact differential with the help of integrating factor.
4. To know how to find out maximum and minimum values of functions with constraints using Lagrange multipliers.

Keywords : Partial derivatives, total derivatives, exact and inexact differentials, integrating factors, maxima and minima problems with constraint.

3.2 ❐ Introduction

We consider a field scalar such as temperature (T) and its distribution in a region of space. We see that temperature may change with x, y and z co-ordinates of space and also with time t if the state is not steady. Thus we see temperature is, in essence, function of several variables x, y, z, t i.e.

$$T = T\ (x,\ y,\ z,\ t) \tag{3.1}$$

Now if we want to find the rate of change of temperature we can find it in various ways. Suppose we want to find the rate of change with x co-ordinate only keeping y, z, and t constants. We use partial derivative $\left(\frac{\partial T}{\partial x}\right)$. If temperature T is function of x- co-ordinates alone we could find the above rate of change by $\frac{dT}{dx}$, the ordinary derivative. Derivatives are also used in finding the maxima or minima of a curve.

Now rates occur very often in physics e.g. time rates, space rate etc. and we have to find these rates in the form of differential equations which we have to solve to find out the rate of functional dependence of the quantity with other.

3.3 ❐ Partial Derivative

Suppose we have a function f, having more than one independent variables (x, y) i.e. $f = f(x, y)$. Now if for $f(x, y)$, keeping y as constant, an ordinary differentiation with respect to x is found, the derivative so obtained is called partial derivative and is denoted by $\left(\frac{\partial f}{\partial x}\right)$ or f_x where $f_x = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x, y)}{\Delta x}$ (3.1)

Similarly treating x as constant, we get $f_y = \frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f\left(x, y+\Delta y\right) - f(x, y)}{\Delta y}$

(3.2)

Now we can determine higher order partial derivatives also which are denoted by f_{xx}, f_{xy}, f_{yy} for second order; f_{xxx}, f_{xxy}, f_{xxz} etc. for third order partial derivatives.

Where $f_{xx} = \frac{\partial^2 f}{\partial x^2}$, $f_{xy} = \frac{\partial^2 f}{\partial x \partial y}$, $f_{yy} = \frac{\partial^2 f}{\partial y^2}$

$$f_{xxx} = \frac{\partial^3 f}{\partial x^3}, \quad f_{xxy} = \frac{\partial^3 f}{\partial^2 x \partial y}, \quad f_{xxz} = \frac{\partial^3 f}{\partial^2 x \partial z} \text{ etc.}$$

A notation which is frequently used in thermodynamics is $\left(\frac{\partial f}{\partial x}\right)_y$, meaning we have to find out partial derivative of $f(x, y)$ with respect to x when y is held constant. Similarly $\left(\frac{\partial f}{\partial y}\right)_x$ is defined.

Now suppose we find out $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$ and also find out $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$. Then a question automatically comes out that could we write $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$?

It can be proved that if the first and second order partial derivatives of $f(x, y)$ are continuous, then only $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$, otherwise not.

In thermodynamics these conditions are usually satisfied and the equality hold.

3.3.1 Total Differential

We consider a function of two variable (x, y) represented by $z = f(x, y)$, which represents a surface. Now the derivatives $\frac{\partial f}{\partial x} = f_x$, $\frac{\partial f}{\partial y} = f_y$, at a point, are the slopes of the two tangent lines to the surface in the x and y directions at that point. The symbols $\Delta x = dx$ and $\Delta y = dy$ represent changes in the independent variables x and y. The quantity Δz means the corresponding change in z along the surface.

We define dz by the equation $dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$ (3.3)

The differential dz is called the total differential of z.

Equation (3.3) may also be interpreted as follows. Any change in z, Δz will be sum of changes due to change in x and changes in y respectively. Now rate of change of z with respect to x and y is given by respectively $\frac{\Delta z}{\Delta x}$ and $\frac{\Delta z}{\Delta y}$.

Therefore we can write change in z as $\Delta z = \left(\dfrac{\Delta z}{\Delta x}\right)_y \Delta x + \left(\dfrac{\Delta z}{\Delta y}\right)_x \Delta y$.

Now in the limit $\Delta x \to 0$ and $\Delta y \to 0$ we can write

$$dz = \left(\frac{\partial z}{\partial x}\right)dx + \left(\frac{\partial z}{\partial y}\right)dy \tag{3.3}$$

3.3.2 Error determination and approximation

If z = $f(x, y, u, ...)$, then the total variation Δz in z due to changes Δx, Δy, Δu, ... in x, y, u, ... is given by [see 3.3A].

$$\Delta z = \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y + \frac{\partial z}{\partial u}\Delta u + ...$$

Here Δx (or δx), Δy, ... are the actual errors in x & y, ... while Δz is the approximate error in z i.e. in $z = f(x, y, u, ...)$

Now Δx, Δy, ... are known as absolute errors in x, y, ... in measurement and $\dfrac{\Delta x}{x}, \dfrac{\Delta y}{y}, ...$ are called the proportional error in x, y, ... etc.

Example of Art : 3.2 and 3.3 :

1. If $u = \dfrac{1}{r}$, where $r = \sqrt{x^2 + y^2 + z^2}$

Show that $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2} = 0$

Solution : $u = \dfrac{1}{\left(x^2 + y^2 + z^2\right)^{\frac{1}{2}}} = \left(x^2 + y^2 + z^2\right)^{-\frac{1}{2}}$

$$\therefore \frac{\partial u}{\partial x} = -\frac{1}{2}\left(x^2 + y^2 + z^2\right)^{-\frac{3}{2}} .\ 2x = -x\left(x^2 + y^2 + z^2\right)^{-\frac{3}{2}}$$

And $$\frac{\partial^2 u}{\partial x^2} = -\left[1.\left(x^2 + y^2 + z^2\right)^{-\frac{3}{2}} + x\left(-\frac{3}{2}\right)\left(x^2 + y^2 + z^2\right)^{-\frac{5}{2}} 2x\right]$$

$$= -\left(x^2+y^2+z^2\right)^{-\frac{5}{2}}\left[x^2+y^2+z^2-3x^2\right]$$

$$= \left(x^2+y^2+z^2\right)^{-\frac{5}{2}}\left(2x^2-y^2-z^2\right) = r^{-5}\left(2x^2-y^2-z^2\right)$$

Similarly $\dfrac{\partial^2 u}{\partial y^2} = r^{-5}\left(2y^2-x^2-z^2\right)$

and $\quad \dfrac{\partial^2 u}{\partial z^2} = r^{-5}\left(2z^2-x^2-y^2\right)$

$$\therefore \quad \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2} = r^{-5}\left[2\left(x^2+y^2+z^2\right)-2\left(x^2+y^2+z^2\right)\right] = 0$$

2) Express two-dimensional Laplace's equation :

$\dfrac{\partial^2 u}{\partial x^2}+\dfrac{\partial^2 u}{\partial y^2} = 0$ in polar co-ordinate.

Solution : Equation of transformation from Cartesian co-ordinates (x, y) to polar co-ordinates (r, θ) is given by : $x = r\cos\theta$, $y = r\sin\theta$

$$\therefore r = \sqrt{x^2+y^2},\ \theta = \tan^{-1}\left(\frac{y}{x}\right) \qquad \text{(i)}$$

Now

$$\left.\begin{aligned} \frac{\partial r}{\partial x} &= \frac{2x}{2\sqrt{x^2+y^2}} = \frac{x}{r} = \cos\theta \\ \frac{\partial r}{\partial y} &= \frac{y}{r} = \sin\theta \end{aligned}\right\} \qquad \text{(ii)}$$

$$\left.\begin{aligned} \frac{\partial\theta}{\partial x} &= \frac{1}{1+\frac{y^2}{x^2}}\cdot\left(-\frac{y}{x^2}\right) = -\frac{y}{x^2+y^2} = -\frac{\sin\theta}{r} \\ \frac{\partial\theta}{\partial x} &= \frac{1}{1+\frac{y^2}{x^2}}\cdot\left(\frac{1}{x}\right) = \frac{x}{x^2+y^2} = \frac{\cos\theta}{r} \end{aligned}\right\} \qquad \text{(iii)}$$

Now
$$\left.\begin{aligned}\frac{\partial}{\partial x}=\frac{\partial}{\partial r}\cdot\frac{\partial r}{\partial x}+\frac{\partial}{\partial \theta}\cdot\frac{\partial \theta}{\partial x}=\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\\ \frac{\partial}{\partial y}=\frac{\partial}{\partial r}\cdot\frac{\partial r}{\partial y}+\frac{\partial}{\partial \theta}\cdot\frac{\partial \theta}{\partial y}=\sin\theta\frac{\partial}{\partial r}+\frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\end{aligned}\right\} \qquad \text{(iv)}$$

$$\therefore \frac{\partial^2}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\right)=\left(\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\left(\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)$$

$$=\cos^2\theta\frac{\partial^2}{\partial r^2}-\frac{2\sin\theta\cos\theta}{r}\frac{\partial^2}{\partial r\,\partial\theta}+\frac{\sin^2\theta}{r^2}\frac{\partial^2}{\partial\theta^2}+\frac{\sin^2\theta}{r}\frac{\partial}{\partial\theta}+\frac{2\sin\theta\cos\theta}{r^2}\frac{\partial}{\partial\theta} \qquad \text{(v)}$$

Similarly
$$\frac{\partial^2}{\partial y^2}=\sin^2\theta\frac{\partial^2}{\partial r^2}+\frac{2\sin\theta\cos\theta}{r}\frac{\partial^2}{\partial r\,\partial\theta}+\frac{\cos^2\theta}{r^2}\frac{\partial^2}{\partial\theta^2}+\frac{\cos^2\theta}{r}\frac{\partial}{\partial\theta}$$
$$-\frac{2\sin\theta\cos\theta}{r^2}\frac{\partial}{\partial\theta} \qquad \text{(vi)}$$

Adding (v) and (vi) we get,

$$\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}=\frac{\partial^2}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}+\frac{1}{r}\frac{\partial}{\partial r}$$

$$\therefore \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=\frac{\partial^2 u}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 u}{\partial\theta^2}+\frac{1}{r}\frac{\partial u}{\partial r}=0$$

3) Prove that if $f(x, y, z) = 0$

I. $\left(\frac{\partial z}{\partial x}\right)_y=\frac{1}{\left(\frac{\partial x}{\partial z}\right)_y}$ and two similar relation

II. $\left(\frac{\partial x}{\partial y}\right)_z\left(\frac{\partial y}{\partial z}\right)_x\left(\frac{\partial z}{\partial x}\right)_y=-1$

Proof : We have, $f(x, y, z) = 0$ i.e. $x = x(y, z)$ and $z = z(x, y)$

$$\therefore\ dx=\left(\frac{\partial x}{\partial y}\right)_z dy+\left(\frac{\partial x}{\partial z}\right)_y dz \qquad \text{(i)}$$

and $$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy \qquad \text{(ii)}$$

Eliminating dz from (i) and (ii) we get

$$dx = \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y \left[\left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy\right] \qquad \text{(iii)}$$

Now dx and dy are independent of each other. So, equating the co-efficient of dx from both side of (iii), we get,

$$\left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial z}{\partial x}\right)_y = 1 \Rightarrow \left(\frac{\partial z}{\partial x}\right)_y = \frac{1}{\left(\frac{\partial x}{\partial z}\right)_y} \quad \text{(I)}$$

Similarly, by equating the co-efficient of dy we obtain,

$$\left(\frac{\partial x}{\partial y}\right)_z + \left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial z}{\partial y}\right)_x = 0$$

or, $$\left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z + 1 = 0$$

or, $$\left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z = -1 \quad \text{(II)}$$

Note : Relation (I) and (II) are extensively used in thermodynamics systems e.g. hydrostatic system given by $f(P, V, T) = 0$.

Example 4 : The rate of flow (V) of a liquid through a capillary tube of radius r and length l at a pressure difference P between it ends is given by : $V = \dfrac{\pi P r^4}{8\eta l}$, where η is the viscosity of the liquid. In an experiments, the errors in the measurements of P, r, l and V are 1%, 1.5%, 0.5% and 2% respectively. Evaluate the error in the measurements of η.

Solution : We have $\eta = \dfrac{\pi P r^4}{8Vl}$

$$\therefore \ln \eta = \ln \pi + \ln P + \ln r^4 - \ln 8 - \ln V - \ln l$$

or, $$\frac{\delta\eta}{\eta} = 0 + \frac{\delta P}{P} + 4\frac{\delta r}{r} - 0 - \frac{\delta V}{V} - \frac{\delta l}{l}$$

or, $$\left(\frac{\delta\eta}{\eta}\right)_{\max} = \frac{\delta P}{P} + 4\frac{\delta r}{r} + \frac{\delta V}{V} + \frac{\delta l}{l}$$, therefore

Percentage error, $$\left(\frac{\delta\eta}{\eta}\right)_{\max} \times 100 = \frac{\delta P}{P}\times 100 + 4\frac{\delta r}{r}\times 100 + \frac{\delta V}{V}\times 100 + \frac{\delta l}{l}\times 100$$

$$= 1 + 4 \times 1.5 + 2 + 0.5 = 9.5\%$$

Notes : During experiments students are asked to find out the maximum proportional error in their measurement. Therefore all the terms are added. For percentage error ; proportional error is multiplied by 100.

Example 5 : If $u = f(x, y, z, ...)$ where $x, y, z, ...$ are all functions of a variable t.

Prove that,

I) $$\frac{du}{dt} = \frac{\partial u}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial u}{\partial y}\cdot\frac{dy}{dt} + ...$$

II) $$\frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$$

Proof : We have $u = f(x, y, ...)$

$$\therefore du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial z}dz + ... \qquad \text{(i)}$$

But from the definition of differential. We obtain : $du = \frac{du}{dt}dt,\ dx = \frac{dx}{dt}dt,\ dy = \frac{dy}{dt}dt$ etc.

From (i) $$\frac{du}{dt} = \frac{\partial u}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial u}{\partial y}\cdot\frac{dy}{dt} + ... \qquad \text{(I)}$$

If $u = f(x, y) = C$ (constant), then y is an implicit function of x.

Now from equation (i) $$\frac{du}{dx} = \frac{\partial u}{\partial x}\cdot\frac{dx}{dx} + \frac{\partial u}{\partial y}\cdot\frac{dy}{dx} = 0$$

$$\Rightarrow \frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\frac{dy}{dx}=0$$

or, $$\frac{dy}{dx}=-\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$$

Exercise : Art 3.2 and 3.3 :

1) If $u = \log(x^3 + y^3 + z^3 - 3xyz)$, show that

$$\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)^2 u=-\frac{9}{\left(x^2+y^2+z^2\right)^2}$$

2) If z varies directly as x and inversely as y and the possible errors in measuring x and y are 1% and 0.5% respectively, find the amount of error in z. Given $z=\frac{1}{3}$ when $x = 3$; $y = 5$.

3) If $z = f(x+ct)+\phi(x-ct)$, show that $\frac{\partial^2 z}{\partial t^2}=c^2\frac{\partial^2 z}{\partial x^2}$, where c is a constant.

Solutions to exercise Art : 3.2 and 3.3 :

Solution (1) : $u = \log(x^3 + y^3 + z^3 - 3xyz)$

$$\therefore \frac{\partial u}{\partial x}=\frac{3x^2-3yz}{x^3+y^3+z^3-3xyz},\ \frac{\partial u}{\partial y}=\frac{3y^2-3xz}{x^3+y^3+z^3-3xyz},\ \frac{\partial u}{\partial z}=\frac{3z^2-3xy}{x^3+y^3+z^3-3xyz}$$

$$\therefore \frac{\partial u}{\partial x}+\frac{\partial u}{\partial y}+\frac{\partial u}{\partial z}=\frac{3\left(x^2+y^2+z^2-xy-yz-xz\right)}{(x+y+z)\left(x^2+y^2+z^2-xy-yz-xz\right)}=\frac{3}{x+y+z}$$

Now $$\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)^2 u=\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)\left(\frac{\partial u}{\partial x}+\frac{\partial u}{\partial y}+\frac{\partial u}{\partial z}\right)$$

$$=\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\right)\left(\frac{3}{x+y+z}\right)$$

$$= -\frac{3}{(x+y+z)^2} - \frac{3}{(x+y+z)^2} - \frac{3}{(x+y+z)^2} = \frac{-9}{(x+y+z)^2}$$

Solution 2 : We have $z \propto x$ and $z \propto \frac{1}{y}$.

$\therefore z = k\frac{x}{y}$, where k is a constant.

Now when $x = 3$, $y = 5$, $z = \frac{1}{3} = k\frac{3}{5}$

$$\therefore k = \frac{5}{9}$$

$$\therefore z = \frac{5x}{9y}$$

$\therefore$ ln z = ln $5x$ – ln $9y$ = ln 5 + ln x – ln 9 – ln y

$$\therefore \frac{\delta z}{z} = 0 + \frac{\delta x}{x} - 0 - \frac{\delta y}{y}$$

$$\therefore \left(\frac{\delta z}{z}\right)_{\max} = \frac{\delta x}{x} + \frac{\delta y}{y}$$

$$\therefore \frac{\delta z}{z}\times 100 = \frac{\delta x}{x}\times 100 + \frac{\delta y}{y}\times 100 = 1 + 0.5 = 1.5\%$$

$\therefore$ maximum proportional error in z is 1.5%.

Now $\frac{\delta z}{z} = \frac{\delta x}{x} + \frac{\delta y}{y}$

$$\therefore \delta z = \left(\frac{\delta x}{x} + \frac{\delta y}{y}\right) z = \left(\frac{1}{100} + \frac{0.5}{100}\right) z = \frac{1.5}{100} z$$

$\therefore$ error in z is 1.5% of z.

Solution (3) : We have $z = f(x+ct) + \phi(x-ct)$

$$\therefore \frac{\partial z}{\partial x} = f'(x+ct)\frac{\partial}{\partial x}(x+ct) + \phi'(x-ct)\frac{\partial}{\partial x}(x-ct)$$

$$= f'(x+ct).1 + \phi'(x-ct).1$$

$$\therefore \frac{\partial^2 z}{\partial x^2} = f''(x+ct) + \phi''(x-ct) \qquad \text{(i)}$$

Again $\frac{\partial z}{\partial t} = f'(x+ct)\cdot\frac{\partial}{\partial t}(x+ct) + \phi'(x-ct)\cdot\frac{\partial}{\partial t}(x-ct) = c\ f'(x+ct) - c\phi'(x-ct)$

And $\frac{\partial^2 z}{\partial t^2} = c^2 f''(x+ct) + c^2\phi''(x-ct) = c^2\left[f''(x+ct) + \phi''(x-ct)\right]$ (ii)

From (i) and (ii) : $\frac{\partial^2 z}{\partial t^2} = c^2\frac{\partial^2 z}{\partial x^2}$

3.4 ❐ Exact and Inexact Differentials

We consider a function $z = f(x, y)$ = constant, which is continuous along with its first order partial derivative. Then the total differential is $dz = \left(\frac{\partial f}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y}\right)dy = 0$

This can be expressed as $dz = M(x, y)dx + N(x, y)dy = 0$ (3.4)

where $M(x, y) = \left(\frac{\partial f}{\partial x}\right)$ and $N(x, y) = \left(\frac{\partial f}{\partial y}\right)$.

Since the function $f(x, y)$ has continuous first order derivative, we can write $\frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^2 f}{\partial y\partial x}$ i.e $\left(\frac{\partial N}{\partial x}\right) = \frac{\partial M}{\partial y}$ (3.5)

This is the necessary and sufficient condition that the expression (3.4) be an exact differential equation and the differential $dz = M(x, y)dx + N(x, y)dy$ be an exact differential.

If dz is an exact differential $z = f(x, y)$ is called a point function or state function.

If however equation (3.5) is not satisfied then the differential dz is called inexact differential and the function $z = f(x, y)$ is called a path function. Conditions for equality of f_{xy} and f_{yx}

a) If (a, b) be a point in the domain of definitions of $f(x, y)$ so that $f_x(x, y)$ and $f_y(x, y)$ are differentiable at (a, b) then $f_{xy}(a, b) = f_{yx}(a, b)$

b) If (a, b) be a point in the domain of definitions of $f(x, y)$ so that $f_x(x, y)$ exist in a certain neighborhood of (a, b) and $f_{xy}(x, y)$ is continuous at (a, b) then $f_{xy}(a, b) = f_{yx}(a, b)$

3.4.1 : Integrating factor : Integrating factor is a function chosen to make an inexact differential to be transformed into an exact differential.

We consider the equation $dz = M(x, y)dx + N(x, y)dy = 0$

Which we suppose, to be not exact. Now if there exists a function $\phi(x, y)$ such that $\phi(x, y)\left[Mdx + Ndy\right] = d\varphi$ for some function $\varphi(x, y)$, then $\phi(x, y)$ is called an integrating factor of equation (3.4)

For example the equation $xdy - ydx = 0$ is not exact, multiplying it by $-\frac{1}{y^2}$, the equation became

$$\frac{-xdy + ydx}{y^2} = \frac{ydx - xdy}{y^2} = 0$$

Or, $$d\left(\frac{x}{y}\right) = 0$$

which is exact and has the general solution $\frac{x}{y}$ = constant.

Thus $-\frac{1}{y^2}$ is integrating factor (I.F) of the inexact differential $xdy - ydx$.

3.4.2 : Rules to find out integrating factor :

Let the differential equation $dz = M(x, y)dx + N(x, y)dy = 0$ is not exact. i.e. $\left(\frac{\partial M}{\partial y}\right) \neq \left(\frac{\partial N}{\partial x}\right)$

Rule I :

i. $Mdx + Ndy = 0$ and $M(x, y)$, $N(x, y)$ are both homogeneous functions of (x, y) of the same degree, then $\frac{1}{Mx + Ny}$ is an integrating factor of the equation

$Mdx + Ndy = 0$

Example 1 : Consider the differential equation $dz = (x^2y - 2xy^2)dx + (3x^2y - x^3)dy = 0$

Here $M(x, y) = x^2y - 2xy^2$ $\quad \therefore \dfrac{\partial M}{\partial y} = x^2 - 4y$

$N(x, y) = 3x^2y - x^3$ $\quad \therefore \dfrac{\partial N}{\partial x} = 6xy - 3x^2$

Therefore, $\left(\dfrac{\partial M}{\partial y}\right) \neq \left(\dfrac{\partial N}{\partial x}\right)$ and dz is not exact differential.

Now $M(x, y) = x^2y - 2xy^2 = x^3\left[\dfrac{y}{x} - 2\left(\dfrac{y^2}{x^2}\right)\right]$

$$= x^3\phi\left(\frac{y}{x}\right)$$

$\therefore$ $M(x, y)$ is a homogeneous function of degree 3 in x and y.

Similarly $N(x, y)$ is a homogeneous function of degree 3 in x and y

Therefore $Mx + Ny = x^2y^2 \neq 0$ .and $\dfrac{1}{x^2y^2}$ is an I.F.

Now multiplying both sides of the given equation by $\dfrac{1}{x^2y^2}$, we get

$$\frac{x^2y - 2xy^2}{x^2y^2}dx + \frac{3x^2y - x^3}{x^2y^2}dy = 0$$

Or, $$\left(\frac{1}{y} - \frac{2}{x}\right)dx + \left(\frac{3}{y} - \frac{x}{y^2}\right)dy = 0$$

Or, $$\frac{1}{y}dx - \frac{x}{y^2}dy - \frac{2dx}{x} + \frac{3dy}{y} = 0$$

Or, $$\frac{ydx - xdy}{y^2} - \frac{2dx}{x} + \frac{3dy}{y} = 0$$

Or, $$d\left(\frac{x}{y}\right)-2d(\log x)+3d(\log y)=0$$

Therefore $\frac{dz}{x^2y^2}=d\left[\frac{x}{y}-2\log x+3\log y\right]$ is an exact differential.

Rule II : Consider the differential equation $M(x, y)dx + N(x, y)dy = 0$. If this equation is not exact, then $\frac{\partial M}{\partial y}\neq\frac{\partial N}{\partial x}$ however if $\frac{\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}}{N}$ is function of x only denoted by $f(x)$, then $\mu(x)=e^{\int f(x)dx}$ will be an integrating factor of the given differential equation.

Rule III : However if $\frac{\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}}{-N}$ is function of y only, denoted by $f(y)$ then $\mu(y)=e^{\int f(y)dy}$ will be an integrating factor.

Example 2 : The differential equation $dz = (3xy - y^2)dx + x(x - y)dy = 0$ is not exact, since $\frac{\partial M}{\partial y}=3x-2y$ and $\frac{\partial N}{\partial x}=2x-y$.

$$\therefore \frac{\partial M}{\partial y}\neq\frac{\partial N}{\partial x}$$

However note that $\frac{\frac{\partial M}{\partial y}-\frac{\partial N}{\partial x}}{N}=\frac{1}{x}$ therefore by Rull II $e^{\int f(x)dx}=e^{\int\frac{dx}{x}}=e^{\ln x}=x$ will be an integrating factor.

Now multiplying both sides of the given equation by $\mu(x) = x$ yields $(3x^2y - xy^2)dx + (x^3 - x^2y)dy = 0$ which is exact because

$$\frac{\partial M}{\partial y}=3x^2-2xy$$

$$\frac{\partial N}{\partial x}=3x^2-2xy$$

$$\therefore \frac{\partial M}{\partial y}=\frac{\partial N}{\partial x}$$

Example 3 : The differential $dz = (x + y) \sin y dx + (x \sin y + \cos y)dy$ is not exact, since

$$\frac{\partial M}{\partial y} = (x+y)\cos y + \sin y$$

$$\frac{\partial N}{\partial x} = \sin y$$

However $\dfrac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{-M} = -\dfrac{\cos y}{\sin y}$ is a function of y alone denoted by $\varphi(y)$

Now $\quad \int \varphi(y)dy = -\int \dfrac{\cos y}{\sin y} dy = -\ln(\sin y)$

Therefore, $\quad$ I.F $= e^{\int \varphi(y)dy} = e^{-\ln(\sin y)}$

$$= e^{\ln(\sin y)^{-1}} = (\sin y)^{-1}$$

$$\therefore \frac{dz}{\sin y} = (x+y)dx + \left(x + \frac{\cos y}{\sin y} \right) dy \text{ which is exact.}$$

Since, $\dfrac{\partial M}{\partial y} = 1, \quad \dfrac{\partial N}{\partial x} = 1$

3.5 ❒ Constrained Maximization using Lagrange's Undermined Multipliers

We discuss a problem of maximum and minimum values of a function with constraint as follows,

Suppose we want to find the maximum or minimum of a function $u(x, y)$, where x and y are related by an equation $\phi(x, y)$ = constant i.e. x, y are not independent. This type of extra relation between the variables are known as constraints.

In such type of cases the points where maxima or minima occur and corresponding maximum or minimum values of the function can be determined by a number of methods e.g. a) method of elimination; b) method of implicit differentiation ; c) method of Lagrange multipliers.

Sometimes methods a) and b) can involve enormous calculation and we can solve the problem in concise form by a process known as method of Lagrange's undetermined multipliers.

3.5.1. Method of Lagrange's Undetermined Multipliers with Functions of Two Independent Variables and one f´ Equations

Now we discuss the method of Lagrange's undetermined multipliers, to find the maximum or minimum points of u = u(x, y) consisting of two independent variables. We set $\frac{du}{dx} = 0$ or $du = 0$. Again since $\phi(x, y)$ = constant, we get $d\phi = 0$. Then u is really a function of one variable, say x. therefore,

$$\left.\begin{aligned} du &= \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0 \\ d\phi &= \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy = 0 \end{aligned}\right\} \tag{3.6}$$

We multiply the $d\phi$ equation by λ and add it to the du equation, then we have,

$$\left(\frac{\partial u}{\partial x} + \lambda \frac{\partial \phi}{\partial x}\right) dx + \left(\frac{\partial u}{\partial y} + \lambda \frac{\partial \phi}{\partial y}\right) dy = 0 \tag{3.7}$$

where λ is undetermined multiplier.

Now we chosen λ so that,

$$\frac{\partial u}{\partial y} + \lambda \frac{\partial \phi}{\partial y} = 0 \tag{3.8}$$

From equation (3.7) and (3.8) we get,

$$\frac{\partial u}{\partial x} + \lambda \frac{\partial \phi}{\partial x} = 0 \tag{3.9}$$

Equation (3.8), (3.9) and $\phi(x, y)$ = constant can now be solved for the three unknowns x, y, λ.

3.5.2 : Method of Lagrange undetermined multiplier with function of three independent variables and one ϕ – equation

Now we discuss the same problem with function of three independent variables (x, y, z). We want to find maximum or minimum values of $u(x, y, z)$, when $\phi(x, y, z)$ = constant.

For maximum or minimum values of u, we set

$$\left.\begin{aligned} du &= \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial z}dz = 0 \\ \textit{and} \quad d\phi &= \frac{\partial \phi}{\partial x}dx + \frac{\partial \phi}{\partial y}dy + \frac{\partial \phi}{\partial z}dz = 0 \end{aligned}\right\} \tag{3.10}$$

Now multiplying $d\phi$ equation by λ and adding it to the du equation, we get

$$\left(\frac{\partial u}{\partial x} + \lambda\frac{\partial \phi}{\partial x}\right)dx + \left(\frac{\partial u}{\partial y} + \lambda\frac{\partial \phi}{\partial y}\right)dy + \left(\frac{\partial u}{\partial z} + \lambda\frac{\partial \phi}{\partial z}\right)dz = 0 \tag{3.11}$$

Since x, y, z are related by ϕ = constant, there are two independent values in this problem when x and y are indepenent, z is determined from the ϕ equation. Similarly, dx and dy may have any values we chose, then dz is determined from $d\phi$ equation.

We chose λ so that

$$\frac{\partial u}{\partial z} + \lambda\frac{\partial \phi}{\partial z} = 0 \tag{3.12}$$

Then from (3.11), for $dy = 0$, we get

$$\frac{\partial u}{\partial x} + \lambda\frac{\partial \phi}{\partial x} = 0 \tag{3.13}$$

and for dx = 0; $\dfrac{\partial u}{\partial y} + \lambda\dfrac{\partial \phi}{\partial y} = 0$ (3.14)

Solving equations (3.12), (3.13), (3.14)

and $\phi(x, y, z)$ = constant, we can find out x, y, z and λ.

3.5.3 : Method of Lagrange's undetermined multiplier with two ϕ equations

Suppose we have two constraint equations : $\phi_1(x, y, z, w)$ = constant (3.15)

And $\phi_2(x, y, z, w)$ = constant (3.16)

And our function is now $u(x, y, z, w)$ (3.17)

There are two independent variables, say x & y.

Therefore, $d\phi_1 = \dfrac{\partial \phi_1}{\partial x}dx + \dfrac{\partial \phi_1}{\partial y}dy + \dfrac{\partial \phi_1}{\partial z}dz = 0$ (3.18)

And $$d\phi_2 = \frac{\partial \phi_2}{\partial x}dx + \frac{\partial \phi_2}{\partial y}dy + \frac{\partial \phi_2}{\partial z}dz = 0 \tag{3.19}$$

And we set $$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial z}dz + \frac{\partial u}{\partial w}dw = 0 \tag{3.20}$$

Multiplying equation (3.18) by λ_1 and equation (3.19) by λ_2 and adding the result to equation (3.20), we get

$$\left(\frac{\partial u}{\partial x} + \lambda_1 \frac{\partial \phi_1}{\partial x} + \lambda_2 \frac{\partial \phi_2}{\partial x}\right)dx + \left(\frac{\partial u}{\partial y} + \lambda_1 \frac{\partial \phi_1}{\partial y} + \lambda_2 \frac{\partial \phi_2}{\partial y}\right)dy$$

$$+\left(\frac{\partial u}{\partial z} + \lambda_1 \frac{\partial \phi_1}{\partial z} + \lambda_2 \frac{\partial \phi_2}{\partial z}\right)dz + \left(\frac{\partial u}{\partial w} + \lambda_1 \frac{\partial \phi_1}{\partial w} + \lambda_2 \frac{\partial \phi_2}{\partial w}\right)dw = 0 \tag{3.21}$$

We determine λ_1 and λ_2 from the equations,

$$\left.\begin{aligned} \frac{\partial u}{\partial z} + \lambda_1 \frac{\partial \phi_1}{\partial z} + \lambda_2 \frac{\partial \phi_2}{\partial z} = 0 \\ \frac{\partial u}{\partial w} + \lambda_1 \frac{\partial \phi_1}{\partial w} + \lambda_2 \frac{\partial \phi_2}{\partial w} = 0 \end{aligned}\right\} \tag{3.22}$$

Then for $dy = 0$, we have

$$\left(\frac{\partial u}{\partial x} + \lambda_1 \frac{\partial \phi_1}{\partial x} + \lambda_2 \frac{\partial \phi_2}{\partial x}\right) = 0 \tag{3.23}$$

And for $dx = 0$, we have,

$$\left(\frac{\partial u}{\partial y} + \lambda_1 \frac{\partial \phi_1}{\partial y} + \lambda_2 \frac{\partial \phi_2}{\partial y}\right) = 0 \tag{3.24}$$

Now solving equations (3.22), (3.23), (3.24) and (3.15), (3.14) we get x, y, z, w, λ_1 & λ_2.

3.5.4 : Working rules for constraint maximization or minimization using Lagrange's multipliers

Rule I : To find the maximum and minimum values of $u(x, y, z)$ if $\phi(x, y, z)$ = *constant*, we form the function $U = u + \lambda\phi$ and set the three partial derivative of U equal to zero. We solve these equations and the equation ϕ = *constant*, for x, y, z and λ.

Rule II : To find maximum or minimum of u subject to the conditions ϕ_1 = *constant* and ϕ_2 = *constant*, we define $U = u + \lambda_1\phi_1 + \lambda_2\phi_2$ and set each of the partial derivatives of U equal to zero. Solve these equation and the ϕ equations for the variables and λ's.

Rule III : For a problem with still more variables and conditions there are more equations but no change in method.

Example of Art 3.6 :

Example I : Using the method of Lagrange's multiplier, find the area of largest rectangle that can be inscribed in a semi-circle of radius R with one of the largest side of the rectangle coinciding with diameter.

Solution : The rectangle to be inscribed in a circle should be symmetric about y-axis. When length and breadth of the rectangle is $2x$ and y, its area is $2xy$. Also $x^2 + y^2 = R^2$. Therefore $U = 2xy + \lambda(x^2 + y^2 - R^2)$

$$\therefore\ \frac{\partial U}{\partial x} = 2y + 2\lambda x$$

$$\frac{\partial U}{\partial y} = 2x + 2\lambda y$$

For $\quad \dfrac{\partial U}{\partial x} = 0,\ 2y + 2\lambda x = 0$ (i)

And $\quad \dfrac{\partial U}{\partial y} = 0,\ 2x + 2\lambda y = 0$ (ii)

From equation (i) and (ii), we get,

$$-\frac{y}{x} = \lambda = -\frac{x}{y}$$

or, $\quad x^2 = y^2 \quad$ i.e. $x = y$ (iii)

$$\therefore R^2 = 2x^2 = 2y^2 \quad \therefore x = y = \frac{R}{\sqrt{2}}$$

$$\therefore\ \text{Area} = 2xy = 2\frac{R}{\sqrt{2}}\cdot\frac{R}{\sqrt{2}} = R^2$$

$\therefore$ Maximum area = R^2

Example 2. Show that the rectangular solid of maximum volume that can be inscribed in a sphere is a cube.

Solution : Let inscribable maximum rectangular solid has got length, breadth and height $2x$, $2y$, and $2z$ respectively.

$$\therefore \quad \text{volume of solid, } V = 8\,xyz \qquad \text{(i)}$$

Also, $x^2 + y^2 + z^2 = r^2$, where r is the radius of the sphere

$$\therefore \ \phi(x, y, z) = x^2 + y^2 + z^2 - r^2 \qquad \text{(ii)}$$

$$\therefore \ U = V + \lambda\phi \qquad \text{(iii)}$$

Now, $$\frac{\partial U}{\partial x} = \frac{\partial V}{\partial x} + \lambda\frac{\partial \phi}{\partial x} = 0;$$

or, $$8yz + \lambda(2x) = 0 \qquad \text{(iv)}$$

$$\frac{\partial U}{\partial y} = \frac{\partial V}{\partial y} + \lambda\frac{\partial \phi}{\partial y} = 0;$$

or, $$8xz + \lambda(2y) = 0 \qquad \text{(v)}$$

$$\frac{\partial U}{\partial z} = \frac{\partial V}{\partial z} + \lambda\frac{\partial \phi}{\partial z} = 0;$$

or, $$8xy + \lambda(2z) = 0 \qquad \text{(vi)}$$

From (iv), $2\lambda x^2 = -8xyz$

From (v), $2\lambda y^2 = -8xyz$

From (vi), $2\lambda z^2 = -8xyz$

$$\therefore 2\lambda x^2 = 2\lambda y^2 = 2\lambda z^2\ ;\ i.e.\ x^2 = y^2 = z^2$$

$$\therefore \ x = y = z$$

Hence in a sphere, the rectangular solid having maximum volume that can be inscribed within it is a cube.

Example 3 : Using the method of Lagrange multiplier find the maximum of $F = 4xyz$ subject to the constraint $x^2 + y^2 + z^2 - a^2 = 0$

Solution : Let $\phi(x, y, z) = x^2 + y^2 + z^2 - a^2$ (i)

We consider $U = F(x, y, z) + \lambda\phi(x, y, z) = 4xyz + \lambda\phi(x, y, z)$ (ii)

$$\therefore \frac{\partial U}{\partial x} = 4yz + 2\lambda x = 0 \quad \text{(iii)}$$

$$\frac{\partial U}{\partial y} = 4yz + 2\lambda y = 0 \quad \text{(iv)}$$

$$\frac{\partial U}{\partial z} = 4yz + 2\lambda z = 0 \quad \text{(v)}$$

From equations (iii) and (iv)

$$\lambda = -\frac{2yz}{x} = -\frac{2zx}{y} \text{ i.e. } y^2 = x^2 \text{ or } y = x$$

Similarly, from equation (iv) and (v), we get $x = z$

$$\therefore x = y = z \quad \text{(vi)}$$

From equation (i), $x = y = z = \pm\dfrac{a}{\sqrt{3}}$

Then $F = 4xyz$ will lie within $-\dfrac{4a}{3\sqrt{3}}$ to $\dfrac{4a}{3\sqrt{3}}$. So, the maximum value of F will be $\dfrac{4a}{3\sqrt{3}}$

Exercise of Art 3.5 :

1) The temperature T at any point (x, y, z) in space is $T = 400\, xyz^2$. Find the highest temperature at the surface of a unit sphere $x^2 + y^2 + z^2 = 1$.

2) Find the volume of the greatest rectangular parallelepiped that can be inscribed in the ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$

3) Find the point on the plane $ax + by + cz = p$ at which the function $f = x^2 + y^2 + z^2$ has a minimum value and find the minimum f.

Solution of exercise of Art 3.5 :

Solution to the problem 1 :

We have $T = 400\ xyz^2$

$$\phi(x, y, z) = x^2 + y^2 + z^2 - 1$$

Therefore $U = T + \lambda\phi$ (i)

and $$\frac{\partial U}{\partial x} = \frac{\partial T}{\partial x} + \lambda\frac{\partial \phi}{\partial x} = 400yz^2 + \lambda(2x) = 0 \quad \text{(ii)}$$

$$\frac{\partial U}{\partial y} = \frac{\partial T}{\partial y} + \lambda\frac{\partial \phi}{\partial y} = 400xz^2 + \lambda(2y) = 0 \quad \text{(iii)}$$

$$\frac{\partial U}{\partial z} = \frac{\partial T}{\partial z} + \lambda\frac{\partial \phi}{\partial z} = 800xyz + \lambda(2z) = 0 \quad \text{(iv)}$$

Now multiplying equation (ii) (iii) and (iv) by x, y & z respectively and adding we get,

$$1600xyz^2 + 2\lambda(x^2 + y^2 + z^2) = 0$$

Or, $\lambda = -800xyz^2$, since $x^2 + y^2 + z^2 = 1$

Putting the value of λ in equation (ii) we get

$400yz^2 + 2x(-800\ xyz^2) = 0$; or, $1 - 4x^2 = 0$; or, $x = \pm\frac{1}{2}$

Similarly $y = \pm\frac{1}{2}$

Putting the value of λ in equation (iv), we get

$$800xyz - 1600xyz^3 = 0$$

or, $1 - 2z^2 = 0$; or, $z = \pm\frac{1}{2}$

now using the values of x, y, z in T, we get $T = 400 \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = 50$

Solution to problem 2 : We take, $2x$, $2y$ and $2z$ as the edges of the parallelopiped whose edges are parallel to the x, y, z – axes respectively. Therefore the volume of the parallelepiped is $v = 8xyz$.

Let $\phi(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$

Therefore following Lagrange's method of undetermined multiplier, we have

$$U(x, y, z) = V + \lambda\phi(x, y, z) \quad \text{(i)}$$

or, $$U(x, y, z) = 8xyz + \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right)$$

$$\therefore \frac{\partial U}{\partial x} = 8yz + \lambda\left(\frac{2x}{a^2}\right) = 0 \quad \text{(ii)}$$

$$\frac{\partial U}{\partial y} = 8xz + \lambda\left(\frac{2y}{b^2}\right) = 0 \quad \text{(iii)}$$

and $$\frac{\partial U}{\partial z} = 8xy + \lambda\left(\frac{2z}{c^2}\right) = 0 \quad \text{(iv)}$$

From equation (ii); $\lambda = -\dfrac{8yz}{\frac{2x}{a^2}}$ and from equation (iii), $\lambda = -\dfrac{8xz}{\frac{2z}{c^2}}$

$$\therefore \frac{8yza^2}{2x} = \frac{8xzc^2}{2z} \quad or \quad \frac{x^2}{a^2} = \frac{y^2}{b^2}$$

Similarly, from (iii) and (iv), we get,

$$\frac{y^2}{b^2} = \frac{z^2}{c^2}$$

$$\therefore \frac{x^2}{a^2} = \frac{y^2}{b^2} = \frac{z^2}{c^2}$$

Again we have $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$; or, $\dfrac{x^2}{a^2} = \dfrac{1}{3}$

$$\therefore x = \frac{a}{\sqrt{3}}, y = \frac{b}{\sqrt{3}}, \text{ and } z = \frac{c}{\sqrt{3}}$$

Thus the volume of the largest rectangular parallelepiped is $V = 8xyz = 8\frac{a}{\sqrt{3}}\frac{b}{\sqrt{3}}\frac{c}{\sqrt{3}} = \frac{8abc}{3\sqrt{3}}$.

Solution to problem (3) :

We have $U(x, y, z) = f(x, y, z) + \lambda\phi(x, y, z)$ (i)

Where $f(x, y, z) = x^2 + y^2 + z^2$ and $\phi(x, y, z) = ax + by + cz - p$

Therefore from (i), differentiating partially,

$$\frac{\partial U}{\partial x} = \frac{\partial f}{\partial x} + \lambda\frac{\partial \phi}{\partial x} = 2x + \lambda a = 0 \text{ or, } x = -\frac{\lambda a}{2}$$

$$\frac{\partial U}{\partial y} = \frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y} = 2y + \lambda b = 0 \text{ or, } y = -\frac{\lambda b}{2}$$

$$\frac{\partial U}{\partial z} = \frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial z} = 2y + \lambda c = 0 \text{ or, } z = -\frac{\lambda c}{2}$$

Substituting the values of x, y, z in equation $ax + by + cz = p$, we get

$$a\left(-\frac{\lambda a}{2}\right) + b\left(-\frac{\lambda b}{2}\right) + c\left(-\frac{\lambda c}{2}\right) = p$$

Or, $$\lambda\left(a^2 + b^2 + c^2\right) = -2p; \text{ or, } \lambda = \frac{-2p}{a^2 + b^2 + c^2}$$

$$\therefore x = \frac{ap}{a^2 + b^2 + c^2},\ y = \frac{bp}{a^2 + b^2 + c^2},\ z = \frac{cp}{a^2 + b^2 + c^2}$$

minimum values of $$f = \frac{a^2p^2}{\left(a^2 + b^2 + c^2\right)^2} + \frac{b^2p^2}{\left(a^2 + b^2 + c^2\right)^2} + \frac{c^2p^2}{\left(a^2 + b^2 + c^2\right)^2}$$

or, $$f = \frac{p^2\left(a^2 + b^2 + c^2\right)}{\left(a^2 + b^2 + c^2\right)^2}$$

or, $$f = \frac{p^2}{a^2 + b^2 + c^2} = f_{minimum}$$

Note : Lagrange' method cannot determine the nature of the stationary points. However it is ascertained from the condition of the problem.

3.6 ❒ Summary

1. Emphasis is given on error calculation in physical measurements with the help partial and total derivatives.
2. Exact and inexact differential are defined and rules are stated to make inexact differentials, exact.
3. Constrained maximization with Lagrange's multipliers discussed.

Unit 4 ❒ Vector Calculus

Structure

4.1 ❐ Objectives

When you go through the article you may be able to learn

1. Definitions of scalars and vectors
2. Vector algebra, which is a little bit different from scalar algebra
3. Some examples of application of vectors in different branches of physics.

4.2 ❐ Introduction

One may ask why we need to study scalars, vectors or in general Tensors ? One of the simple reasons may be that physical laws can be expressed effectively in concise form and without any ambiguity with the help of scalars, vectors and tensors. But this is not all. More logical reason for using scalars, vectors and tensors lies in the fact that physical laws must obey the principle of Galilean Invariance (in non-relativistic domain) which states that physical phenomena appear to be the same for all observers moving in inertial frames with constant relative velocity with respect to each other in respect of translation and rotation of the co-ordinate system. In other words physical laws must be invariant in all inertial frames of references.

In view of the above mathematical formulation of the physical laws must contain those entities which have such invariance properties and these entities are scalars, vector and Tensors. This is why a student of physics and science in general, must learn the properties of scalars, vectors and tensors.

4.3 ❒ Vectors and Scalars

4.3.1 Familiarities with Vectors and Scalars

Measurable physical entities which have both magnitude and direction and obey parallelogram law of addition are called Vectors. This is geometrical or graphical representation of vectors. On the other hand physical quantities which have magnitude only are called scalars. Both the vectors and scalars have their respective units.

Vectors can also be defined as a set of three numbers (in three dimensional spaces) which we call its components with respect to a co-ordinate system in vector space. This is algebraic definition of vector.

A physical scalar is a quantity which remains invariant under all co-ordinate systems.

4.3.2 Example of graphical representation

Graphically or geometrically a vector is represented by a line with an arrow head. The length of the line is its magnitude and the arrow points towards its direction. Beginning of the line is termed as origin or tail and the arrow head is called terminus.

$$\overrightarrow{\qquad\qquad\qquad\qquad}$$
O **P** A

Vector **P** is represented by the Line $\overrightarrow{OA}$ with an arrow, O is its tail and A is terminus. The modulous or magnitude of **P** = | **P** | = **P** is given by the length OA. This representation is independent of the origin of any co-ordinate system. Vectors are represented by **bold face** letters and their magnitudes by ordinary letters.

4.3.3 Vectors in terms of components

We consider a rectangular co-ordinate system as in fig (4A.1). Let the vector ***i*** be a unit vector in the positive *x* direction and let **j** and **k** be unit vectors in the positive *y* and *z* directions. If A_x and A_y are the scalar components of a vector in the (*x*, *y*) plane, $\mathbf{i}A_x$ and $\mathbf{j}A_y$ are its vector components and their sum is the vector **A** (fig. 4A.2).

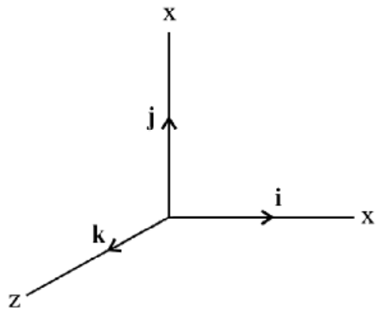
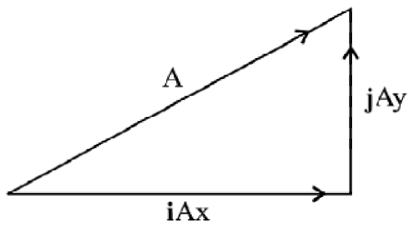


Fig. (4A.1)



(Fig. 4A.2)

Similarly, in three dimension, $\mathbf{A} = \mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z$

However these two ways of representing a vector are not completely equivalent, for the algebraic definition requires a co-ordinate system, but the geometrical representation does not require any co-ordinate system.

This difficulty is removed by making the algebraic representation also independent of any particularly co-ordinate system by defining a vector in the following way.

A vector in three dimensions is a set of three numbers, called its components, which transform under a rotation of co-ordinate system according to the following transformation equation.

$$Y_i = b_{i1}x_1 + b_{i2}x_2 + b_{i3}x_3$$

$$= \sum_{k=1}^{3} b_{ik}x_k \qquad (4A.1)$$

Y_i are the components of ***x*** in the new co-ordinate system and x_k are the component of ***x*** in the old co-ordinate system. The co-efficient b_{ik} etc are the numbers which are determined by the given co-ordinate rotation, they do not depend on ***x***.

We note that translation have no effect on the components of vector which are numbers but not scalars, because they do not remain invariant under rotation of co-ordinate system. However exception to the definition of vector given in equation (4A.1) is position vector ***r*** = ***i****x* + ***j****y* + ***k****z* which is defined with respect to specific origin.

Tensors are quantities that do not have any specified directions but have different values in different directions. Examples are moment of inertia tensor, dielectric susceptibility tensor etc. Tensors however are defined only through their transformation under changes of co-ordinate system.

A physical entity which has only one component is called tensor of zero rank or a scalar. If it has more than one component but less than or equal to four, it is called a vector or a tensor of rank 1. A tensor of rank 2 has nine components.

4.3.4 Examples of scalars, vector and tensors

Scalar : A scalar field is created by simply assigning scalar quantities (numbers) to each point in space.

Temperature of a body or potential of gravitational or electrostatic field are examples of scalar fields. Mass, volume, density, length etc. are scalar quantities.

Vector : A vector field is created by assigning vectors to each point in space. An electrostatic field, a gravitational field, electromagnetic field are examples.

Vectors usually possess both magnitude and direction. Force, momentum, electric dipole moment, magnetic dipole moment etc. are examples of vectors.

Tensor. A tensor cannot be visualised geometrically hence it is defined in terms of field or transformation properties under rotation of co-ordinate system. A tensor field has a tensor corresponding to each point space. An example is the stress on a material. Other examples of tensors include the strain tensor, the conductivity tensor and the inertia tensor.

4.3.5 Equal vectors and null vectors

Two vectors are said to be equal when their magnitudes as well as direction are identical i.e. $\boldsymbol{A} = \boldsymbol{B}$ i.e. $\boldsymbol{A} - \boldsymbol{B} = 0$. The right hand side of this vector equation is also a vector called null vector with arbitrary direction.

4.3.6 Unit Vectors

A vector having unit scalar magnitude is defined as unit vector. Any vector $\boldsymbol{A}$ can be written as $A = A\boldsymbol{a}$ (4A.2), where $\boldsymbol{a} = \dfrac{\boldsymbol{A}}{A}$ is a unit vector in the direction of the vector $\boldsymbol{A}$.

In the Cartesian co-ordinate system, unit vectors $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ in the direction of X, Y, Z-axes respectively are commonly used.

4.3.7 Position vector or radius vector

The position vector $\boldsymbol{r}$ of an object located at $P(x, y, z)$ is given by $\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$ (4A.3). Now

$$OP^2 = OB^2 + BP^2 = OA^2 + AB^2 + BP^2 = x^2 + y^2 + z^2$$

$$\therefore\ OP = r = |\boldsymbol{r}| = \sqrt{x^2 + y^2 + z^2} \qquad (4A.4)$$

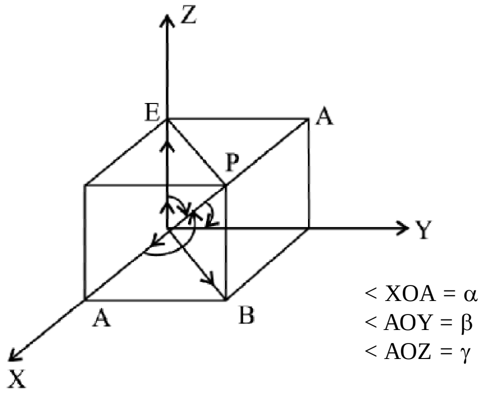


Fig. (4A.3)

In equation (4A.3) $\boldsymbol{i}x, \boldsymbol{j}y, \boldsymbol{k}z$ are the components of the vector $\boldsymbol{r}$ in X,Y,Z direction respectively.

If the vector $\boldsymbol{r}$ makes an angle α, β, γ respectively with X,Y,Z -axes,

then $x = r\cos\alpha,\ y = r\cos\beta,\ z = r\cos\gamma$

therefore, $x^2 + y^2 + z^2 = r^2\left(\cos^2\alpha + \cos^2\beta + \cos^2\gamma\right)$

or comparing with equation (4A.4), we get $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$ (4A.5)

equation (4A.5) is known as vector direction cosine law and $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are called direction cosine of $\boldsymbol{r}$ with respect to x, y, z axes respectively. Sometimes we represent cos α by l, cos β by m, cos γ by n, so that

$$l^2 + m^2 + n^2 = 1 \qquad (4A.6)$$

So any vector $\boldsymbol{A}$ with components $\boldsymbol{i}A_x$, $\boldsymbol{j}A_y$ and $\boldsymbol{k}A_z$ along X, Y and Z directions respectively can be written as

$$\boldsymbol{A} = \boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z$$

$$= \boldsymbol{A}_x + \boldsymbol{A}_y + \boldsymbol{A}_z \tag{4A.7}$$

If the vector $\boldsymbol{A}$ makes an angle α, β, γ with X, Y and Z-axes respectively,

$$A_x = A\cos\alpha,\ A_y = A\cos\beta,\ A_z = A\cos\gamma$$

$$A^2 = A_x{}^2 + A_y{}^2 + A_z{}^2 = |\boldsymbol{A}|^2 \tag{4A.8}$$

4.3.8 Addition of vectors [Graphical representation]

The sum of vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ is a vector $\boldsymbol{C}$ by placing the origin of $\boldsymbol{B}$ on the terminus of $\boldsymbol{A}$ and joining the initial point of $\boldsymbol{A}$ to the terminus of $\boldsymbol{B}$.

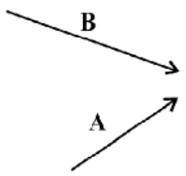


Fig. 4A.4(a)

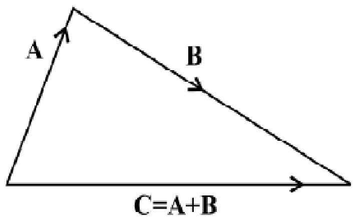


Fig. 4A.4(b)

We write $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{C}$. This definition is equivalent to the parallelogram law for vector addition as indicated in fig 4A.4(c)

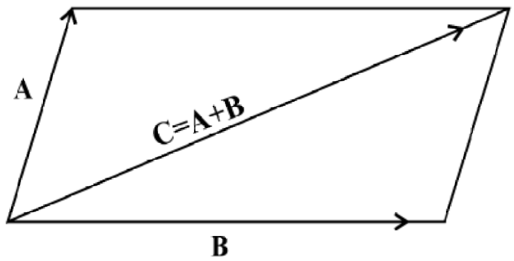


Fig. 4A.4(c)

The law of vector addition therefore is the parallelogram law of addition which states that the sum of two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ is given in magnitude and direction by the diagonal of the parallelogram formed by the sides representing the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$.

In the same way any number of vectors can be added. Fig. 4A.5 shows how to obtain the sum of resultant $\boldsymbol{R}$ of the vectors $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$.

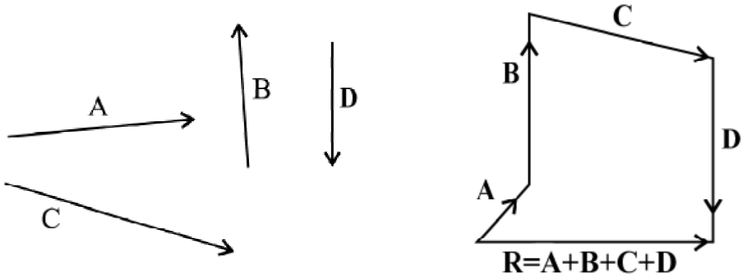


Fig. 4A.5

4.3.9 Subtraction of vectors

Subtraction of the vector ***B*** from the vector ***A*** is defined as the addition of the negative vector **–*B*** to ***A***. Thus ***A*** + (– ***B***) = ***A*** – ***B***.

4.3.10 Addition and subtraction of vector [algebraic or co-ordinate representation method]

To find sum or difference of two vectors, we add or subtract like components together as follows.

Let $\boldsymbol{A} = \boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z$, where $A = |\boldsymbol{A}| = \sqrt{A_x{}^2 + A_y{}^2 + A_z{}^2}$;

and $\boldsymbol{B} = \boldsymbol{i}B_x + \boldsymbol{j}B_y + \boldsymbol{k}B_z$, where $B = |\boldsymbol{B}| = \sqrt{B_x{}^2 + B_y{}^2 + B_z{}^2}$

Therefore $\boldsymbol{A} \pm \boldsymbol{B} = \left(A_x + B_x\right)\boldsymbol{i} + \left(A_y + B_y\right)\boldsymbol{j} + \left(A_z + B_z\right)\boldsymbol{k}$

and $|\boldsymbol{A} + \boldsymbol{B}| = \left[\left(A_x \pm B_x\right)^2 + \left(A_y \pm B_y\right)^2 + \left(A_z \pm B_z\right)^2\right]^{1/2}$

Likewise for the sum and the difference of a large number of vectors.

4.3.11 Multiplication of vectors by scalars

If ***A*** is a vector and *m* is any positive real number, and then *m**A*** is defined to be a vector having magnitude equal to *m* times that of the given vector ***A*** in the same direction.

Likewise – *m**A*** is a vector in the direction opposite to that of ***A*** and having magnitude equal to *m* times that of ***A***.

4.4 ❒ Laws of Vector Algebra

Vector addition is commutative and associative i.e. $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$

And $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$, respectively also multiplication of vectors by scalars is commutative, associative and distributive i.e.

$$m\boldsymbol{A} = \boldsymbol{A}m$$

$$m\,(n\boldsymbol{A}) = n\,(m\boldsymbol{A})$$

$$m\,(\boldsymbol{A} + \boldsymbol{B}) = m\boldsymbol{A} + m\boldsymbol{B}, \text{ respectively.}$$

m and n are two different scalars.

4.4.1 Linear dependence of vectors

Let $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3 ...$ be vectors and $\alpha_1, \alpha_2, \alpha_3, ...$ are scalars, not all of which are zero. If there exists a relation of the type $\sum_i \alpha_i \boldsymbol{A}_i = \alpha_1\boldsymbol{A}_1 + \alpha_2\boldsymbol{A}_2 + \alpha_3\boldsymbol{A}_3 + ... = 0$ (4A.9)

Then the system of the vectors $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3 ...$ is said to be linearly dependent.

If the system of vectors $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3$ are not linearly dependent, then $\alpha_1, \alpha_2, \alpha_3, ...$ are all zero i.e.

$$\alpha_1 = \alpha_2 = \alpha_3 = ... = 0 \qquad (4A.10)$$

The system of vectors in this case is said to be linearly independent.

If $\boldsymbol{r} = \alpha_1\boldsymbol{A}_1 + \alpha_2\boldsymbol{A}_2 + \alpha_3\boldsymbol{A}_3 + ...$ i.e. $-\boldsymbol{r} + \alpha_1\boldsymbol{A}_1 + \alpha_2\boldsymbol{A}_2 + \alpha_3\boldsymbol{A}_3 + ... = 0$, then system of vectors $\boldsymbol{r}, \boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3, ...$ is linearly dependent.

Any vector $\boldsymbol{r}$ coplanar with any two non-collinear vectors $\boldsymbol{A}$ & $\boldsymbol{B}$ can be uniquely expressed as a linear combination of the given vectors i.e. $\boldsymbol{r} = \alpha\boldsymbol{A} + \beta\boldsymbol{B}$ (4A.11)

Where α, β are scalars.

From equation (4A.11) we can write $-\ \boldsymbol{r} = \alpha\boldsymbol{A} + \beta\boldsymbol{B} = 0$ i.e. $\boldsymbol{r}, \boldsymbol{A}, \boldsymbol{B}$ vectors are linearly dependent. It is to be noted that necessary and sufficient condition that three vectors be linearly dependent is that they may be coplanar.

Example of Art 4.3.7 to 4.4.1 :

Example 1 : Position vectors of three points P, Q, R and $2\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}$, $\boldsymbol{i} - 3\boldsymbol{j} - 5\boldsymbol{k}$ and $3\boldsymbol{i} - 4\boldsymbol{j} - 4\boldsymbol{k}$ respectively. Find the vectors $\overrightarrow{PQ}$ and $\overrightarrow{QR}$ and their magnitudes.

Solutions : $\overrightarrow{PQ} = (\boldsymbol{i} - 3\boldsymbol{j} - 5\boldsymbol{k}) - (2\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}) = -\boldsymbol{i} - 2\boldsymbol{j} - 6\boldsymbol{k}$

$$\therefore \left|\overrightarrow{PQ}\right| = PQ = \sqrt{(-1)^2 + (-2)^2 + (-6)^2} = \sqrt{41}$$

$$\overrightarrow{QR} = (3\boldsymbol{i} - 4\boldsymbol{j} - 4\boldsymbol{k}) - (-\boldsymbol{i} - 3\boldsymbol{j} - 5\boldsymbol{k}) = 2\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}$$

$$\therefore\ QR = \left|\overrightarrow{QR}\right| = \sqrt{(2)^2 + (-1)^2 + (1)^2} = \sqrt{6}$$

Example 2 : Prove that, the line joining the mid points of two sides of a triangle is parallel and half to the third.

Solution : In the triangle ΔPQR, the position vectors of P, Q, R be $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ respectively. If the mid points of PQ and PR be D & E respectively, then the position vectors of $D = \dfrac{\boldsymbol{a}+\boldsymbol{b}}{2}$ and the position vector of E is $\dfrac{\boldsymbol{a}+\boldsymbol{c}}{2}$.

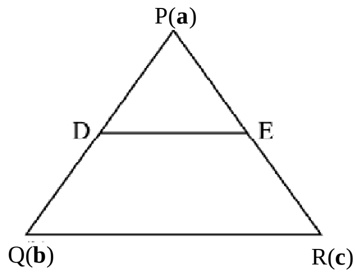


Fig. Example (2)

$$\therefore \overrightarrow{DE} = (\textit{position vector of } E) - (\textit{position vector of } D)$$

$$= \frac{\boldsymbol{a}+\boldsymbol{c}}{2} - \frac{\boldsymbol{a}+\boldsymbol{b}}{2} = \frac{(\boldsymbol{c}-\boldsymbol{b})}{2} = \frac{1}{2}\overrightarrow{BC}$$

Hence $\overrightarrow{DE}$ is parallel to $\overrightarrow{BC}$ and half of BC.

Example 3 : Prove that (i) if $\boldsymbol{P}$ and $\boldsymbol{Q}$ are two non-collinear vectors and $\alpha\boldsymbol{P} + \beta\boldsymbol{Q} = 0$, then show that, $\alpha = \beta = 0$

(ii) If $\boldsymbol{P}$, $\boldsymbol{Q}$, $\boldsymbol{R}$ are non-coplanar vector and $\alpha\boldsymbol{P} + \beta\boldsymbol{Q} + \gamma\boldsymbol{R} = 0$, then show that $\alpha = \beta = \gamma = 0$

Solution : (i) suppose, $\alpha \neq 0$

$\therefore \boldsymbol{P} = \dfrac{-\beta}{a}\boldsymbol{Q} = m\boldsymbol{Q}$ where m is a scalar.

Therefore $\boldsymbol{P}$ and $\boldsymbol{Q}$ collinear if $\alpha \neq 0$.

Again, $\beta \neq 0$; $\therefore \boldsymbol{Q} = \dfrac{-\alpha}{\beta}\boldsymbol{P}$

Or, $\boldsymbol{Q} = n\boldsymbol{P}$ where n is scalar.

Therefore again $\boldsymbol{Q}$ and $\boldsymbol{P}$ are collinear when $\beta \neq 0$.

Therefore we see that $\boldsymbol{P}$ and $\boldsymbol{Q}$ are non-collinear when $\alpha = \beta = 0$

(ii) Let $\alpha \neq 0$, $\therefore \boldsymbol{P} = \frac{-\beta}{\alpha}\boldsymbol{Q} - \frac{\gamma}{\alpha}\boldsymbol{R}$

Or, $\boldsymbol{P} = m\boldsymbol{Q} + n\boldsymbol{R}$ where m and n are two scalars. Therefore $\boldsymbol{P}$, $\boldsymbol{Q}$, $\boldsymbol{R}$ are coplanar when $\alpha \neq 0$. Similarly we can prove that $\boldsymbol{P}$, $\boldsymbol{Q}$, $\boldsymbol{R}$ are coplanar when $\beta \neq 0$ and $\gamma \neq 0$.

Therefore for $\alpha = \beta = \gamma = 0$, $\boldsymbol{P}$, $\boldsymbol{Q}$ and $\boldsymbol{R}$ are non-coplanar.

Example 4 : If $\boldsymbol{A} = 2\boldsymbol{i} + 3\boldsymbol{j} - \boldsymbol{k}$ and $\boldsymbol{B} = 3\boldsymbol{i} - \boldsymbol{j} + 5\boldsymbol{k}$. Find the value of $\boldsymbol{A} \pm \boldsymbol{B}$

Solution : We have

$$\boldsymbol{A} = A_x\boldsymbol{i} + A_y\boldsymbol{j} + A_z\boldsymbol{k} = 2\boldsymbol{i} + 3\boldsymbol{j} - \boldsymbol{k}$$

$$\boldsymbol{B} = B_x\boldsymbol{i} + B_y\boldsymbol{j} + B_z\boldsymbol{k} = 3\boldsymbol{i} - \boldsymbol{j} + 5\boldsymbol{k}$$

$$\therefore \quad \boldsymbol{A} + \boldsymbol{B} = (A_x + B_x)\boldsymbol{i} + (A_y + B_y)\boldsymbol{j} + (A_z + B_z)\boldsymbol{k}$$

$$= (2 + 3)\boldsymbol{i} + (3 - 1)\boldsymbol{j} + (-1 + 5)\boldsymbol{k}$$

$$= 5\boldsymbol{i} + 2\boldsymbol{j} + 4\boldsymbol{k}$$

Similarly $\boldsymbol{A} - \boldsymbol{B} = -\boldsymbol{i} + 4\boldsymbol{j} - 6\boldsymbol{k}$

4.4.2 Product of Vectors

We often come across in physics, certain combination of vectors which have the properties of products. The products of two vectors may be a scalar or a vector quantity depending upon how the product is defined.

For example work done is the scalar product of two vectors namely force ($\boldsymbol{F}$) and displacement ($\boldsymbol{d}$), whereas angular momentum of a particle about the origin is the vector product of position vector ($\boldsymbol{r}$) and its linear momentum (mv). Work done being a scalar quantity but angular momentum is vector.

4.4.3 Scalar product of two vector :

We define scalar product or dot product of two vector as follows :

$$\boldsymbol{A}.\boldsymbol{B} = AB \cos\theta \tag{4A.12}$$

Where $|\boldsymbol{A}| = A$ and $|\boldsymbol{B}| = B$ and θ is the acute angle between $\boldsymbol{A}$ and $\boldsymbol{B}$, clearly, the scalar product is the product of the magnitude of one vector and the projection of the other on it.

From the fig (4A.6) B cos θ is the length of the resolved part of $\overrightarrow{OC}$ i.e OD along $\boldsymbol{A}$, i.e. $\boldsymbol{A}.\boldsymbol{B} = A(B \cos \theta)$ = A (resolved part of modulous of $\boldsymbol{B}$ along $\boldsymbol{A}$).

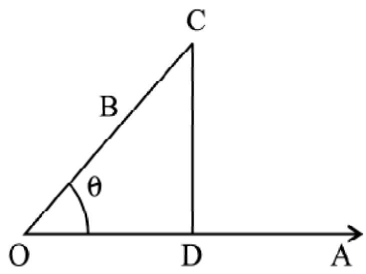


Fig. 4A.6

Similarly we can write

$\boldsymbol{A}.\boldsymbol{B} = B(A \cos \theta) = B$(resolved part of modulus $\boldsymbol{A}$ along the direction $\boldsymbol{B}$). From equation (4A.12) it is clear that if $\boldsymbol{A}$ and $\boldsymbol{B}$ are two non-zero vectors, then their dot product will be zero only when the direction of the vectors are perpendicular. The dot product will be equal to the product of their moduli, when their directions are parallel.

For the unit vectors $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ along the rectangular co-ordinate system,

$$\boldsymbol{i}.\boldsymbol{i} = \boldsymbol{j}.\boldsymbol{j} = \boldsymbol{k}.\boldsymbol{k} = 1 \tag{4A.13}$$

and $$\boldsymbol{i}.\boldsymbol{j} = \boldsymbol{j}.\boldsymbol{k} = \boldsymbol{k}.\boldsymbol{i} = 0 \tag{4A.14}$$

Hence the scalar or dot product of two vectors :

$$\boldsymbol{A} = \boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z \tag{4A.15}$$

$$\boldsymbol{B} = \boldsymbol{i}B_x + \boldsymbol{j}B_y + \boldsymbol{k}B_z \tag{4A.16}$$

will be written as $$\boldsymbol{A}.\boldsymbol{B} = A_xB_x + A_yB_y + A_zB_z \tag{4A.17}$$

scalar product of two vectors obey commutative law i.e. $\boldsymbol{A}.\boldsymbol{B} = \boldsymbol{B}.\boldsymbol{A}$

also scalar product of two vectors obey distributive law i.e. for three vector $\boldsymbol{A}$, $\boldsymbol{B}$ & $\boldsymbol{C}$ we can write

$$\boldsymbol{A}.(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A}.\boldsymbol{B} + \boldsymbol{A}.\boldsymbol{C} \tag{4A.18}$$

For two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, we can have the following relations

1. $(\boldsymbol{A} + \boldsymbol{B}).(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{A}.\boldsymbol{A} + \boldsymbol{B}.\boldsymbol{A} - \boldsymbol{A}.\boldsymbol{B}. - \boldsymbol{B}.\boldsymbol{B}. = \boldsymbol{A}^2 + \boldsymbol{B}^2$
2. $(\boldsymbol{A} + \boldsymbol{B})^2 = (\boldsymbol{A} + \boldsymbol{B}).(\boldsymbol{A} + \boldsymbol{B}) = \boldsymbol{A}.\boldsymbol{A}. + \boldsymbol{A}.\boldsymbol{B}. + \boldsymbol{B}.\boldsymbol{A}. + \boldsymbol{B}.\boldsymbol{B}. = \boldsymbol{A}^2 + \boldsymbol{B}^2 + 2\boldsymbol{A}.\boldsymbol{B}$

4.5 ❒ Vector Product

The vector product or cross product $\boldsymbol{C}$ of two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ is defined as

$$\boldsymbol{A} \times \boldsymbol{B} = AB \sin \theta \, \boldsymbol{n} = \boldsymbol{C} = C\boldsymbol{n} \tag{4A.19}$$

Where $A = |\boldsymbol{A}|$, $B = |\boldsymbol{B}|$ and $C = |\boldsymbol{C}|$

And θ is the acute angle between $\boldsymbol{A}$ & $\boldsymbol{B}$ when joined tail to tail, and $\boldsymbol{n}$ is a unit vector in the direction of $\boldsymbol{C}$. Direction of $\boldsymbol{C}$ is normal to the plane containing $\boldsymbol{A}$ & $\boldsymbol{B}$ and in a sense such that the vectors $\boldsymbol{A}$, $\boldsymbol{B}$ & $\boldsymbol{C}$ form a right handed system (fig 4A.7)

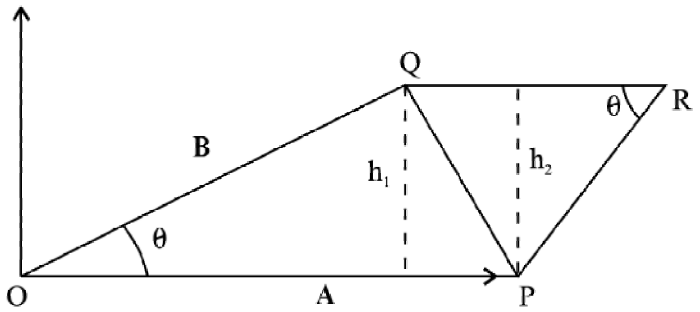


Fig. 4A.7

$\boldsymbol{A} = \overrightarrow{(OP)},\ \boldsymbol{B} = \overrightarrow{OQ}$

Geometrically the magnitude of the cross product of two vectors represents the area of the parallelogram having the two vectors as its side.

From the fig 4A.7, the area of the parallelogram $\square\, OPRQ$ = area of the triangle ΔQOP + area of the triangle ΔPQR

$$= \frac{1}{2}(OP)h_1 + \frac{1}{2}(QR)h_2$$

Now $h_1 = QO\sin\theta,\ h_2 = PR\sin\theta$

$$\square\, OPRQ = \frac{1}{2}(OP)(OQ)\sin\theta + \frac{1}{2}(QR)(PR)\sin\theta$$

$$= \frac{1}{2}|\boldsymbol{A}||\boldsymbol{B}|\sin\theta + \frac{1}{2}|\boldsymbol{A}||\boldsymbol{B}|\sin\theta$$

[Since $QR = OP = |\boldsymbol{A}|$ and $PR = OQ = |\boldsymbol{B}|$]

$$= |\boldsymbol{A}||\boldsymbol{B}|\sin\theta$$

$$\therefore \square\, OPRQ = |\boldsymbol{A}||\boldsymbol{B}|\sin\theta = |\boldsymbol{A}\times\boldsymbol{B}|$$

If $\boldsymbol{n}$ be a unit vector normal to the plane of the parallelogram then it gives = $|\boldsymbol{A} \times \boldsymbol{B}|\mathbf{n}$. This suggests that it may be useful to represent area by vectors. Following properties of vector products are easily verified.

1. $\boldsymbol{A} \times \boldsymbol{B} = -\, \boldsymbol{B} \times \boldsymbol{A}$
2. $\boldsymbol{A} \times [\boldsymbol{B} + \boldsymbol{C} + \boldsymbol{D} + \ldots] = \boldsymbol{A} \times \boldsymbol{B} + \boldsymbol{A} \times \boldsymbol{C} + \boldsymbol{A} \times \boldsymbol{D} + \ldots$
3. $\boldsymbol{A} \times \boldsymbol{A} = 0$
4. $\boldsymbol{i} \times \boldsymbol{j} = \boldsymbol{k}$; $\boldsymbol{j} \times \boldsymbol{k} = \boldsymbol{i}$; $\boldsymbol{k} \times \boldsymbol{i} = \boldsymbol{j}$
5. $\boldsymbol{i} \times \boldsymbol{i} = \boldsymbol{j} \times \boldsymbol{j} = \boldsymbol{k} \times \boldsymbol{k} = 0$
6. $\boldsymbol{A} \times \boldsymbol{B} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}$

4.5.1 Kronecker delta and Levicivita symbols

$$\delta_{ij} \begin{cases} = 1\,; if\ \ i = j \\ = 0\,; if\ \ i \neq j \end{cases}$$

Example : $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are unit vectors along Cartesian co-ordinate axes *X*, *Y* and *Z* respectively. We can define $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}$

The definition of the Levicivita symbol is $\varepsilon_{ijk} = 1$; if *i*, *j*, *k* = 1,2,3; 2,3,1; 3,1,2

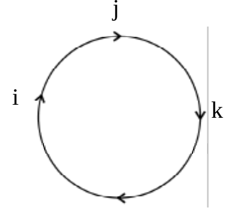
= – 1; if *i*, *j*, *k* = 3,2,1; 2,1,3; 1,3,2

= 0 if any indices are repeated

We say that ε_{ijk} is anti-symmetric with respect to every pair of indices, since each exchange of indices produce a change in sign.

If you read the indices *i*, *j*, *k* cyclically, then if the indices read in the direction 1, 2, 3; 1, 2,3; 1,... the result is +1; if the indices read in the opposite direction the result is – 1

i, *j*, *k*; *j*, *k*, *i* ; *k*, *i*, *j*



We now show that the components of the cross-product of two vectors can be written as $(\boldsymbol{B}\times\boldsymbol{C})_i = \varepsilon_{ijk}\boldsymbol{B}_j\boldsymbol{C}_k$

We have $\boldsymbol{B} \times \boldsymbol{C} = \boldsymbol{i}(B_yC_z - B_zC_y) + \boldsymbol{j}(B_zC_x - B_xC_z) + \boldsymbol{k}(B_xC_y - B_yC_x)$

We replace x, y, z by 1, 2, 3

Now the first component of $(\boldsymbol{B}\times\boldsymbol{C})_1 = \varepsilon_{1\,jk}B_jC_k$

Now if j, k = 2, 3 or 3, 2

$$(\boldsymbol{B}\times\boldsymbol{C})_1 = \varepsilon_{123}B_2C_3 + \varepsilon_{132}B_3C_2$$

Now $\varepsilon_{123} = +1$ and $\varepsilon_{123} = -1$

$$\therefore \quad (\boldsymbol{B} \times \boldsymbol{C})_1 = B_2C_3 - B_3C_2$$

If we take j, k = 1, 3 or 1, 2; $\varepsilon_{1\,jk} = 0$

Similarly other components of $\boldsymbol{B} \times \boldsymbol{C}$ can be found out $(\boldsymbol{B}\times\boldsymbol{C})_2 = \varepsilon_{2\,jk}B_jC_k$

With j, k = 3, 1 or 1, 3

$$(\boldsymbol{B}\times\boldsymbol{C})_2 = \varepsilon_{231}B_3C_1 + \varepsilon_{213}B_1C_3$$

$$= B_3C_1 - B_1C_3$$

Since $\varepsilon_{231} = +1$, $\varepsilon_{213} = -1$;

It is to be noted that the formulae in vector analysis can be written in the form using δ_{ij} and ε_{ijk}.

4.5.2 : Multiple products of vectors

With the help of dot and cross products of two vectors, it is possible to build multiple products involving several vectors. We shall discuss here two kinds of triple products which are specially important. One is called triple scalar product and the other is called triple vector product.

4.5.3 Triple Scalar Product

We consider three vectors ***A***,***B***,***C*** and arrange them in anticlockwise direction (right handed) as shown below in fig(4A.7)

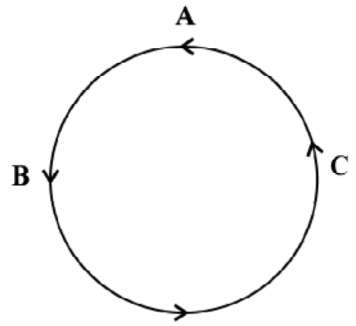


Fig (4A.7)

Then triple scalar product [***ABC***] is defined as [***ABC***] = ***A***.***B*** × ***C*** = ***B***.***C*** × ***A*** = ***C***.***A*** × ***B*** (4A.20)

Thus if a cyclic change (right handed) is made in the sequence of ***B***, ***C***; the triple scalar product remains the same.

However for left handed (clockwise) cyclic change as shown below in fig. (4A.8)

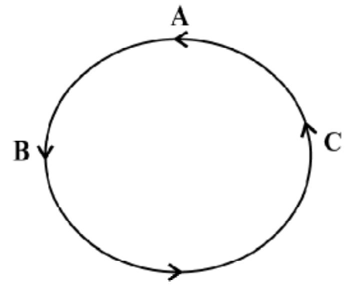


Fig. (4A.8)

We have [***ACB***] = ***A.C*** × ***B*** = ***C.B*** × ***A*** = ***B.A*** × ***C***. Thus triple scalar product depends on the handedness of the vectors ***A***,***B*** and ***C***. In writing ***A.B*** × ***C*** etc. no bracket is necessary. It is then seen that [***ABC***] = – [***ACB***]

Properties :

1. $[\boldsymbol{ABC}] = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}$

where A_x etc. are components of the respective vectors.

2. In a triple scalar product 'dot' and 'cross' can be interchanged, implying

$$[\boldsymbol{ABC}] = \boldsymbol{A} \times \boldsymbol{B}.\boldsymbol{C} = \boldsymbol{B} \times \boldsymbol{C}.\boldsymbol{A} = \boldsymbol{C} \times \boldsymbol{A}.\boldsymbol{B}$$

3. If any of the two non-vanishing vectors $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ be parallel or equal, $[\boldsymbol{ABC}] = 0$

4. $[\boldsymbol{ABC}] = 0$, the vectors $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ are coplanar and also linearly dependent.

5. $[\boldsymbol{ABC}]$ gives the volume of a parallelopiped having $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ as coterminous edges. This is the geometric interpretation of triple scalar product.

6. Triple scalar product is distributive i.e. $[\boldsymbol{A}\ \boldsymbol{B} + \boldsymbol{C}\ \boldsymbol{D} - \boldsymbol{E}] = [\boldsymbol{ABD}] + [\boldsymbol{ACD}] - [\boldsymbol{ABE}] - [\boldsymbol{ACE}]$.

7. The volume of the tetrahedron ABCD is the numerical value of $\frac{1}{6}\left[\overrightarrow{AB}\,\overrightarrow{AC}\,\overrightarrow{AD}\right]$, Fig 4A.9

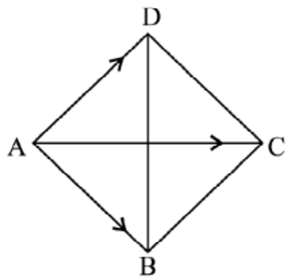


Fig. (4A.9)

8. For an orthonormal right handed vector triad $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ we have

$[\boldsymbol{ijk}] = [\boldsymbol{jki}] = [\boldsymbol{kij}] = 1$ and for left handed triad $[\boldsymbol{ikj}] = [\boldsymbol{kji}] = [\boldsymbol{jik}] = -1$

Geometrical interpretation :

The vectors $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ are represented by $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ respectively. The magnitude of the vector $\boldsymbol{B} \times \boldsymbol{C}$ is the area of the parallelogram OBDC and its direction is along $\overrightarrow{OP}$, perpendicular to the plane OBDC.

Drop perpendicular from A on OP which is AM. So OM is the height of the parallelepiped.

Then $\boldsymbol{A}.\boldsymbol{B} \times \boldsymbol{C}$ = (projection of $\boldsymbol{A}$ *on* $\boldsymbol{B} \times \boldsymbol{C}$) × *magnitude of* $\boldsymbol{B} \times \boldsymbol{C}$

= *height of the parallelepiped* × *area of the base of the parallelepiped*

= *volume of the parallelepiped*

By taking a various faces in turn we find that $\boldsymbol{A}.\boldsymbol{B} \times \boldsymbol{C} = \boldsymbol{B}.\boldsymbol{C} \times \boldsymbol{A} = \boldsymbol{C}.\boldsymbol{A} \times \boldsymbol{B}$
= volume of the parallelepiped with three adjacent side
*as the magnitude of **A**, **B** and **C***

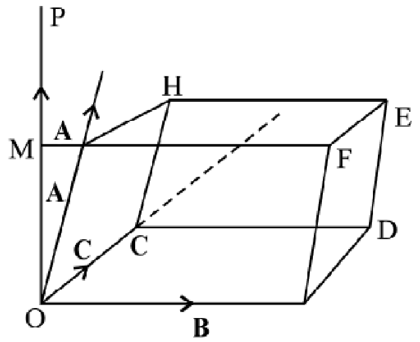


Fig. (4A.10)

4.5.4 : Triple Vector Product

If ***A, B*** and ***C*** are the three vectors, then triple vector product is defined as either $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C})$ or $(\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C}$, parentheses is essential, since $\boldsymbol{A} \times \boldsymbol{B} \times \boldsymbol{C}$ is meaningless.

We have

$$\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = (\boldsymbol{A}.\boldsymbol{C})\boldsymbol{B} - (\boldsymbol{A}.\boldsymbol{B})\boldsymbol{C} \qquad (4A.21)$$

The value of a triple vector product is a linear combination of the two vectors in the parentheses, e.g. ***B*** and ***C***; the co-efficient of each vector is the dot product of the other two, the middle vector in the triple product, e.g. ***B***; always has the positive sign and the other vector in the parentheses e.g. ***C***; always has the negative sign. Thus the vector $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C})$ lies in the plane of ***B*** and ***C***.

From the discussion as above

$$(\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C} = (\boldsymbol{A}.\boldsymbol{C})\boldsymbol{B} - (\boldsymbol{B}.\boldsymbol{C})\boldsymbol{A} \qquad (4A.22)$$

Now $(\boldsymbol{B} \times \boldsymbol{C}) \times \boldsymbol{A} = (\boldsymbol{B}.\boldsymbol{A})\boldsymbol{C} - (\boldsymbol{A}.\boldsymbol{C})\boldsymbol{B}$

$$= -[(\boldsymbol{A}.\boldsymbol{C})\boldsymbol{B} - (\boldsymbol{A}.\boldsymbol{B})\boldsymbol{C}]$$

$$= -\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C})$$

Proof of equation (4A.21) :

$\boldsymbol{B} \times \boldsymbol{C}$ is a vector perpendicular to the plane of ***B*** and ***C.*** thus $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C})$ is some vector in the plane of ***B*** and ***C***.

Therefore we can write

$$\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = l\boldsymbol{B} + m\boldsymbol{C}\ ; \qquad (4A.23)$$

l, m are scalars. Making dot product with $\boldsymbol{A}$ both sides $\boldsymbol{A}.\ \boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = l(\boldsymbol{A}.\boldsymbol{B}) + m(\boldsymbol{A}.\boldsymbol{C})$

or, $\quad 0 = l(\boldsymbol{A}.\boldsymbol{B}) + m(\boldsymbol{A}.\boldsymbol{C})$

Using property (3) of triple scalar product

Or, $\quad \dfrac{l}{\boldsymbol{A.C}} = -\dfrac{m}{\boldsymbol{A.B}} = n$ (say)

Substituting the values of l and m in equation (4A.23) ,

$$\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = n\ (\boldsymbol{A}.\boldsymbol{C})\boldsymbol{B} - n(\boldsymbol{A}.\boldsymbol{B})\boldsymbol{C} \qquad (4A.24)$$

Since vector equations are independent of co-ordinate system, we can take, to facilitate our calculations, but without any loss of generality,

$$\boldsymbol{A} = \boldsymbol{C} = \boldsymbol{i},\ \boldsymbol{B} = \boldsymbol{j}$$

Therefore from equation (4A.24)

$$\boldsymbol{i} \times (\boldsymbol{j} \times \boldsymbol{i}) = n\boldsymbol{j}$$

Or $\quad \boldsymbol{i} \times (-\boldsymbol{k}) = n\boldsymbol{j}$

Or $\quad \boldsymbol{k} \times \boldsymbol{i} = n\boldsymbol{j} = \boldsymbol{j}$

$\therefore \quad n = 1$

Therefore $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = (\boldsymbol{A}.\boldsymbol{C})\boldsymbol{B} - (\boldsymbol{A}.\boldsymbol{B})\boldsymbol{C} \qquad (24.21)$

proof of equation (4A.22) is left as an exercise.

4.5.5 Product Of Four Vectors

Scalar product of four vectors :

Scalar product of four vectors ***A, B, C*** and ***D*** is defined as $(\boldsymbol{A} \times \boldsymbol{B}).\ (\boldsymbol{C} \times \boldsymbol{D})$

Now let's suppose $\boldsymbol{C} \times \boldsymbol{D} = \boldsymbol{N}$

Then $(\boldsymbol{A} \times \boldsymbol{B}).(\mathrm{C} \times \boldsymbol{D}) = (\boldsymbol{A} \times \boldsymbol{B}).\boldsymbol{N}$

$$= \boldsymbol{A}.\boldsymbol{B} \times \boldsymbol{N} = \boldsymbol{A}.\boldsymbol{B} \times (\boldsymbol{C} \times \boldsymbol{D})$$

$$= \boldsymbol{A}.\ [(\boldsymbol{B}.\boldsymbol{D})\boldsymbol{C} - (\boldsymbol{B}.\boldsymbol{C})\boldsymbol{D}]$$

$$= (\boldsymbol{A}.\boldsymbol{C})(\boldsymbol{B}.\boldsymbol{D}) - (\boldsymbol{A}.\boldsymbol{D})\ (\boldsymbol{B}.\boldsymbol{C})$$

$$= \begin{vmatrix} \boldsymbol{A.C} & \boldsymbol{A.D} \\ \boldsymbol{B.C} & \boldsymbol{B.D} \end{vmatrix} \qquad (4A.24A)$$

Vector product of four vectors :

Vectors product of four vectors ***A***, ***B***, ***C*** and ***D*** is defined as,

$(\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C} \times \boldsymbol{D})$

now let $\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{N}$

$\therefore\ (\boldsymbol{A} \times \boldsymbol{B}) \times (\boldsymbol{C} \times \boldsymbol{D}) = \boldsymbol{N} \times (\boldsymbol{C} \times \boldsymbol{D})$

$= (\boldsymbol{N}.\boldsymbol{D})\boldsymbol{C} - (\boldsymbol{N}.\boldsymbol{C})\boldsymbol{D}$

$= (\boldsymbol{A} \times \boldsymbol{B}.\boldsymbol{D})\boldsymbol{C} - (\boldsymbol{A} \times \boldsymbol{B}.\boldsymbol{C})\boldsymbol{D}$

$= l\boldsymbol{C} - m\boldsymbol{D}$ (4A.24B)

where l and m are scalar.

Therefore $(\boldsymbol{A} \times \boldsymbol{B}) \times (\boldsymbol{C} \times \boldsymbol{D})$ lies in the plane of ***C*** and ***D***.

Now let $\boldsymbol{C} \times \boldsymbol{D} = \boldsymbol{N}$

$\therefore\ (\boldsymbol{A} \times \boldsymbol{B}) \times (\boldsymbol{C} \times \boldsymbol{D}) = (\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{N}$

$= (\boldsymbol{A}.\boldsymbol{N})\boldsymbol{B} - (\boldsymbol{B}.\boldsymbol{N})\boldsymbol{A}$

$= (\boldsymbol{A}.\boldsymbol{C} \times \boldsymbol{D})\boldsymbol{B} - (\boldsymbol{B}.\boldsymbol{C} \times \boldsymbol{D})\boldsymbol{A}$

$= p\boldsymbol{B} - q\boldsymbol{A}$ (4A.25)

where p and q are scalars. Therefore $(\boldsymbol{A} \times \boldsymbol{B}) \times (\boldsymbol{C} \times \boldsymbol{D})$ can also be expressed as a linear combination of the vectors ***B*** and ***A***.

Example of Art 4.4.2 to 4.5.5 :

Example 5: If ***A*** and ***B*** are two vectors, show that $|\boldsymbol{A} \times \boldsymbol{B}| = \sqrt{(\boldsymbol{A}.\boldsymbol{A})(\boldsymbol{B}.\boldsymbol{B}) - (\boldsymbol{A}.\boldsymbol{B})^2}$

Solution : R.H.S : $= \sqrt{A^2B^2 - A^2B^2\cos^2\theta} = AB\sqrt{1-\cos^2\theta} = AB\sin\theta = |\boldsymbol{A} \times \boldsymbol{B}|$

Example 6 : Find the angle between the vectors $\boldsymbol{A} = 4\boldsymbol{i} + 3\boldsymbol{j} + \boldsymbol{k}$ and $\boldsymbol{B} = 2\boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k}$. Also find a unit vector perpendicular to both ***A*** and ***B***. Use concept of dot product only.

Solution : Using the definition of dot product,

$\cos\theta = \dfrac{\boldsymbol{A}.\boldsymbol{B}}{|\boldsymbol{A}||\boldsymbol{B}|}$, where θ is the angle between the ***A*** and ***B***.

Now $|\boldsymbol{A}| = \sqrt{(4)^2 + (3)^2 + (1)^2} = \sqrt{26}$

$|\boldsymbol{B}| = \sqrt{(2)^2 + (-1)^2 + (2)^2} = \sqrt{9} = 3$

And $\boldsymbol{A}.\boldsymbol{B} = (4\boldsymbol{i} + 3\boldsymbol{j} + \boldsymbol{k}).\ (2\boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k}) = 8 - 3 + 2 = 7$

$$\therefore\ \cos\theta = \frac{7}{3\sqrt{26}};\quad \theta = \cos^{-1}\left(\frac{7}{3\sqrt{26}}\right)$$

Now let $\hat{\boldsymbol{r}}$ be a unit vector perpendicular to both $\boldsymbol{A}$ and $\boldsymbol{B}$ so that $\hat{\boldsymbol{r}} = \alpha\boldsymbol{i} + \beta\boldsymbol{j} + \gamma\boldsymbol{k}$

Therefore $\alpha^2 + \beta^2 + \gamma^2 = 1$ (i)

Again $\hat{\boldsymbol{r}}.\boldsymbol{B} = 2\alpha - \beta + \gamma = 0$ (ii)

And $\hat{\boldsymbol{r}}.\boldsymbol{A} = 4\alpha - 3\beta + \gamma = 0$ (iii)

Solving equation (ii) and (iii) by cross-multiplication

$$\frac{\alpha}{4} = \frac{\beta}{-2} = \frac{\gamma}{-10} = \frac{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}{\sqrt{(4)^2 + (-2)^2 + (-10)^2}} = \frac{1}{\sqrt{120}}$$

$$\therefore \alpha = \frac{4}{\sqrt{120}} = \frac{2}{\sqrt{30}};\ \beta = \frac{-1}{\sqrt{30}};\ \gamma = -\frac{5}{\sqrt{30}}$$

therefore $\hat{\boldsymbol{r}} = \frac{1}{\sqrt{30}}\left[2\boldsymbol{i} - \boldsymbol{j} - 5\boldsymbol{k}\right]$

Example 7 : Given $\boldsymbol{A} = \boldsymbol{i} + \boldsymbol{j} + \mathbf{k}$ and $\boldsymbol{C} = \boldsymbol{j} - \boldsymbol{k}$. Find a vector $\boldsymbol{B}$ such that, $\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{C}$ and $\boldsymbol{A}.\boldsymbol{B} = 3$

Solution : Suppose $\boldsymbol{B} = \alpha\boldsymbol{i} + \beta\boldsymbol{j} + \gamma\boldsymbol{k}$

Now $\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{C}$ gives,

$$\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & 1 & 1 \\ \alpha & \beta & \gamma \end{vmatrix} = \boldsymbol{j} - \boldsymbol{k}$$

Or, $\quad \boldsymbol{i}(\gamma-\beta)-\boldsymbol{j}(\gamma-\alpha)+\boldsymbol{k}(\beta-\alpha)=\boldsymbol{j}-\boldsymbol{k}$

Corresponding co-efficient of $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ from both sides,

We have $\gamma-\beta=0,\ \gamma-\alpha=-1,$ and $\beta-\alpha=-1$

as $\beta=\gamma, \alpha=1+\gamma,\ \beta=\alpha-1;\ \alpha=1+\beta$ (i)

Again, $\boldsymbol{A}.\boldsymbol{B}$ = 3 gives $\alpha+\beta+\gamma=3$ (ii)

Solving equations (i) and (ii) we get

$$\beta=\gamma=\frac{2}{3} \text{ and } \alpha=1+\beta=\frac{5}{3}$$

Therefore $B=\frac{5}{3}\boldsymbol{i}+\frac{2}{3}\boldsymbol{j}+\frac{2}{3}\boldsymbol{k}=\frac{1}{3}\left[5\boldsymbol{i}+2\boldsymbol{j}+2\boldsymbol{k}\right]$

Example 8 : If $\boldsymbol{a}$ = 4$\boldsymbol{i}$ + 3$\boldsymbol{j}$ + $\boldsymbol{k}$, $\boldsymbol{b}$ = 2$\boldsymbol{i}$ – $\boldsymbol{j}$ + 2$\boldsymbol{k}$, find a unit vector $\hat{\boldsymbol{n}}$ perpendicular to vector $\boldsymbol{a}$ and $\boldsymbol{b}$ such that $\boldsymbol{a}$, $\boldsymbol{b}$, $\hat{\boldsymbol{n}}$ form a right handed system. Find the angle between the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$.

Solution : We have, $a\times b=\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 4 & 3 & 1 \\ 2 & -1 & 2 \end{vmatrix}=7\boldsymbol{i}-6\boldsymbol{j}-10\boldsymbol{k}$

And $\quad |\boldsymbol{a}\times\boldsymbol{b}|=\sqrt{(7)^2+(-6)^2+(-10)^2}=\sqrt{185}$

Therefore $\hat{\boldsymbol{n}}=\dfrac{\boldsymbol{a}\times\boldsymbol{b}}{|\boldsymbol{a}|\times|\boldsymbol{b}|}=\dfrac{7\boldsymbol{i}-6\boldsymbol{j}-10\boldsymbol{k}}{\sqrt{185}}$

Also $\quad |\boldsymbol{a}|=\sqrt{4^2+3^2+1^2}=\sqrt{126}$

$$|\boldsymbol{b}|=\sqrt{2^2+1^2+2^2}=3$$

If θ be the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$, then

$|\boldsymbol{a}\times\boldsymbol{b}|=|\boldsymbol{a}|\,|\boldsymbol{b}|\sin\theta,$ then $\sin\theta=\dfrac{|\boldsymbol{a}\times\boldsymbol{b}|}{|\boldsymbol{a}|\times|\boldsymbol{b}|}=\dfrac{\sqrt{185}}{3\sqrt{26}}=\sin 62^0 42';\ \therefore \theta=62^0 42'$

Example 9 : Find the value of λ for which vector :

$$\boldsymbol{A} = 2\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}$$

$$\boldsymbol{B} = \boldsymbol{i} + 2\boldsymbol{j} - 3\boldsymbol{k}$$

$\boldsymbol{C} = 3\boldsymbol{i} + \lambda\boldsymbol{j} + 5\boldsymbol{k}$, are coplanar.

Solution : There vectors ***A***, ***B***, ***C*** are coplanar if $[\boldsymbol{ABC}] = 0$

or, $$\begin{vmatrix} 2 & -1 & 1 \\ 1 & 2 & -3 \\ 3 & \lambda & 5 \end{vmatrix} = 0$$

Or, $$2(10+3\lambda)+(5+9)+(\lambda-6)=0$$

Or, $$7\lambda+28=0$$

$$\therefore \lambda=-4$$

Example 10 : Show that, $(\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C} = \boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C})$ only when ***A*** and ***C*** are collinear or $(\boldsymbol{A} \times \boldsymbol{C}) \times \boldsymbol{B} = 0$

Solution : Given $(\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C} = \boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C})$

This is possible if, $\boldsymbol{B}(\boldsymbol{A}.\boldsymbol{C}) + \boldsymbol{C}(\boldsymbol{A}.\boldsymbol{B}) - \boldsymbol{B}(\boldsymbol{A}.\boldsymbol{C}) - \boldsymbol{A}(\boldsymbol{B}.\boldsymbol{C}) = 0$

or, if $\boldsymbol{C}(\boldsymbol{A}.\boldsymbol{B}) - \boldsymbol{A}(\boldsymbol{B}.\boldsymbol{C}) = 0$

or, if $(\boldsymbol{A} \times \boldsymbol{C}) \times \boldsymbol{B} = 0$

This show that either $\boldsymbol{B} = 0$ or $\boldsymbol{A} \times \boldsymbol{C} = 0$, but $\boldsymbol{B} \neq 0$, hence $\boldsymbol{A} \times \boldsymbol{C} = 0$, hence ***A*** and ***C*** are collinear.

Example 11 : If ***A***, ***B*** and ***C*** satisfy the condition $(\boldsymbol{A} \times \boldsymbol{B}) + (\boldsymbol{B} \times \boldsymbol{C}) + (\boldsymbol{C} \times \boldsymbol{A}) = 0$, show that the vector are coplanar.

Solution : We have $(\boldsymbol{A} \times \boldsymbol{B}) + (\boldsymbol{B} \times \boldsymbol{C}) + (\boldsymbol{C} \times \boldsymbol{A}) = 0$

$$\therefore \quad [(\boldsymbol{A} \times \boldsymbol{B}) + (\boldsymbol{B} \times \boldsymbol{C}) + (\boldsymbol{C} \times \boldsymbol{A})].\boldsymbol{A} = 0$$

or, $$\boldsymbol{A} \times \boldsymbol{B}.\boldsymbol{A} + \boldsymbol{B} \times \boldsymbol{C}.\boldsymbol{A} + \boldsymbol{C} \times \boldsymbol{A}.\boldsymbol{A} = 0$$

or, $$\boldsymbol{B} \times \boldsymbol{C}.\boldsymbol{A} = 0 \text{ ; or } [\boldsymbol{ABC}] = 0$$

Therefore, ***A***, ***B***, ***C*** are coplanar

4.6 ❒ Reciprocal System of Vectors

The concept of reciprocal vectors finds applications in solid state physics in connection with reciprocal lattice.

Let there be a set of three non-coplanar vectors ***a, b***, ***c***. The set of other three vectors ***A***, ***B***, ***C*** defined by the equation

$$\boldsymbol{A} = \frac{\boldsymbol{b}\times\boldsymbol{c}}{[\boldsymbol{abc}]}, \boldsymbol{B} = \frac{\boldsymbol{c}\times\boldsymbol{a}}{[\boldsymbol{abc}]},\ \boldsymbol{C} = \frac{\boldsymbol{a}\times\boldsymbol{b}}{[\boldsymbol{abc}]} \qquad (4A.26)$$

Are called reciprocal vectors triads to the vectors ***a***, ***b*** and ***c***.

The vector triads ***a***, ***b*** and ***c*** and its reciprocal triads ***A***, ***B***, ***C*** are either both right handed or both left handed. The vector triads (***a***, ***b***, ***c***) and (***A***, ***B***, ***C***) are mutually reciprocal.

i.e. $$\boldsymbol{a} = \frac{\boldsymbol{B}\times\boldsymbol{C}}{[\boldsymbol{ABC}]}, \boldsymbol{b} = \frac{\boldsymbol{C}\times\boldsymbol{A}}{[\boldsymbol{ABC}]},\ \boldsymbol{c} = \frac{\boldsymbol{A}\times\boldsymbol{B}}{[\boldsymbol{ABC}]}$$

Where ***A***, ***B***, ***C*** are non-co-planar vectors given by $[\boldsymbol{ABC}] \neq 0$

Properties of reciprocal system :

1. If ***a***, ***b***, ***c*** and ***A***, ***B***, ***C*** be reciprocal triads of vectors, then

$$\boldsymbol{a}.\boldsymbol{A} = \boldsymbol{b}.\boldsymbol{B} = \boldsymbol{c}.\boldsymbol{C} = 1 \qquad (4A.27)$$

Proof : We have $\boldsymbol{A} = \dfrac{\boldsymbol{b}\times\boldsymbol{c}}{[\boldsymbol{abc}]}$

$$\therefore \boldsymbol{a}.\boldsymbol{A} = \frac{\boldsymbol{a}.\boldsymbol{b}\times\boldsymbol{c}}{[\boldsymbol{abc}]} = \frac{[\boldsymbol{abc}]}{[\boldsymbol{abc}]} = 1$$

Similarly $\boldsymbol{b}.\boldsymbol{B} = \boldsymbol{c}.\boldsymbol{C} = 1$;

Then $\boldsymbol{a}.\boldsymbol{A} + \boldsymbol{b}.\boldsymbol{B} + \boldsymbol{c}.\boldsymbol{C} = 3$

And $\boldsymbol{A} = \dfrac{1}{\boldsymbol{a}},\ \boldsymbol{B} = \dfrac{1}{\boldsymbol{b}},\ \boldsymbol{C} = \dfrac{1}{\boldsymbol{c}}$

2. If ***a***, ***b***, ***c*** and ***A***, ***B***, ***C*** are reciprocal triad of vectors, then

$$\boldsymbol{a}.\boldsymbol{B} = \boldsymbol{a}.\boldsymbol{C} = 0$$

$$\boldsymbol{b}.\boldsymbol{A} = \boldsymbol{b}.\boldsymbol{C} = 0 \qquad (4A.28)$$

$$\boldsymbol{c}.\boldsymbol{A} = \boldsymbol{c}.\boldsymbol{B} = 0$$

Proof : $\boldsymbol{a}.\boldsymbol{B} = \boldsymbol{a}.\dfrac{\boldsymbol{c}\times\boldsymbol{a}}{[\boldsymbol{abc}]} = \dfrac{[\boldsymbol{acb}]}{[\boldsymbol{abc}]} = 0$

Since [***aca***] = 0; and similar for other relation.

3. The triple scalar product of any three non-co-planar vectors ***a***, ***b***, ***c*** is reciprocal to the corresponding triple scalar product of reciprocal vectors ***A***, ***B***, ***C***.

Proof :

$$[\boldsymbol{ABC}] = \boldsymbol{A}.\boldsymbol{B}\times\boldsymbol{C} = (\boldsymbol{b}\times\boldsymbol{c}).\frac{[(c\times a)\times(a\times b)]}{[abc]^3}$$, using equation (4A.26)

Now $(\boldsymbol{c} \times \boldsymbol{a}) \times (\boldsymbol{a} \times \boldsymbol{b}) = (\boldsymbol{c} \times \boldsymbol{a}) \times \boldsymbol{N}$ when $\boldsymbol{N} = \boldsymbol{a} \times \boldsymbol{b}$

$$= \boldsymbol{a}(\boldsymbol{c}.\boldsymbol{N}) - \mathrm{c}(\boldsymbol{a}.\boldsymbol{N})$$

$$= \boldsymbol{a}(\boldsymbol{c}.\boldsymbol{a} \times \boldsymbol{b}) - \boldsymbol{c}(\boldsymbol{a}.\boldsymbol{a} \times \boldsymbol{b})$$

$$= \boldsymbol{a}[\boldsymbol{abc}] \text{ since } \boldsymbol{a}.\boldsymbol{a} \times \boldsymbol{b} = 0$$

$$\therefore [\boldsymbol{ABC}] = \frac{(\boldsymbol{b}\times\boldsymbol{c}).\boldsymbol{a}[\boldsymbol{abc}]}{[\boldsymbol{abc}]^3} = \frac{[\boldsymbol{abc}]^2}{[\boldsymbol{abc}]^3} = \frac{1}{[\boldsymbol{abc}]} \qquad (4A.29)$$

Example of Art 4.6 :

Example 12 : Show that the orthonormal vector triads (***i***, ***j***, ***k***) is self reciprocal.

Solution : Let $(\boldsymbol{i}', \boldsymbol{j}', \boldsymbol{k}')$ be the set of vectors reciprocal to (***i***, ***j***, ***k***) then,

$$\boldsymbol{i}' = \frac{\boldsymbol{j}\times\boldsymbol{k}}{[\boldsymbol{ijk}]},\ \boldsymbol{j}' = \frac{\boldsymbol{k}\times\boldsymbol{i}}{[\boldsymbol{ijk}]},\ \boldsymbol{k}' = \frac{\boldsymbol{i}\times\boldsymbol{j}}{[\boldsymbol{ijk}]}$$

$$\therefore \boldsymbol{i}' = \boldsymbol{i}\ ;\ \boldsymbol{j}' = \boldsymbol{j}\ ;\ \boldsymbol{k}' = \boldsymbol{k}$$

Therefore the orthonormal set of vector triads ***i***, ***j***, ***k*** is self-reciprocal.

Example in Machanics :

Example 13 : A particle being acted on by constant force (4***i*** + ***j*** – 3***k***) and (3***i*** + ***j*** – ***k***) is displaced from the point (***i*** + 2***j*** + 3***k***) to the point (5***i*** + 4***j*** – ***k***). Find the total work done by the forces.

Solution : The displacement $\boldsymbol{d}$ is given by $\boldsymbol{d} = (5\boldsymbol{i} + 4\boldsymbol{j} - \boldsymbol{k}) - (\boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}) = 4\boldsymbol{i} + 2\boldsymbol{j} - 4\boldsymbol{k}$

Resultant forces $\boldsymbol{F} = (4\boldsymbol{i} + \boldsymbol{j} - 3\boldsymbol{k}) + (3\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k}) = 7\boldsymbol{i} + 2\boldsymbol{j} - 4\boldsymbol{k}$

$\therefore$ Work done $= \boldsymbol{F}.\boldsymbol{d} = 28 + 4 + 16 = 48$ units of work

Example 14 : A rigid body is spinning with angular velocity 27 radians per sec about an axis parallel to $2\boldsymbol{i} + \boldsymbol{j} - 2\boldsymbol{k}$ passing through the point $\boldsymbol{i} + 3\boldsymbol{j} - \boldsymbol{k}$. Find the velocity of the point of the body where position vector is $4\boldsymbol{i} + 8\boldsymbol{j} + \boldsymbol{k}$.

Solution : Unit vector along the direction of the angular velocity ω is

$$\frac{2i + j - 2k}{\sqrt{4+1+4}} = \frac{1}{3}(2\boldsymbol{i} + \boldsymbol{j} - 2\boldsymbol{k})$$

Or, $$\omega = \frac{27}{3}(2\boldsymbol{i} + \boldsymbol{j} - 2\boldsymbol{k}) = 9(2\boldsymbol{i} + \boldsymbol{j} - 2\boldsymbol{k})$$

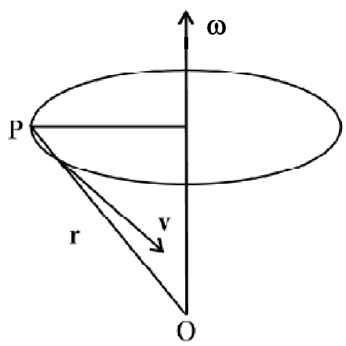


Fig. Example (4)

Let O be the point having position vector $\boldsymbol{i} + 3\boldsymbol{j} + \boldsymbol{k}$ and the point P of the body has the position vector, $4\boldsymbol{i} + 8\boldsymbol{j} + \boldsymbol{k}$. Then $\boldsymbol{r} = \overrightarrow{\boldsymbol{OP}} = (4\boldsymbol{i} + 8\boldsymbol{j} + \boldsymbol{k}) - (\boldsymbol{i} + 3\boldsymbol{j} - \boldsymbol{k})$

Or, $\boldsymbol{r} = 3\boldsymbol{i} + 5\boldsymbol{j} + 2\boldsymbol{k}$

$\therefore$ Linear velocity of P is, $\boldsymbol{v} = \boldsymbol{w} \times \boldsymbol{r}$

$$= 9\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 2 & 1 & 2 \\ 3 & 5 & 2 \end{vmatrix} = 9(12\boldsymbol{i} - 10\boldsymbol{j} + 7\boldsymbol{k})$$

Example 15 : Find the torque about O (3, –1, 3) of a Force $\boldsymbol{F}$(4, 2, 1) passing through the point A(5, 2, 4)

Solution : Position vector of A(5, 2, 4) relative to O (3, –1, 3) is $\boldsymbol{r} = 2\boldsymbol{i} + 3\boldsymbol{j} + \boldsymbol{k}$

Again the force $\boldsymbol{F} = 4\boldsymbol{i} + 2\boldsymbol{j} + \boldsymbol{k}$

$$\therefore \text{Torque} = \boldsymbol{r} \times \boldsymbol{F} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 2 & 3 & 1 \\ 4 & 2 & 1 \end{vmatrix} = \boldsymbol{i} + 2\boldsymbol{j} - 8i$$

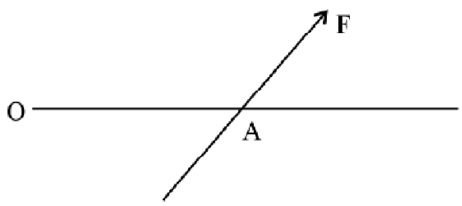


Fig. Example (15)

4.7 ❒ Properties of Vectors Under Rotation

A vector is a mathematical object that transforms in a particular way under rotation.

We consider a point $P(x, y)$ in a two dimensional co-ordinate system OX and OY, let now the co-ordinate frame rotate in anticlockwise direction by an angle θ so that OX´, OY´ are the new positions of the axes.

If the co-ordinate of the point P in the new co-ordinate frame by (x´, y´), then from fig(4A.11), we find

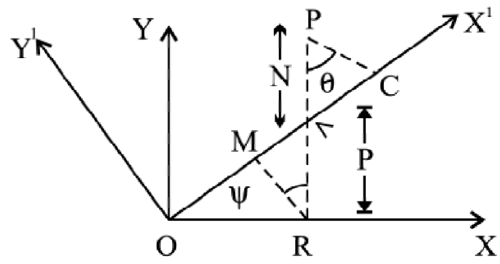


Fig. (4A.11)

$$x' = OM + MN + NC$$

$$= x\cos\theta + p\sin\theta + q\sin\theta = x\sin\theta + y\sin\theta \qquad (4A.30)$$

Since $p + q = y$ and $OR = x$

$$y' = PC = q\cos\theta = (PR - RN)\cos\theta$$

$$= (y - p)\cos\theta = y\cos\theta - p\cos\theta$$

$$= y\cos\theta - MR = y\cos\theta - x\sin\theta$$

$$= -x\sin\theta + y\cos\theta \qquad (4A.31)$$

Since in $\Delta MRN, \cos\theta = \frac{MR}{RN} = \frac{MR}{p} \quad \therefore MR = p\cos\theta$

Again in $\Delta ORM, \sin\theta = \frac{MR}{OR} = \frac{MR}{x} \quad \therefore MR = x\sin\theta$

The transformation given by equation (4A.30) and (4A.31) are called orthogonal transformation.

We take ***i, j*** unit vectors along OX and OY axes and ***i´,j´*** along rotated axes OX´, OY´ respectively. Then the position vector

$$\boldsymbol{r} = \overrightarrow{OP} = \boldsymbol{i}x + \boldsymbol{j}y \text{ (in old coordinate axes)}$$

$$= \boldsymbol{i}'x' + \boldsymbol{j}'y' \text{ (in new co-ordinate axes)} \qquad (4A.32)$$

Therefore component, of $\boldsymbol{r}$ in new co-ordinate axes are $\boldsymbol{i}'x'$ and $\boldsymbol{j}'y'$ which is different from $\boldsymbol{i}x$ and $\boldsymbol{j}y$ respectively. But magnitude of $\boldsymbol{r} = |\boldsymbol{r}| = \sqrt{(x^2 + y^2)} = \sqrt{(x'^2 + y'^2)}$ is same in both co-ordinate system ; since

$$x'^2 + y'^2 = (x\cos\theta + y\sin\theta)^2 + (-x\sin\theta + y\cos\theta)^2$$

$$= x^2\cos^2\theta + y^2\sin^2\theta + 2xy\sin\theta\cos\theta + x^2\sin^2\theta + y^2\cos^2\theta - 2xy\sin\theta\cos\theta$$

$$= x^2\left(\cos^2\theta + \sin^2\theta\right) + y^2\left(\cos^2\theta + \sin^2\theta\right) = x^2 + y^2 \qquad (4A.33)$$

Thus we have following observations :

$$\boldsymbol{r} = \overrightarrow{OP}$$

1. The vector $\boldsymbol{r}$ remains the same in the two co-ordinate systems, though its components change.
2. The length of the vector $\boldsymbol{r} = \overrightarrow{OP}$ remains the same in both the co-ordinate system.

The results can be generalised for any vector in three dimensions also i.e.

$$\boldsymbol{A} = \boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z = \boldsymbol{i}'A'_x + \boldsymbol{j}'A'_y + \boldsymbol{k}'A'_z \qquad (4A.34)$$

and $$|A| = \sqrt{A_x{}^2 + A_y{}^2 + A_z{}^2} = \sqrt{A'_x{}^2 + A'_y{}^2 + A'_z{}^2} \qquad (4A.35)$$

Thus our conclusions are vectors and scalars remain invariant under rotation of co-ordinate systems.

4.7.1 Scalar product of two vectors under rotation of coordinate system

Consider two vectors $\boldsymbol{A} = \boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z$ and $\boldsymbol{B} = \boldsymbol{i}B_x + \boldsymbol{j}B_y + \boldsymbol{k}B_z$ and their scalar products $\boldsymbol{A}.\boldsymbol{B} = \boldsymbol{A}_x\boldsymbol{B}_x + \boldsymbol{A}_y\boldsymbol{B}_y + \boldsymbol{A}_z\boldsymbol{B}_z$ with respect to a rectangular Cartesian co-ordinate system fig (4A.11). We consider rotation of the co-ordinate system about Z-axis which is perpendicular to the plane of the paper and passing through the origin O. Under anticlockwise rotation through angle θ, the new co-ordinate system becomes x', y' and $z' = z$ and the components of the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ are transformed according to equation (4A.30) and (4A.31) as

$$\left.\begin{aligned} A'_x &= A_x\cos\theta + A_y\sin\theta \\ A'_y &= -A_x\cos\theta + A_y\sin\theta \\ A'_z &= A_z \end{aligned}\right\} \qquad (4A.36)$$

$$\left.\begin{aligned} B'_x &= B_x \cos\theta + B_y \sin\theta \\ B'_y &= -B_x \cos\theta + B_y \sin\theta_y \\ B'_z &= B_z \end{aligned}\right\} \tag{4A.37}$$

In the new co-ordinate system the scalar product become

$$\boldsymbol{A}'.\boldsymbol{B}' = A'_x B'_x + A'_y B'_y + A'_z B'_z \tag{4A.38}$$

Now substituting equation (4A.36) and (4A.37) in equation (4A.38), we get

$$\boldsymbol{A}'.\boldsymbol{B}' = \left(A_x \cos\theta + A_y \sin\theta\right)\left(B_x \cos\theta + B_y \sin\theta\right)$$

$$+\left(-A_x \sin\theta + A_y \cos\theta\right)\left(-B_x \sin\theta + B_y \cos\theta\right) + A_z B_z$$

$$= A_x B_x \cos^2\theta + A_x B_y \cos\theta \sin\theta + A_y B_x \sin\theta \cos\theta + A_y B_y \sin^2\theta + A_x B_x \sin^2\theta$$

$$-A_x B_y \sin\theta \cos\theta - A_y B_x \cos\theta \sin\theta + A_y B_y \cos^2\theta + A_z B_z$$

$$= A_x B_x \left(\cos^2\theta + \sin^2\theta\right) + A_y B_y \left(\cos^2\theta + \sin^2\theta\right) + A_z B_z$$

$$= A_x B_x + A_y B_y + A_z B_z = \boldsymbol{A}.\boldsymbol{B} \tag{4A.39}$$

equation (4A.39) shows that scalar product of two vectors remains invariant under rotation of co-ordinate systems or conversely since $\boldsymbol{A}.\boldsymbol{B}$ remains invariant it must be scalar.

4.7.2 Vector Product of Two Vectors Under Rotation of Coordinate Axes

Consider two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ and their vector product $\boldsymbol{A} \times \boldsymbol{B}$. Under rotation of coordinate axes through angle θ counter clockwise about z axis, let the cross product become $\boldsymbol{A}' \times \boldsymbol{B}'$.

By equation (4A.30) and (4A.31), we can write

$$(\boldsymbol{A}' \times \boldsymbol{B}')_{x'} = (\boldsymbol{A} \times \boldsymbol{B})_x \cos\theta + (\boldsymbol{A} \times \boldsymbol{B})_y \sin\theta \tag{4A.40}$$

$$(\boldsymbol{A}' \times \boldsymbol{B}')_{y'} = -(\boldsymbol{A} \times \boldsymbol{B})_x \sin\theta + (\boldsymbol{A} \times \boldsymbol{B})_y \cos\theta \tag{4A.41}$$

$$(\boldsymbol{A}' \times \boldsymbol{B}')_{z'} = (\boldsymbol{A} \times \boldsymbol{B})_z \tag{4A.42}$$

Now squaring equations (4A.40) and (4A.41) both sides and adding, we get

$$(\boldsymbol{A}'\times\boldsymbol{B}')^2{}_{x'}+(\boldsymbol{A}'\times\boldsymbol{B}')^2{}_{y'}+(\boldsymbol{A}'\times\boldsymbol{B}')^2{}_{z'}=(\boldsymbol{A}\times\boldsymbol{B})^2{}_{x}+(\boldsymbol{A}\times\boldsymbol{B})^2{}_{y}+(\boldsymbol{A}\times\boldsymbol{B})^2{}_{z}$$

or, $$|\boldsymbol{A}'\times\boldsymbol{B}'|^2=|\boldsymbol{A}\times\boldsymbol{B}|^2$$

$$|\boldsymbol{A}'\times\boldsymbol{B}'|=|\boldsymbol{A}\times\boldsymbol{B}| \qquad (4A.42)$$

The invariance of dot and cross products imply that both the magnitude of the vectors and the angle between them remains unchanged in a rotation.

Exercise of Arts 4.6 to 4.7.2 :

1) Find a unit vector parallel to the sum of vectors $\boldsymbol{A}_1 = 2\boldsymbol{i} + 4\boldsymbol{j} - 5\boldsymbol{k}$ and $\boldsymbol{A}_2 = \boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}$.

2) Find the position vector of the centroid of a triangle ABC, when the position vector ***A***, ***B*** and ***C*** are ***a, b, c***.

3) If $|\boldsymbol{A} + \boldsymbol{B}| = |\boldsymbol{A} - \boldsymbol{B}|$, then show the ***A*** and ***B*** are perpendicular.

4) If $\boldsymbol{A}.\boldsymbol{B} = \boldsymbol{A}.\boldsymbol{C}$, does it necessarily follow that ***B*** and ***C*** are equal.

5) If $|\boldsymbol{A}| = |\boldsymbol{B}|$, prove that $\boldsymbol{A} + \boldsymbol{B}$ is perpendicular to $\boldsymbol{A} - \boldsymbol{B}$.

6) Define direction cosine of a vector. If l_1, m_1, n_1 and l_2, m_2, n_2 be the direction cosine of the vectors, show that the angle θ between them is $\cos\theta = l_1l_2 + m_1m_2 + n_1n_2$.

7) a) ***A***, ***B*** and ***C*** are three vectors. If $\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C} = 0$, show that $\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{B} \times \boldsymbol{C} = \boldsymbol{C} \times \boldsymbol{A}$.

   b) Is the converse true ?

8) Show that the components of a vector ***G*** along and perpendicular to a vector ***H*** in the planes of ***G*** and ***H*** are respectively given by $\dfrac{(\boldsymbol{H}.\boldsymbol{G})\boldsymbol{H}}{|\boldsymbol{H}|^2}$ and $\dfrac{(\boldsymbol{H}\times\boldsymbol{G})\times\boldsymbol{H}}{|\boldsymbol{H}|^2}$.

9) Prove that :

   a) $(\boldsymbol{A} \times \boldsymbol{B}).(\boldsymbol{C} \times \boldsymbol{D}) = (\boldsymbol{A}.\boldsymbol{C})(\boldsymbol{B}.\boldsymbol{D}) - (\boldsymbol{A}.\boldsymbol{D})(\boldsymbol{B}.\boldsymbol{C})$

   b) $(\boldsymbol{A} \times \boldsymbol{B}).(\boldsymbol{C} \times \boldsymbol{D}) + (\boldsymbol{B}.\boldsymbol{C}).(\boldsymbol{A} \times \boldsymbol{D}) + (\boldsymbol{C} \times \boldsymbol{A}).(\boldsymbol{B} \times \boldsymbol{D}) = 0$

10) If ***A***,***B***,***C***,***D*** are such that $\boldsymbol{A} \times \boldsymbol{C} = \boldsymbol{B} \times \boldsymbol{D}$ and $\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{C} \times \boldsymbol{D}$, then show that the vectors $(\boldsymbol{A} - \boldsymbol{D})$ and $(\boldsymbol{B} - \boldsymbol{C})$ are collinear.

Solution of exercise of art 4.6 to 4.7 :

Solution (1) :

Required unit vector $= \dfrac{\boldsymbol{A}_1 + \boldsymbol{A}_2}{|\boldsymbol{A}_1 + \boldsymbol{A}_2|}$

Now $\boldsymbol{A}_1 + \boldsymbol{A}_2 = 3\boldsymbol{i} + 6\boldsymbol{j} - 2\boldsymbol{k}$

$$\therefore |\boldsymbol{A}_1 + \boldsymbol{A}_2| = \sqrt{3^2 + 6^2 + (-2)^2} = \sqrt{49} = 7$$

$\therefore$ required unit vector $= \dfrac{3i + 6j - 2k}{7}$

Solution (2) : The centroid $\boldsymbol{G}$ of ΔABC divides the median of the triangle in the ratio 2:1

Now position vector of $D = \dfrac{\boldsymbol{b}+\boldsymbol{c}}{2}$

Therefore position vector of

$$\boldsymbol{G} = \frac{1\boldsymbol{a} + \frac{2(\boldsymbol{b}+\boldsymbol{c})}{2}}{1+2} = \frac{\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}}{3}$$

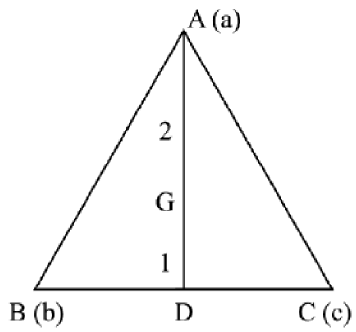


Fig. Solution (2)

Solution (3) : We have $|\boldsymbol{A} + \boldsymbol{B}| = |\boldsymbol{A} - \boldsymbol{B}|$

Squaring both sides,

$$|\boldsymbol{A} + \boldsymbol{B}|^2 = |\boldsymbol{A} - \boldsymbol{B}|^2$$

Or $\quad \boldsymbol{A}^2 + \boldsymbol{B}^2 + 2\boldsymbol{A}.\boldsymbol{B} = \boldsymbol{A}^2 + \boldsymbol{B}^2 - 2\boldsymbol{A}.\boldsymbol{B}$

Or, $\quad 4\boldsymbol{A}.\boldsymbol{B} = 0$

$\boldsymbol{A}.\boldsymbol{B} = 0$ i.e. $\boldsymbol{A}$ is perpendicular to $\boldsymbol{B}$ and *vice-versa*.

Solution (4) :

If $\boldsymbol{A}.\boldsymbol{B} = \boldsymbol{A}.\boldsymbol{C}$, then $\boldsymbol{A}.(\boldsymbol{B} - \boldsymbol{C}) = 0$, i.e. either $\boldsymbol{A}$ is perpendicular to $\boldsymbol{B} - \boldsymbol{C}$ or $\boldsymbol{B} - \boldsymbol{C} = \boldsymbol{0}$ where $\boldsymbol{0}$ is a null vector. Hence, it does not necessarily follows that $\boldsymbol{B}$ and $\boldsymbol{C}$ are equal.

Solution (5) :

We have $|\boldsymbol{A}| = |\boldsymbol{B}|$

Or, $|\boldsymbol{A}|^2 = |\boldsymbol{B}|^2$; or $\boldsymbol{A}.\boldsymbol{A} = \boldsymbol{B}.\boldsymbol{B}$

Now, $(\boldsymbol{A} + \boldsymbol{B}).(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{A}.\boldsymbol{A} - \boldsymbol{A}.\boldsymbol{B} + \boldsymbol{B}.\boldsymbol{A} - \boldsymbol{B}.\boldsymbol{B} = 0$

Since $\boldsymbol{A}.\boldsymbol{A} = \boldsymbol{B}.\boldsymbol{B}$

$\therefore$ $\boldsymbol{A} + \boldsymbol{B}$ is perpendicular to $\boldsymbol{A} - \boldsymbol{B}$.

Solution (6) :

The cosine of the angles made by a vector with X, Y and Z axes of a rectangular Cartesian co-ordinate axes, are called direction cosine l, m, n of the vector respectively. If the vector makes angle α, β and γ with respect to X, Y and Z axes respectively, then $l = \cos\alpha$, $m = \cos\beta$ and $n = \cos\gamma$.

Let $\boldsymbol{A} = \boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z$, then $\dfrac{\boldsymbol{A}}{|\boldsymbol{A}|} = \boldsymbol{i}\dfrac{A_x}{|\boldsymbol{A}|} + \boldsymbol{j}\dfrac{A_y}{|\boldsymbol{A}|} + \boldsymbol{k}\dfrac{A_z}{|\boldsymbol{A}|} = \boldsymbol{i}l + \boldsymbol{j}m + \boldsymbol{k}n$ (i)

Now $\cos\alpha = \dfrac{A_x}{|\boldsymbol{A}|}$; $\cos\beta = \dfrac{A_y}{|\boldsymbol{A}|}$; $\cos\gamma = \dfrac{A_z}{|\boldsymbol{A}|}$

$$\therefore \cos^2\alpha + \cos^2\beta + \cos^2\gamma = \frac{\left(A_x^2 + A_y^2 + A_z^2\right)}{|\boldsymbol{A}|^2} = \frac{|\boldsymbol{A}|^2}{|\boldsymbol{A}|^2} = 1 \qquad \text{(ii)}$$

If there are two vectors $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, then from equation (i),

$$\frac{\boldsymbol{A}_1}{|\boldsymbol{A}_1|} = \boldsymbol{i}l_1 + \boldsymbol{j}m_1 + \boldsymbol{k}n_1 \qquad \text{(iii)}$$

where l_1, m_1, n_1 are direction cosines of the vector $\boldsymbol{A}_1$.

and $$\frac{\boldsymbol{A}_2}{|\boldsymbol{A}_2|} = \boldsymbol{i}l_2 + \boldsymbol{j}m_2 + \boldsymbol{k}n_2 \qquad \text{(iv)}$$

where l_2, m_2, n_2 are direction cosine of the vector $\boldsymbol{A}_2$.

Now if θ be the angle between $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, we have

$$\cos\theta = \frac{\boldsymbol{A}_1.\boldsymbol{A}_2}{|\boldsymbol{A}_1||\boldsymbol{A}_2|} = \frac{|\boldsymbol{A}_1|\left(\boldsymbol{i}l_1 + \boldsymbol{j}m_1 + \boldsymbol{k}n_1\right)\cdot|\boldsymbol{A}_2|\left(\boldsymbol{i}l_2 + \boldsymbol{j}m_2 + \boldsymbol{k}n_2\right)}{|\boldsymbol{A}_1|.|\boldsymbol{A}_2|}$$

using equations (iii) and (iv).

$$\therefore \cos\theta = l_1 l_2 + m_1 m_2 + n_1 n_2$$

Solution (7) :

a) We have $\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C} = 0$, then $\boldsymbol{A} + \boldsymbol{B} = -\boldsymbol{C}$, now $\boldsymbol{B} \times (\boldsymbol{A} + \boldsymbol{B}) = -\boldsymbol{B} \times \boldsymbol{C}$

Or, $\boldsymbol{B} \times \boldsymbol{A} + \boldsymbol{B} \times \boldsymbol{B} = -\boldsymbol{B} \times \boldsymbol{C}$, or, $-\boldsymbol{A} \times \boldsymbol{B} = -\boldsymbol{B} \times \boldsymbol{C}$

Or, $\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{B} \times \boldsymbol{C}$ (i)

Similarly, $\boldsymbol{B} \times \boldsymbol{C} = \boldsymbol{C} \times \boldsymbol{A}$ (ii)

Therefore, $\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{B} \times \boldsymbol{C} = \boldsymbol{C} \times \boldsymbol{A}$

b) We have $\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{B} \times \boldsymbol{C}$

$\therefore$ $-\boldsymbol{A} \times \boldsymbol{B} = -\boldsymbol{B} \times \boldsymbol{C}$ $\therefore$ $\boldsymbol{B} \times \boldsymbol{A} = \boldsymbol{C} \times \boldsymbol{B}$

or, $\boldsymbol{B} \times \boldsymbol{A} + \boldsymbol{B} \times \boldsymbol{B} = \boldsymbol{C} \times \boldsymbol{B} = -\boldsymbol{B} \times \boldsymbol{C}$

or, $\boldsymbol{B} \times \boldsymbol{A} + \boldsymbol{B} \times \boldsymbol{B} + \boldsymbol{B} \times \boldsymbol{C} = 0$

or, $\boldsymbol{B} \times (\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C}) = 0$ now $\boldsymbol{B} \neq 0$

therefore $\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C} = 0$

Solution (8) :

Let $\overrightarrow{OA} = \boldsymbol{H}, \overrightarrow{OB} = \boldsymbol{G}$ and OM, the projection of $\boldsymbol{G}$ on $\boldsymbol{H}$.

The component of $\boldsymbol{G}$ along $\boldsymbol{H}$ $= G \cdot \dfrac{\boldsymbol{H}}{|\boldsymbol{H}|}$

where $\dfrac{\boldsymbol{H}}{|\boldsymbol{H}|}$ is the unit vector along $\overrightarrow{OA}$. Therefore vector component $\boldsymbol{G}$ along $\boldsymbol{H}$ is

$$\left(\boldsymbol{G} \cdot \frac{\boldsymbol{H}}{|\boldsymbol{H}|}\right)\frac{\boldsymbol{H}}{|\boldsymbol{H}|} = \frac{(\boldsymbol{G}.\boldsymbol{H})\boldsymbol{H}}{|\boldsymbol{H}|^2},$$ vector

component of $\boldsymbol{G}$ perpendicular to $\boldsymbol{H}$

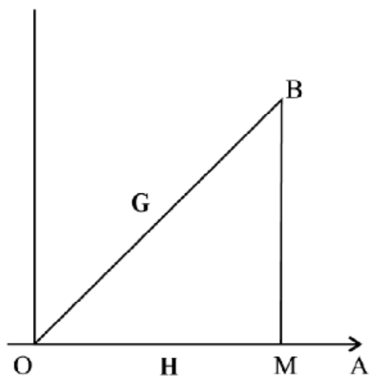


Fig. Solution (8)

$$= \overrightarrow{OB} - \overrightarrow{OM} = \boldsymbol{G} - \frac{(\boldsymbol{G}.\boldsymbol{H})\boldsymbol{H}}{|\boldsymbol{H}|^2} = \frac{(\boldsymbol{H}.\boldsymbol{H})\boldsymbol{G} - (\boldsymbol{G}.\boldsymbol{H})\boldsymbol{H}}{|\boldsymbol{H}|^2} = \frac{(\boldsymbol{H} \times \boldsymbol{G}) \times \boldsymbol{H}}{|\boldsymbol{H}|^2}$$

Solution (9) :

a) Let $\boldsymbol{P} = \boldsymbol{A} \times \boldsymbol{B}$, then $(\boldsymbol{A} \times \boldsymbol{B}).(\boldsymbol{C} \times \boldsymbol{D}) = \boldsymbol{P}.(\boldsymbol{C} \times \boldsymbol{D})$

Now $\quad \boldsymbol{P}.(\boldsymbol{C} \times \boldsymbol{D}) = (\boldsymbol{C} \times \boldsymbol{D}).\boldsymbol{P} = \boldsymbol{C}.\boldsymbol{D} \times \boldsymbol{P} = \boldsymbol{C}.\boldsymbol{D} \times (\boldsymbol{A} \times \boldsymbol{B})$

$= \boldsymbol{C}.[(\boldsymbol{D}.\boldsymbol{B})\boldsymbol{A} - (\boldsymbol{D}.\boldsymbol{A})\boldsymbol{B}] = (\boldsymbol{C}.\boldsymbol{A})(\boldsymbol{D}.\boldsymbol{B}) - (\boldsymbol{C}.\boldsymbol{B})(\boldsymbol{D}.\boldsymbol{A})$

$= (\boldsymbol{A}.\boldsymbol{C})(\boldsymbol{B}.\boldsymbol{D}) - (\boldsymbol{A}.\boldsymbol{D})(\boldsymbol{B}.\boldsymbol{C})$

b) Now, $(\boldsymbol{A} \times \boldsymbol{B}).(\boldsymbol{C} \times \boldsymbol{D}) = (\boldsymbol{A}.\boldsymbol{C})(\boldsymbol{B}.\boldsymbol{D}) - (\boldsymbol{A}.\boldsymbol{D})(\boldsymbol{B}.\boldsymbol{C})$

Similarly $(\boldsymbol{B} \times \boldsymbol{C}).(\boldsymbol{A} \times \boldsymbol{D}) = (\boldsymbol{B}.\boldsymbol{A})\ (\boldsymbol{C}.\boldsymbol{D}) - (\boldsymbol{B}.\boldsymbol{D})\ (\boldsymbol{A}.\boldsymbol{C})$

$(\boldsymbol{C} \times \boldsymbol{A}).(\boldsymbol{B} \times \boldsymbol{D}) = (\boldsymbol{C}.\boldsymbol{B})\ (\boldsymbol{A}.\boldsymbol{D}) - (\boldsymbol{C}.\boldsymbol{D})\ (\boldsymbol{A}.\boldsymbol{B})$

$\therefore \quad (\boldsymbol{A} \times \boldsymbol{B}).(\boldsymbol{C} \times \boldsymbol{D}) + (\boldsymbol{B} \times \boldsymbol{C}).(\boldsymbol{A} \times \boldsymbol{D}) + (\boldsymbol{C} \times \boldsymbol{A}).(\boldsymbol{B} \times \boldsymbol{D}) = 0$

Solution (10) :

We have

$(\boldsymbol{A} - \boldsymbol{D}) \times (\boldsymbol{B} - \boldsymbol{C}) = (\boldsymbol{A} - \boldsymbol{D}) \times \boldsymbol{B} - (\boldsymbol{A} \times \boldsymbol{D}) \times \boldsymbol{C}$

$= \boldsymbol{A} \times \boldsymbol{B} - \boldsymbol{D} \times \boldsymbol{B} - \boldsymbol{A} \times \boldsymbol{C} + \boldsymbol{D} \times \boldsymbol{C}$

$= \boldsymbol{A} \times \boldsymbol{B} + \boldsymbol{B} \times \boldsymbol{D} - \boldsymbol{A} \times \boldsymbol{C} - \boldsymbol{C} \times \boldsymbol{D} = 0$

Therefore $(\boldsymbol{A} - \boldsymbol{D})$ and $(\boldsymbol{B} - \boldsymbol{C})$ is collinear.

4.8 ❑ Polar, Axial Vectors and Pseudo Scalars

A polar vector remains invariant under reversal of co-ordinate axes from x to $-x$, y to $-y$ and z to $-z$, since, the displacement vector $\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z = -\boldsymbol{i}(-x) - \boldsymbol{j}(-y) - \boldsymbol{k}(-z)$, on reversal of co-ordinate axes.

Let us consider the reflection of OX axis by a mirror as shown in the figure (4A.12).

On mirror reflection OX is to the left, OY remains unchanged.

Therefore $\boldsymbol{d} = \boldsymbol{i}x + \boldsymbol{j}y$ (before reflection)

And $\boldsymbol{d} = -\boldsymbol{i}(-x) + \boldsymbol{j}y = \boldsymbol{i}x + \boldsymbol{j}y$ (after reflection)

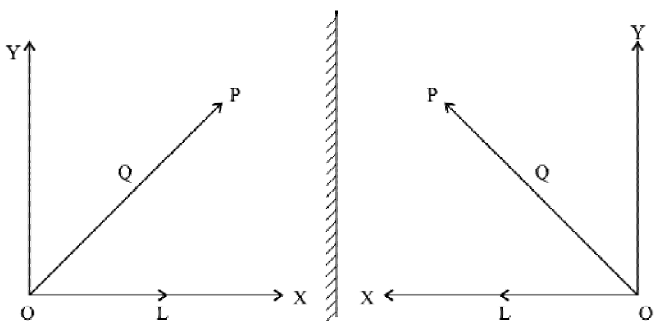


Fig. (4A.12)

Therefore displacement vector ***d*** remains invariant under mirror reflection. Such vectors as ***r*** or ***d*** are called polar vectors or true vectors.

Again in polar vectors only linear action in a particular direction is involved and hence does not depend on the frame of reference. Example of polar vectors are force, linear momentum etc.

Now vectors such as angular velocity, angular momentum etc. which are defined in terms of cross or vector product of two vectors, are called axial vectors for, some kind of rotation about an axis is involved in these vectors. The sense of direction in these vectors depend on the handedness (right or left) of the reference frame.

In the figure (4A.13) below, we have shown that an axial vector reverse its direction on mirror reflection.

If ***A*** and ***B*** are two polar vectors then ***A*** × ***B*** must be an axial vector or pseudo vector.

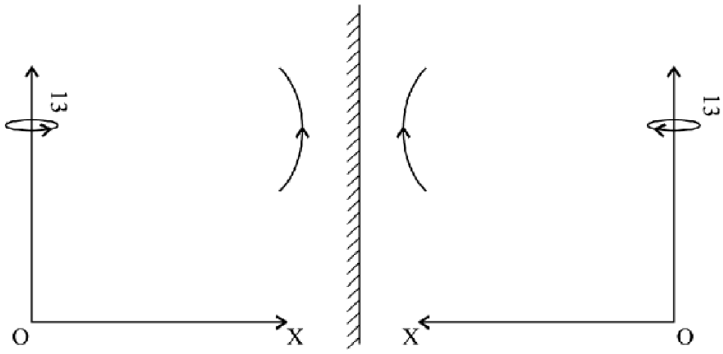


Fig. (4A.13)

Again for three polar vectors ***A,B,C*** the triple scalar product ***A***.***B*** ×***C*** changes its sign through reversal of axes *x* to –*x*, *y* to –*y* and *z* to –*z*. Such scalars are called pseudo scalars.

4.8.1 Scalar and vector fields

Physical entities may have different values at different points in a region of space and in this sense we can say that physical entities are functions of the space coordinates *x,y,z*. Suppose we have a physical quantity $\phi = \phi(x, y, z)$ so that ϕ is single valued, finite continuous function of *x,y,z* and possessing continuous first space derivatives, in the region under consideration. Then the region is called a field of ϕ. If ϕ is a scalar, then the field is called a scalar field and ϕ is called field scalar. Alternatively if $\vec{\phi}$ is a vector quantity, then the region is called a vector field and $\vec{\phi}$ is called the field vector.

Examples of scalar point functions are the temperature, electrostatic potential due to a charged body, gravitational potential energy of a massive body etc. and are called field scalars and the corresponding fields are called scalar fields.

In electric, magnetic and gravitational field, the intensity of the field in general varies from point to point and is function of the co-ordinates. Hence these intensities are field vectors and the corresponding fields are called vector fields.

It is to be noted that with the aid of certain differential operators, it is possible to associate a vector field with each scalar field. This association is of fundamental importance in mathematical physics.

A scalar field may be drawn geometrically by a series of surfaces on which the field scalar does not vary e.g. isothermal surfaces, equipotential surfaces etc. On which temperature, potential remain constant respectively. Such surfaces are called level surfaces. Obviously level surfaces cannot intersect each other, for if they do, there will be two values of ϕ at their common line of intersection which contradicts the very definition of scalar field.

4.8.2 Classification of vector fields

A vector field $\boldsymbol{A}$ is characterised by its divergence and curl and the field is determined completely, if its divergence and curl are known.

Absence or presence of curl and divergence of a vector field can be pictorially represented as follows :

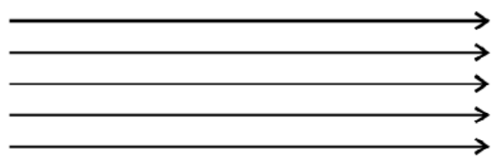


Fig. (4A.14A)

In case of parallel flow of an incompressible field with constant velocity as shown in (fig.4A.14A), div $\boldsymbol{A}$ and curl $\boldsymbol{A}$ are both zero. Such vector fields are called solenoidal and irrotational respectively.

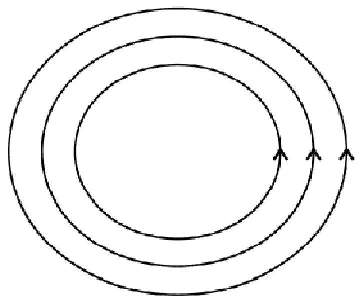


Fig. (4A.14B)

Vortex as shown in (fig.4A.14B) is formed in a moving field where curl $\boldsymbol{A} \neq 0$ at the centre of such vortex and the vector field is characterised as rotational. In this case net inward or outward flow is zero and div $\boldsymbol{A}$ = 0 and we call the vector field as solenoidal.

In case fluid is compressible, there can be excess of outflow over the inward in addition to the flow being rotational. This is shown in (fig. 4A.14C). In this case div $\boldsymbol{A} \neq 0$ and curl $\boldsymbol{A} \neq 0$.

Again when there is no rotation of a compressible fluid as shown in fig. (4A.14D,14E), there can be excess of outward flow over the inward [fig. 4A.14D] or excess of inward flow over the outward [fig. 4A.14E], we say curl $\boldsymbol{A}$ = 0, but div $\boldsymbol{A} \neq 0$.

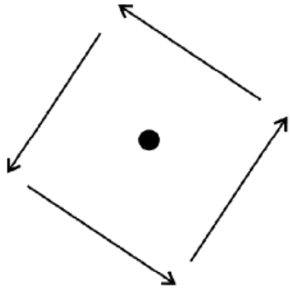


Fig. (4A.14C)

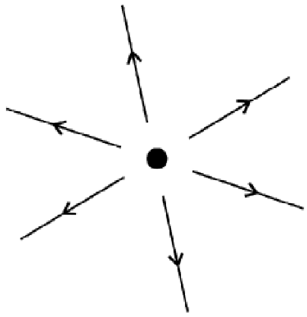


Fig. (4A.14D)

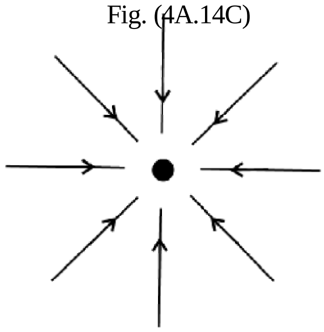


Fig. (4A.14E)

Now when the vector field is irrotational i.e. $\nabla \times \boldsymbol{A} = 0$, $\boldsymbol{A} = \nabla \phi$; and when it is solenoidal i.e. $\nabla \cdot \boldsymbol{A} = 0$, or $\nabla \cdot \nabla \phi = 0$ or $\nabla^2 \phi = 0$ (Laplace equation).

When $\nabla \times \boldsymbol{A} = 0$ but $\nabla \cdot \boldsymbol{A} \neq 0$, we get $\nabla \cdot \nabla \phi \neq 0$ or $\nabla^2 \phi \neq 0$ (Poisson's equation)

4.9 ❐ Summary - I

1) Invariance properties of scalar vectors are discussed.
2) Various types of product of vectors are discussed with reference to the example of machanics.
3) Use of kronecker delta, Levicivita symbol and classification vector field discussed for curious students.

4.10 ❐ Vector Differentiation

From the definition of derivatives applied to vector functions, different space and time derivative of vectors(with their physical meaning) have been explained and to obtain idea of special type of vector differential operators, for example, del or nabla applied to scalar and vector functions like gradient, divergence and curl. Also an important objective of this chapter is to fully realise the physical or geometrical interpretation of gradient, divergence and curl. Different vector identities are listed with or without del.

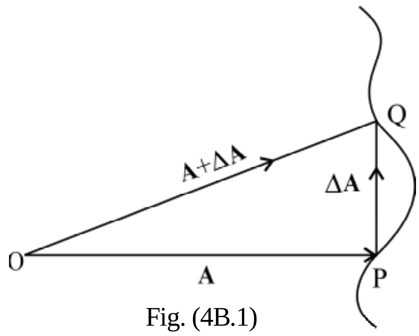


Fig. (4B.1)

Operations of differentiation of vectors are important in the sense that this concept is necessary for defining the various operators useful in vector analysis.

Consider a vector $\boldsymbol{A}(u)$ which is a continuous function of a continuous scalar variable u. As u change, a curve is traced by the terminus of $\boldsymbol{A}(u)$ (fig. 4B.1). In analogy with scalar functions, we define the derivative $\frac{d\boldsymbol{A}}{du}$ as

$$\frac{d\boldsymbol{A}}{du} = \lim_{\Delta u \to 0}\left(\frac{\Delta \boldsymbol{A}}{\Delta u}\right) = \lim_{\Delta u \to 0}\frac{\boldsymbol{A}(u+\Delta u)-\boldsymbol{A}(u)}{\Delta u} \tag{4B.1}$$

The derivative $\frac{d\boldsymbol{A}}{du}$ is a vector.

whose direction is the limiting direction of $\Delta\boldsymbol{A}$ as $\Delta u \to 0$. That is the direction of the derivative lies along the tangent to the curve at the point P as Δu $\to$ 0 in the sense of increasing u.

If $\boldsymbol{r}(u)$ be the position vector of the point $P(x,y,z)$ with respect to a set of rectangular axes with origin O, then $\boldsymbol{r}(u) = \boldsymbol{i}x(u) + \boldsymbol{j}y(u) + \boldsymbol{k}z(u)$. And $\frac{d\boldsymbol{r}}{du} = \lim_{\Delta u \to 0}\frac{\Delta \boldsymbol{r}}{\Delta u}$ $= \lim_{\Delta u \to 0}\frac{\boldsymbol{r}(u+\Delta u)-\boldsymbol{r}(u)}{\Delta u}$.

This $\frac{d\boldsymbol{r}}{du}$ is a vector in the direction of the tangent to the space curve at (x, y, z) and is given by

$$\frac{d\boldsymbol{r}}{du} = \boldsymbol{i}\left(\frac{dx}{du}\right) + \boldsymbol{j}\left(\frac{dy}{du}\right) + \boldsymbol{k}\left(\frac{dz}{du}\right) \tag{4B.2}$$

The derivative of a constant vector is a null vector.

From equation (4B.2) we see that derivative of a vector ***r*** means a vector whose components are the derivatives of the components of ***r***, since, ***i, j, k***, the unit basis vectors in rectangular Cartesian system are constant in magnitude and direction. But in other co-ordinate system, for example plane polar co-ordinate in two dimensions and spherical or cylindrical co-ordinate system in three dimensions, the unit basis vectors change with direction though their magnitudes are constant. Therefore in calculating the derivative of a vector written in these co-ordinate systems, we must differentiate the basis vectors as well as the components.

Now consider the vector ***A*** is a derivable function of other scalar s and *s* is a derivable function of another scalar *u*, then $\frac{d\boldsymbol{A}}{du} = \lim_{\Delta u \to 0} \frac{\Delta \boldsymbol{A}}{\Delta s} \cdot \frac{\Delta s}{\Delta u} = \frac{d\boldsymbol{A}}{ds} \cdot \frac{ds}{du}$

As $\frac{d\boldsymbol{A}}{du}$ is a vector and in general a function of *u*, we may find $\frac{d^2\boldsymbol{A}}{du^2}, \frac{d^3\boldsymbol{A}}{du^3}$ etc.

Rules for differentiation :

If ***A*** and ***B*** be two derivable vectors, each being function of the scalar variable *u* and *s* be another scalar. Then

i> $\frac{d}{du}[\boldsymbol{A} \pm \boldsymbol{B}] = \frac{d\boldsymbol{A}}{du} \pm \frac{d\boldsymbol{B}}{du}$

ii> $\frac{d}{du}(s\boldsymbol{A}) = s\frac{d\boldsymbol{A}}{du} + \boldsymbol{A}\frac{ds}{du}$

iii> $\frac{d}{du}(\boldsymbol{A} \cdot \boldsymbol{B}) = \boldsymbol{A} \cdot \frac{d\boldsymbol{B}}{du} + \boldsymbol{B} \cdot \frac{d\boldsymbol{A}}{du}$ [dot product being associated, the order of the vectors, may be changed].

iv> $\frac{d}{du}(\boldsymbol{A} \times \boldsymbol{B}) = \frac{d\boldsymbol{A}}{du} \times \boldsymbol{B} + \boldsymbol{A} \times \frac{d\boldsymbol{B}}{du}$ [vector products does not obey commutative law, the order of the vectors cannot be changed].

v> $\frac{d}{du}\left(\boldsymbol{A} \times \frac{d\boldsymbol{A}}{du}\right) = \frac{d\boldsymbol{A}}{du} \times \frac{d\boldsymbol{A}}{du} + \boldsymbol{A} \times \frac{d^2\boldsymbol{A}}{du^2}$, since cross product of equal vectors is zero.

4.11 ❒ Constant vector function: Constancy in direction and magnitude

4..11.1 : The necesssary and sufficient condition for a vector function $\boldsymbol{A}(u)$ to be a constant is $\frac{d}{du}\boldsymbol{A}(u) = 0$

Proof : Condition is necessary

If $\boldsymbol{A}(u)$ is constant, then for every change in Δu of the scalar variable u

$$\boldsymbol{A}(u+\Delta u) - \boldsymbol{A}(u) = 0$$

$$\therefore \lim_{\Delta u \to 0} \frac{\boldsymbol{A}(u+\Delta u) - \boldsymbol{A}(u)}{\boldsymbol{A}(u)} = \frac{d\boldsymbol{A}}{du} = 0$$

Condition is sufficient :

Consider $\boldsymbol{A}(u) = \boldsymbol{i}A_1(u) + \boldsymbol{j}A_2(u) + \boldsymbol{k}A_3(u)$

where A_1, A_2, A_3 are three scalar functions of u.

Then $\frac{d\boldsymbol{A}}{du} = \boldsymbol{i}\frac{dA_1}{du} + \boldsymbol{j}\frac{dA_2}{du} + \boldsymbol{k}\frac{dA_3}{du} = 0$

Implies $\frac{dA_1}{du} = 0$; $\frac{dA_2}{du} = 0$; and $\frac{dA_3}{du} = 0$

Therefore A_1, A_2, A_3 are constant and hence $\boldsymbol{A}(u)$ is a constant vector. Thus the condition is sufficient.

4.11.2 : The necessary and sufficient condition for a vector $A(u)$ to have constant magnitude is A. $\frac{d\boldsymbol{A}}{du} = 0$

Proof : Condition is necessary :

We have $\boldsymbol{A}(u).\boldsymbol{A}(u) = |\boldsymbol{A}(u)^2|$

Therefore $\boldsymbol{A}(u).\frac{d\boldsymbol{A}(u)}{dt} = |\boldsymbol{A}(u)| \cdot \frac{d}{du}|\boldsymbol{A}(u)|$

And $\quad \boldsymbol{A}(u).\dfrac{d\boldsymbol{A}(u)}{dt}=0$ if and only if $\dfrac{d}{du}\left|\boldsymbol{A}(u)\right|=0$

Or $\quad \boldsymbol{A}(u) = \textit{constant}.$

Therefore condition is necessary.

Condition is sufficient :

Assume that $\boldsymbol{A}(u)$ has a constant magnitude $|\boldsymbol{A}(u)|$. Then definitely $\dfrac{d\boldsymbol{A}(u)}{du}=0$. So that, $\boldsymbol{A}(u).\dfrac{d\boldsymbol{A}(u)}{du}=0$. Thus the condition is also sufficient.

4.11.3 The condition for a vector $\boldsymbol{A}(u)$ to have constant direction is $\boldsymbol{A}(u)\times\dfrac{d\boldsymbol{A}(u)}{du}=0$

Proof : Consider $\hat{a}(u)$ to be a unit vector in the direction of $\boldsymbol{A}(u)$, the $\boldsymbol{A}(u)$ = $\left|\boldsymbol{A}(u)\right|\hat{\boldsymbol{a}}(u)=A(u)\hat{\boldsymbol{a}}(u)=A\hat{\boldsymbol{a}}$

Therefore $\dfrac{d\boldsymbol{A}(u)}{du}=\dfrac{dA(u)}{du}\hat{\boldsymbol{a}}(u)+\dfrac{d\hat{\boldsymbol{a}}(u)}{du}A(u)$

$$\therefore \boldsymbol{A}\times\frac{d\boldsymbol{A}}{du}=\boldsymbol{A}\times\left[\frac{dA}{du}\hat{\boldsymbol{a}}+\frac{d\hat{\boldsymbol{a}}}{du}A\right]$$

$$=A\hat{\boldsymbol{a}}\times\left[\frac{dA}{du}\hat{\boldsymbol{a}}+\frac{d\hat{\boldsymbol{a}}}{du}A\right]$$

$$=A\frac{dA}{du}\hat{\boldsymbol{a}}\times\hat{\boldsymbol{a}}+A^2\hat{\boldsymbol{a}}\times\frac{d\hat{\boldsymbol{a}}}{du}=0$$

Now $\hat{\boldsymbol{a}}(u)$ is constant, $\dfrac{d\hat{\boldsymbol{a}}}{du}=0$; and $\hat{\boldsymbol{a}}\times\hat{\boldsymbol{a}}=0$

4.12 ❒ Derivative of Triple Scalar Product

Consider S = ***A.*** ***B*** × ***C***, where ***A***, ***B*** and ***C*** are vector functions of the scalar variable *u*.

Then $\quad \dfrac{dS}{du}=\dfrac{d}{du}\ [\boldsymbol{A}.\ \boldsymbol{B}\times\boldsymbol{C}]$

$$= \frac{d\boldsymbol{A}}{du} \cdot \boldsymbol{B} \times \boldsymbol{C} + \boldsymbol{A} \cdot \frac{d\boldsymbol{B}}{du} \times \boldsymbol{C} + \boldsymbol{A}.\boldsymbol{B} \times \frac{d\boldsymbol{C}}{du}$$

The cyclic order of the factor on each term is maintained.

4.13 ❐ Derivative of Triple Vector Product

Consider S = ***A*** × (***B*** × ***C***), where ***A***, ***B***, ***C*** are vector functions of the scalar variable *u*.

Then $\frac{dS}{du} = \frac{d\boldsymbol{A}}{du} \times (\boldsymbol{B} \times \boldsymbol{C}) + \boldsymbol{A} \times \left(\frac{d\boldsymbol{B}}{du} \times \boldsymbol{C} \right) + \boldsymbol{A} \times \left(\boldsymbol{B} \times \frac{d\boldsymbol{C}}{du} \right)$

The order of the factor in each term is maintained.

4.14 ❐ Velocity and Acceleration of a Particle

The position vector of a particle at time *t* is given by

$$\boldsymbol{r}(t) = \boldsymbol{i}x(t) + \boldsymbol{j}y(t) + \boldsymbol{k}z(t)$$

The displacement $\Delta\boldsymbol{r}(t)$ in time Δt is given by

$$\Delta\boldsymbol{r}(t) = \boldsymbol{r}(t + \Delta t) - \boldsymbol{r}(t)$$

Now time rate of change of displacement of a particle is its velocity $\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt} = \lim_{\Delta t \to 0} \frac{\Delta\boldsymbol{r}}{\Delta t}$ and this velocity ***v*** is in the direction of the tangent to the path of the particle at (*x*, *y*, *z*) in time *t*.

The acceleration of the moving particle being the time rate of change of ***v*** and we have acceleration of the particle $\boldsymbol{f} = \lim_{\Delta t \to 0} \frac{\Delta\boldsymbol{v}}{\Delta t} = \frac{d\boldsymbol{v}}{dt} = \frac{d^2\boldsymbol{r}}{dt^2}$

Exxample of 4.4:

Example (1): A particle moving in a plane. Find the radial and transverse components of velocity and acceleration of the particle in plane polar co-ordinate.

Solution : At any time *t*, let the position vector of the particle at a point (r, θ), $\boldsymbol{r} = r\hat{\boldsymbol{r}} = \boldsymbol{i}x + \boldsymbol{j}y$; (i) or, $\boldsymbol{r} = r\hat{\boldsymbol{r}} = \boldsymbol{i}r\cos\theta + \boldsymbol{j}r\sin\theta = r(\boldsymbol{i}\cos\theta + \boldsymbol{j}\sin\theta)$, where $x = r\cos\theta;\ y = r\sin\theta$ and $\hat{\boldsymbol{r}}$ = unit vector in the direction vector $\boldsymbol{r} = \boldsymbol{i}\cos\theta + \boldsymbol{j}\sin\theta$; (i) $\hat{\boldsymbol{\theta}}$ is unit vector in a direction perpendicular to $\boldsymbol{r}$.

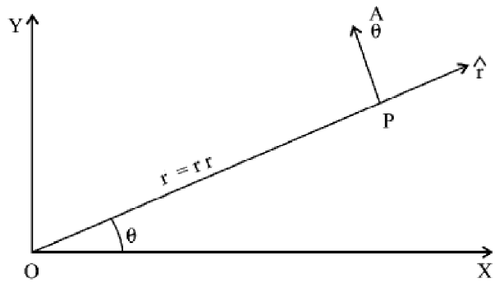


Fig. Solution 1

$$\therefore \hat{\boldsymbol{\theta}} = \boldsymbol{i}\cos\left(\theta + \frac{\pi}{2}\right) + \boldsymbol{j}\sin\left(\theta + \frac{\pi}{2}\right)$$

or, $$\hat{\boldsymbol{\theta}} = -\boldsymbol{i}\sin\theta + \boldsymbol{j}\cos\theta \qquad \text{(ii)}$$

Now the velocity of the particle is given by,

$$\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt} = \frac{d}{dt}(r\hat{\boldsymbol{r}}) = \frac{dr}{dt}\hat{\boldsymbol{r}} + r\frac{d\hat{\boldsymbol{r}}}{dt}$$

Now, $$\frac{d\hat{\boldsymbol{r}}}{dt} = (-\boldsymbol{i}\sin\theta + \boldsymbol{j}\cos\theta)\frac{d\theta}{dt} = \hat{\boldsymbol{\theta}}\frac{d\theta}{dt}$$

And $$\frac{d\hat{\boldsymbol{\theta}}}{dt} = (-\boldsymbol{i}\cos\theta - \boldsymbol{j}\cos\theta)\frac{d\theta}{dt} = -\hat{\boldsymbol{r}}\frac{d\theta}{dt}$$

Therefore $$\boldsymbol{v} = \frac{dr}{dt}\hat{\boldsymbol{r}} + r\frac{d\theta}{dt}\hat{\boldsymbol{\theta}} = \dot{r}\hat{\boldsymbol{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}}$$

where the dot (.) indicate differentiation with time.

Thus radial and transverse components of velocity are : $v_r = \dot{r};\ v_\theta = r\dot{\theta}$ so that

$$\boldsymbol{v} = v_r\hat{r} + v_\theta\hat{\boldsymbol{\theta}}$$

Acceleration of the particle is given by,

$$\boldsymbol{a} = \frac{d\boldsymbol{v}}{dt} = \frac{d}{dt}\left[\dot{r}\hat{\boldsymbol{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}}\right] = \ddot{r}\hat{\boldsymbol{r}} + \dot{r}\frac{d\hat{\boldsymbol{r}}}{dt} + \dot{r}\dot{\theta}\hat{\boldsymbol{\theta}} + r\ddot{\theta}\hat{\boldsymbol{\theta}} + r\dot{\theta}\frac{d\hat{\boldsymbol{\theta}}}{dt}$$

where $$\ddot{r} = \frac{d^2r}{dt^2} \text{ and } \ddot{\theta} = \frac{d^2\theta}{dt^2}.$$

Now substituting the values of $\frac{d\hat{\boldsymbol{r}}}{dt}$ and $\frac{d\hat{\boldsymbol{\theta}}}{dt}$, we get

$$a = \ddot{r}\hat{\boldsymbol{r}} + \dot{r}\frac{d\theta}{dt}\hat{\boldsymbol{\theta}} + \dot{r}\dot{\theta}\hat{\boldsymbol{\theta}} + r\ddot{\theta}\hat{\boldsymbol{\theta}} + r\dot{\theta}\left(-\hat{\boldsymbol{r}}\frac{d\theta}{dt}\right) = \left(\ddot{r} - r\dot{\theta}^2\right)\hat{\boldsymbol{r}} + \left(2\dot{r}\dot{\theta} + r\ddot{\theta}\right)\hat{\boldsymbol{\theta}} = a_r\hat{\boldsymbol{r}} + a_\theta\hat{\boldsymbol{\theta}}$$

where a_r is the radial component and a_θ is the transverse component of acceleration.

4.15 ❑ Relative Velocity and Acceleration

Consider two particles at P_1 and P_2 moving along the curve C_1 and C_2 and having respectively position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ at time t

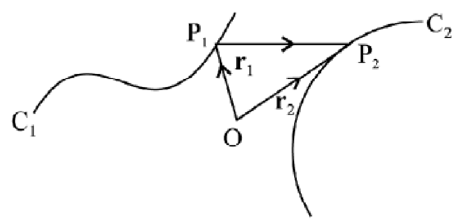


Fig. (4B.2)

Then $\overrightarrow{P_1P_2} = \boldsymbol{r}_2 - \boldsymbol{r}_1 = \boldsymbol{r}$

Differentiating with respect to time t, we get

$\frac{d\boldsymbol{r}}{dt} = \frac{d\boldsymbol{r}_2}{dt} - \frac{d\boldsymbol{r}_1}{dt}$ = relative velocity of particle P_2 with respect to that at P_1. Similarly relative velocity of the particle at P_1 with respect to that at P_2 is given by

$$\frac{d\boldsymbol{r}_1}{dt} - \frac{d\boldsymbol{r}_2}{dt} = -\frac{d\boldsymbol{r}}{dt}$$

Relative acceleration of the particle at P_2 with respect to P_1 is given by $\frac{d^2\boldsymbol{r}}{dt^2}$ and that of P_1 with respect to P_2 is $-\frac{d^2\boldsymbol{r}}{dt^2}$.

4.16 ❑ Gradient of a Scalar Field

The gradient of a scalar field $\phi(x, y, z)$ at a point (x_0, y_0, z_0) is a vector, denoted by the symbol $\nabla\phi$ (read "del ϕ") and is defined by

$$\text{grad } \phi = \nabla\phi = \boldsymbol{i}\left(\frac{\partial\phi}{\partial x}\right)_{x_0y_0z_0} + \boldsymbol{j}\left(\frac{\partial\phi}{\partial y}\right)_{x_0y_0z_0} + \boldsymbol{k}\left(\frac{\partial\phi}{\partial z}\right)_{x_0y_0z_0} \tag{4B.3}$$

where the operator 'del' or 'nabla' is a vector differential operator given by

$$\boldsymbol{\nabla} \equiv \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z} \tag{4B.4}$$

Let $\phi(\boldsymbol{r})$ be some field scalar at some point $\boldsymbol{r}(x, y, z)$ and $\phi(\boldsymbol{r} + d\boldsymbol{r})$ be the value of ϕ at $\boldsymbol{r} + d\boldsymbol{r}(x + dx, y + dy, z + dz)$; then $\phi(\boldsymbol{r} + d\boldsymbol{r}) - \phi(\boldsymbol{r}) = d\phi$

When $$d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz$$

Or, $$d\phi = \left(\boldsymbol{i}\frac{\partial\phi}{\partial x} + \boldsymbol{j}\frac{\partial\phi}{\partial y} + \boldsymbol{k}\frac{\partial\phi}{\partial z}\right).$$

$$= (\boldsymbol{i}dx + \boldsymbol{j}dy + \boldsymbol{k}dz)$$

$$= \nabla\phi.d\boldsymbol{r} = grad\,\phi.dr \qquad (4B.5)$$

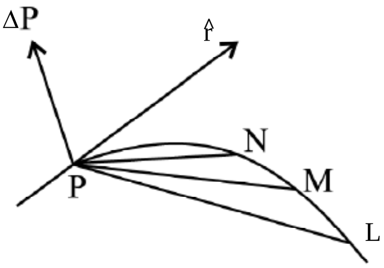


Fig. (4B.3)

Now the points (x, y, z) satisfying $\phi(x, y, z) = k$ = constant, in general defines a surface in region of space. This surface $\phi(x, y, z) = k$ is called a level surface, since at every point of the surface $\phi(x, y, z)$ has a value equal to k [fig. 4B.3].

Now differentiating equation (4B.5) with respect to r, we get $\frac{\partial\phi}{\partial r} = \nabla\phi.\hat{\boldsymbol{r}}$ (4B.6), where $\hat{\boldsymbol{r}}$ is a unit vector in the direction of $\boldsymbol{r}$.

Now suppose $\hat{\boldsymbol{r}}$ is tangent to the surface ϕ = constant at the point P. Consider $\frac{\Delta\phi}{\Delta r}$ for path PL, PM, PN etc approaching the tangent $\hat{\boldsymbol{r}}$. Since ϕ = constant on the surface and L, M, N, P etc are all on the surface, $\Delta\phi = 0$ and $\frac{\Delta\phi}{\Delta r} = 0$ for such path. But $\frac{\Delta\phi}{\Delta r}$ in the tangent direction is the limit of $\frac{\Delta\phi}{\Delta r}$ as $\Delta r \to 0$ (that is as PL, PM, PN etc approaches $\hat{\boldsymbol{r}}$, so $\frac{\Delta\phi}{\Delta r}$ in the direction $\hat{\boldsymbol{r}}$ is zero also).

Thus for $\hat{\boldsymbol{r}}$ along the tangent to ϕ = constant $\nabla\phi.\hat{\boldsymbol{r}} = 0$, this means that $\nabla\phi$ is perpendicular to $\hat{\boldsymbol{r}}$. Since this is true for any $\hat{\boldsymbol{r}}$ tangent to the surface at the point, then at that point, the vector $\nabla\phi$ is perpendicular to the level surface ϕ = constant.

Again from equation (4B.5), since ϕ = constant, $d\phi = 0 = \nabla\phi.\hat{\boldsymbol{r}}dr$. Since $dr \neq 0$, $\nabla\phi.\hat{\boldsymbol{r}} = 0$ i.e. $\nabla\phi$ is perpendicular to $\hat{\boldsymbol{r}}$.

4.17 ❒ Directional Derivative

Suppose rate of change $\phi(x, y, z)$ with distance is to be evaluated at a given point $P_0(x_0, y_0, z_0)$ and in a given directional $\overrightarrow{P_0Q}$ as shown in fig [4B.4]. Let R be the distance in the direction $\hat{\boldsymbol{R}}$ where $\hat{\boldsymbol{R}} = \boldsymbol{i}l + \boldsymbol{j}m + \boldsymbol{k}n$, is the unit vector in that direction; *l*, *m*, *n* being the direction cosine of the directed line $\overrightarrow{P_0Q}$. Then, $\boldsymbol{r} - \boldsymbol{r}_0 = R\,\hat{\boldsymbol{R}} = R(\boldsymbol{i}l + \boldsymbol{j}m + \boldsymbol{k}m)$ where (x, y, z) is the position co-ordinate of the terminus of the vector $\boldsymbol{R}$ and (x_0, y_0, z_0) is that of it's tail. Therefore, we write,

$$x = x_0 + Rl$$
$$y = y_0 + Rm \qquad (4B.7)$$
$$z = z_0 + Rn$$

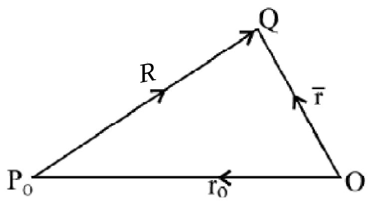


Fig. (4B.4)

From equation (4B.7) we see that $\phi\,(x, y, z)$ is function of just one variable *R*, the distance along the line measurement from (x_0, y_0, z_0). Then

$$\frac{d\phi}{dR} = \frac{\partial\phi}{\partial x}\frac{dx}{dR} + \frac{\partial\phi}{\partial y}\frac{dy}{dR} + \frac{\partial\phi}{\partial z}\frac{dz}{dR}$$

$$= l\frac{\partial\phi}{\partial x} + m\frac{\partial\phi}{\partial y} + n\frac{\partial\phi}{\partial z}$$

$$= \nabla\phi.\hat{\boldsymbol{R}} \qquad (4B.8)$$

where $\nabla\phi = \boldsymbol{i}\frac{\partial\phi}{\partial x} + \boldsymbol{j}\frac{\partial\phi}{\partial y} + \boldsymbol{k}\frac{\partial\phi}{\partial z}$

Equation (4B.8) gives the directional derivative

if we take $(x_0, y_0, z_0) = (0, 0, 0)$ i.e. the origin of co-ordinate system, then $\hat{\boldsymbol{R}} = \hat{\boldsymbol{r}}$ and vector $\boldsymbol{R}$ becomes positions vector $\boldsymbol{r}$ and therefore equation (4B.8) can be written as

$$\frac{d\phi}{dr} = \nabla\phi.\hat{\boldsymbol{r}} \qquad (4B.6)$$

4.18 ❒ Normal Derivative

We consider two neighbouring level surface defined by ϕ and $\phi + d\phi$. Shorted distance between surfaces at the point *P* is $dn = dr\cos\theta = \boldsymbol{dr}.\hat{\boldsymbol{n}}$ as shown in the figure (4B.6). Therefore $dr = \dfrac{dn}{\cos\theta}$.

Now rate of increase of ϕ at P in the direction $\overrightarrow{PQ}$ is $\dfrac{d\phi}{dr} = \dfrac{d\phi}{dn}\cos\theta$

Or, $\left(\dfrac{d\phi}{dr}\right)_{\max} = \dfrac{d\phi}{dn}$ [now as discussed previously $\left(\dfrac{\partial\phi}{\partial r}\right)_{\max} = |\nabla\phi|$]

since $|\nabla\phi|$ is the value of the directional derivative in the direction normal to the surface it is often called normal derivative and written $|\nabla\phi| = \dfrac{d\phi}{dn}$ (4B.12)

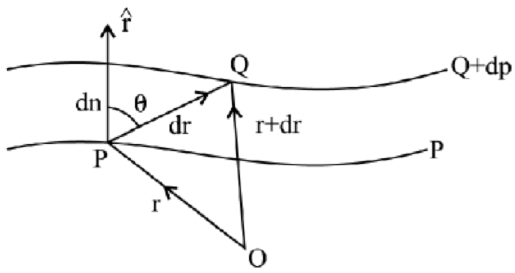


Fig. (4B.6)

Example of Art 4B.8 and 4B.9 :

Example 1: Find directional derivative of $\phi = x^2 - 2y^2 + 4z^2$ at the point (1, 1, – 1) in the direction $2\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k}$. In what direction the directional derivative at that point is maximum and what is its value ?

Solution 1 : $\nabla\phi = (2x\boldsymbol{i} - \boldsymbol{j}4y + \boldsymbol{k}8z) = 2\boldsymbol{i} - 4\boldsymbol{j} - 8\boldsymbol{k}$ at the point (1, 1, –1)

Unit vector along the direction $2\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k}$ is $\dfrac{1}{\sqrt{6}}(2\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k})$

$\therefore$ Required directional derivative is, $(2\boldsymbol{i} - 4\boldsymbol{j} - 8\boldsymbol{k}).\dfrac{1}{\sqrt{6}}(2\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k}) = \dfrac{4-4+8}{\sqrt{6}} = \dfrac{8}{\sqrt{6}}$

Directional derivative is maximum along $\nabla\phi$ and its maximum value is $|\nabla\phi| = \sqrt{84} = 2\sqrt{21}$

Example 2 : Find the equation of tangent plane and normal to the surface $z = x^2 + y^2$ at the point (2, – 1, 5).

Solution 2 : $\nabla\phi = \boldsymbol{i}2x + \boldsymbol{j}2y - \boldsymbol{k} = 4\boldsymbol{i} - 2\boldsymbol{j} - \boldsymbol{k}$ at (2, – 1, 5)

Position vector of the point (2, – 1, 5) is, $\boldsymbol{r}_0 = 2\boldsymbol{i} - \boldsymbol{j} + 5\boldsymbol{k}$

Therefore equation of tangent plane is given by $(\boldsymbol{r} - \boldsymbol{r}_0).\nabla\phi = 0$

Where $\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$ is any point on the tangent plane.

i.e. $(\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z - 2i + \boldsymbol{j} - 5\boldsymbol{k}).(4\boldsymbol{i} - 2\boldsymbol{j} - \boldsymbol{k}) = 0$

or, $(x - 2).4 - 2(y + 1) - (z - 5) = 0$

i.e. $4x - 2y - z = 5$

equation to the normal to the surface is $(\boldsymbol{r} - \boldsymbol{r}_0) \times \nabla\phi = 0$

where $\boldsymbol{r}$ is any point on the nromal.

i.e. $$\{(x-2)\boldsymbol{i} + (y+1)\boldsymbol{j} + (z-5)\boldsymbol{k}\} \times (4\boldsymbol{i} - 2\boldsymbol{j} - \boldsymbol{k}) = 0$$

or, $$\frac{x-2}{4} = \frac{y+1}{-2} = \frac{z-5}{-1}$$

Exercise of Art 4.14, 4.15 and 4.18 :

1) Find unit vector normal to the surface $x^2 + y - z = 1$ at the point (1, 0, 0).

2) For the function $\phi(x, y) = \dfrac{x}{x^2 + y^2}$, find the magnitude of the directional derivative along a line making an angle 30^0 with the positive x – axis at (0, 2).

3) The velocity of a boat relative to water is represented by $3\boldsymbol{i} + 4\boldsymbol{j}$, and that of water relative to earth is $\boldsymbol{i} - 3\boldsymbol{j}$. What is the velocity of the boat relative to earth, if $\boldsymbol{i}$ and $\boldsymbol{j}$ represent 1 km an hour east and north respectively.

4) Find the angle between the surface $\boldsymbol{x}^2 + y^2 + z^2 = 4$ and $z = x^2 + y^2 - 5$ at the point (1, –1, 2).

5) Find $\nabla\phi$, (i) when ϕ = ln $|\boldsymbol{r}|$ and (ii) prove that $\nabla r^n = nr^{n-2}\boldsymbol{r} = nr^{n-1}\hat{\boldsymbol{r}}$

6) Find the normal derivative of $f = xy + yz + zx$ at (1, 1, 3).

Solution :

Solution (1) : $\phi = (x, y, z) = x^2 + y^2 - z^2$

$$\therefore \nabla\phi = \boldsymbol{i}2x + \boldsymbol{j}2y - \boldsymbol{k}2z = 2\boldsymbol{i} + 2\boldsymbol{k} \text{ at the point } (1, 0, -1).$$

Now unit normal to $\phi = (x, y, z)$ at (1, 0, –1) is,

$$\hat{\boldsymbol{n}} = \frac{\nabla\phi(at\ the\ point\ (1,0,-1))}{|\nabla\phi|(at\ the\ point\ (1,0,-1))} = \frac{2(\boldsymbol{i}+\boldsymbol{k})}{\sqrt{2^2+2^2}} = \frac{\boldsymbol{i}+\boldsymbol{k}}{\sqrt{2}}$$

Solution 2 : We have $\phi(x, y) = \dfrac{x}{x^2+y^2}$

$$\therefore \nabla\phi = \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right)\left(\frac{x}{x^2+y^2}\right)$$

$$= \boldsymbol{i}\frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) + \boldsymbol{j}\frac{\partial}{\partial y}\left(\frac{x}{x^2+y^2}\right) + \boldsymbol{k}.0$$

$$= \boldsymbol{i}\left[\frac{1}{x^2+y^2} - \frac{x(2x)}{\left(x^2+y^2\right)^2}\right] - \boldsymbol{j}\frac{x(2y)}{\left(x^2+y^2\right)^2}$$

$\nabla\phi$ at the point (0, 2) is $\dfrac{i}{4}$.

Now as the line makes an angle 30^0 with the x – axis : $\boldsymbol{r} = \boldsymbol{i}(r \cos 30^0) + \boldsymbol{j}(r \sin 30^0)$

$$= r\left[\boldsymbol{i}\frac{\sqrt{3}}{2} + \frac{\boldsymbol{j}}{2}\right]$$

$$\therefore \frac{\boldsymbol{r}}{r} = \hat{\boldsymbol{r}} = \boldsymbol{i}\frac{\sqrt{3}}{2} + \frac{\boldsymbol{j}}{2}$$

$\therefore$ directional derivative along $\hat{\boldsymbol{r}}$ is,

$$\nabla\phi\Big|_{(0,2)}.\hat{\boldsymbol{r}} = \frac{\boldsymbol{i}}{4}\cdot\left(\boldsymbol{i}\frac{\sqrt{3}}{2} + \frac{\boldsymbol{j}}{2}\right) = \frac{\sqrt{3}}{8}$$

Solution 3 : We assume $\boldsymbol{v}_B$ & $\boldsymbol{v}_w$ be the velocity of the boat and that of water relative to earth respectively.

Therefore, $3\boldsymbol{i} + 4\boldsymbol{j} = \boldsymbol{v}_B - \boldsymbol{v}_w$, where $\boldsymbol{v}_w = \boldsymbol{i} - 3\boldsymbol{j}$

$$\therefore \boldsymbol{v}_B = 3\boldsymbol{i} + 4\boldsymbol{j} - \boldsymbol{v}_w = 3\boldsymbol{i} + 4\boldsymbol{j} + \boldsymbol{i} - 3\boldsymbol{j} = 4\boldsymbol{i} + \boldsymbol{j}$$

$$\therefore |\boldsymbol{v}_B| = \sqrt{4^2+1^2} = \sqrt{17} \text{ km/hr}$$

If θ be the angle between $\mathbf{v}_B$ and east direction then $\tan\theta = \frac{1}{4}$ or $\theta = \tan^{-1}\frac{1}{4}$

Solution 4 : We suppose $\phi_1(x, y, z) = x^2 + y^2 + z^2 - 4 = 0$ and $\phi_2(x, y, z)$ $= x^2 + y^2 - z^2 - 5 = 0$

Now, the angle between two surfaces is the angle between normals at a point $P(1, -1, 2)$.

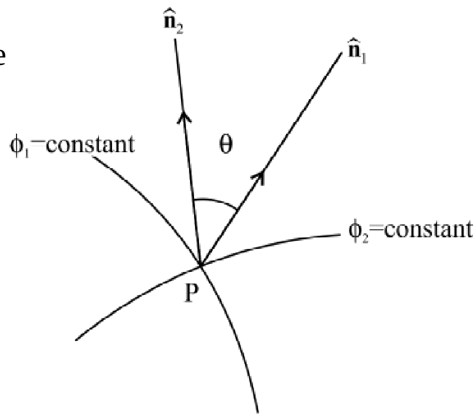


Fig. Solution (4)

Now, $\hat{\boldsymbol{n}}_1 = \dfrac{\nabla\phi_1}{|\nabla\phi_1|}$ at the point (1, –1, 2)

And $\hat{\boldsymbol{n}}_2 = \dfrac{\nabla\phi_2}{|\nabla\phi_2|}$ at the point (1, –1, 2)

$$\therefore \hat{\boldsymbol{n}}_1.\hat{\boldsymbol{n}}_2 = \cos\theta = \frac{\nabla\phi_1 \cdot \nabla\phi_2}{|\nabla\phi_1||\nabla\phi_2|} \text{ at } (1, -1, 2)$$

Where $\hat{\boldsymbol{n}}_1$ and $\hat{\boldsymbol{n}}_2$ are the unit normal on the surface ϕ_1 = constant and ϕ_2 = constant at the point P.

Now, $\nabla\phi_1\Big|_{(1,\,-1,2)} = 2\boldsymbol{i} - 2\boldsymbol{j} + 4\boldsymbol{k}$ and $\nabla\phi_2\Big|_{(1,\,-1,2)} = 2\boldsymbol{i} - 2\boldsymbol{j} - \boldsymbol{k}$

$$\therefore \cos\theta = \frac{(2i - 2j + 4k).(2i - 2j - k)}{\sqrt{2^2 + (-2)^2 + 4^2}\,\sqrt{2^2 + (-2)^2 + (-1)^2}} = \frac{4 + 4 - 4}{\sqrt{12}\sqrt{9}} = \frac{4}{6\sqrt{3}} = \frac{2}{3\sqrt{3}}$$

Solution 5 :

i) We have $\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$

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

$$\therefore \phi = \ln\sqrt{x^2 + y^2 + z^2} = \frac{1}{2}\ln\left(x^2 + y^2 + z^2\right)$$

$$\therefore \nabla\phi = \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right)\frac{1}{2}\ln\left(x^2 + y^2 + z^2\right)$$

$$= \frac{1}{2}\left[\frac{2x}{x^2+y^2+z^2}\boldsymbol{i} + \frac{2y}{x^2+y^2+z^2}\boldsymbol{j} + \frac{2z}{x^2+y^2+z^2}\boldsymbol{k}\right] = \frac{\boldsymbol{i}x+\boldsymbol{j}y+\boldsymbol{k}z}{x^2+y^2+z^2}$$

$$= \frac{\boldsymbol{r}}{r^2} = \frac{\hat{\boldsymbol{r}}}{r}$$

ii) We have $r^n = \left(x^2+y^2+z^2\right)^{\frac{n}{2}}$

$$\therefore \nabla r^n = \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right)\left(x^2+y^2+z^2\right)^{\frac{n}{2}}$$

$$= \boldsymbol{i}\frac{\partial}{\partial x}\left(x^2+y^2+z^2\right)^{\frac{n}{2}} + \boldsymbol{j}\frac{\partial}{\partial y}\left(x^2+y^2+z^2\right)^{\frac{n}{2}} + \boldsymbol{k}\frac{\partial}{\partial z}\left(x^2+y^2+z^2\right)^{\frac{n}{2}}$$

Or

$$\nabla r^n = \boldsymbol{i}\left[\frac{n}{2}\left(x^2+y^2+z^2\right)^{\frac{n}{2}-1}.2x\right] + \boldsymbol{j}\left[\frac{n}{2}\left(x^2+y^2+z^2\right)^{\frac{n}{2}-1}.2y\right]$$

$$+\boldsymbol{k}\left[\frac{n}{2}\left(x^2+y^2+z^2\right)^{\frac{n}{2}-1}.2z\right]$$

$$= n\left(x^2+y^2+z^2\right)^{\frac{n}{2}-1}(\boldsymbol{i}x+\boldsymbol{j}y+\boldsymbol{k}z)$$

$$= n\left(r^2\right)^{\frac{n}{2}-1}\boldsymbol{r} = nr^{n-2}\boldsymbol{r} = nr^{n-1}\hat{\boldsymbol{r}}$$

Solution 6 : From equation (4B.12), normal derivative $\frac{d\phi}{dn} = |\nabla\phi|$

Now $\phi = f = xy + yz + zx$

Now $\nabla f = \boldsymbol{i}\dfrac{\partial f}{\partial x} + \boldsymbol{j}\dfrac{\partial f}{\partial y} + \boldsymbol{k}\dfrac{\partial f}{\partial z} = \boldsymbol{i}(y+z) + \boldsymbol{j}(z+x) + \boldsymbol{k}(x+y)$

$\therefore \nabla f$ at the point (1, 1, 2) is $4\boldsymbol{i} + 4\boldsymbol{j} + 2\boldsymbol{k}$

$\therefore |\nabla f| = \sqrt{4^2+4^2+2^2} = \sqrt{16+16+4} = \sqrt{36} = 6$

4.19 Geometrical and Physical Meanings of Grad $\nabla\phi$:

From equations (4B.8) using the definition of dot product, since $|\hat{\boldsymbol{R}}| = 1$

$$\frac{d\phi}{dR} = |\nabla\phi||\hat{\boldsymbol{R}}|\cos\theta = |\nabla\phi|\cos\theta \tag{4B.7}$$

where θ is the angle between $\nabla\boldsymbol{\phi}$ & $\hat{\boldsymbol{R}}$. Therefore $\left(\dfrac{d\phi}{dR}\right)_{\max} = |\nabla\phi|$ i.e. maximum value of $\dfrac{d\phi}{dR}$ is $|\nabla\phi|$ and it is in the direction of $\nabla\phi$ [fig. 4B5)

When θ = 180°, we get largest of decrease of ϕ i.e. $\dfrac{d\phi}{dR} = -|\nabla\phi|$

When $\theta = \dfrac{\pi}{2}$, $\hat{\boldsymbol{R}}$, is tangent to the surface ϕ (x, y, z) = constant at the point P Fig. [4B.5] and $\nabla\phi\cdot\hat{\boldsymbol{R}} = 0$ i.e. $\nabla\phi$ is perpendicular to the tangent $\hat{\boldsymbol{R}}$ at the point P, since this is the true for any $\hat{\boldsymbol{R}}$ tangent to the surface at the point P then $\nabla\phi$ is perpendicular to the surface $\phi(x, y, z)$ constant.

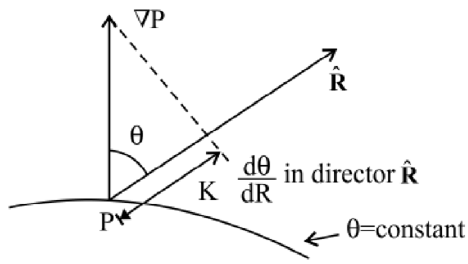


Fig. (4B.5)

4.20 : The 'Del' or 'Nabla' operator

When we write grad $\phi = \left[\boldsymbol{i}\dfrac{\partial}{\partial x} + \boldsymbol{j}\dfrac{\partial}{\partial y} + \boldsymbol{k}\dfrac{\partial}{\partial z}\right]\phi$, we call the bracket Del or Nabla or ∇. Thus 'Del' is a differential operator, has no meaning by itself like other operator, e.g. $\dfrac{d}{dx}$ or sine or $\log_e$ etc but has vector properties.

So far we have discussed $\nabla\phi$ where ϕ is a scalar. But ∇ can operate on vectors too.

Suppose $\boldsymbol{A}(x, y, z)$ is a continuously differentiable point function with components A_x, A_y and A_z in the X, Y and Z direction repectively and $\boldsymbol{A}$ can vary in magnitude and direction from point to point.

We can now form two useful combination of ∇ and $\boldsymbol{A}$. We define the divergence of $\boldsymbol{A}$ i.e. div $\boldsymbol{A}$ or $\nabla .\boldsymbol{A}$ by

$$\nabla \cdot \boldsymbol{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \tag{4B.9}$$

And curl $\boldsymbol{A}$ or rot $\boldsymbol{A}$ or $\nabla \times \boldsymbol{A}$ by

$$\nabla \times \boldsymbol{A} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}$$

$$= \boldsymbol{i}\left(\frac{\partial A_z}{\partial y} + \frac{\partial A_y}{\partial z}\right) + \boldsymbol{j}\left(\frac{\partial A_x}{\partial z} + \frac{\partial A_z}{\partial x}\right) + \boldsymbol{k}\left(\frac{\partial A_y}{\partial x} + \frac{\partial A_x}{\partial y}\right) \tag{4B.10}$$

Equation (4B.9) follows from the definition of dot product of two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ and equation (4B.10) follows from definition of cross product. So we can say that in the above two equation ∇ behave almsot like a vector.

Now $\nabla\phi$ is a vector function and we can write $\boldsymbol{A} = \nabla\phi$, where ϕ is a scalar function. Now $\nabla . \boldsymbol{A}$ becomes $\nabla \cdot \nabla\phi$ = div gard ϕ. This important expression is called Laplacian of ϕ and is written as

$$\nabla^2\phi = \frac{\partial}{\partial x}\left(\frac{\partial\phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial\phi}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial\phi}{\partial z}\right)$$

$$= \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} \tag{4B.11}$$

Several important equations in mathematical physics involving Laplacian are :

$$\nabla^2\phi = -\frac{\rho}{\varepsilon_0} \text{ (Poisson's equation)}$$

$$\nabla^2\phi = 0 \text{ (Laplace's equation)}$$

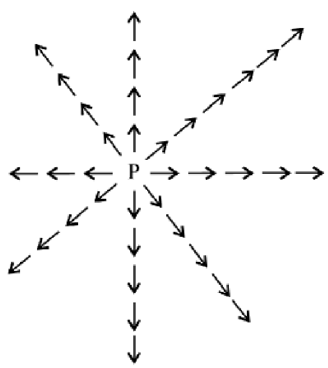
$$\nabla^2\phi = \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} \text{ (Wave equation)}$$

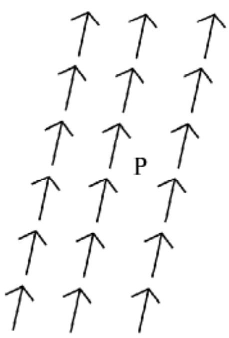
$$\nabla^2\phi = \frac{1}{h^2}\frac{\partial\phi}{\partial t} \text{ [diffusion, heat conduction equation etc.]}$$

where h is constant and t is time.

4.20.1 Divergence of a vector field

Divergence of a vector field $\boldsymbol{A}$ measures how much the vector $\boldsymbol{A}$ spreads out from the point in question. The vector function in fig (4B.6) has a positive divergence at the point P, if it is spreading out from there. If the arrows pointed inwards, it would be a negative divergence on the other hand the function in fig (4B.7) has zero divergence at P, as it is not spreading outwards or inwards at all.





The points, at which $\nabla \cdot \boldsymbol{A} > 0$ are called sources, while the points at which $\nabla \cdot \boldsymbol{A} < 0$ are called sinks of vector field $\boldsymbol{A}$. But if at all points $\nabla \cdot \boldsymbol{A} = 0$, then the vector field is said to be solenoidal. Thus a solenoidal vector field is without a source or sink.

Physical significance of a divergence is that it gives the net rate of outflow per unit volume evaluated at a point. This is 'outflow' of actual substance for liquids, gases or particles and 'flux' for electric and magnetic fields.

We consider an element of volume [fig (4B.8)] in a region through which water is flowing and take the vector point function $\boldsymbol{A}$ to be equal to $\rho\boldsymbol{v}$, where ρ is the density and $\boldsymbol{v}$, the velocity of water flow at that point such that $\boldsymbol{A} = \rho\boldsymbol{v}$, Obviosuly $\rho\boldsymbol{v}$ gives the mass of water following per unit area per second in the direction of $\boldsymbol{v}$. If water is flowing in the direction $\boldsymbol{v}$ making an angle θ with the normal $\hat{\boldsymbol{n}}$ to a surface, then amount of water crossing unit area of the surface in unit time is $\boldsymbol{A}.\hat{\boldsymbol{n}}$, if $\hat{\boldsymbol{n}}$ is a unit vector.

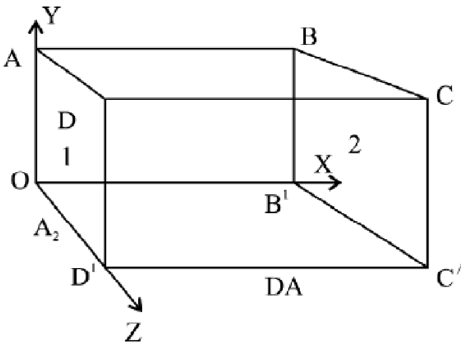


Fig. (4B.8)

The rate at which water flows into element of volume *dxdydz* through surface 1 is $(\boldsymbol{A}.\boldsymbol{i})dydz = A_x dxdydz$, where $\boldsymbol{A} = \boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z$. Similarly the rate at which water flows out through the surface 2 is $\left[A_x + \frac{\partial A_x}{\partial x}dx\right]dydz$, since the value of x-component of $\boldsymbol{A}$ at the surface 2 is $\left[A_x + \frac{\partial A_x}{\partial x}dx\right]$, for constant y & z. Therefore the net outflow through these two surface is

$$[(A_x \text{ at surface 1}) - (A_x \text{ at surface 2})] \text{ dydz}$$

$$= \frac{\partial A_x}{\partial x} dxdydz$$

We get similar expression for the net outflow through the other two pairs of opposite surfaces ; namely

$$\frac{\partial A_y}{\partial y} dxdydz \quad (\textit{through ABCD and OB}'C'D' \textit{ surfaces})$$

And $$\frac{\partial A_z}{\partial z} dxdydz \quad (\textit{through OB}'BA \textit{ and CC}'D'D \textit{ surfaces})$$

Then the total net rate of outflow of water from the volume element *dxdydz* is

$$\left(\frac{\partial A_x}{\partial x}+\frac{\partial A_y}{\partial y}+\frac{\partial A_z}{\partial z}\right)dxdydz=\nabla\cdot\boldsymbol{A}\,dxdydz \qquad (4B.12)$$

From definition of divergence in equation (4B.9). Therefore net rate of outflow of water per unit volume is $\nabla\cdot\boldsymbol{A}$. This is the physical meaning of divergence.

Equation of continuity :

If some physical entity is generated within a certain region of a field, then that region is termed as source. On the other hand, if the physical entity is absorbed then the region is called a sink.

Clearly, if there are no source or sink presents in the field, then the net outflow of the incompressible physical entity over any part of the region is zero.

If the total strength of the sources is greater than that of the sinks, the net outflow is said to be positive, otherwise it is zero or negative.

Now from the physical significance of divergence, we see that, divergence is somewhat like density, since like density, divergence is evaluated per unit volume and may vary from point to point.

Therefore from the above discussions we see that divergence of a vector field $\boldsymbol{A}$ will be different from zero due to i) non-equality of source and sink strength and ii) time variation of density, in case of compressible fluids.

Now we consider a region of volume *dxdydz* is which water is flowing and where there is source and sink. From the principle of conservation of mass :

Rate of increase of mass in *dxdydz* = Rate of creation of mass minus Rate of outwards flow in *dxdydz*

Or $$\frac{\partial\rho}{\partial t}dxdydz=\phi dxdydz-\nabla\cdot\boldsymbol{A}\,dxdydz \qquad (4B.13)$$

Where ϕ is the net mass of fluid being created per unit time per unit volume, which is source density minus sink density and ρ is the mass per unit volume or density of the fluid and $\frac{\partial\rho}{\partial t}$ is the rate of increase of mass per unit volume per unit time. Since $dxdydz\neq 0$, we get

$$\frac{\partial\rho}{\partial t}=\phi-\nabla\cdot\boldsymbol{A} \qquad (4B.14)$$

If there are no source or sink or source strength equals sink strength, then ϕ = 0 and the resulting equation (4B.15) is called equation of continuity : $\nabla \cdot \boldsymbol{A} + \frac{\partial \rho}{\partial t} = 0$ (4B.15)

In case of incompressible fluid, $\frac{\partial \rho}{\partial t} = 0$ and we get $\nabla \cdot \boldsymbol{A} = \phi$ (4B.16)

In case of electric field, $\boldsymbol{D}$, the electric displacement vector, and the sources and sink are electric free charges and we get $\nabla \cdot \boldsymbol{D} = charge \;\; density = \phi$.

For magnetic field $\boldsymbol{B}$, the source and sinks are magnetic free poles, which does not exist. Therefore $\nabla \cdot \boldsymbol{B} = 0$.

4.20.2 Integral form of divergence

We consider $\hat{\boldsymbol{n}}$ to be the unit vector normal to $d\sigma$, a small area from the surface of a small volume element of fig (4B.9) then the mass of fluid flowing out through $d\sigma$ is $\boldsymbol{A}.\hat{\boldsymbol{n}}\; d\sigma$ and the total outflow from the volume enclosed by the surface is $\oint\!\!\oint \boldsymbol{A}.\hat{\boldsymbol{n}}\; d\sigma$.

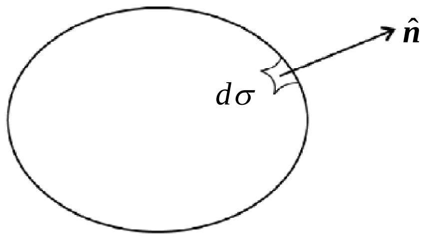


Fig. (4B.9)

Again to the volume element $d\tau$ = $dxdydz$, the total outflow from $d\tau$ is $\nabla \cdot \boldsymbol{A}\, d\tau$

Therefore, $\nabla \cdot \boldsymbol{A}\, d\tau = \oint\!\!\oint \boldsymbol{A}.\hat{\boldsymbol{n}}\; d\sigma$

Or, $\nabla \cdot \boldsymbol{A} = \lim_{d\tau \to 0} \frac{1}{d\tau} \oint\!\!\oint \boldsymbol{A}.\hat{\boldsymbol{n}}\; d\sigma$ (4B.17)

Equation (4B.17) gives the integral definition of divergence.

Examples of Art 4.20 :

Example 1 : Prove that $div\; r^n \boldsymbol{r} = (n+3) r^n$

Solution 1 : Let $r^n \boldsymbol{r} = r^n \left(\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z \right)$, $r = \left(x^2 + y^2 + z^2 \right)^{\frac{1}{2}} = \boldsymbol{i}\phi_1 + \boldsymbol{j}\phi_2 + \boldsymbol{k}\phi_3$

$$\therefore div\; r^n \boldsymbol{r} = \frac{\partial \phi_1}{\partial x} + \frac{\partial \phi_2}{\partial y} + \frac{\partial \phi_3}{\partial z}$$

now $\phi_1 = xr^n$

$$\therefore \frac{\partial \phi_1}{\partial x} = r^n + nxr^{n-1}\frac{\partial r}{\partial x} = r^n + nxr^{n-1}\frac{x}{r} = r^n + nr^{n-2}x^2$$

similar expressions for $\frac{\partial \phi_2}{\partial y}, \frac{\partial \phi_3}{\partial z}$ can be obtained and are given by

$$\frac{\partial \phi_2}{\partial y} = r^n + nr^{n-2}y^2$$

$$\frac{\partial \phi_3}{\partial z} = r^n + nr^{n-2}z^2$$

$$\therefore div\ r^n\boldsymbol{r} = 3r^n + nr^{n-2}\left(x^2 + y^2 + z^2\right) = \left(3r^n + nr^n\right) = r^n(n+3)$$

Note : When n = – 3, $div\ r^n\boldsymbol{r} = \nabla \cdot \frac{r}{r^3} = 0.$ Therefore the vector $\frac{r}{r^3}$ is solenoidal.

Example 2 : Show $\nabla^2(r^n) = n(n+1)r^{n-2}$

Solution 2 : $\nabla^2 r^n = \nabla \cdot \nabla r^n = \nabla \cdot \left(nr^{n-2}\boldsymbol{r}\right) = n\nabla \cdot r^{n-2}\boldsymbol{r}$

Let $\phi = r^{n-2}$. Now using formula (9) in Art 4B.14.2, we can write

$$\nabla^2 r^n = n\nabla \cdot \phi\boldsymbol{r} = n\phi\nabla \cdot \boldsymbol{r} + n\nabla\phi \cdot \boldsymbol{r}$$

Now $\nabla \cdot \boldsymbol{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z}$ = 1 + 1 + 1 = 3, where $\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$ and

$$\nabla\phi = \nabla r^{n-2} = (n-2)r^{n-3}\hat{\boldsymbol{r}}$$

$$= (n-2)\ r^{n-4}\,\boldsymbol{r}$$

$$\therefore \nabla^2 r^n = 3nr^{n-2} + n(n-2)r^{n-4}\boldsymbol{r}.\boldsymbol{r} = 3nr^{n-2} + n(n-2)r^{n-2}$$

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

Example 3 : Prove that $\nabla^2 f(r) = \frac{d^2 f}{dr^2} + \frac{2}{r}\frac{df}{dr}$

Solution 3 : $\nabla^2 f(r) = \nabla \cdot \nabla f(r)$

Now, $\nabla f(r) = \boldsymbol{i}\frac{\partial f(r)}{\partial x} + \boldsymbol{j}\frac{\partial f(r)}{\partial y} + \boldsymbol{k}\frac{\partial f(r)}{\partial z} = \boldsymbol{i}\frac{\partial f(r)}{\partial r}\cdot\frac{\partial r}{\partial x} + \boldsymbol{j}\frac{\partial f(r)}{\partial r}\cdot\frac{\partial r}{\partial y} + \boldsymbol{k}\frac{\partial f(r)}{\partial r}\cdot\frac{\partial r}{\partial z}$

$$= \boldsymbol{i}\frac{\partial f(r)}{\partial r}\cdot\frac{x}{r} + \boldsymbol{j}\frac{\partial f(r)}{\partial r}\cdot\frac{y}{r} + \boldsymbol{k}\frac{\partial f(r)}{\partial r}\cdot\frac{z}{r} = \frac{1}{r}\frac{\partial f(r)}{\partial r}(\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z) = \frac{\partial f(r)}{\partial r}\frac{\boldsymbol{r}}{r}$$

Now $\nabla^2 f(r) = \nabla \cdot \frac{1}{r}\frac{df(r)}{dr}\boldsymbol{r}$

$$= \frac{\partial}{\partial x}\left[\frac{x}{r}\frac{df(r)}{dr}\right] + \frac{\partial}{\partial y}\left[\frac{y}{r}\frac{df(r)}{dr}\right] + \frac{\partial}{\partial z}\left[\frac{z}{r}\frac{df(r)}{dr}\right]$$

$$= \frac{\partial}{\partial x}\left[\frac{x}{r}f'(r)\right] + \frac{\partial}{\partial y}\left[\frac{y}{r}f'(r)\right] + \frac{\partial}{\partial z}\left[\frac{z}{r}f'(r)\right]$$

Now, $\frac{\partial}{\partial x}\left[\frac{x}{r}f'(r)\right] = \frac{1}{r}f'(r) + x\frac{\partial}{\partial x}\left(\frac{1}{r}\right)f'(r) + \frac{x}{r}\frac{\partial}{\partial x}f'(r)$

$$= \frac{f'(r)}{r} + x\frac{\partial}{\partial r}\left(\frac{1}{r}\right)\frac{\partial r}{\partial x}f'(r) + \frac{x}{r}\frac{\partial}{\partial r}f'(r)\frac{\partial r}{\partial x}$$

$$= \frac{f'(r)}{r} - \frac{x^2}{r^3}f'(r) + \frac{x}{r}f''(r)\frac{x}{r} = \frac{f'(r)}{r} - \frac{x^2}{r^3}f'(r) + \frac{x}{r}f''(r)\frac{x}{r}$$

$$= \frac{f'(r)}{r} - \frac{x^2}{r^3}f'(r) + \frac{x^2}{r^2}f''(r) \qquad \text{(i)}$$

Similarly $\frac{\partial}{\partial y}\left[\frac{y}{r}f'(r)\right] = \frac{f'(r)}{r} - \frac{y^2}{r^3}f'(r) + \frac{y^2}{r^2}f''(r)$ (ii)

And $\frac{\partial}{\partial z}\left[\frac{z}{r}f'(r)\right] = \frac{f'(r)}{r} - \frac{z^2}{r^3}f'(r) + \frac{z^2}{r^2}f''(r)$ (iii)

Adding equations (i), (ii) and (iii), we get

$$\nabla^2 f(r) = \frac{3}{r} f'(r) - \frac{x^2+y^2+z^2}{r^3} f'(r) + \frac{x^2+y^2+z^2}{r^2} f''(r)$$

$$= \frac{3}{r} f'(r) - \frac{f'(r)}{r} + f''(r) = f''(r) + \frac{2}{r} f'(r)$$

Exercise of Art 4.20 :

1) A rigid body is rotating with constant angular velocity ω. Show that the linear velocity is solenoidal.

2) Prove that $\nabla^2\left(\frac{1}{r}\right) = 0$

3) Prove that $\nabla.\boldsymbol{F} = 6(x+y+z)$, when $\boldsymbol{F} = \nabla\left(x^3+y^3+z^3-3xyz\right)$

4) A vector field is defined by $\boldsymbol{F} = \frac{\vec{r}}{r^2}$, Evaluate $\boldsymbol{F}$.

5) For what value of a the vector, $\boldsymbol{A} = (x + 3y)\boldsymbol{i} + (y - 2z)\boldsymbol{j} + (x + az)\boldsymbol{k}$ is solenoidal.

Solution :

Solution (1) : $\boldsymbol{\omega} = \boldsymbol{i}\omega_x + \boldsymbol{j}\omega_y + \boldsymbol{k}\omega_z$ $= constant\ vector$

Now linear velocity $\boldsymbol{v} = \boldsymbol{\omega}\times\boldsymbol{r} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \omega_x & \omega_y & \omega_z \\ x & y & z \end{vmatrix}$

or, $\boldsymbol{v} = \boldsymbol{i}\left(\omega_y z - \omega_z y\right) + \boldsymbol{j}\left(\omega_z x - \omega_x z\right) + \boldsymbol{k}\left(\omega_x y - \omega_y x\right)$

Now, $\nabla\cdot\boldsymbol{v} = \frac{\partial}{\partial x}\left(\omega_y z - \omega_z y\right) + \frac{\partial}{\partial y}\left(\omega_z x - \omega_x z\right) + \frac{\partial}{\partial z}\left(\omega_x y - \omega_y x\right) = 0+0+0 = 0$

Since, $\nabla\cdot\boldsymbol{v} = 0, \boldsymbol{v}$ is solenoidal.

Solution 2 : From example (2) of Art 4B.11

We have $\nabla^2(r^n) = n(n+1)r^{n-2}$

Putting $n = -1$, we get $\nabla^2\left(\frac{1}{r}\right) = 0$

Solution (3) : $\boldsymbol{F} = \nabla\left(x^3 + y^3 + z^3 - 3xyz\right)$

or, $$\boldsymbol{F} = \boldsymbol{i}\frac{\partial}{\partial x}\left(x^3 + y^3 + z^3 - 3xyz\right) + \boldsymbol{j}\frac{\partial}{\partial y}\left(x^3 + y^3 + z^3 - 3xyz\right)$$

$$+\boldsymbol{k}\frac{\partial}{\partial z}\left(x^3 + y^3 + z^3 - 3xyz\right)$$

$$= \boldsymbol{i}\left(3x^2 - 3yz\right) + \boldsymbol{j}\left(3y^2 - 3xz\right) + \boldsymbol{k}\left(3z^2 - 3xy\right)$$

Now, $$\nabla \cdot \boldsymbol{F} = \frac{\partial}{\partial x}\left(3x^2 - 3yz\right) + \frac{\partial}{\partial y}\left(3y^2 - 3xz\right) + \frac{\partial}{\partial z}\left(3z^2 - 3xy\right)$$

$$= 6x + 6y + 6z = 6(x + y + z)$$

Solution (4) : $\boldsymbol{F} = \frac{\vec{r}}{r^2} = \boldsymbol{i}\frac{x}{r^2} + \boldsymbol{j}\frac{y}{r^2} + \boldsymbol{k}\frac{z}{r^2}$ where $r^2 = x^2 + y^2 + z^2$

Now, $$\nabla \cdot \boldsymbol{F} = \frac{\partial}{\partial x}\left(\frac{x}{r^2}\right) + \frac{\partial}{\partial y}\left(\frac{y}{r^2}\right) + \frac{\partial}{\partial z}\left(\frac{z}{r^2}\right)$$

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

$$= \frac{1}{r^2} - \frac{2x^2}{r^4}$$

Similarly, $$\frac{\partial}{\partial y}\left(\frac{y}{r^2}\right) = \frac{1}{r^2} - \frac{2y^2}{r^4}$$

and $$\frac{\partial}{\partial z}\left(\frac{z}{r^3}\right)=\frac{1}{r^2}-\frac{2z^2}{r^4}$$

$$\therefore \nabla\cdot\boldsymbol{F}=\frac{3}{r^2}-\frac{2\left(x^2+y^2+z^2\right)}{r^4}=\frac{3}{r^2}-\frac{2}{r^2}=\frac{1}{r^2}$$

Solution 5 : $\boldsymbol{A} = (x+3y)\boldsymbol{i} + (y-2z)\boldsymbol{j} + (x+az)\boldsymbol{k}$

Now, $$\nabla\cdot\boldsymbol{A}=\frac{\partial}{\partial x}(x+3y)+\frac{\partial}{\partial y}(y-2z)+\frac{\partial}{\partial z}(x+az)=1+1+a=2+a$$

For $\boldsymbol{A}$ to be solenoidal $\nabla\cdot\boldsymbol{A}=0$

i.e. $2 + a = 0$

i.e. $a = -2$

4.21 ❒ Curl of A Vector Field

We have shown earlier that if a rigid body rotates, about an axis with a constant angular velocity $\boldsymbol{\omega}$ in the direction of the axis, it is related to the linear velocity $\boldsymbol{v}$ of a particle in rigid body by $\boldsymbol{\omega}=\frac{1}{2}(\nabla\times\boldsymbol{v})$ (4B.18)

Thus curl $\boldsymbol{v}$ relates to angular velocity of rotation and gives us a clue to the name curl $\boldsymbol{v}$ or rotation $\boldsymbol{v}$ or rot v of the vector del × $\boldsymbol{v}$.

In case of flow of fluid the value of $\nabla\times\boldsymbol{v}$ at a point is the measure of the angular velocity of the fluid in the neighbourhood of the point.

When $\nabla\times\boldsymbol{v}=0$ everywhere in some region, the velocity field $\boldsymbol{v}$ is called irrotational, or in general for any vector field $\boldsymbol{A}$, if $\nabla\times\boldsymbol{A}=0$ in a region of space the field $\boldsymbol{A}$ is called irrotational or conservative field or lamellar vector field. In such vector fields line integral around a closed path is zero.

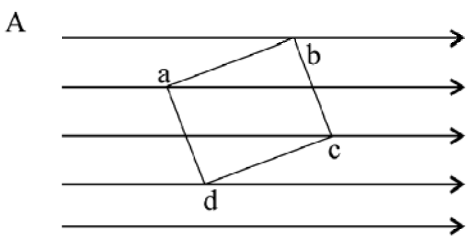


Fig (4B.10)

But for non-lamellar vector fields the line integral around a closed path is not zero and curl of a vector field plays an important role.

We consider a vector field $\boldsymbol{A}$, if we define circulation as the line integral $\oint \boldsymbol{A}.d\boldsymbol{r}$ around a closed plane curve. For flow of water we can take $\boldsymbol{A} = \rho \boldsymbol{v}$ and we can get a physical meaning of the circulation as follows.

We consider a closed rectangular path *abcd* in a vector field $\boldsymbol{A}$ as shown in figure (4B.10). If the plane of this rectangle is perpendicular to the vector field $\boldsymbol{A}$, then the circulation along this path is zero, since $\boldsymbol{A}$ is perpendicular to every element $d\boldsymbol{r}$ of *abcd*.

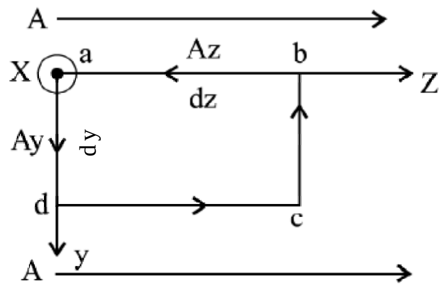


Fig (4B.11)
[X-axis being perpendicular to the plane of the paper at a]

If however the plane of the rectangle is made parallel to the direction of $\boldsymbol{A}$, the line integral along *ad* and *bc* will be zero ; but the line integral along *ab* and *cd* will have some definite value [fig 4B.11] and therefore circulation will not be zero.

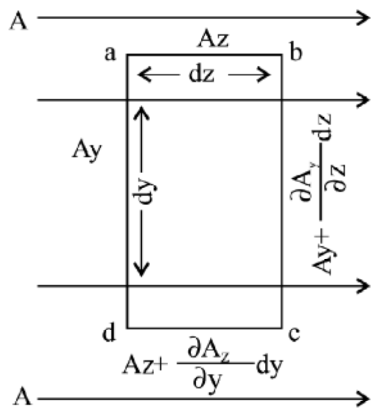


Fig (4B.12)

Let $\boldsymbol{A}$ be the vector point function having components A_x, A_y, A_z along *x, y, z* co-ordinate. We consider the plane are *dydz*. Fig (4B.11).

The value of the components of vector $\boldsymbol{A}$ along side of rectangle *adcb* are (fig. 4B.12)

(along ad) : A_y

(along bc) : $A_y + \dfrac{\partial A_y}{\partial z} dz$

(along ab) : A_z

(along dc) : $A_z + \dfrac{\partial A_z}{\partial y} dy$

Hence the line integral around the closed path adcda is

$$-A_z dz - \left(A_y + \frac{\partial A_y}{\partial z} dz \right) dy + \left(A_z + \frac{\partial A_z}{\partial y} dy \right) dz + A_y dy$$

[since direction of cd and da are in opposite sense]. In evaluating the line integral we must go around the area element *dydz* as in Fig (4B.11) keeping the area to ours left/right hand rule,

$$=\left(\frac{\partial A_y}{\partial z}-\frac{\partial A_z}{\partial y}\right)dydz$$

Since the path of the integration are either parallel or antiparallel to the vector components, the line integral will have maximum value.

Now the area enclosed by the path of integration being dydz, the maximum value of the line integral per unit area is $\left(\frac{\partial A_y}{\partial z}-\frac{\partial A_z}{\partial y}\right)$, which according to the definition (4B.10) of curl is scalar component of $\nabla\times\boldsymbol{A}$ along the unit positive normal to the area i.e. $\boldsymbol{i}$.

$$\therefore\ (\nabla\times\boldsymbol{A})\cdot\boldsymbol{i}=\left(\frac{\partial A_y}{\partial z}-\frac{\partial A_z}{\partial y}\right)=\frac{1}{dydz}\oint_{around\ dydz}\boldsymbol{A}.d\boldsymbol{r}$$

Similarly considering rectangular path whose planes are perpendicular to *y* and *z* axes, it can be shown that

$$(\nabla\times\boldsymbol{A})\cdot\boldsymbol{j}=\left(\frac{\partial A_x}{\partial z}-\frac{\partial A_z}{\partial x}\right)=\frac{1}{dzdx}\oint_{around\ dzdx}\boldsymbol{A}.d\boldsymbol{r}$$

and $$(\nabla\times\boldsymbol{A})\cdot\boldsymbol{k}=\left(\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}\right)=\frac{1}{dxdy}\oint_{around\ dxdy}\boldsymbol{A}.d\boldsymbol{r}$$

$$\therefore\ (\nabla\times\boldsymbol{A})=\boldsymbol{i}\left(\frac{\partial A_y}{\partial z}-\frac{\partial A_z}{\partial y}\right)+\boldsymbol{j}\left(\frac{\partial A_x}{\partial z}-\frac{\partial A_z}{\partial x}\right)+\boldsymbol{k}\left(\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}\right)$$

Therefore $|\nabla\times\boldsymbol{A}|$ is related to the maximum value of circulation per unit area and can be written as

$$\nabla\times\boldsymbol{A}\cdot\hat{\boldsymbol{n}}=\lim_{d\sigma\to 0}\frac{1}{d\sigma}\oint_{around\ d\sigma}\boldsymbol{A}.d\boldsymbol{r} \qquad (4B.18)$$

Where $d\sigma$ is any arbitrary area with positive unit outward normal $\hat{\boldsymbol{n}}$ so that when $\hat{\boldsymbol{n}}$ = $\boldsymbol{i}$, $d\sigma$ = *dydz* etc. Equation (4B.18) can be written as a definition of $\nabla\times\boldsymbol{A}$.

Example of Art. 4.20 :

Example (1) : If a rigid body rotates about on axis passing through the origin with angular velocity $\boldsymbol{\omega}$ and with linear velocity $\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$, then prove that $\boldsymbol{\omega} = \frac{1}{2}(\nabla \times \boldsymbol{v})$.

Solution 1 : $\boldsymbol{\omega} = \boldsymbol{i}\omega_x + \boldsymbol{i}\omega_y + \boldsymbol{j}\omega_x + \boldsymbol{k}\omega_z = constant$

and $\quad \boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$

Therefore $\nabla \times \boldsymbol{v} = \nabla \times \ (\boldsymbol{\omega} \times \boldsymbol{r}) = (\nabla \cdot \boldsymbol{r})\boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \nabla)\boldsymbol{r}$ (i)

using equation (4A.21).

Since $\boldsymbol{\omega}$ is a constant vector $\nabla \cdot \boldsymbol{\omega} = 0$; therefore we have used $(\boldsymbol{\omega} \cdot \nabla)$ as co-efficient of $\boldsymbol{r}$ in equation (i)

$$\nabla \times \boldsymbol{v} = 3\boldsymbol{\omega} - \left(\omega_x \frac{\partial}{\partial x} + \omega_y \frac{\partial}{\partial y} + \omega_z \frac{\partial}{\partial z}\right)\boldsymbol{r} = 3\boldsymbol{\omega} - \left(\omega_x \frac{\partial \boldsymbol{r}}{\partial x} + \omega_y \frac{\partial \boldsymbol{r}}{\partial y} + \omega_z \frac{\partial \boldsymbol{r}}{\partial z}\right)$$

$$= 3\boldsymbol{\omega} - \left(\omega_x \boldsymbol{i} + \omega_y \boldsymbol{j} + \omega_z \boldsymbol{k}\right) = 3\boldsymbol{\omega} - \boldsymbol{\omega} = 2\boldsymbol{\omega} \Rightarrow \boldsymbol{\omega} = \frac{1}{2}(\nabla \times \boldsymbol{v})$$

Example 2 : Find the constant *a, b, c* so that the vector,

$\boldsymbol{A} = (x + 2y + az)\boldsymbol{i} + (bx - 3y - z)\boldsymbol{j} + (4x + cy + 2z)\boldsymbol{k}$, is irrotational.

Solution 2 : Since $\boldsymbol{A}$ is irrotational, $\nabla \times \boldsymbol{A} = 0$

Or, $$\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x+2y+az & bx-3y-z & 4x+cy-2z \end{vmatrix} = 0$$

Or, $\quad (c + 1)\boldsymbol{i} + \boldsymbol{j}(a - 4) + \boldsymbol{k}(b - 2) = 0$

Therefore, $c + 1 = 0 \Rightarrow c = -1$

$a - 4 = 0 \Rightarrow a = 4$

$b - 2 = 0 \Rightarrow b = 2$

Example 3 : Find the corresponding scalar potential of the vector ***A*** in example 2.

Solution 3 : Since $\nabla \times \boldsymbol{A} = 0$ and ***A*** is defined in a simply connected region, vector field ***A*** is said to be conservative and hence it can be expressed as a gradient of a scalar potential, say, ϕ.

$$\therefore \boldsymbol{A} = -\nabla\phi = -\boldsymbol{i}\frac{\partial\phi}{\partial x} - \boldsymbol{j}\frac{\partial\phi}{\partial y} - \boldsymbol{k}\frac{\partial\phi}{\partial z}$$

Or, if A_x, A_y and A_z are the components of the vector ***A***, we can write,

$$\boldsymbol{A} = \boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z = -\boldsymbol{i}\frac{\partial\phi}{\partial x} - \boldsymbol{j}\frac{\partial\phi}{\partial y} - \boldsymbol{k}\frac{\partial\phi}{\partial z}$$

$$\therefore \frac{\partial\phi}{\partial x} = -A_x = -(x + 2y + 4z) \tag{i}$$

$$\frac{\partial\phi}{\partial y} = -A_y = -(2x - 3y - z) \tag{ii}$$

$$\frac{\partial\phi}{\partial z} = -A_z = -(4x - y + 2z) \tag{iii}$$

Integrating equation (i) with respect to x, keeping y, z constant, we get,

$$\phi = -\left(\frac{x^2}{2} + 2yx + 4zx + f_1(y, z)\right) \tag{iv}$$

Similarly integrating equation (ii), with respect to y; keeping x and z as constant, we get

$$\phi = -\left(2xy - \frac{3y^2}{2} - zy + f_2(x, z)\right) \tag{v}$$

And from equation (iii), we get,

$$\phi = -\left(4xz - yz + z^2 + f_3(x, y)\right) \tag{vi}$$

$f_1(y, z)$, $f_2(x, z)$ and $f_3(x, y)$ are constants of integration. Now f_1, f_2 and f_3 are to be suitably chosen in order that function ϕ were identical in all these three equations.

By inspection, we find $f_1(y, z) = -zy - \dfrac{3y^2}{2} + z^2$

$$f_2(x, z) = 4zx + x^2/2 + z^2, \quad f_3(x, y) = \frac{x^2}{2} + 2yz - \frac{3y^2}{2}$$

$$\therefore \phi = -\left(\frac{x^2}{2} - \frac{3y^2}{2} + z^2 + 2yx - yz + 4xz\right) + c$$

where c is a constant independent of x, y, z.

Example 4 : What do you mean by an exact differential? Show that a necessary and sufficient condition that $F_1dx + F_2dy + F_3dz$ be an exact differential is that $\nabla \times \boldsymbol{F} = 0$ where $\boldsymbol{F} = \boldsymbol{i}F_1 + \boldsymbol{j}F_2 + \boldsymbol{k}F_3$.

Solution 4 : If $P = P(x, y, z)$; $Q = Q(x, y, z)$; $R = R(x, y, z)$, then $Pdx + Qdy + Rdz$ is an exact differential if there exists a function $\phi(x, y, z)$ such that

$$d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz$$

is equal to $Pdx + Qdy + Rdz$.

$Pdx + Qdy + Rdz$, is an exact differential, if the following conditions hold good :

$$\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = \left(\frac{\partial Q}{\partial z} - \frac{\partial R}{\partial y}\right) = \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right) = 0$$

The necessary condition that $F_1dx + F_2dy + F_3dz$ is an exact differential :

Let $F_1dx + F_2dy + F_3dz$ be an exact differential. Then $F_1dx + F_2dy + F_3dz = d\phi$

$$= \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz$$

$$\therefore F_1 = \frac{\partial\phi}{\partial x};\ F_2 = \frac{\partial\phi}{\partial y};\ F_3 = \frac{\partial\phi}{\partial z}$$

$$\therefore \boldsymbol{F} = \boldsymbol{i}F_1 + \boldsymbol{j}F_2 + \boldsymbol{k}F_3 = \boldsymbol{i}\frac{\partial \phi}{\partial x} + \boldsymbol{j}\frac{\partial \phi}{\partial y} + \boldsymbol{k}\frac{\partial \phi}{\partial z} = \nabla\phi$$

$$\therefore \nabla \times \boldsymbol{F} = \nabla \times \nabla\phi = 0$$

The sufficient condition that $F_1dx + F_2dy + F_3dz$ is an exact differential :

Let $\nabla \times \boldsymbol{F} = 0$

or, $$\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = 0$$

or, $$\boldsymbol{i}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \boldsymbol{j}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \boldsymbol{k}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) = 0$$

or, $$\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) = \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) = \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) = 0$$

$\therefore$ $F_1dx + F_2dy + F_3dz$ = exact differential

also, when $\nabla \times \boldsymbol{F} = 0$, $F = \nabla\phi$, so $\boldsymbol{F}.\boldsymbol{dr} = \nabla\phi \cdot \boldsymbol{dr} = d\phi$

or, $(\boldsymbol{i}F_1 + \boldsymbol{j}F_2 + \boldsymbol{k}F_3).(\boldsymbol{i}\,dx + \boldsymbol{j}\,dy + \boldsymbol{k}\,dz) = d\phi$

or, $F_1dx + F_2dy + F_3dz = d\phi$ = an exact differential

Exercise of Art 4.21 :

1) If $\nabla.\boldsymbol{E} = 0, \nabla.\boldsymbol{H} = 0, \nabla \times \boldsymbol{E} = -\mu_0\frac{\partial H}{\partial t}, \nabla \times \boldsymbol{H} = \varepsilon_0\frac{\partial E}{\partial t}$, where μ_0, ε_0 are constants. Show that $\boldsymbol{E}$ and $\boldsymbol{H}$ satisfy the wave equation $\nabla^2\boldsymbol{u} = \varepsilon_0\mu_0\frac{\partial^2\boldsymbol{u}}{\partial t^2}$.

2) If $\boldsymbol{A}$ and $\boldsymbol{B}$ are irrotational, show that $\boldsymbol{A} \times \boldsymbol{B}$ is solenoidal.

3) If $u\boldsymbol{F} = \nabla v$, where u, v are scalar field and $\boldsymbol{F}$ is a vector field, show that $\boldsymbol{F} \cdot \nabla \times \boldsymbol{F} = 0$.

4) Prove the following ; (i) If $\boldsymbol{A}$ is irrotational ($\boldsymbol{A} \times \boldsymbol{r}$) is solenoidal, (ii) If $\nabla \times \boldsymbol{A} = \frac{\partial B}{\partial t}$, then show that $\nabla \cdot \boldsymbol{B}$ is independent of t.

5) Prove that (i) $\nabla \times (\boldsymbol{A} \times \boldsymbol{r}) = 2\boldsymbol{A}$ and (ii) $(\boldsymbol{A} \times \nabla) \times \boldsymbol{r} = -2\boldsymbol{A}$, where $\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$ where $\boldsymbol{A}$ is a constant vector.

Solution of exercise of Art 4B.12 :

Solution (1) : We have

$$\nabla \times (\nabla \times \boldsymbol{E}) = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E} \quad \text{(i)}$$

See Art 4B.13 item no (3)

Now substituting $\nabla \times \boldsymbol{E} = -\mu_0 \dfrac{\partial \boldsymbol{H}}{\partial t}$ in equation (i)

$$-\mu_0 \nabla \times \frac{\partial \boldsymbol{H}}{\partial t} = -\nabla^2 \boldsymbol{E} \qquad \because \nabla \cdot \boldsymbol{E} = 0$$

or, $\mu_0 \dfrac{\partial}{\partial t}(\nabla \times \boldsymbol{H}) = \nabla^2 \boldsymbol{E}$

or, $\varepsilon_0 \mu_0 \dfrac{\partial^2 \boldsymbol{E}}{\partial t^2} = \nabla^2 \boldsymbol{E}$ (ii), substituting $\nabla \times \boldsymbol{H} = \varepsilon_0 \dfrac{\partial \boldsymbol{E}}{\partial t}$

Similarly we can arrive at

$\varepsilon_0 \mu_0 \dfrac{\partial^2 \boldsymbol{H}}{\partial t^2} = \nabla^2 \boldsymbol{H}$ (iii), by replacing $\boldsymbol{E}$ by $\boldsymbol{H}$,

Using common term $\boldsymbol{u}$ for $\boldsymbol{E}$ and $\boldsymbol{H}$,

We get : $\varepsilon_0 \mu_0 \dfrac{\partial^2 \boldsymbol{u}}{\partial t^2} = \nabla^2 \boldsymbol{u}$

Solution 2 : We have $\nabla \times \boldsymbol{A} = 0, \nabla \times \boldsymbol{B} = 0$

$$\nabla \cdot (\boldsymbol{A} \times \boldsymbol{B}) = \boldsymbol{B} \cdot \nabla \times \boldsymbol{A} - \boldsymbol{A} \cdot \nabla \times \boldsymbol{B} = 0$$

(See Art 4B.14.2 item no. 12)

$\therefore$ $\boldsymbol{A} \times \boldsymbol{B}$ is solenoidal.

Solution 3 : Since $\boldsymbol{F} = \dfrac{1}{u} \nabla v$, u and v are scalars,

$$\therefore \nabla \times \boldsymbol{F} = \nabla \times \left(\frac{1}{u} \nabla v \right) = \nabla \left(\frac{1}{u} \right) \times \nabla v + \frac{1}{u} \nabla \times \nabla v$$

(see Art 4B.14.2 item no. 15)

Now $\nabla\times\nabla v=0$

$$\therefore \nabla\times\boldsymbol{F}=\nabla\left(\frac{1}{u}\right)\times\nabla v$$

$$\therefore \boldsymbol{F}\cdot\nabla\times\boldsymbol{F}=\frac{1}{u}\nabla v\cdot\nabla\left(\frac{1}{u}\right)\times\nabla v=0$$

Since any two vector in a triple scalar product is equal, implies the product is zero.

Solution 4 :

i) We have $\nabla\times\boldsymbol{A}=0$

Now $\nabla\cdot\left(\boldsymbol{A}\times\boldsymbol{r}\right)=\boldsymbol{r}\cdot\nabla\times\boldsymbol{A}-\boldsymbol{A}\cdot\nabla\times\boldsymbol{r}$

Now $\nabla\times\boldsymbol{r}=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z}\\ x & y & z\end{vmatrix}$

$$=\boldsymbol{i}\left(\frac{\partial z}{\partial y}-\frac{\partial y}{\partial z}\right)+\boldsymbol{j}\left(\frac{\partial x}{\partial z}-\frac{\partial z}{\partial x}\right)+\boldsymbol{k}\left(\frac{\partial y}{\partial x}-\frac{\partial x}{\partial y}\right)=\boldsymbol{i}.0+\boldsymbol{j}.0+\boldsymbol{k}.0=0$$

Since *x*, *y*, *z* are independent.

Therefore $\nabla\cdot\left(\boldsymbol{A}\times\boldsymbol{r}\right)=\boldsymbol{r}\cdot\mathbf{0}\times\boldsymbol{A}\cdot\mathbf{0}=\mathbf{0}$ and ($\boldsymbol{A}\times\boldsymbol{r}$) is solenoidal.

ii) $\nabla\times\boldsymbol{A}=\dfrac{\partial B}{\partial t}$

taking divergence on both sides $\nabla\cdot\nabla\times\boldsymbol{A}=\dfrac{\partial}{\partial t}\left(\nabla\cdot\boldsymbol{B}\right)=0$

since $\nabla\cdot\nabla\times\boldsymbol{A}=0$

therefore $\nabla\cdot\boldsymbol{B}$ is independent of *t*.

Solution 5 :

i) $$\nabla\times(\boldsymbol{A}\times\boldsymbol{r})=(\nabla\cdot\boldsymbol{r})\boldsymbol{A}-(\boldsymbol{A}\cdot\nabla)\boldsymbol{r}=3\boldsymbol{A}-\left(A_x\frac{\partial}{\partial x}+A_y\frac{\partial}{\partial y}+A_z\frac{\partial}{\partial z}\right)\boldsymbol{r}$$

$$=3\boldsymbol{A}-\left[A_x\frac{\partial \boldsymbol{r}}{\partial x}+A_y\frac{\partial \boldsymbol{r}}{\partial y}+A_z\frac{\partial \boldsymbol{r}}{\partial z}\right]$$

$$=3\boldsymbol{A}-\left[A_x\boldsymbol{i}+A_y\boldsymbol{j}+A_z\boldsymbol{k}\right]=3\boldsymbol{A}-\boldsymbol{A}=2\boldsymbol{A}$$

where $\boldsymbol{A}=\boldsymbol{i}A_x+\boldsymbol{j}A_y+\boldsymbol{k}A_z$

ii) We have $$\boldsymbol{A}\times\nabla=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ A_x & A_y & A_z\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\end{vmatrix}$$

$$=\boldsymbol{i}\left[A_y\frac{\partial}{\partial z}-A_z\frac{\partial}{\partial y}\right]+\boldsymbol{j}\left[A_z\frac{\partial}{\partial x}-A_x\frac{\partial}{\partial z}\right]+\boldsymbol{k}\left[A_x\frac{\partial}{\partial y}-A_y\frac{\partial}{\partial x}\right]$$

Now,

$$(\boldsymbol{A}\times\nabla)\times\boldsymbol{r}=\begin{vmatrix}i & j & k\\ A_y\frac{\partial}{\partial z}-A_z\frac{\partial}{\partial y} & A_z\frac{\partial}{\partial x}-A_x\frac{\partial}{\partial z} & A_x\frac{\partial}{\partial y}-A_y\frac{\partial}{\partial x}\\ x & y & z\end{vmatrix}$$

$$=\boldsymbol{i}\left[A_z\frac{\partial}{\partial x}-A_x\frac{\partial}{\partial z}\right]z-\boldsymbol{i}\left[A_x\frac{\partial}{\partial y}-A_y\frac{\partial}{\partial x}\right]y+\boldsymbol{j}\left[A_x\frac{\partial}{\partial y}-A_y\frac{\partial}{\partial x}\right]x$$

$$-\boldsymbol{j}\left[A_y\frac{\partial}{\partial z}-A_z\frac{\partial}{\partial y}\right]z+\boldsymbol{k}\left[A_y\frac{\partial}{\partial z}-A_z\frac{\partial}{\partial y}\right]y-\boldsymbol{k}\left[A_z\frac{\partial}{\partial x}-A_x\frac{\partial}{\partial z}\right]x$$

$$=-\boldsymbol{i}A_x-\boldsymbol{i}A_x-\boldsymbol{j}A_y-\boldsymbol{j}A_y-\boldsymbol{k}A_z-\boldsymbol{k}A_z$$

$$=-2[\boldsymbol{i}A_x+\boldsymbol{j}A_y+\boldsymbol{k}A_z]=-2\boldsymbol{A}$$

4.22 ❒ Vector Identities

In various application of vector analysis expressions involving ∇ and scalar or vector functions are involved. We can verify these expressions by writing out components. These verifications become however easier if we treated ∇ an ordinary vector remembering that it is also a vector differential operator.

1 > Curl (grad ϕ) = 0

Proof : $\text{curl}(grad\ \phi) = \nabla \times (\nabla \phi)$

$$= \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} & \frac{\partial \phi}{\partial z} \end{vmatrix}$$

$$= \boldsymbol{i}\left(\frac{\partial^2 \phi}{\partial y \partial z} - \frac{\partial^2 \phi}{\partial z \partial y}\right) + \boldsymbol{j}\left(\frac{\partial^2 \phi}{\partial z \partial x} - \frac{\partial^2 \phi}{\partial x \partial z}\right) + \boldsymbol{k}\left(\frac{\partial^2 \phi}{\partial x \partial y} - \frac{\partial^2 \phi}{\partial y \partial x}\right)$$

= $\boldsymbol{i}(0) + \boldsymbol{j}(0) + \boldsymbol{k}(0) \quad = 0$

2> div(curl $\boldsymbol{A}$) = 0

proof : div (curl $\boldsymbol{A}$) $= \nabla \cdot \nabla \times \boldsymbol{A} = \nabla \times \nabla \cdot \boldsymbol{A} = 0$

interchanging “dot” and “cross” in a triple scalar product and treating ∇ as a normal vector.

3> curl (curl $\boldsymbol{A}$) $= \nabla \times (\nabla \times \boldsymbol{A})$

$$= \nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}$$

Proof : $\nabla \times (\nabla \times \boldsymbol{A}) = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla \cdot \nabla \boldsymbol{A} = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}$

Using the formula for $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C})$, treating ∇ as a normal vector and also an operator which differentiate $\boldsymbol{A}$. Laplacian of a vector, $\nabla^2 \boldsymbol{A}$, simply means a vector whose components are $\nabla^2 \boldsymbol{A}_x, \nabla^2 \boldsymbol{A}_y, \nabla^2 \boldsymbol{A}_z$

4> $\nabla \cdot \phi \boldsymbol{A} = \phi \nabla \boldsymbol{A} + \nabla_{\phi} \cdot \boldsymbol{A}$, where ϕ is a scalar.

Proof : $\nabla \cdot \phi \boldsymbol{A} = \nabla_{\phi} \cdot (\phi \boldsymbol{A}) + \nabla_{A} \cdot (\phi \boldsymbol{A})$

Where the subscript on ∇ indicate which function is to be differentiated i.e. ∇_{ϕ} will diferentiate ϕ, keeping $\boldsymbol{A}$ constant.

Now $\nabla_{\phi} \cdot \phi \boldsymbol{A} = \nabla_{\phi} \phi \cdot \boldsymbol{A} = \boldsymbol{A} \cdot \nabla \phi$, since ϕ is a scalar, we can put it on either side of dot. On the last step we have removed the subscript since $\boldsymbol{A}$ no longer appear after. Again $\nabla_{\phi} \cdot \phi \boldsymbol{A} = \phi \nabla \cdot \boldsymbol{A}$ since ϕ is a scalar and is not differentiated, it may be treated as a constant. Therefore, collecting all the terms we have :

$$\nabla \cdot \phi \boldsymbol{A} = \boldsymbol{A} \cdot \nabla \phi + \phi(\nabla \cdot \boldsymbol{A})$$

5> grad $(\boldsymbol{A}.\boldsymbol{B}) = \nabla(\boldsymbol{A}.\boldsymbol{B}) = (\boldsymbol{A}.\nabla)\boldsymbol{B} + (\boldsymbol{B}.\nabla)\boldsymbol{A} + \boldsymbol{A} \times (\nabla \times \boldsymbol{B}) + \boldsymbol{B} \times (\nabla \times \boldsymbol{A})$

Pooof : consisting only the x-component

$$grad_x(\boldsymbol{A}.\boldsymbol{B}) = \frac{\partial}{\partial x}\left(A_x B_x + A_y B_y + A_z B_z\right)$$

$$= A_x \frac{\partial B_x}{\partial x} + B_x \frac{\partial A_x}{\partial x} + A_y \frac{\partial B_y}{\partial x} + B_y \frac{\partial A_y}{\partial x} + A_z \frac{\partial B_z}{\partial x} + B_z \frac{\partial A_z}{\partial x}$$

$$= \left(A_x \frac{\partial B_x}{\partial x} + A_y \frac{\partial B_x}{\partial y} + A_z \frac{\partial B_x}{\partial z} \right) - A_y \frac{\partial B_x}{\partial y} - A_z \frac{\partial B_x}{\partial z} + \left(B_x \frac{\partial A_x}{\partial x} + B_y \frac{\partial A_x}{\partial y} + B_z \frac{\partial A_x}{\partial z} \right)$$

$$-B_y \frac{\partial A_x}{\partial y} - B_z \frac{\partial A_x}{\partial z} + A_y \frac{\partial B_y}{\partial x} + B_y \frac{\partial A_y}{\partial x} + A_z \frac{\partial B_z}{\partial x} + B_z \frac{\partial A_z}{\partial x}$$

$$= \boldsymbol{A} \cdot \nabla B_x + \boldsymbol{B}.\nabla A_x + A_y \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right) + B_y \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) - A_z \left(\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x} \right)$$

$$- B_z \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right)$$

$$= \boldsymbol{A} \cdot \nabla B_x + \boldsymbol{B}.\nabla A_x + \left(\boldsymbol{A} \times curl\ \boldsymbol{B}\right)_x + \left(\boldsymbol{B} \times curl\ \boldsymbol{A}\right)_x$$

Hence, considering y and z components of L.H.S and adding them all, we get

$$\nabla(\boldsymbol{A}.\boldsymbol{B}) = (\boldsymbol{A}.\nabla)\boldsymbol{B} + (\boldsymbol{B}.\nabla)\boldsymbol{A} + (\boldsymbol{A}\times \boldsymbol{curl\ B}) + (\boldsymbol{B}\times \boldsymbol{curl\ A})$$

Now we find out $(\boldsymbol{A}\times \boldsymbol{curl\ B})_x = A_y(\boldsymbol{curl\ B})_z - A_z(\boldsymbol{curl\ B})_y$

$$= A_y\left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right) - A_z\left(\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x}\right)$$

Similarly, $(\boldsymbol{B} \times \boldsymbol{curl\ A})_x = B_y(\boldsymbol{curl\ A})_z - B_z(\boldsymbol{curl\ A})_y$

$$= B_y\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) - B_z\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)$$

6> $div(\boldsymbol{A}\times\boldsymbol{B}) = \boldsymbol{B}\cdot\nabla\times\boldsymbol{A} - \boldsymbol{A}\cdot\nabla\times\boldsymbol{B}$

Proof : $div\ (\boldsymbol{A}\times\boldsymbol{B}) = \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial j} + \boldsymbol{k}\frac{\partial}{\partial z}\right)\cdot(\boldsymbol{A}\times\boldsymbol{B})$

$$= \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right)\cdot\left[\boldsymbol{i}\left(A_yB_z - A_zB_y\right) + \boldsymbol{j}\left(A_zB_x - A_xB_z\right) + \boldsymbol{k}\left(A_xB_y - A_yB_x\right)\right]$$

$$= \frac{\partial}{\partial x}\left(A_yB_z - A_zB_y\right) + \frac{\partial}{\partial y}\left(A_zB_x - A_xB_z\right) + \frac{\partial}{\partial z}\left(A_xB_y - A_yB_x\right)$$

$$= B_x\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) + B_y\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) + B_z\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)$$

$$-\left[A_x\left(\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z}\right) + A_y\left(\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x}\right) + A_z\left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right)\right]$$

$$= \boldsymbol{B}\cdot\nabla\times\boldsymbol{A} - \boldsymbol{A}\cdot\nabla\times\boldsymbol{B}$$

7> $curl(\boldsymbol{A}\times\boldsymbol{B}) = \boldsymbol{A}\nabla\cdot\boldsymbol{B} - \boldsymbol{B}\nabla\cdot\boldsymbol{A} + (\boldsymbol{B}\cdot\nabla)\boldsymbol{A} - (\boldsymbol{A}\cdot\nabla)\boldsymbol{B}$

$= \boldsymbol{A\ div\ B - B\ div\ A + B\ grad\ A - A\ grad\ B}$

Proof :

$$curl(\boldsymbol{A}\times\boldsymbol{B}) = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ A_yB_z - A_zB_y & A_zB_x - A_xB_z & A_xB_y - A_yB_x \end{vmatrix}$$

Considering the x-component only

$curl_x(\boldsymbol{A}\times\boldsymbol{B}) = \dfrac{\partial}{\partial y}\left(A_xB_y - A_yB_x\right) - \dfrac{\partial}{\partial z}\left(A_zB_x - A_xB_z\right)$ Adding and subtracting $\dfrac{\partial}{\partial x}\left(A_xB_x\right)$, we get

$curl_x(\boldsymbol{A}\times\boldsymbol{B})$

$$= \frac{\partial}{\partial x}\left(A_xB_x\right) + \frac{\partial}{\partial y}\left(A_xB_y\right) + \frac{\partial}{\partial z}\left(A_xB_z\right)$$

$$-\left[\frac{\partial}{\partial x}\left(A_xB_x\right) + \frac{\partial}{\partial y}\left(A_yB_x\right) + \frac{\partial}{\partial z}\left(A_zB_x\right)\right]$$

$$= A_x\left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z}\right) + \left(B_x\frac{\partial A_x}{\partial x} + B_y\frac{\partial A_y}{\partial y} + B_z\frac{\partial A_z}{\partial z}\right)$$

$$-B_x\left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right) - \left(A_x\frac{\partial B_x}{\partial x} + A_y\frac{\partial B_y}{\partial y} + A_z\frac{\partial B_z}{\partial z}\right)$$

$$= \left(\boldsymbol{B}.\nabla\right)A_x + A_x\left(\nabla.\boldsymbol{B}\right) - \left(\boldsymbol{A}.\nabla\right)B_x - B_x\left(\nabla.\boldsymbol{A}\right)$$

Considering components in y and z direction and adding. We get

$curl(\boldsymbol{A}\times\boldsymbol{B})$

$$= \left(\boldsymbol{B}.\nabla\right)\left(\boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z\right) - \left(\boldsymbol{A}.\nabla\right)\left(\boldsymbol{i}B_x + \boldsymbol{j}B_y + \boldsymbol{k}B_z\right)$$

$$+ \left(\boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z\right)\left(\nabla.\boldsymbol{B}\right) - \left(\boldsymbol{i}B_x + \boldsymbol{j}B_y + \boldsymbol{k}B_z\right)\left(\nabla.\boldsymbol{A}\right)$$

$$\therefore\ \nabla\times(\boldsymbol{A}\times\boldsymbol{B}) = \left(\boldsymbol{B}.\nabla\right)\boldsymbol{A} - \left(\boldsymbol{A}.\nabla\right)\boldsymbol{B} + \boldsymbol{A}\left(\nabla.\boldsymbol{B}\right) - \boldsymbol{B}\left(\nabla.\boldsymbol{A}\right)$$

8> $curl(\phi \boldsymbol{A}) = \phi\ curl\ \boldsymbol{A} - \boldsymbol{A} \times grad\ \phi,\ \phi$ is a scalar

Proof : $curl\ \phi \boldsymbol{A} = \nabla \times (\phi \boldsymbol{A}) = \phi \nabla \times \boldsymbol{A} - \boldsymbol{A} \times \nabla \phi$

$$= \left(\boldsymbol{i} \frac{\partial}{\partial x} + \boldsymbol{j} \frac{\partial}{\partial y} + \boldsymbol{k} \frac{\partial}{\partial z} \right) \times \left[\boldsymbol{i}(\phi A_x) + \boldsymbol{j}(\phi A_y) + \boldsymbol{k}(\phi A_z) \right]$$

$$= \boldsymbol{i} \left[\frac{\partial(\phi A_z)}{\partial y} - \frac{\partial(\phi A_y)}{\partial z} \right] + \boldsymbol{j} \left[\frac{\partial(\phi A_x)}{\partial z} - \frac{\partial(\phi A_z)}{\partial x} \right] + \boldsymbol{k} \left[\frac{\partial(\phi A_y)}{\partial x} - \frac{\partial(\phi A_x)}{\partial y} \right]$$

Considering only x-components

$$curl_x(\phi \boldsymbol{A}) = \left[\phi \left(\frac{\partial(A_z)}{\partial y} - \frac{\partial(A_y)}{\partial z} \right) + A_z \frac{\partial \phi}{\partial y} - A_y \frac{\partial \phi}{\partial z} \right] \boldsymbol{i}$$

Similarly $curl_y(\phi \boldsymbol{A}) = \left[\phi \left(\frac{\partial(A_x)}{\partial z} - \frac{\partial(A_z)}{\partial x} \right) + A_x \frac{\partial \phi}{\partial z} - A_z \frac{\partial \phi}{\partial x} \right] \boldsymbol{j}$

And $curl_z(\phi \boldsymbol{A}) = \left[\phi \left(\frac{\partial(A_y)}{\partial x} - \frac{\partial(A_x)}{\partial y} \right) + A_y \frac{\partial \phi}{\partial x} - A_x \frac{\partial \phi}{\partial y} \right] \boldsymbol{k}$

Adding we get,

$\therefore\ curl\ \boldsymbol{A} = \boldsymbol{i}\ curl_x(\phi \boldsymbol{A}) + \boldsymbol{j}\ curl_y(\phi \boldsymbol{A}) + \boldsymbol{k}\ curl_z(\phi \boldsymbol{A})$

$= \phi \nabla \times \boldsymbol{A} + \nabla \phi \times \boldsymbol{A}$

$= \phi \nabla \times \boldsymbol{A} - \boldsymbol{A} \times \nabla \phi$

4.23 ❒ List of Vector Relations

4B.14 We list below some useful vector relations and vector equations which are frequently used in many areas of physics. Students should carefully go through these relations for their benefit.

We have grouped the vector identities into two categories – one involving 'del' operator and the other not involving it.

Similarly we grouped the vector equations in physics into two categories – one involving Laplacian and the other not. The symbols in equations have usual meaning.

4.23.1 : Vector relations not involving 'del' operator

1. $\boldsymbol{A} \times (\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A} \times \boldsymbol{B} + \boldsymbol{A} \times \boldsymbol{C}$
2. $(\boldsymbol{B} + \boldsymbol{C}) \times \boldsymbol{A} = \boldsymbol{B} \times \boldsymbol{A} + \boldsymbol{C} \times \boldsymbol{A}$
3. $\boldsymbol{A} \times \boldsymbol{B}.\boldsymbol{C} = \boldsymbol{A}.\boldsymbol{B} \times \boldsymbol{C}$
4. $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{B}(\boldsymbol{A}.\boldsymbol{C}) - \boldsymbol{C}(\boldsymbol{A}.\boldsymbol{B})$
5. $\boldsymbol{A} \times \boldsymbol{B} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}$
6. $\boldsymbol{A} \cdot \boldsymbol{B} \times \boldsymbol{C} = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}$

4.23.2. : Vector relations involving 'del' operator

1. $\nabla \cdot \boldsymbol{r} = 3$
2. $\nabla \times \boldsymbol{r} = 0$
3. $\nabla r^n = nr^{n-1}\hat{\boldsymbol{r}} = nr^{n-2}\boldsymbol{r}$
4. $\nabla^2 r^n = n(n+1)r^{n-2}$
5. $\nabla^2 \left(\frac{1}{r}\right) = 0$
6. $\nabla(m\phi + n\psi) = m\nabla\phi + n\nabla\psi$
7. $\nabla(\phi\psi) = \phi\nabla\psi + \psi\nabla\phi$
8. $\nabla \cdot (m\boldsymbol{A} + n\boldsymbol{B}) = m(\nabla \cdot \boldsymbol{A}) + n(\nabla \cdot \boldsymbol{B})$
9. $\nabla \cdot (\phi\boldsymbol{A}) = \phi(\nabla \cdot \boldsymbol{A}) + \nabla\phi \cdot \boldsymbol{A}$
10. $\nabla \times (m\boldsymbol{A} + n\boldsymbol{B}) = m(\nabla \times \boldsymbol{A}) + n(\nabla \times \boldsymbol{B})$
11. $\nabla \times (\phi\boldsymbol{A}) = \phi(\nabla \times \boldsymbol{A}) + (\nabla\phi) \times \boldsymbol{A}$
12. $\nabla \cdot (\boldsymbol{A} \times \boldsymbol{B}) = \boldsymbol{B} \cdot (\nabla \times \boldsymbol{A}) - \boldsymbol{A} \cdot (\nabla \times \boldsymbol{B})$

13. $\nabla(\boldsymbol{A}\cdot\boldsymbol{B}) = (\boldsymbol{A}\cdot\nabla)\boldsymbol{B} + (\boldsymbol{B}\cdot\nabla)\boldsymbol{A} + \boldsymbol{A}\times(\nabla\times\boldsymbol{B}) + \boldsymbol{B}\times(\nabla\times\boldsymbol{A})$

14. $\nabla\cdot(\nabla\phi) = \nabla^2\phi = \dfrac{\partial^2\phi}{\partial x^2} + \dfrac{\partial^2\phi}{\partial y^2} + \dfrac{\partial^2\phi}{\partial x^2}$

15. $\nabla\times(\nabla\phi) = 0$

16. $\nabla\cdot(\nabla\times\boldsymbol{A}) = 0$

17. $\nabla\times(\nabla\times\boldsymbol{A}) = \nabla(\nabla\cdot\boldsymbol{A}) - \nabla^2\boldsymbol{A}$

4.23.3. : Vector equation of variation branches of physics

1. Lorentz forces : $\boldsymbol{F} = q[\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B}]$
2. Maxwell's field equation (in vacuum)

$$\nabla\cdot\boldsymbol{D} = 0; \quad \nabla\times\boldsymbol{E} = -\frac{\partial\boldsymbol{B}}{\partial t}$$

$$\nabla\cdot\boldsymbol{B} = 0; \quad \nabla\times\boldsymbol{H} = -\frac{\partial\boldsymbol{D}}{\partial t}$$

Where $\boldsymbol{D} = \varepsilon_0\boldsymbol{E}$ *and* $\boldsymbol{B} = \mu_0\boldsymbol{H}$

In vacuum $\rho = 0$ *and* $\boldsymbol{J} = 0$

3. Equation of continuity : $\dfrac{\partial\rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{v}) = 0$

4.23.4. : Vector equation involving Laplacian

1. Poisson's equation $\nabla^2\phi = -\dfrac{\rho}{\varepsilon_0}$
2. Laplace's equation : $\nabla^2\phi = 0$
3. Wave equation : $\nabla^2\phi = \dfrac{1}{c^2}\dfrac{\partial^2\phi}{\partial t^2}$
4. Diffusion (or heat conduction) equation : $\nabla^2\phi = \dfrac{1}{h}\dfrac{\partial\phi}{\partial t}$

Keywords :

Gradient of a scalar function, divergence and curl of vector function, directional and normal derivatives, integral forms of divergence and curl.

4.24 ❑ Summary - II

- Constancy of direction and magnitude of vector function have been discussed with respect to derivatives.
- Derivatives of triple scalar and vector products have been discussed.
- Velocity, acceleration and relative velocity of a particle in terms of time derivatives have been exemplified.
- Gradient, divergence and curl have been defined, obtained their geometrical meanings.
- Physical meaning of directional derivative has been discussed.

4.25 ❑ Vector Integration

There are plenty of uses of integration in physics. Relevant integrals are set up to represent physical quantities such as volume, mass, moment of inertia etc. and then evaluated by suitable methods. The basic idea behind setting up of the evaluating is that an integral is the limit of a sum.

Objective

This unit deals with definition of multiple integrals with special reference to double and triple integrals with examples. Algebraic method suitable to find the elements of area, volume etc. in different co-ordinate systems is developed. Line, Surface and volume integrals of vector fields are discussed. Applications of Gauss's divergence theorem, Stoke's theorem and Green's theorem in a plane are discussed with examples.

4.26 ❑ Double and Triple Integral

In case of single variable x, we define the definite integral $\int_a^b f(x)\,dx$ as the limit of the sum of the areas of the rectangle as shown in the fig (4C.1) and use $\int_a^b f(x)\,dx$ to calculate the area of the curve :

$$\int_a^b f(x)\,dx = \lim_{h\to 0} h\sum_{r=1}^{n} f(a+rh),\ \ h=\frac{b-a}{n}$$

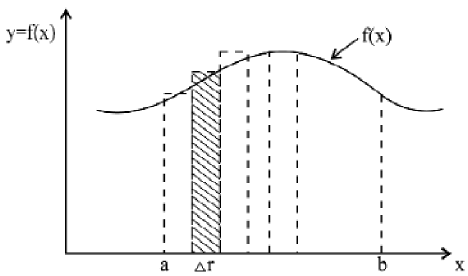


Fig (4C.1)

(4C.1)

From the geometry of fig (4C.1) it is apparent that the sum of the areas of the rectangles will tend to the area under the curve in the limit $n \to \infty$,

Where $\int_a^b f(x)\,dx = \lim\limits_{\substack{n\to\infty \\ \Delta r\to 0}} \sum_{r=1}^{n} f(x_r)\,\delta r$, and Δr is the width of the rectangle.

(4C.2)

We define the double integral of $f(x,y)$ over the area A is the (x, y) plane.

[Fig 4C.2] as a limit of the sum and write it as,

$$\iint\limits_A f(x,y)dA = \lim_{\substack{n\to\infty \\ \Delta A_r\to 0}} \sum_{r=1}^{n} f\left(x_r, y_r\right)\delta A_r \qquad (4C.3)$$

Where the elementary area dA can be chosen accordingly. In Cartesian co-ordinate elementary area $\delta A_r = \delta x_r \delta y_r$ $\quad \therefore dA = \lim\limits_{\substack{\delta x\to 0 \\ \delta y\to 0}} \delta x \delta y = dxdy$. In polar co-ordinate elementary area

$$\delta A = r\delta\theta.\delta r = r\delta\theta\delta r$$

$$\therefore\ dA = \lim_{\substack{\delta r\to 0 \\ \delta\theta\to 0}} r\delta\theta\delta r = rd\theta dr$$

In fig 4C.2 we have divided the (x, y) plane into little rectangles of area $\Delta A_r = \Delta x_r \Delta y_r$. Above each $\Delta x_r \Delta y_r$ is a box reaching up to the surface. We can approximate the volume of this cylinder by a sum of those boxes as represented by the double integral (4C.3).

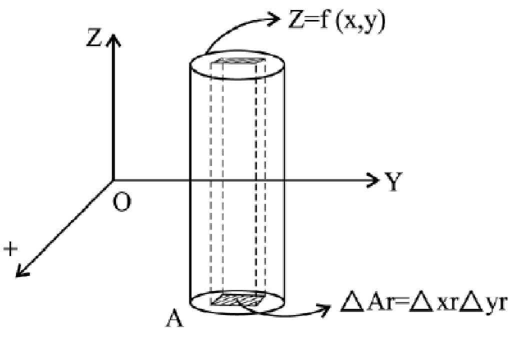


Fig. (4C. 2)

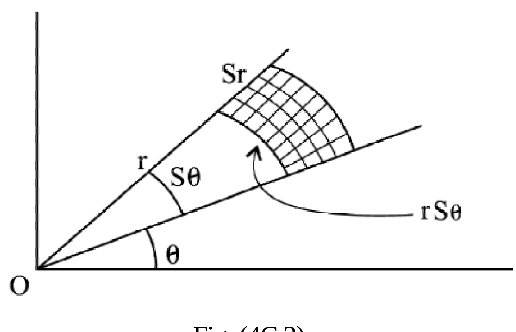


Fig. (4C.3)

Multiple integrals are usually evaluated by using interated (repeated) integrals.

A triple integral of $f(x, y, z)$ over a volume V, written $\iiint_V f(x, y, z)dxdydz,$, is also defined as the limit of a sum and is evaluated by an iterated integral. We consider a function $f(x, y, z)$ to be defined at every point in a region bounded by a volume V. If (x_r, y_r, z_r) be any point in the r^{th} element of volume δV_r , then the limit of the sum

$$\sum_{r=1}^{\infty} f\left(x_r, y_r, z_r\right)\delta V_r$$

If it exits as $r \to \infty$ and $\Delta V_r \to 0$ then,

$$\lim_{\substack{r\to\infty \\ \Delta V_r\to 0}} \sum_{r=1}^{n} f\left(x_r, y_r, z_r\right)\delta V_r$$

$$= \iiint_V f(x, y, z)\, dV = \iiint_V f(x, y, z)\, dx\, dy\, dz \tag{4C.4}$$

is known as the triple integral of $f(x, y, z)$ over the volume V.

Now volume element in Cartesian co-ordinate system is $dV = dxdydz$ but in cylindrical and spherical co-ordinate systems are respectively $dV = r\ drd\theta dz$ and $uV = r^2 \sin\theta\, d\theta\, d\phi\, dr$, we summarize below expression for line element, surface element and volume element in different co-ordinate systems.

Cylindrical co-ordinates (r, θ, z) :

$$\left.\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}\right\}$$ (transformation equations from rectangular to cyliderical co-ordinates)

$dV = r\ drd\theta dz\}$ element of volume.

$dS^2 = (dr)^2 + r^2(d\theta)^2 + (dz)^2\}$ line element

$dA = rd\theta dz\}$ surface element

Spherical co-ordinate (r, θ, ϕ):

$$\left.\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned}\right\}$$ transformation equations from rectangular to spherical co-ordinates

$$dV = r^2 \sin\theta d\theta d\phi dr\Big\}\text{ volume element}$$

$$dS^2 = (dr)^2 + r^2(d\theta)^2 + r^2\sin^2\theta(d\phi)^2\Big\}\text{ line element}$$

$$dA = r^2 \sin\theta d\theta d\phi\Big\}\text{ surface element.}$$

4.26.1 : Examples of double integration

Problem 1: Evaluated $\iint_A xy\ dx\ dy$, where A is the domain bounded by X-axis, ordinate $x = 2a$ and the curve $x^2 = 4ay$.

Solution 1:

Point of intersection Q [fig 4C.4] is given by (2a,a). The domain of intersection is OPQ.

$$\therefore \iint\limits_A xy\,dxdy = \int_0^{2a}\int_0^{\frac{x^2}{4a}} xy\ \ dxdy$$

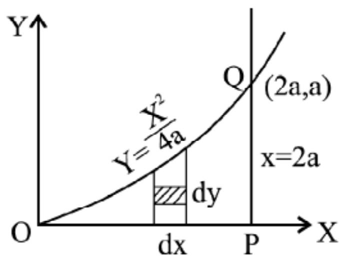


Fig. (4C.4)

First method :

Here first we integrate with respect to y, treating x as constant between the limits 0 to $\frac{x^2}{4a}$. The limits of integrations are $x = 0$ to $2a$ and $y = 0$ to $\frac{x^2}{4a}$.

$$\therefore\ \iint_A xy\ dxdy = \int_0^{2a} x\left[\int_0^{\frac{x^2}{4a}} ydy\right]dx = \int_0^{2a} x\left[\frac{y^2}{2}\right]_0^{\frac{x^2}{4a}} dx$$

$$= \int_0^{2a}\frac{x^5}{32a^2}dx = \frac{1}{32a^2}\int_0^{2a} x^5 dx = \frac{1}{32a^2}\left[\frac{x^6}{6}\right]_0^{2a}$$

$$= \frac{1}{192a^2}\left[(2a)^6 - 0\right] = \frac{64a^6}{192a^2} = \frac{a^4}{3}$$

Second method :

We now integrated w.r.t. x first, treating y as constant between the limits $\sqrt{4ay}$ to 2a [fig 4C.5]. The limits of integration are $y = 0$ to a and $x = \sqrt{4ay}$ to 2a.

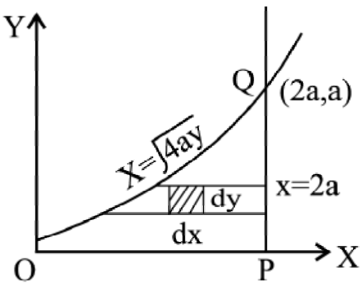


Fig. (4C.5)

$$\therefore \iint_A xy\,dxdy = \int_0^a y\left[\int_{\sqrt{4ay}}^{2a} x\,dx\right]dy$$

$$= \int_0^a y\left[\frac{x^2}{2}\right]_{\sqrt{4ay}}^{2a} dy$$

$$= \int_0^a y(2a^2 - 2ay)\,dy = 2a\int_0^a \left(ay - y^2\right)dy = \frac{a^4}{3}$$

Problem 2 : Evaluate $\int_{x=0}^{2}\int_{y=0}^{x} x^2(x+y)\,dx\,dy$

Solution 2 : $\int_{x=0}^{2} dx \int_{y=0}^{x} (x^3 + x^2 y)dy = \int_{x=0}^{2} dx\left[x^3 y + \frac{x^2 y^2}{2}\right]_0^x$

$$= \int_{x=0}^{2}\left[x^4 + \frac{x^4}{2}\right]dx = \frac{48}{5}$$

Problem 3 : Find by double integration, the area inside the circle $r = a\sin\theta$ and the cardioid $r = a(1-\cos\theta)$.

Solution 3 : Points of intersection of the circle : $r = a\sin\theta$ and the cardioid : $r = a(1-\cos\theta)$ are (0, 0) and $\left(a, \frac{\pi}{2}\right)$ since when

$$r = a,\ \theta = \frac{\pi}{2};\ r = 0,\ \theta = 0$$

$\therefore$ required area within the circle and cardioid is :

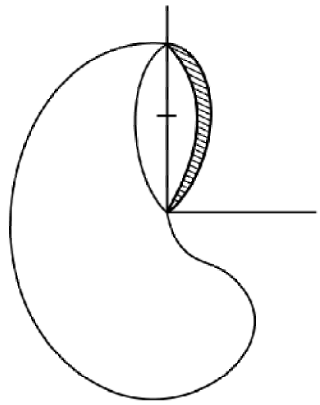


Fig. (4C.6)

$$\int_{\theta=0}^{\theta=\frac{\pi}{2}}\int_{r=a(1-\cos\theta)}^{r=a\sin\theta} r\,d\theta\,dr=\int_{\theta=0}^{\theta=\frac{\pi}{2}} d\theta\int_{r=a(1-\cos\theta)}^{r=a\sin\theta} r\,dr$$

$$=\frac{1}{2}\int_{\theta=0}^{\theta=\frac{\pi}{2}} d\theta\left[r^2\right]_{a(1-\cos\theta)}^{a\sin\theta}$$

$$=\frac{1}{2}\int_{\theta=0}^{\theta=\frac{\pi}{2}} d\theta\left[a^2\sin^2\theta-a^2(1-\cos\theta)^2\right]$$

$$=\frac{a^2}{2}\int_0^{\frac{\pi}{2}}\left[\sin^2\theta-(1-\cos\theta)^2\right]d\theta$$

$$=\frac{a^2}{2}\int_0^{\frac{\pi}{2}}\left[\sin^2\theta-\left(1-2\cos\theta+\cos^2\theta\right)\right]d\theta$$

$$=\frac{a^2}{2}\int_0^{\frac{\pi}{2}}\left[-1+2\cos\theta+\sin^2\theta-\cos^2\theta\right]d\theta$$

$$=\frac{a^2}{2}\int_0^{\pi/2}\left[2\cos\theta-2\cos^2\theta\right]d\theta$$

$$=a^2\int_0^{\pi/2}\left(\cos\theta-\cos^2\theta\right)d\theta$$

$$=a^2\left(1-\pi/2\right)$$

Problem 4 : Find the volume bounded by the cylinder $x^2 + y^2 = 4$ and the planes $y+z = 4$ and $z = 0$

Solution 4 : From fig (4C.7) it is evident that the required volume V,

$V=\iint_A z\,dx\,dy$, where A is the circle $x^2 + y^2 = 4$ in XOY plane.

$$\therefore V=\iint(4-y)\,dx\,dy$$

Or

$$V = \int_{y=-2}^{y=+2} dy \int_{x=-\sqrt{4-y^2}}^{x=\sqrt{4-y^2}} dx$$

$$= 2\int_{y=-2}^{y=+2} dy \int_{x=0}^{x=\sqrt{4-y^2}} (4-y)dx$$

$$= 2\int_{y=-2}^{y=+2} dy \left[4x - yx\right]_0^{\sqrt{4-y^2}}$$

$$= 2\int_{y=-2}^{y=+2} dy \left[4\sqrt{4-y^2} - y\sqrt{4-y^2}\right]$$

$$= 2\int_{y=-2}^{y=+2} 4\sqrt{4-y^2}\,dy - 2\int_{y=-2}^{y=+2} y\sqrt{4-y^2}\,dy$$

$= 16\pi$ units

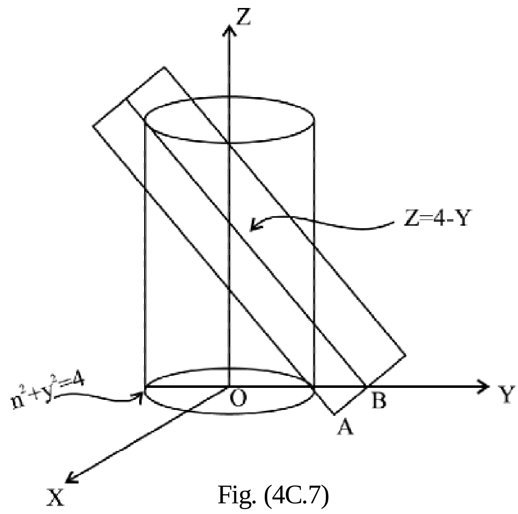


Fig. (4C.7)

Now to evaluate $2\int_{y=-2}^{y=+2} 4\sqrt{4-y^2}\,dy$ we put $y = 2\cos\theta \;\therefore\; y^2 = 4\cos\theta$ and dy

$= -2\sin\theta\, d\theta$ when $y = 2, \cos\theta = 1 \;\therefore \theta = 0$ and when y

$= -2, \cos\theta = -1$ and $\theta = \pi$

$\therefore$ the given integral $2\int_{\pi}^{0} 8\sin\theta(-2\sin\theta)d\theta = -32\int_{\pi}^{0} \sin^2\theta\, d\theta$

$$= 16\int_{0}^{\pi} (1-\cos 2\theta)\, d\theta = 16\pi$$

similarly it can be shown, $2\int_{y=-2}^{y=+2} 4\sqrt{4-y^2}\,dy = 0$

4.26.2 Change of order of integration

In case of double integration we have two method of evaluting the double integration by using iterated integrals. It is often seen that one of the methods we use is more convenient; we choose the easier method. It is common experience that if we change the order, the corresponding limits of the variables are to be changed.

Example 1 : Change the order of integration in the integral :

$$\int_{x=0}^{1}\int_{y=x^2}^{\sqrt{x}} f(x, y)dy\,dx$$

Solution 1 : In the given integral, integration with respect to y should be given first preference. So, the strip AB parallel to y- axis with thickness dx is considered. Finally the strip AB moves from $x = 0$ to $x = 1$ and the area is obtained.

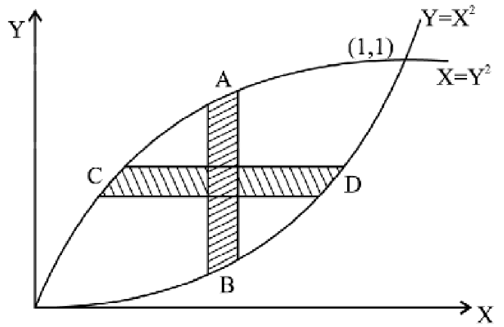


Fig. (4C.8)

While changing the order, let us consider the strips CD with thickness dy, parallel to the x-axis is considered. Now C is on $x = y^2$ and D is on $x = \sqrt{y}$ and y moves from 0 to 1, we write the above integral as :

$$\int_{y=0}^{y=1} dy \int_{x=y^2}^{x=\sqrt{y}} f(x, y)\,dx$$

Example 2 : Evaluated the following integral by changing the order of integration :

$$\int_0^a \int_{\frac{x^2}{a}}^{2a-x} xy\,dx\,dy$$

Solution 2 : In the given order, first we integrate with respect to y, from $y = \dfrac{x^2}{a}$ to $y = 2a - x$ along the vertical strip MN; then the strip MN moves parallel to x-axis from $x = 0$ to $x = a$ as shown in fig (4C.9).

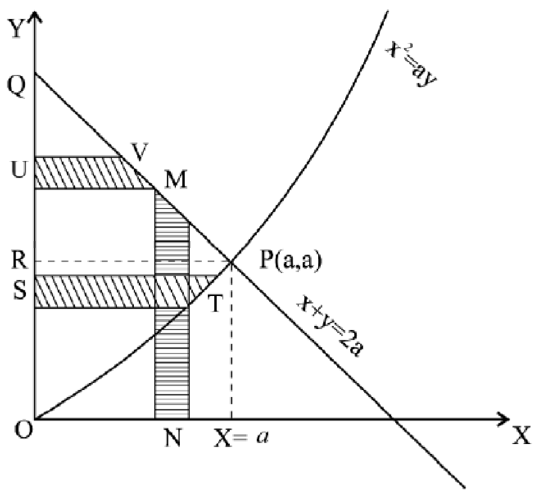


Fig. (4C.9)

When the order of integration is reversed, we have to consider the total area OPQ as a sum of two similar areas PQR and POR.

On the area OPR, we consider the strip ST parallel to x-axis and integrated with respect to x from x = 0 to $x=\sqrt{ay}$ and ST slides from y = 0 to y = a.

Again on the area RPQ, we consider the strips UV and integrated from x = 0 to x = 2a – y and finally the strips slides from y = a to y = 2a.

$$\therefore I = \int_{y=0}^{a}\left[\int_{x=0}^{\sqrt{ay}} xy\,dx\right]dy + \int_{y=0}^{2a}\left[\int_{x=0}^{2a-y} xy\,dx\right]dy$$

$$= \int_{y=0}^{a}\left[\frac{x^2y}{2}\right]_0^{\sqrt{ay}} dy + \int_{y=a}^{y=2a}\left[\frac{x^2y}{2}\right]_0^{2a-y} dy$$

$$= \frac{1}{2}\int_{y=0}^{a} ay^2 dy + \frac{1}{2}\int_{y=a}^{y=2a} (2a-y)^2 y\,dy$$

$$= \frac{a}{2}\left[\frac{y^3}{3}\right]_0^a + \frac{1}{2}\left[2a^2y^2 - \frac{4ay^3}{3} + \frac{y^4}{4}\right]_a^{2a} = \frac{3}{8}a^4$$

4.26.3 Examples of triple integral

1. **Volume as triple integral :** The volume of a solid is obtained by evaluating the triple integral fig (4C.10),

$$V = \iiint \delta V = \iiint \delta x\,\delta y\,\delta z$$

If ρ is the density of the solid, it's mass,

$$M = \iiint \rho\,\delta x\,\delta y\,\delta z$$

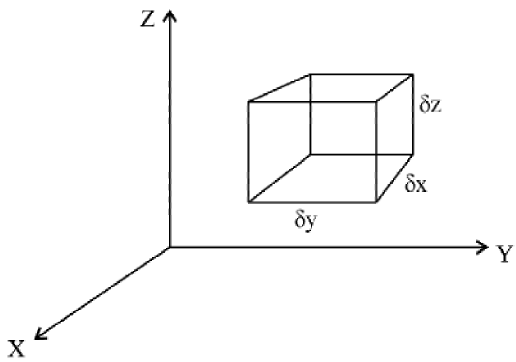


Fig. (4C.10)

2. **Volume as volume of revolution :** We consider an element of area $\delta x\,\delta y$ on a plane area A. The revolution of the element about X-axis will generate a ring of volume (4C.11) :

$$\delta V = \pi\left[(y+\delta y)^2 - y^2\right]\delta x$$

$$= \pi\left[y^2 + (\delta y)^2 + 2y\delta y - y^2\right]\delta x$$

$$= \pi\left[2y\delta y + (\delta y)^2\right]\delta x$$

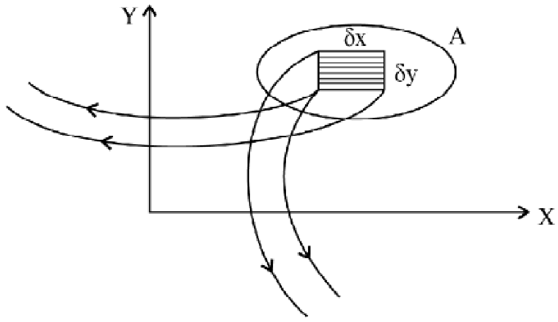


Fig. (4C.11)

Since δy is small, $(\delta y)^2$ is smaller, and we can neglect it in comparison to $2y\delta y$.

$$\therefore \delta V \approx 2\pi y\delta y\delta x$$

therefore the whole volume generated by the entire area about the X-axis is

$$V_1 = \iint_A 2\pi\ y\ dy\ dx$$

Similarly, δV about Y-axis is

$$\delta V = \pi\left[(x+\delta x)^2 - x^2\right]\delta y$$

$$\approx 2\pi x\ \delta x\ \delta y, \text{ neglecting } (\delta x)^2$$

$$V_2 = \iint_A 2\pi x\ dx\ dy$$

4.27 ❑ Change of Variables : Jacobian

It is convenient to develop an algebraic method suitable to find the elements of area, volume etc. in different co-ordinate systems and also at the same time for any change of variables in a multiples integral.

In two dimensions suppose x and y are given functions of two new variables u and v by the transformation equations.

$$x = x(u, v) \ \& \ y = y(u, v)$$

The Jacobian of (x, y) with respect to (u, v) is the determinant in equation (4C.5)

$$J = J\left(\frac{x, y}{u, v}\right) = \frac{\partial(x,y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} \qquad (4C.5)$$

Then the area element $dy\ dx$ is replaced in (u, v) system by the area element

$$dA = |J|\ du\ dv \qquad (4C.6)$$

where $|J|$ is the absolute value of the Jacobian given in equation (4C.5).

In three dimensions, suppose $= x(u, v, w)$; $y = y(u, v, w)$ and $z = z(u, v, w)$ then the Jacobian of transformation from (x, y, z) to (u, v, w) is given by

$$J = \frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} & \dfrac{\partial y}{\partial w} \\ \dfrac{\partial z}{\partial u} & \dfrac{\partial z}{\partial v} & \dfrac{\partial z}{\partial w} \end{vmatrix} \qquad (4C.7)$$

Then the volume element $dx\ dy\ dz$ is replaced in (u, v, w) system by the volume element

$$dV = |\ J\ |\ du\ dv\ dw \qquad (4C.8)$$

Evaluation of double and triple integral becomes easier by change of variables. Two important formulae are listed below for the purpose.

1. Double integral :

$$\iint_{R_{xy}} f(xy)dx\,dy = \iint_{R_{uv}} f\left[x(u,v), y(u,v)\right] |\,J\,|\,du\,dv \qquad (4C.9)$$

R_{xy} and R_{uv} are symbols of region in xy-plane and uv-plane respectively.

Where $|J| = \left|\dfrac{\partial(x, y)}{\partial(u,v)}\right|$

2. Triple integral :

$$\iiint\limits_{R_{xyz}} f(x, y, z)dx\,dy\,dz$$

$$= \iiint\limits_{R_{uvw}} f\left[x(u,v,w),\, y(u,v,w),\, z(u,v,w)\right]|J|\, du\, dv\, dw \qquad (4C.10)$$

where $|J| = \left|\dfrac{\partial(x,y,z)}{\partial(u,v,w)}\right|$

Special cases :

A> **Cartesian to polar co-ordinate system :** $(x, y) \to (r, \theta)$

Transformation equations are $x = r\cos\theta,\ y = r\sin\theta$

$$J = \frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta$$

$$= r\left(\cos^2\theta + \sin^2\theta\right) = r$$

Therefore, element of area $dA = dx\,dy = |J|\,dr\,d\theta = r\,dr\,d\theta$

And $\iint_{R_{xy}} f(x,y)\,dx\,dy = \iint_{R_{r\theta}} f(r\cos\theta, r\sin\theta)|J|\,dr\,d\theta$

$$= \iint\limits_{R_{r\theta}} f(r\cos\theta, r\sin\theta)\, r\, dr\, d\theta$$

B> Cartesian to cylindrical co-ordinate system : $(x, y, z) \to (r, \theta, z)$

Transformation equations are $x = r\cos\theta,\ \ y = r\sin\theta,\ \ z = z$

$$\therefore J = \frac{\partial(x,y,z)}{\partial(r,\theta,z)} = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} & \dfrac{\partial x}{\partial z} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} & \dfrac{\partial y}{\partial z} \\ \dfrac{\partial z}{\partial r} & \dfrac{\partial z}{\partial \theta} & \dfrac{\partial z}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

Or, $J = r$

Therefore

$$\iiint_{R_{xyz}} f(x, y, z)\, dx\, dy\, dz = \iiint_{R_{r\theta z}} f(r\cos\theta, r\sin\theta, z)\left|J\right| dr\, d\theta\, dz$$

$$= \iiint_{R_{r\theta z}} f(r\cos\theta, r\sin\theta, z) r\, dr\, d\theta\, dz$$

C> **Cartesian to spherical polar co-ordinates :** $(x, y, z) \rightarrow (r, \theta, \phi)$

Transformation equations are

$x = r\sin\theta\cos\phi$

$y = r\sin\theta\sin\phi$

$z = r\cos\theta$

$$\therefore J = \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix} = \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}$$

$$= r^2\sin\theta\left[-\sin^2\phi(-\sin^2\theta - \cos^2\theta) - \cos^2\phi\left(-\sin^2\theta - \cos^2\theta\right)\right]$$

$$= r^2\sin\theta$$

$\therefore$ Volume element $dV = dxdydz = \left|J\right| drd\theta d\phi = r^2\sin\theta\, dr\, d\theta\, d\phi$

Exercise of 4.26 and 4.27 :

1) Change the order of the integration in $\int_0^\infty \int_x^\infty \frac{e^{-y}}{y} dxdy$ and hence find its value.

2) Evaluate the integral $\iiint_V \left(x^2 + y^2 + z^2\right) dx\, dy\, dz$ where V is the volume of the sphere with center at the origin and radius R.

3) Evaluate $\iint_R (x+y)^2\, dx\, dy$, where R is the parallelogram in the xy – plane with vertices (1, 0), (3, 1), (2, 2) and (0,1) using the transformation, $u = x + y$, $v = x - 2y$.

4) Transform the integral $\int_{-1}^{+1}\int_{x}^{\frac{1}{x}} V\, dx\, dy\, dz$ by the substitution i.e. $u = 1 + x$ and $v = xy$

5) Given the transformation $x = u^2 - v^2$, $y = 2uv(a)$ compute its Jacobian; (b) Evaluate $\left(\frac{\partial u}{\partial x}\right)_y$ and $\left(\frac{\partial v}{\partial x}\right)_y$.

Solution :

Solution 1 : The region of integration is bounded by $y = x$, $x = 0$ and infinity boundary.

We take a strip parallel to x – axis to change the order.

The extremities of the strips lies on $x = 0$ and $y = x$. Therefore limits of x are from $x = 0$ to $x = y$ and the limits of y are from $y = 0$ to $y = \infty$

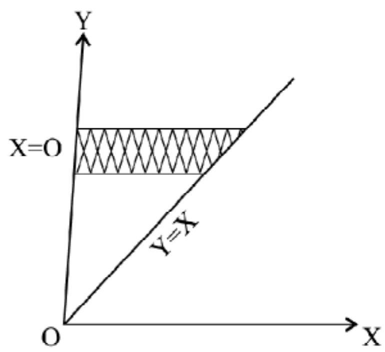


Figure : Solution (1)

$$\therefore \int_0^{\infty}\int_0^{y} \frac{e^{-y}}{y}\, dx\, dy = \int_0^{\infty}\int_0^{y} \frac{e^{-y}}{y}\, dy\, dx$$

$$= \int_0^{\infty} \frac{e^{-y}}{y}[x]_0^y\, dy = \int_0^{\infty} e^{-y} dy = \left[-e^{-y}\right]_0^{\infty} = 1$$

Solution 2 : Using spherical polar co-ordinate

$$x = r\sin\theta\cos\phi,\ y = r\sin\theta\sin\phi,\ z = r\cos\theta$$

$$\left(x^2+y^2+z^2\right)dx\, dy\, dz = r^2 r^2 \sin\theta\, d\theta\, d\phi\, dr = r^4 \sin\theta\, d\theta\, d\phi\, dr$$

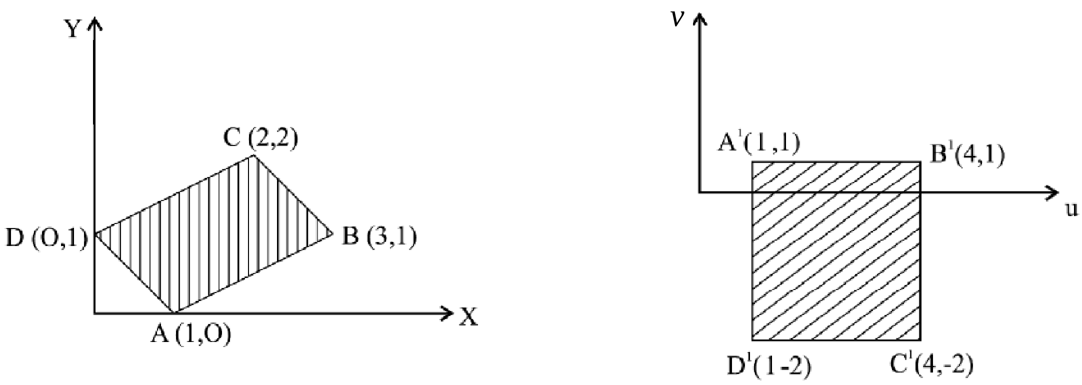
$$\therefore \iiint_V \left(x^2+y^2+z^2\right)dx\, dy\, dz = \int_{\theta=0}^{\pi}\int_{\theta=0}^{2\pi}\int_{r=0}^{R} r^4 \sin\theta\, d\theta\, d\phi\, dr$$

$$= \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta \, d\theta \int_0^{R} r^4 dr = \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta \, d\theta \left(\frac{R^5}{5} \right)$$

$$= \frac{R^5}{5} \cdot 2\pi \left[-\cos\theta \right]_0^{\pi} = \frac{2\pi R^5}{5} [-\cos\pi + \cos 0]$$

$$= \frac{2\pi R^5}{5} [1 - \cos\pi] = \frac{2\pi R^5}{5} [1 - (-1)] = \frac{4\pi R^5}{5}$$

Solution 3 : The region R, i.e. the parallelogram ABCD in xy – plane because they became the region R', i.e. the rectangle A´B´C´D´ in the uv – plane.



(Figure solution 3)

Where $u = x + y$, $v = x - 2y$

$$\therefore x = \frac{1}{3}(2u + v), \; y = \frac{1}{3}(u - v)$$

Now Jacobian of transformation,

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} = -\frac{1}{3}$$

$\therefore$ the given integral is,

$$\iint_{R'} u^2 |J| \, du \, dv = \int_{u=1}^{4} \int_{v=-2}^{1} u^2 \frac{1}{3} du \, dv = \int_{u=1}^{4} u^2 \frac{1}{3} du \, [v]_{-2}^{1}$$

$$= \int_{u=1}^{4} u^2 \frac{1}{3} du.3 = \int_{u=1}^{4} u^2 du$$

$$= \left[\frac{u^3}{3}\right]_1^4 = \frac{4^3}{3} - \frac{1^3}{3} = \frac{1}{3}(64-1) = \frac{63}{3} = 21$$

Solution 4 : $u = 1 + x$; $v = xy$

$$\therefore x = u - 1,\ y = \frac{v}{u-1}$$

$$\therefore\ jacobian \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & -\frac{v}{(u-1)^2} \\ 0 & \frac{1}{u-1} \end{vmatrix} = \frac{1}{u-1}$$

so, $dx\, dy = |J|\, du\, dv = \frac{1}{u-1} du\, dv$

Now, the limits of y are x to $\frac{1}{x}$ also $v = xy$, so v varies from x^2 to 1 i.e. $(u - 1)^2$ to 1. Limits of x varies from – 1 to + 1.

So, limits is u varies from 0 to 2.

Hence $\int_{-1}^{+1} \int_{x}^{\frac{1}{x}} V\, dx\, dy = \int_{u=0}^{u=2} \int_{v=(u-1)^2}^{v=1} V' \frac{1}{u-1} du\, dv$

where V' is the function V changed in u and v.

Solution 5 :

a) The Jacobian $J(u, v)$ of u, v with respect to x, y is,

$$J(u, v) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 2u & -2v \\ 2v & 2u \end{vmatrix} = 4(u^2 + v^2)$$

b) We have $u^2 - v^2 = x$ and $2uv = y$

$$\therefore \left(u^2+v^2\right)^2 = \left(u^2-v^2\right)^2 + 4u^2v^2 = x^2+y^2$$

$$\therefore u^2 = \frac{1}{2}\left(x+\sqrt{x^2+y^2}\right)$$

Differentiating partially with respect to x

$$2u\left(\frac{\partial u}{\partial x}\right)_y = \frac{1}{2}\left(1+\frac{x}{\sqrt{x^2+y^2}}\right) = \frac{1}{2}\left(\frac{x+\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}}\right) = \frac{u^2}{\sqrt{x^2+y^2}}$$

$$\therefore \left(\frac{\partial u}{\partial x}\right)_y = \frac{u}{2\sqrt{x^2+y^2}}$$

Again $v^2 = \frac{1}{2}\left(\sqrt{x^2+y^2}-x\right)$

$$\therefore 2v\left(\frac{\partial u}{\partial x}\right)_y = \frac{1}{2}\left(\frac{x}{\sqrt{x^2+y^2}}-1\right) = -\frac{v^2}{u^2+v^2}$$

$$\therefore \left(\frac{\partial v}{\partial x}\right)_y = -\frac{v}{2(u^2+v^2)}$$

4.28 ❐ Ordinary integrals of vector

Vectors which are functions of single variable, are integrated in the same way as scalar. Thus if $\boldsymbol{V}(t) = \boldsymbol{i}V_x(t) + \boldsymbol{j}V_y(t) + \boldsymbol{k}V_z(t)$, then

$$\int \boldsymbol{V}(t)dt = \boldsymbol{i}\int V_x(t)dt + \boldsymbol{j}\int V_y(t)dt + \boldsymbol{k}\int V_z(t)\, dt \qquad (4C.11)$$

(4C.11) is indefinite integral of $\boldsymbol{V}(t)$

However if $\boldsymbol{V}(t) = \frac{d}{dt}\boldsymbol{R}(t)$ then,

$\int \boldsymbol{V}(t)dt = \boldsymbol{R}(t) + \boldsymbol{C}$, where $\boldsymbol{C}$ is a constant vector. The definite integral of $\boldsymbol{V}(t)$ between the limits $t = T_0$ to $t = T$, is given by,

$$\int_{T_0}^{T} \boldsymbol{V}(t)dt = \int_{T_0}^{T} \frac{d}{dt}\boldsymbol{R}(t)dt = \left[\boldsymbol{R}(t) + \boldsymbol{C}\right]_{T_0}^{T} = \boldsymbol{R}(t) + \boldsymbol{R}(T_0) \qquad (4C.12)$$

4.28.1 Line integral of a vector field

We consider a vector field with field vector $\boldsymbol{V}$. We draw a continuous curve C in the field (Fig. 4C.12) where $\boldsymbol{V}$ is defined at every point on it. Let us choose $(n–1)$ points C_i, which divide the curve C into n segments $\Delta \boldsymbol{C}_i$ and form the sum $\sum_{i=0}^{n-1} \boldsymbol{V}_i . \Delta \boldsymbol{C}_i$ (4C.13), where $\mathbf{V}_i$ is the value of $\boldsymbol{V}$ at C_i and $\Delta \boldsymbol{C}_i$ is the vector whose rectangular components are $\boldsymbol{i}\Delta x_i$, $\boldsymbol{j}\Delta y_i$, $\boldsymbol{k}\Delta z_i$ and which joins C_{i-1} and C_i.

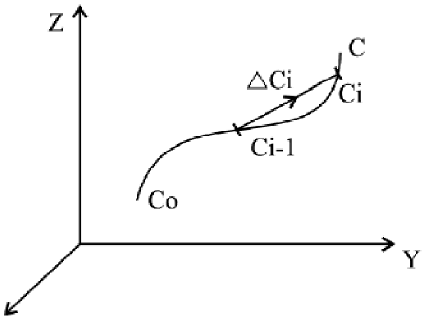


Fig. (4C.13)

Now taking the limit of the sum (4C.13), e.g.

$\lim\limits_{\substack{i\to\infty \\ \Delta C_i - 0}} \sum_{i=0}^{n-1} \boldsymbol{V}_i \cdot \Delta \boldsymbol{C}_i$ (4C.13) and if then the sum approaches a definite limit, then this limit is defined as the line integral $\int_C \boldsymbol{V}.d\boldsymbol{r}$ along the curve C. i.e.

$$\lim_{\substack{i\to\infty \\ \Delta C_i - 0}} \sum_{i=0}^{n-1} \boldsymbol{V}_i \cdot \Delta \boldsymbol{C}_i = \int_C \boldsymbol{V}.d\boldsymbol{r} \qquad (4C.14)$$

In general the value of the line integral depends upon $\boldsymbol{V}$ and the path C joining the end points C_0 and C.

We note the followings :

1. If $\boldsymbol{V}(\boldsymbol{r})$ be the force acting on a particle which moves along the curve from C_0 to C, then the line integral (4C.14) represents the work done by the force. If this line integral along a closed path $\oint_C \boldsymbol{V}.d\boldsymbol{r}$ is zero, then $\boldsymbol{V}(\boldsymbol{r})$ is called a conservative force field and in that case $\boldsymbol{V}(\boldsymbol{r})$ is also called irrotational.
2. If $\boldsymbol{V}(\boldsymbol{r})$ is a conservative force filed, then the line integral (4C.14) does not depend on the choice of path connecting any two points on the curve i.e. it becomes

independent of path, between any two points on the curve and in that case $\boldsymbol{V}(\boldsymbol{r})$ can be expressed as a gradient of a scalar, called potential function (say ϕ) of the force field i.e. $\boldsymbol{V}(\boldsymbol{r}) = -\nabla\phi$

3. If $\boldsymbol{V}(\boldsymbol{r})$ represents the velocity of fluid flow, then the integral is called the circulation of $\boldsymbol{V}$ along the closed curve C. When the circulation of $\boldsymbol{V}(\boldsymbol{r})$ along a closed curve is zero the $\boldsymbol{V}(\boldsymbol{r})$ is called irrotational.

   Essentially line integral is an integration along a curve and there is only one independent variable. Therefore to evaluate a line integral we have to transform the integrand in terms of a single variable using the equation of the curve along which integration is required to be evaluated.

Example Art of 4.28.1 :

Example 1 : Prove that if the closed line integral of a field vector $\boldsymbol{A}$ vanishes, then it is the gradient of a field scalar .

Solution 1 : Let P_1 and P_2 be two points connected by any two curves P_1QP_2 and P_1RP_2, forming a closed curve $P_1QP_2RP_1$. Now

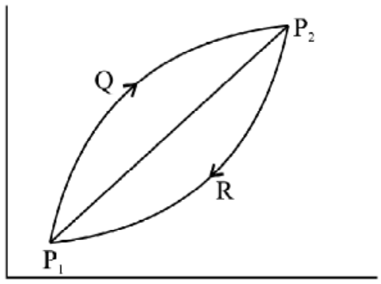


Fig. Example (1)

$$\oint_C \boldsymbol{A}.\boldsymbol{dr} = \int_{P_1}^{P_2} \boldsymbol{A}.\boldsymbol{dr} = 0$$

Or, $$\oint_C \boldsymbol{A}.\boldsymbol{dr} = \int_{P_1QP_2} \boldsymbol{A}.\boldsymbol{dr} + \int_{P_2RP_1} \boldsymbol{A}.\boldsymbol{dr} = 0$$

Or, $$\int_{P_1QP_2} \boldsymbol{A}.\boldsymbol{dr} = -\int_{P_2RP_1} \boldsymbol{A}.\boldsymbol{dr} = \int_{P_1RP_2} \boldsymbol{A}.\boldsymbol{dr} \qquad \text{(i)}$$

From (i) we see that line integral of a vector $\boldsymbol{A}$ is independent of path connecting two points P_1 and P_2 and depends only on the co-ordinates of ends points P_1 and P_2

therefore $$\int_{p_1}^{p_2} \boldsymbol{A}.\boldsymbol{dr} = \phi_{p_2} - \phi_{p_1} \qquad \text{(ii)}$$

where ϕ is some scalar field.

$\therefore \boldsymbol{A}.\boldsymbol{dr} = d\phi$. If P_1 and P_2 are two very close points in the field. Then $\boldsymbol{A}.\boldsymbol{dr} = \nabla\phi \cdot \boldsymbol{dr}$

or, $\left(A-\nabla\phi\right)\cdot \boldsymbol{dr}=0$

which is true for any ***dr*.**

Therefore $\boldsymbol{A}=\nabla\phi$ (iii)

Example 2 : If $\boldsymbol{F}$ = $(5xy - 6x^2)\boldsymbol{i} + (2y - 4x)\boldsymbol{j}$, then calculate the line integral $\oint_C \boldsymbol{F}.\boldsymbol{dr}$ along the curve C in the xy plane given by $y = x^3$ from the point (1, 1) to (2,8)

Solution 2 :

$\boldsymbol{F.dr} = (5xy - 6x^2)dx + (2y - 4x)dy$

We convert ***F.dr*** in terms of x only by substituting $y = x^3$, $dy = 3x^2\ dx$

$\therefore\ \boldsymbol{F.dr} = (5x.x^3 - 6x^2)dx + (2x^3 - 4x)3x^2dx$

$= (5x^4 - 6x^2)dx + (6x^5 - 12x^3)dx$

$$\therefore \int_C \boldsymbol{F}.\boldsymbol{dr} = \int_{x=1}^{x=2}\left(6x^5+5x^4-12x^3-6x^2\right)dx = \left[x^6+x^5-3x^4-2x^3\right]_1^2$$

$= (64 - 1) + (32 - 1) - 3(16 - 1) - 2(8 - 1) = 65 + 31 - 45 - 14 = 35$

Example 3 : If $\boldsymbol{A}=\nabla\phi$, where ϕ is some scalar functions of position, show that the line integral of ***A*** along a curve ***C*** linking two points A and B is independent of the choice of the curve ***C***.

Solution 3 : We consider a close line integral along APBQA of the vector ***A***, i.e. $\oint_{APBQA}\boldsymbol{A.dr}$

Therefore $\oint_{APBQA}\boldsymbol{A.dr} = \oint_{APBQA}\nabla\phi.\boldsymbol{dr} = \oint_{APBQA} d\phi = 0$ (i)

$$\therefore P\int_A^B \boldsymbol{A.dr} + Q\int_B^A \boldsymbol{A.dr} = 0$$

Or, $$P\int_A^B \boldsymbol{A.dr} = -\,Q\int_B^A \boldsymbol{A.dr} = Q\int_A^B \boldsymbol{A.dr} \quad \text{(ii)}$$

For equation (ii) we see that line integral is independent of path connecting A and B.

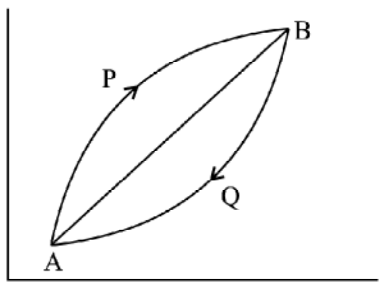


Fig. Example (3)

Example 4 : Calculate the work done when a force $\boldsymbol{F} = 3xy\boldsymbol{i} - y^2\boldsymbol{j}$ moves a particle in the xy – plane from (0, 0) to (1,2) along the parabola $y = 2x^2$

Solution 4 : $\boldsymbol{F} = 3xyi - y^2j = 3x(2x^2)\boldsymbol{i} - (2x^2)^2\boldsymbol{j} = 6x^3\boldsymbol{j} - 4x^4\boldsymbol{j}$ now, $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}$

or, $\boldsymbol{dr} = dx\boldsymbol{i} + dy\boldsymbol{j} = dx\boldsymbol{i} + 4xdx\boldsymbol{j}$

$$\therefore \int_C \boldsymbol{F}.\boldsymbol{dr} = \int_0^1 (6x^3dx - 16x^5)dx = \left[\frac{3}{2}x^4 - \frac{8}{2}x^6\right]_0^1 = -\frac{7}{6}\, unit$$

Example 5 : A particle of constant mass m is moving in a conservative force field $\boldsymbol{F} = -\nabla\phi$. If A and θ be two points in space, prove that $\phi(A) + \frac{1}{2}mv_A^2 = \phi(B) + \frac{1}{2}mv_B^2$, where v_A and v_B are the magnitudes of velocities of the particle at A and B respectively.

Solution 5 : The work done by the force $\boldsymbol{F}$ is

$$dw = \boldsymbol{F}.\boldsymbol{dr} = m\frac{d^2r}{dt^2}\cdot\boldsymbol{dr} = m\frac{d}{dt}\left(\frac{dr}{dt}\right)\cdot\left(\frac{dr}{dt}\right)dt$$

$$= \frac{1}{2}m\frac{d}{dt}\left(\frac{dr}{dt}\right)^2 dt = \frac{1}{2}m\, d(v^2)$$

where $\boldsymbol{v} = \frac{dr}{dt}$

$\therefore$ total work done, $W = \int_A^B \frac{1}{2}m\, d(v^2) = \frac{1}{2}m\left(v_B^2 - v_A^2\right)$ (i)

but $\boldsymbol{F} = -\nabla\phi$

$$W = \int_A^B \boldsymbol{F}.\boldsymbol{dr} = -\int_A^B \nabla\phi.dr = \int_A^B d\phi = \phi_A - \phi_B \qquad \text{(ii)}$$

from (i) and (ii) : $\frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2 = \phi_A - \phi_B$

or, $\phi_B + \frac{1}{2}mv_B^2 = \phi_A + \frac{1}{2}mv_A^2$

where ϕ_A and $\frac{1}{2}mv_A^2$ are respectively potential energy and kinetic energy of the particle at A. Thus total energy at A and B are equal (conservation of energy). This is known as work-energy theorem in mechanics.

Exercise of Art 4.28.1 :

1) If $\boldsymbol{F} = (3x^2 + 6y)\boldsymbol{i} - 14yz\boldsymbol{j} + 20\,xz^2\boldsymbol{k}$. Evaluate $\int \boldsymbol{F}.d\boldsymbol{r}$ along the straight line from (0,0,0) to (1,1,1).

2) Find the work done in going around a unit circle in the xy plane, (i) counter clockwise from 0 to π (ii) clockwise from 0 to $-\pi$ against a force field given by,

$$F = -\frac{y}{x^2+y^2}\boldsymbol{i} + \frac{x}{x^2+y^2}\boldsymbol{j}$$

Solution :

Solution 1 : We take the parameter t such that $x = t$, $y = t$, $z = t$ varies from 0 to 1.

Now $\boldsymbol{F}.d\boldsymbol{r} = (3x^2 + 6y)dx - 14yz\,dy + 20xz^2\,dz$

$= (3t^2 + 6t)dt - 14t^2dt + 20t^3dt$

$= (3t^2 + 6t - 14t^2 + 20t^3)dt$

$$\therefore \int_C \boldsymbol{F}.d\boldsymbol{r} = \int_{t=0}^{1}\left(3t^2 + 6t - 14t^2 + 20t^3\right)dt = \int_{t=0}^{1}\left(-11t^2 + 6t + 20t^3\right)dt = \frac{13}{3}$$

Solution 2 : We have $\boldsymbol{F} = \frac{-y\boldsymbol{i} + x\boldsymbol{j}}{x^2+y^2}$

Now for unit circle $x^2 + y^2 = 1$

$$\therefore \boldsymbol{F} = \frac{-y\boldsymbol{i} + x\boldsymbol{j}}{x^2+y^2} = -y\boldsymbol{i} + x\boldsymbol{j}$$

Now $dr = \boldsymbol{i}dx + \boldsymbol{j}dy$

$\therefore$ $\boldsymbol{F.dr} = -ydx + x\,dy$, also $x = \cos\theta$, $y = \sin\theta$, $dx = -\sin\theta\, d\theta$, $dy = \cos\theta\, d\theta$

i) Let the counter clockwise path be $C_1(0 \text{ to } \pi)$.

$\therefore$ work done along the counter clockwise is

$$\int_{C_1}(-y\,dx + x\,dy) = \int_0^{\pi}(-\sin\theta - \sin\theta\, d\theta + \cos\theta \cdot \cos\theta\, d\theta)$$

$$= \int_0^{\pi}\left(\sin^2\theta + \cos^2\theta\right)d\theta = \int_0^{\pi} d\theta = \pi$$

ii) The work done along the clockwise path C_2 is

$$\int_{C_2}(-y\,dx + x\,dy) = \int_0^{-\pi}(\sin^2\theta + \cos^2\theta)d\theta = \int_0^{-\pi} d\theta = -\pi$$

Therefore we see that work depends on path. Therefore $\boldsymbol{F}$ is not conservative. Now

$$\oint \boldsymbol{F.dr} = \int_0^{\pi}\boldsymbol{F.dr} + \int_{\pi}^{2\pi}\boldsymbol{F.dr} = \int_0^{\pi} d\theta + \int_{\pi}^{2\pi} d\theta = \pi + 2\pi - \pi = 2\pi$$

Or, using the results of (i) and (ii)

$$\oint_C \boldsymbol{F.dr} = \int_{C_1}\boldsymbol{F.dr} + \int_{C_2}\boldsymbol{F.dr} = \int_0^{\pi}\boldsymbol{F.dr} - \int_0^{-\pi}\boldsymbol{F.dr} = \pi - (-\pi) = 2\pi$$

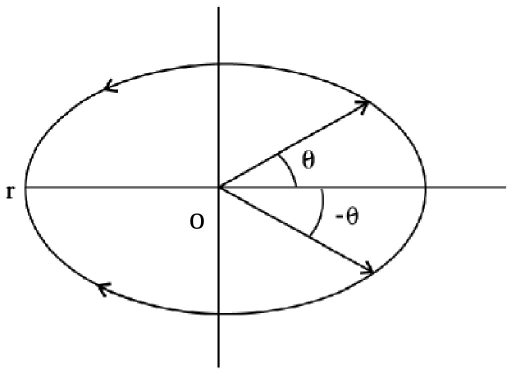


Fig. : Solution (2)
for c_1 : $\theta = 0$ to π
for c_2 : $\theta = 0$ to $-\pi$
Origin 0 is within circle.

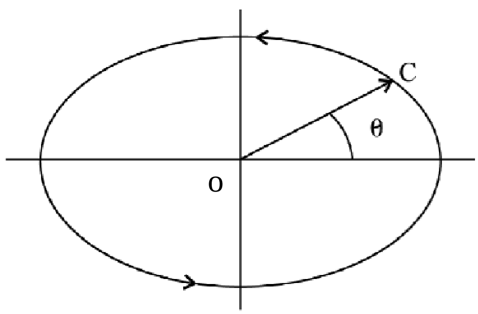


Fig : Solution (2)
For c : $\theta = 0$ to 2π
Origin 0 is within circle

Note : However if the origin o is outside the circle, work done $\oint \boldsymbol{F}.\boldsymbol{dr} = 0$ as shown in fig. below.

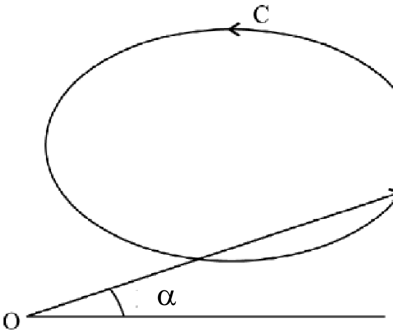


Fig. : Solution (2)
Origin O is outside the circle.

In this case $\oint \boldsymbol{F}.\boldsymbol{dr} = \int_{\alpha}^{\alpha} \boldsymbol{F}.\boldsymbol{dr} = 0$

4.28.2 Surface integral of a vector field

We consider a surface defined *by* $z = f(x, y)$ having continuous first order partial derivatives. Let ds be a small area of the surface and $\hat{\boldsymbol{n}}$ the limit normal vector in the outward direction to this small area. Then the area vector corresponding to this small portion of the surface is $d\boldsymbol{\sigma} = \hat{\boldsymbol{n}}ds$.

The normal surface integral of a continuous vector point function $\boldsymbol{V}(\boldsymbol{r})$ is defined as

$$\iint_S V \cdot d\boldsymbol{\sigma} = \iint_S V \cdot \hat{\boldsymbol{n}}ds \qquad (4C.15)$$

Now the projection of a vector area $\hat{\boldsymbol{n}}ds$ on the xy – plane [Fig 4C.13] whose unit normal is $\boldsymbol{k}$ is given by $\hat{\boldsymbol{n}}ds \cdot \boldsymbol{k} = \hat{\boldsymbol{n}} \cdot \boldsymbol{k}\ ds$

But the projection of ds on xy plane is $dxdy$. Hence $(\hat{\boldsymbol{n}} \cdot \boldsymbol{k})ds = dxdy$,

Therefore $ds = \dfrac{dxdy}{\hat{\boldsymbol{n}} \cdot \boldsymbol{k}}$ and equation (4C.15) can be written as

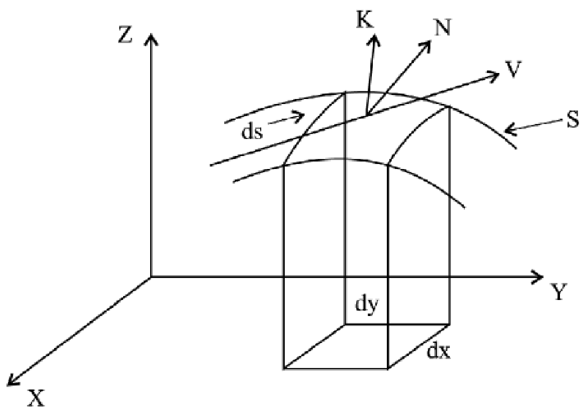


Fig (4C.13)

$$\iint_S \boldsymbol{V}\cdot d\sigma = \iint_s \boldsymbol{V}\cdot\hat{\boldsymbol{n}}ds = \iint_S \boldsymbol{V}\cdot\hat{\boldsymbol{n}}\frac{dxdy}{\hat{\boldsymbol{n}}\cdot\boldsymbol{k}} \quad (4C.16)$$

Similarly, considering the projection of vector area on *yz* plane and zx plane, the surface integral can be expressed as :

$$\iint_S \boldsymbol{V}\cdot\hat{\boldsymbol{n}}\frac{dzdy}{\hat{\boldsymbol{n}}\cdot\boldsymbol{i}} \text{ and } \iint_S \boldsymbol{V}\cdot\hat{\boldsymbol{n}}\frac{dxdz}{\hat{\boldsymbol{n}}\cdot j} \text{ respectively}$$

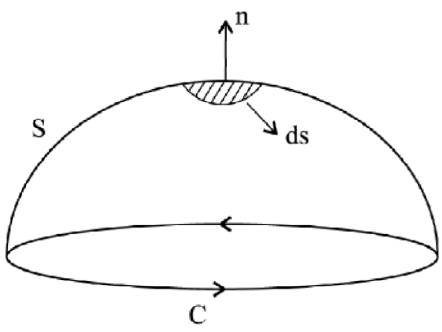


Fig (4C.14)

When a surface enclosed a cetain volume it is called a closed surface. The positive normal to the closed surface is drawn outward from the surface and that is the direction of elementary surface *dS* on the closed surface S.

For open surface the direction of the positive normal is determined from sense of traversing its boundary. If it is right handed then the positive normal is outward. When left handed it is inward as shown in figures 4C.14 and 4C.15 respectively.

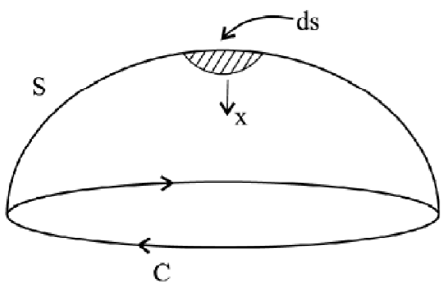


Fig (4C.15)

Flux of a vector field :

The quantity $\iint_S \boldsymbol{V}d\boldsymbol{\sigma} = \iint_S \boldsymbol{V}.\hat{\boldsymbol{n}}\, ds$ is called the flux of the vector filed ***V***. In most case flux cannot be equated to any physical concept. However in the following case we can relate flux to specific physical quantity.

1. When $\boldsymbol{V} = \boldsymbol{v}$, the velocity vector of flow of liquid, $\iint_S v.\hat{\boldsymbol{n}}\, ds$ gives the volume of he liquid crossing the surface *S* per second normally.

2. When $\boldsymbol{V} = \rho\boldsymbol{v}$, where ρ is the density of the following liquid and ***v*** its velocity of flow, $\iint_S \rho\, \boldsymbol{v}.\hat{\boldsymbol{n}}\, ds$ represents the mass of liquid crossing per second normally through the surface.

3. In case of electric and magnetic field flux represents the total number of lines of force crossing the surface normally.

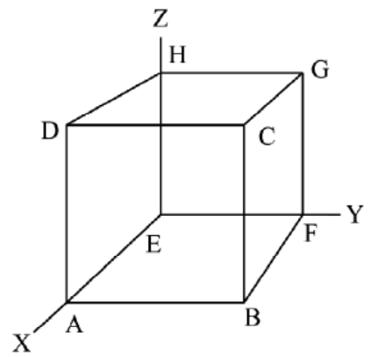
We have also flux of particles and flux of heat, defined similarly.

Exercise of 4.28.2

1. Evaluate $\oiint \boldsymbol{A}.\hat{\boldsymbol{n}}\, ds$ where S is the surface of unit cube bounded by $x = 0$, $x = 1$; $y = 0$, $y = 1$; $z = 0$, $z = 1$ or bounded by co-ordinate planes and the planes $x = 1$, $y = 1$, $z = 1$ when i) $\boldsymbol{A} = \boldsymbol{r}$; ii) $\boldsymbol{A} = 4xz\boldsymbol{i} - y^2\boldsymbol{j} + yz\boldsymbol{k}$

2. Evaluate $\iint_S \boldsymbol{A}.\hat{\boldsymbol{n}}\, ds$, over the entire surface S of the region bounded by the cylinder $x^2 + z^2 = 9$, $x = 0$, $y = 0$, $z = 0$ and $y = 8$, if $\boldsymbol{A} = 6z\boldsymbol{i} + (2x + y)\boldsymbol{j} - x\boldsymbol{k}$

Solution 1 :



(fig. solution 1)

We have $\oint_S \boldsymbol{A}.d\boldsymbol{S} = \oint_S \boldsymbol{A}.\left(\boldsymbol{i}\, ds_x + \boldsymbol{j}\, ds_y + \boldsymbol{k}\, ds_z\right)$

or, $\oint\limits_S \boldsymbol{A}.d\boldsymbol{s} = \oint\limits_S \boldsymbol{A}.\left(\boldsymbol{i}\, dy\, dz + \boldsymbol{j}\, dx\, dz + \boldsymbol{k}\, dx\, dy\right)$

$$= \iint\limits_{ABCD} \boldsymbol{A}.\boldsymbol{i}\, dy\, dz + \iint\limits_{EFGH} \boldsymbol{A}.-\boldsymbol{i}\, dy\, dz + \iint\limits_{BFGC} \boldsymbol{A}.\boldsymbol{j}\, dx\, dz$$

$$+ \iint\limits_{AEHD} \boldsymbol{A}.-\boldsymbol{j}\, dx\, dz + \iint\limits_{DCGH} \boldsymbol{A}.\boldsymbol{k}\, dx\, dy + \iint\limits_{ABFE} \boldsymbol{A}.-\boldsymbol{k}\, dx\, dy$$

i) When $\boldsymbol{A} = \boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$

For the surface ABCD, $x = 1$

$$\therefore \iint\limits_{ABCD} (\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z).i\, dy\, dz = \int_0^1\int_0^1 dx\, dz = 1$$

For the surface EFGH, $x = 0$

$$\therefore \iint_{EFGH} (\boldsymbol{i}x+\boldsymbol{j}y+\boldsymbol{k}z).-\boldsymbol{i}\ dy\ dz = \int_0^1\int_0^1 (0)dy\ dz = 0$$

Similarly

$$\iint_{BFGC} \boldsymbol{A}.\boldsymbol{j}\ dx\ dz = 1 \text{ and } \iint_{AEHD} \boldsymbol{A}.-\boldsymbol{j}\ dx\ dz = 0$$

and $$\iint_{DCGH} \boldsymbol{A}.\boldsymbol{k}\ dx\ dy = 1 \text{ and } \iint_{ABEF} \boldsymbol{A}.-\boldsymbol{k}\ dx\ dy = 0$$

$$\therefore \oiint \boldsymbol{A}\cdot d\boldsymbol{s} = (1+0)+(1+0)+(1+0) = 3$$

ii) When $\boldsymbol{A} = 4xz\boldsymbol{i} - y^2\boldsymbol{j} + yz\boldsymbol{k}$

For the surface ABCD, $x = 1$

$$\therefore \iint_{ABCD} (4xz\boldsymbol{i}-y^2\boldsymbol{j}+yz\boldsymbol{k}).\boldsymbol{i}\ dydz = \int_0^1\int_0^1 4z\ dz\ dy = \int_0^1 4z\ dz\int_0^1 dy$$

$$= 2\left[z^2\right]_0^1 \left[y\right]_0^1 = 2.1.1 = 2$$

For the surface EFGH, $x = 0$

$$\therefore \iint_{EFGH} (4xz\boldsymbol{i}-y^2\boldsymbol{j}+yz\boldsymbol{k}).-\boldsymbol{i}\ dy\ dz = \int_0^1\int_0^1 (0)dy\ dz = 0$$

For the surface BFGC, $y = 1$

$$\therefore \iint_{BFGH} (4xz\boldsymbol{i}-y^2\boldsymbol{j}+yz\boldsymbol{k}).\boldsymbol{j}\ dxdz = \int_0^1\int_0^1 -y^2\ dx\ dz = -\int_0^1 dx\int_0^1 dz = -1$$

For the surface AEHD, $y = 0$

$$\therefore \iint_{AEHD} (4xz\boldsymbol{i}-y^2\boldsymbol{j}+yz\boldsymbol{k}).-\boldsymbol{j}\ dx\ dz = \int_0^1\int_0^1 (0)dx\ dz = 0$$

For the surface DCGH, $z = 1$

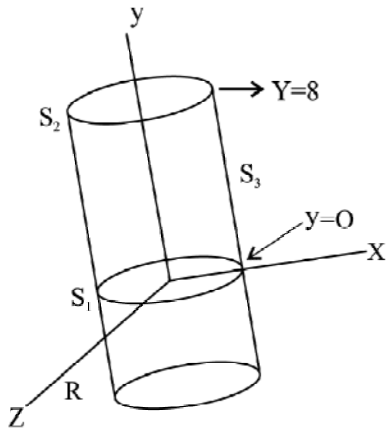
$$\therefore \iint_{DCGH} \left(4xz\boldsymbol{i} - y^2\boldsymbol{j} + yz\boldsymbol{k}\right)\cdot\boldsymbol{k}\ dx\ dy = \int_0^1\int_0^1 ydx\ dy = \int_0^1 y\ dy \int_0^1 dx = \frac{1}{2}\cdot 1 = \frac{1}{2}$$

For the surface ABFE, z = 0

$$\therefore \iint_{AEHD} \left(4xz\boldsymbol{i} - y^2\boldsymbol{j} + yz\boldsymbol{k}\right)\cdot -\boldsymbol{k}\ dxdy = \int_0^1\int_0^1 (0)dx\ dy = 0$$

$$\therefore \oint_S \boldsymbol{A}\cdot\boldsymbol{ds} = (2+0)+(-1+0)+\left(\frac{1}{2}+0\right) = \frac{3}{2}$$

Solution 2 :



(fig. solution 2)

Here the surface S is composed of the bottom circular surface S_1 top circular surface S_2, curved surface S_3.

Therefore $\oiint_S \boldsymbol{A}.\hat{\boldsymbol{n}}ds = \iint_{S_1} \boldsymbol{A}\cdot\hat{\boldsymbol{n}}\ ds + \iint_{S_2} \boldsymbol{A}\cdot\hat{\boldsymbol{n}}\,ds + \iint_{S_3} \boldsymbol{A}\cdot\hat{\boldsymbol{n}}\ ds$

But $\iint_{S_1} \boldsymbol{A}\cdot\hat{\boldsymbol{n}}\ ds = \iint_{S_1}\left[6z\boldsymbol{i} + (2x+y)\boldsymbol{j} - x\boldsymbol{k}\right].(-\boldsymbol{j})ds = \iint_{S_1} -(2x+y)ds$ (i)

Now $\iint_{S_2} \boldsymbol{A}\cdot\hat{\boldsymbol{n}}\ ds = \iint_{S_2}\left\{6z\boldsymbol{i} + \left(2x+y\right)\boldsymbol{j} - xk\right\}\cdot\boldsymbol{j}\ ds = \iint_{S_2}(2x+y)ds$ (ii)

Evaluation of integral (i) :

$$\iint_{S_1} -(2x+y)ds = \iint_{S_1} -2x\, dx\, dz \qquad \because y = 0 \ in \ s_1,\ ds = dx\, dz$$

$$\iint_{S_1} -(2x+y)ds = -2\int_{x=0}^{3}\int_{z=0}^{\sqrt{9-x^2}} x\, dx\, dz = -2\int_{x=0}^{3}\left(\sqrt{9-x^2}\right)xdx$$

$$= -2\left[x\int\sqrt{9-x^2}\,dx - \int 1.\int\sqrt{9-x^2}\,dx\right]_0^3$$

$$= -2\left[x\left\{\frac{x\sqrt{9-x^2}}{2}+\frac{9}{2}\sin^{-1}\frac{x}{3}\right\}_0^3\right]+\left[2\left\{\frac{x\sqrt{9-x^2}}{2}+\frac{9}{2}\sin^{-1}\frac{x}{3}\right\}_0^3\right]$$

$$= -2.3.\frac{9}{2}.\frac{\pi}{2}+2.\frac{9}{2}.\frac{\pi}{2} = -9\pi$$

Evaluation of integral (ii)

$$\iint_{S_2}(2x+y)dx\, dz \qquad \because ds = dx\, dz \ and \ \hat{\boldsymbol{n}} = \boldsymbol{j}$$

$$= 2\int_{x=0}^{3}\int_{z=0}^{\sqrt{9-x^2}} x\, dx\, dz + 8\int_{x=0}^{3}\int_{z=0}^{\sqrt{9-x^2}} dx\, dz \qquad \because y = 8 \ in \ S_2$$

$$= 2\int_{x=0}^{3}\sqrt{9-x^2}\,x\, dx + 8\int_0^3\sqrt{9-x^2}\ dx = 9\pi + 8\left[\frac{x\sqrt{9-x^2}}{2}+\frac{9}{2}\sin^{-1}\frac{x}{3}\right]_0^3$$

$$= 9\pi + 8\left[\frac{9}{2}\times\frac{\pi}{2}\right] = 9\pi + 18\pi = 27\pi$$

again $\displaystyle\iint_{S_3}\boldsymbol{A}\cdot\hat{\boldsymbol{n}}\ ds = \iint_R \boldsymbol{A}\cdot\hat{\boldsymbol{n}}\ \frac{dy\, dz}{\hat{\boldsymbol{n}}.\boldsymbol{i}}$. Now $\phi(x, z) = x^2 + z^2 = 9 =$ constant

$$\therefore \hat{\boldsymbol{n}} = \frac{\nabla\phi}{|\nabla\phi|} = \frac{x\boldsymbol{i} + z\boldsymbol{k}}{3} \qquad \therefore \hat{\boldsymbol{n}}.\boldsymbol{i} = \frac{x}{3}$$

$$\therefore \boldsymbol{A}.\hat{\boldsymbol{n}} = \frac{5}{3}xz \qquad \therefore \iint_{S_3} \boldsymbol{A}.\hat{\boldsymbol{n}}\, ds = \iint \frac{5}{3}xz \frac{dydz}{\frac{x}{3}} = \iint 5z\, dy\, dz$$

$$\therefore \iint_{S_3} \boldsymbol{A}.\hat{\boldsymbol{n}}\, ds = 5 \int_{z=-3}^{+3} \int_{y=0}^{8} z\, dy\, dz = 40 \int_{-3}^{+3} z\, dz = 0$$

$$\therefore \oiint \boldsymbol{A}.\hat{\boldsymbol{n}}\, ds = -9\pi + 27\pi + 0 = 18\pi$$

4.28.3 Volume integral of a vector field

We consider a closed surface in space enclosing a volume V. If $\boldsymbol{V}(\boldsymbol{r})$ be a continuous vector point function, then volume integral is defined as,

$$\iiint_V \boldsymbol{V}(r)\, dV$$

In Cartesian co-ordinate it is written as,

$$\iiint_V \left(\boldsymbol{i}V_x + \boldsymbol{j}V_y + \boldsymbol{k}V_z\right) dx\, dy\, dz$$

$$= \boldsymbol{i}\iiint_V (V_x)\, dx\, dy\, dz + \boldsymbol{j}\iiint_V (V_y)\, dx\, dy\, dz + \boldsymbol{k}\iiint_V (V_z)\, dx\, dy\, dz$$

4.29 ❐ Green's theorem in plane

Statement : It states that if M and N are continuous functions of x and y having continuous derivative in a region R of the xy-plane bounded by a closed curve C, then

$$\oint_C (M\ dx + N\ dy) = \iint_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx\, dy \qquad (4C.17)$$

The line integral in counter clockwise around the boundary of area $\boldsymbol{R}$

Proof : Let us assume that the region R, bounded by a closoed curved C be such that the vertical lines at $x = p$ and at $x = q$ as shown in Fig (4C.16) cuts the curve C in at most two points at P and Q. Similarly the horizontal lines at $y = m$ and $y = n$ cuts the curve C in at most two points at M and N.

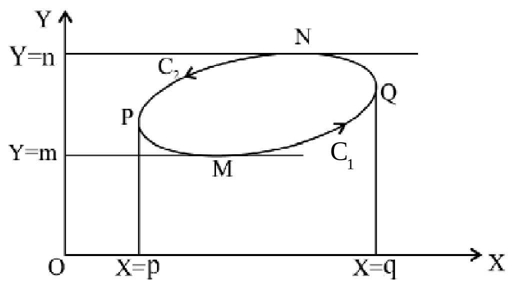


Fig (4C.16)

Let the equation to the curve PMQ be $y = \phi_1(x)$ and that to the curve QNP be $y = \phi_2(x)$ which is denoted by C_1 and C_2 respectively. Now

$$\iint_R \frac{\partial M}{\partial y} dx\, dy = \int_{x=p}^{x=q}\left[\int_{y=\phi_1(x)}^{y=\phi_2(x)} \frac{\partial M}{\partial y}\, dy\right] dx = \int_{x=p}^{x=q}\left[M(x, y)\right]_{\phi_1(x)}^{\phi_2(x)} dx$$

$$= \int_{x=p}^{x=q}\left[M\left(x, \phi_2(x)\right) - M\left(x, \phi_1(x)\right)\right] dx$$

$$= \int_{\substack{x=p \\ (along\ C_2)}}^{x=q} M\left(x, \phi_2(x)\right) dx - \int_{\substack{x=p \\ (along\ C_1)}}^{x=q} M\left(x, \phi_1(x)\right) dx$$

$$= -\int_q^p M\left[x, \phi_2(x)\right] dx - \int_p^q M\left[x, \phi_1(x)\right] dx = -\oint_C M(x, y) dx \qquad (4C.18)$$

Again, let the equation to the curve NPM be $x = \phi_3(y)$ and that to MQN be $x = \phi_4(y)$

$$\therefore \iint_R \frac{\partial N}{\partial x} dx\, dy = \int_{y=m}^{n}\left[\int_{\phi_3(y)}^{\phi_4(y)} \frac{\partial N}{\partial x}\, dx\right] dy = \int_{y=m}^{n}\left[N(x, y)\right]_{\phi_3(y)}^{\phi_4(y)} dy$$

$$= \int_{y=m}^{y=n}\left[N\{\phi_4(y), y\} - N\{\phi_3(y), y\}\right] dy$$

$$= \int_{y=m}^{y=n} N\left[\phi_4(y), y\right] dy + \int_{y=n}^{y=m} N\left[\phi_3(y), y\right] dy$$

$$= \oint_c N(x, y) dy \qquad (4C.18A)$$

Combining (4C.18) and (4C.18A) we get,

$$\iint_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dx\, dy = \oint_C \left[M(x, y)dx + N(x, y)dy \right] \qquad (4C.17)$$

In vector form equation (4C.17) can be re-written as :

$$\oint_C \boldsymbol{V}.d\boldsymbol{r} = \iint_R \nabla \times \boldsymbol{V}.\boldsymbol{k}\; dx\, dy$$

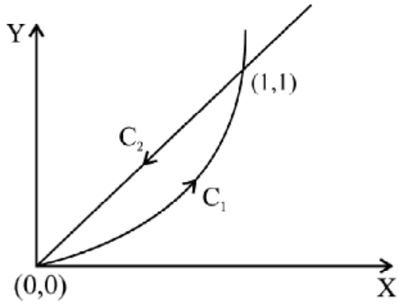
Where $\boldsymbol{V}$ = M$\boldsymbol{i}$ + $N\boldsymbol{j}$, $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}$, $d\boldsymbol{r} = \boldsymbol{i}\; dx + \boldsymbol{j}\; dy$

$\therefore\; \boldsymbol{V}.d\boldsymbol{r} = M\; dx + N\; dy$ and $\nabla \times \boldsymbol{V}.\boldsymbol{k} = \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right)$

Example of Art 4.29 :

Example 1: Verify Green's theorem for $\oint_C \left[\left(xy + y^2 \right) dx + x^2 dy \right]$, where C is bounded by y = x, and $y = x^2$.

Solution 1 : The curve C_1 and the line C_2 intersect at (0, 0) and (1,1). The positive direction in traversing C is as shown in figure.



C_1 is the curve y = x^2
C_2 is the curve y = x
Fig Example (1)

Along the curve C_1, the given integral become

$$I_1 = \int_0^1 \left[\left(xx^2 + x^4 \right) dx + x^2 2x\, dx \right]$$

$$= \int_0^1 \left(3x^3 + x^4 \right) dx = \frac{19}{20}$$

Along the curved C_2, the integral become

$$I_2 = \int_1^0 \left[\left(x^2 + x^2 \right) dx + x^2 dx \right] = \int_1^0 3\, x^2 dx = -1$$

$$\therefore \oint_C \left[\left(xy + y^2 \right) dx + x^2 dy \right] = I_1 + I_2 = \frac{19}{20} - 1 = -\frac{1}{20} \qquad \text{(i)}$$

Now applying Green's theorem, we set that

$$M = xy + y^2,\ N = x^2.\ \text{Now}\ \frac{\partial N}{\partial x} = 2x,\ \frac{\partial M}{\partial y} = x + 2y$$

$$\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = 2x - x - 2y = x - 2y$$

Now $\oint_C \left[\left(xy + y^2\right)dx + x^2 dy\right]$

$$= \iint \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx\, dy = \iint (x - 2y)\, dx\, dy = \int_0^1 dx \int_{x^2}^{x} (x - 2y)\, dy$$

[We have $\oint_C = f_{C_1} + f_{C_2}$. Along C_1, $y = x^2$, x varies from 0 to 1. Along C_2, $y = x$ and x varies from 1 to 0]

$$= \int_0^1 dx\left(xy - y^2\right)_{x^2}^{x} = \int_0^1 dx\left[x^2 - x^2 - x^3 + x^4\right] = \int_0^1 dx\left(x^4 - x^3\right) = -\frac{1}{20} \qquad \text{(ii)}$$

Since (i) and (ii) equal, Green's theorem is verified.

Example 2 : Apply Green's theorem in the plane to evaluate the integral $\oint_C [(2x - y^3)dx - xy\, dy]$ over the boundary of the region bounded by the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 9$.

Solution 2 : Here $M = 2x - y^3$ $\quad N = -xy$ $\quad \therefore \dfrac{\partial M}{\partial y} = -3y^2$ *and* $\dfrac{\partial N}{\partial x} = -y$

$$\therefore \oint_C \left[\left(2x - y^3\right)dx - xy\, dy\right] = \iint \left(-y + 3y^2\right)dx\, dy$$

Now in plane polar co-ordinate $x = r\cos\theta,\ y = r\sin\theta,\ dxdy = r\, dr\, d\theta$. Here r varies from 1 to 3 and θ varies from 0 to 2π.

$$\therefore \oint_C \left[\left(2x - y^3\right)dx - xy\, dy\right] = \int_{r=1}^{3} \int_{\theta=0}^{2\pi} \left(-r\sin\theta + 3r^2\sin^2\theta\right) r\, dr\, d\theta$$

$$= \int_{\theta=0}^{2\pi} \int_{r=1}^{3} \left[-r^2 \sin\theta \, d\theta \, dr + 3r^2 \sin^2\theta \, d\theta \, dr \right]$$

$$= \int_{0}^{2\pi} \sin\theta \, d\theta \int_{1}^{3} -r^2 \, dr + \int_{0}^{2\pi} \sin^2\theta \, d\theta \int_{1}^{3} 3r^3 \, dr$$

$$= 0 + \pi \left[\frac{3r^4}{4} \right]_1^3 = \frac{3\pi}{4}\left[3^4 - 1^4\right] = \frac{3\pi}{4}.80 = 60\pi$$

Since $\int_{0}^{2\pi} \sin\theta \, d\theta = 0 \;\; and \;\; \int_{0}^{2\pi} \sin^2\theta \, d\theta = \pi$

Exercise of Art (4..29) :

1) If C be the boundary of the rectangle (in xy plane) defined by $y = 0$, $x = a$; $y = b$, $x = 0$; evaluate the integral $\oint_C \boldsymbol{F}.\boldsymbol{dr}$, where $\boldsymbol{F} = (x^2 + y^2)\boldsymbol{i} - 2xy\boldsymbol{i}$ by applying Green's theorem.

2) Verify Green's theorem in the plane to evaluate the integral $\oint_C \left[\left(xy - x^2\right)dx + x^2 y \, dy\right]$ over the triangle bounded by the line $y = 0$, $x = 1$ and $y = x$.

3) Apply Green's theorem to prove that the area enclosed by a plane curve is $\frac{1}{2}\oint_C (x \, dy - y \, dx)$. Hence find the area of an ellipse whose semi-major and minor axes are of lengths a and b.

Solution 1 : Now applying Green's theorem we have $M = x^2 + y^2$; $N = -2xy$

$$\therefore \frac{\partial N}{\partial x} = -2y \, ; \frac{\partial M}{\partial y} = 2y$$

$$\therefore \oint_C \left[\left(x^2 + y^2\right)dx - 2xy \, dy\right] = \iint (-2y - 2y) dx \, dy$$

$$= \iint -4y\, dx\, dy = \int_0^a dx \int_{y=0}^{y=b} -4y\, dy = [a]\left[-2b^2\right] = -2ab^2$$

Solution 2 : Along *OA : y = 0 dy = 0*

along AB : $x = 1$, $dx = 0$

along *BO* : $y = x$, $dy = dx$

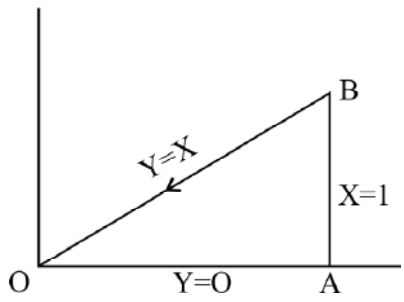


Fig. Solution (2)

Now $\oint_C \left[\left(xy - x^2\right)dx + x^2 y\, dy\right]$

$$= \int_{OA} -x^2 dx + \int_{AB} y\, dy + \int_{BO} x^3\, dx$$

$$= \int_0^1 -x^2 dx + \int_0^1 y\, dy + \int_1^0 x^3 dx = -\frac{1}{3} + \frac{1}{2} - \frac{1}{4} = -\frac{1}{12} \qquad \text{(i)}$$

Now $\quad M = xy - x^2 \qquad N = x^2 y$

$$\therefore \frac{\partial M}{\partial y} x \qquad \frac{\partial N}{\partial x} = 2xy$$

Therefore using Green's theorem

$$\oint_{OABO} \left[\left(xy - x^2\right)dx + x^2 y\, dy\right] = \iint (2xy - x)dx\, dy = \int_0^1 dx \int_{y=0}^{y=x} (2xy - x)dy$$

$$= \int_0^1 dx\left(xy^2 - xy\right)_0^x = \int_0^1 dx\left(x^3 - x^2\right) = \left[\frac{x^4}{4} - \frac{x^3}{3}\right]_0^1 = \frac{1}{4} - \frac{1}{3} = -\frac{1}{12} \qquad \text{(ii)}$$

From equation (i) and (ii) we see that Green's theorem is verified.

Solution 3 : We have from Green's theorem,

$$\oint_C [Mdx + Ndy] = \iint_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx\, dy \qquad \text{(i)}$$

Now we put $N = x$ and $M = -y$ $\qquad \therefore \dfrac{\partial N}{\partial x} = 1, \; \dfrac{\partial M}{\partial y} = -1$

$$\therefore \oint_C [-y\,dx + x\,dy] = \iint_R (1+1)\,dx\,dy$$

or, $\displaystyle \oint_C [x\,dy - y\,dx] = 2\iint_R dx\,dy = 2A$

where A is the area of the plane curve.

$$\therefore A = \frac{1}{2}\oint_C [x\,dy - y\,dx] \qquad \text{(ii)}$$

Now equation of the ellipse : $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$

Therefore $x = a\cos\theta$ and $y = b\sin\theta$, is the parametric equation of the ellipse.

$$\therefore A = \frac{1}{2}\oint_C \left[a\cos\theta . b\cos\theta\,d\theta - b\sin\theta . a(-\sin\theta)d\theta\right]$$

$$= \frac{1}{2}\oint_C ab\left[\cos^2\theta\,d\theta + \sin^2\theta\,d\theta\right]$$

$$= \frac{1}{2}\oint_C ab\left(\cos^2\theta + \sin^2\theta\right)d\theta = \frac{1}{2}ab\int_0^{2\pi} d\theta = \pi ab$$

Now for a circle $a = b$ $\qquad \therefore A = \pi a^2$

4.30 ❐ Gauss's Divergence Theorem

Statement : This theorem states that, the volume integral of the divergence of the vector $\boldsymbol{V}$ taken over any volume τ in its field is equal to the surface integral of $\boldsymbol{V}$ over the closed surface enclosing the volume τ. In vector notation, the theorem is written as

$$\iiint \nabla . \boldsymbol{V}\,d\tau = \iint_{(surface\ enclosing\ \tau)} \boldsymbol{V} . \hat{\boldsymbol{n}}\,d\sigma \qquad \text{(4C.19)}$$

($\hat{\boldsymbol{n}}$ *points out of the closed surface* σ)

Proof : We consider a small volume element $d\tau_i$ of the total volume τ shown in Fig (4C.17) in a vector field ***V***.

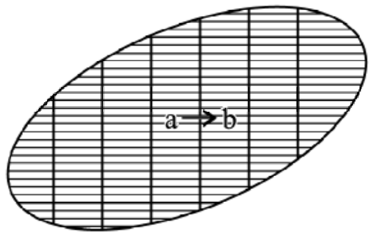


Fig (4C.17)

From the definition of divergence of vector fields, we get net outflow from each $d\tau_i$ as $\nabla.\boldsymbol{V}\, d\tau_i$ and then adding the total outflow from the entire volume τ as

$$\sum_i \nabla.\boldsymbol{V}\, d\tau_i \qquad (4C.20)$$

which is explained below. From Fig (4C.17) it is seen that an outflow from a to b is an inflow from *b* to *a*, so that such outflows across interior faces cancel. The total sum in (4C.20) then equals just the total outflow from the entire volume τ.

When $d\tau_i \to 0$, the sum (4C.20) is converted into a triple integral over the volume τ :

$$\lim_{\substack{i\to\infty \\ d\tau_i\to 0}} \sum_i \nabla.\boldsymbol{V}\, d\tau_i = \iiint_\tau \nabla.\boldsymbol{V}\, d\tau \qquad (4C.21)$$

Now consider the Fig (4C.18) below, outflow or flux of vector field ***V*** through $d\sigma$ is $\boldsymbol{V}.\hat{\boldsymbol{n}}\, d\sigma$ and the total outflow from the volume enclosed by the surface is

$$\iint \boldsymbol{V}.\hat{\boldsymbol{n}}\, d\sigma \qquad (4C.22)$$

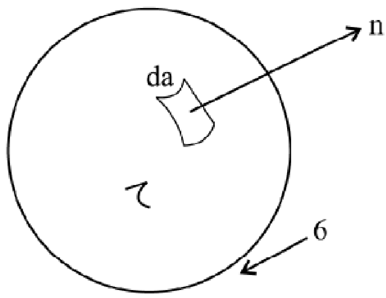


Fig (4C.18)

Where $\hat{\boldsymbol{n}}$ is the unit normal to the surface element $d\sigma$ and pointing outwards.

(σ is the surface enclosing volume τ)

Thus both the equations (4C.21) and (4C.22) give the total outflow from the total volume τ and hence they are equal to each other and we get equation (4C.19)

Example 4.28 and 4.20

Example 1: Evaluate $\oiint_S \boldsymbol{r}.\hat{\boldsymbol{n}}\, d\sigma$

Solution 1: Using divergence theorem,

$$\oiint_S \boldsymbol{r}.\hat{\boldsymbol{n}}\, ds = \int_V \nabla.\boldsymbol{r}\, dV = 3\int_V dv = 3V$$

Example 2 : Using divergence theorem evaluate $\oiint_S \nabla \boldsymbol{r}^2 . \boldsymbol{ds}$

Solution 2 : We have $\oiint_S \nabla \boldsymbol{r}^2 . \boldsymbol{ds} = \int_V \nabla \,. \nabla \boldsymbol{r}^2 dV$

Now, $\nabla r^2 = \boldsymbol{i}\frac{\partial}{\partial x}\left(r^2\right) + \boldsymbol{j}\frac{\partial}{\partial y}\left(r^2\right) + \boldsymbol{k}\frac{\partial}{\partial z}\left(r^2\right)$

where $r^2 = x^2 + y^2 + z^2$.

$\therefore \nabla r^2 = 2x\boldsymbol{i} + 2y\boldsymbol{j} + 2z\boldsymbol{k}$

$$\therefore \nabla . \nabla r^2 = \frac{\partial}{\partial x}(2x) + \frac{\partial}{\partial y}(2y) + \frac{\partial}{\partial z}(2z) = 2 + 2 + 2 = 6$$

$$\therefore \oiint_S \nabla r^2 . \boldsymbol{ds} = \int_V \nabla . \nabla r^2 \, dV = 6\int_V dV = 6V$$

Example 3 : Prove that $\oiint_S (\hat{\boldsymbol{n}} \times \boldsymbol{B}) ds = \iiint_V (\nabla \times \boldsymbol{B}) dV$

Solution 3: We have from Gauss's divergence theorem,

$$\oiint_S \boldsymbol{A} . \hat{\boldsymbol{n}} ds = \iiint_V \nabla . \boldsymbol{A} \, dV$$

Let $\boldsymbol{A} = \boldsymbol{C} \times \boldsymbol{B}$, where $\boldsymbol{C}$ is a constant vector.

$$\therefore \int_V \nabla . \boldsymbol{C} \times \boldsymbol{B} \, dV = \oiint_S \boldsymbol{C} \times \boldsymbol{B} . \hat{\boldsymbol{n}} ds$$

Now $\nabla . \boldsymbol{C} \times \boldsymbol{B} = \boldsymbol{B} . \nabla \times \boldsymbol{C} - \boldsymbol{C} . (\nabla \times \boldsymbol{B}) = -\, C . (\nabla \times \boldsymbol{B})$, since $\boldsymbol{C}$ is a constant vector, $\nabla \times \boldsymbol{C} = 0$

$$\therefore -\boldsymbol{C} \int_V \nabla \times \boldsymbol{B} \, dV = \boldsymbol{C} . \oiint_S \boldsymbol{B} \times \hat{\boldsymbol{n}} \, ds = -\boldsymbol{C} . \oiint_S \hat{\boldsymbol{n}} \times \boldsymbol{B} \, ds$$

$$\therefore \int_V \nabla \times \boldsymbol{B} \, dV = \oiint_S \hat{\boldsymbol{n}} \times \boldsymbol{B} \, ds \qquad (1)$$

[now if the vector ***B*** is always normal to a given closed surface *S*, then ***B*** and ***ds*** is parallel and ***B*** × ***ds*** = **0** and hence from equation (1) $\int_V \nabla\times\boldsymbol{B}\,dV = -\oiint_S \boldsymbol{B}\times\hat{\boldsymbol{n}}\,ds$ $= -\oiint_S \boldsymbol{B}\times\boldsymbol{ds} = 0$]

Example 4 : Prove that $\oiint_S (\nabla\phi\times\nabla\psi).\boldsymbol{ds} = 0$

Solution 4 : We put $\boldsymbol{A} = \nabla\phi\times\nabla\psi$ then,

$$\oiint_S \boldsymbol{A}.\boldsymbol{ds} = \int_V (\nabla.\boldsymbol{A})\,dV = \int_V [\nabla.\nabla\phi\times\nabla\psi]\,dV$$

Now, $\nabla.\nabla\phi\times\nabla\psi = \nabla\psi.\nabla\times\nabla\phi - \nabla\phi.\nabla\times\nabla\psi = 0$

Since $\nabla\times\nabla\phi = 0$ and $\nabla\times\nabla\psi = 0$

$$\therefore \oiint_S (\nabla\phi\times\nabla\psi).\boldsymbol{ds} = 0$$

Example 5 :

$$\int_V \left(\phi\nabla^2\psi - \psi\nabla^2\phi\right) dV = \int_V (\phi\nabla\psi - \psi\nabla\phi).\boldsymbol{ds}$$

Where *V* is the volume bounded by the surface *S* and ϕ,ψ are scalar fields.

Solution 5 : Let $\boldsymbol{A} = \phi\nabla\psi$ Now by divergence theorem

$$\int_V \nabla.\boldsymbol{A}\,dV = \oint_S \boldsymbol{A}.\hat{\boldsymbol{n}}\,ds \quad \text{or} \quad \int_V \nabla.(\phi\nabla\psi)\,dV = \oint_S \phi\nabla\psi.\hat{\boldsymbol{n}}\,ds.$$

But $\nabla.\phi\nabla\psi = \phi\nabla.\nabla\psi + \nabla\phi.\nabla\psi = \phi\nabla^2\psi + \nabla\phi.\nabla\psi$

therefore, $\int_V \left[\phi\nabla^2\psi + \nabla\phi.\nabla\psi\right] dV = \oint_S (\phi\nabla\psi).\boldsymbol{ds}$ (i)

Now interchanging ϕ and ψ, we get

$$\int_V \left[\psi\nabla^2\phi + \nabla\psi.\nabla\phi\right] dV = \oint_S \psi\nabla\phi.\boldsymbol{ds} \qquad \text{(ii)}$$

Subtracting equation (ii) from equation (i), we get

$$\int_V \left[\phi\nabla^2\psi - \psi\nabla^2\phi\right]dV = \oint_S (\phi\nabla\psi - \psi\nabla\phi).\boldsymbol{ds}$$

Example 6 : Applying Gauss's divergence theorem evaluate $\iint_S \frac{\boldsymbol{r}.\hat{n}}{r^3}ds$, where *S* represents any closed surface enclosing volume *V*. When the origin is outside *S*.

Solution 6 :

Let $\boldsymbol{A} = \frac{\boldsymbol{r}}{r^3}$. Now $\boldsymbol{A}$ is continuously differentiable throughout the volume enclosed by S. By Gauss's divergence theorem,

$$\oint_S \boldsymbol{A}.\hat{\boldsymbol{n}}\, ds = \int_V \nabla.\boldsymbol{A}\, dV$$

Now $\nabla.\boldsymbol{A} = \nabla.\frac{\boldsymbol{r}}{r^3} = \frac{1}{r^3}(\nabla.\boldsymbol{r}) + \nabla\frac{1}{r^3}\boldsymbol{r}$

Now $\nabla\frac{1}{r^3} = -3r^{-4}\hat{\boldsymbol{r}} = -\frac{3\boldsymbol{r}}{r^5}$

$$\therefore \nabla.\boldsymbol{A} = \frac{3}{r^3} - \frac{3\boldsymbol{r}.\boldsymbol{r}}{r^5} = \frac{3}{r^3} - \frac{3}{r^3} = 0$$

$$\therefore \oiint_S \frac{(\boldsymbol{r}.\hat{\boldsymbol{n}})}{r^3}ds = \int_V \nabla.\frac{r}{r^3}dV = 0$$

Exercise of Art 4.28 and 4.30 :

1) Evaluate $\int \boldsymbol{F}.\hat{\boldsymbol{n}}\, dS$, where $\boldsymbol{F} = 4xz\boldsymbol{i} - y^2\boldsymbol{j} + yz\boldsymbol{k}$ and *S* is the surface of the cube bounded by $x = 0$, $x = 1$, $y = 0$, $y = 1$, $z = 0$ & $z = 1$ by using divergence theorem.

2) Evaluate $\int x^2 y\, dV$, where *V* is the closed region bounded by the planes $4x + 2y + z = 8$, $x = 0$, $y = 0$ & $z = 0$.

Solution 1 :

$$\oint_S \boldsymbol{F}.\hat{\boldsymbol{n}}\, ds = \int_V \nabla.\boldsymbol{F}\; dV$$

now $\nabla.\boldsymbol{F} = \dfrac{\partial}{\partial x}(4xz) + \dfrac{\partial}{\partial y}\left(-y^2\right) + \dfrac{\partial}{\partial z}(yz)$

or, $\nabla.\boldsymbol{F} = 4z - 2y + y = 4z - y$

$$\therefore \int_V \nabla.\boldsymbol{F}\; dV = \iiint (4z-y)dxdy\, dz = \int_0^1\int_0^1 \left(2z^2 - yz\right)_0^1 dx\, dy = \int_0^1\int_0^1 (2-y)\, dx\, dy$$

$$= \int_0^1 \left(2y - \frac{y^2}{2}\right)_0^1 dx = \int_0^1 \frac{3}{2} dx = \frac{3}{2}$$

Solution 2 : In the problem, z varies from 0 to $8 - 4x - 2y$, y varies from 0 to $4 - 2x$ and x varies from 0 to 2.

$\therefore$ the given integral

$$\iiint_V x^2 y\; dV = \int_{x=0}^{2}\int_{y=0}^{4-2x}\int_{z=0}^{8-4x-2y} x^2 y\; dx\, dy\, dz = \int_{x=0}^{2}\int_{y=0}^{4-2x} x^2 y(8-4x-2y)\; dx\, dy$$

$$= \int_{x=0}^{2}\left[\int_{y=0}^{4-2x} x^2 y(8-4x-2y)dy\right]dx = \int_{x=0}^{2}\left[8x^2\frac{y^2}{2} - 4x^3\frac{y^2}{2} - 2x^2\frac{y^3}{3}\right]_0^{4-2x} dx$$

$$= \int_{x=0}^{2}\left[4x^2y^2 - 2x^3y^2 - \frac{2}{3}x^2y^3\right]_0^{4-2x} dx$$

$$= \int_{x=0}^{2}\left[4x^2(4-2x)^2 - 2x^3(4-2x)^2 - \frac{2}{3}x^2(4-2x)^3\right]dx$$

$$= \int_{x=0}^{2}\left[\frac{64}{3}x^2 - 32x^3 + 16x^4 - \frac{8}{3}x^5\right]dx = \left[\frac{64}{3}\frac{x^3}{3} - 32\frac{x^4}{4} + 16\frac{x^5}{5} - \frac{8}{3}\frac{x^6}{6}\right]_0^2 = \frac{128}{45}$$

4.31 ❑ Stoke's Theorem

Statements : It states that if σ is an open two-sided surface bounded by a simple closed curve and if ***V*** be continuously differentiable point function, then

$$\oint_{\text{curved bounding } \sigma} \boldsymbol{V}.\boldsymbol{dr} = \iint_{\text{open surface } \sigma} \nabla \times \boldsymbol{V}.\hat{\boldsymbol{n}}\, \boldsymbol{d\sigma} \tag{4C.23}$$

where the boundary is traversed in the counter clockwise direction. $\hat{\boldsymbol{n}}$ is the outward drawn unit normal to the surface element $d\sigma$.

Proof : We consider an open surface which is two sided and whose bounding curve is simple (i.e. it must not cross itself) and closed (Fig 4C.19). We consider the surface to be divided into a large number of elementary surfaces, $d\sigma$ with a unit vector $\hat{\boldsymbol{n}}$ normal to each area element and lying on the same side of the surface (Fig. 4C.20).

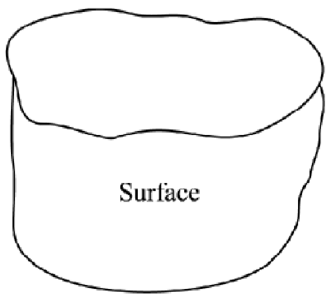


Fig. (4C.19)

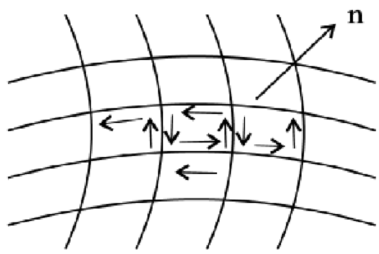


Fig. 4C. 20)

Now from the difinition of curl

$$(\nabla \times \boldsymbol{V}).\hat{\boldsymbol{n}} = \lim_{d\sigma \to 0} \frac{1}{d\sigma} \oint_{\text{around } d\sigma} \boldsymbol{V}.\boldsymbol{dr}$$

or, $$\oint_{\text{around } d\sigma} \boldsymbol{V}.\boldsymbol{dr} = \iint_{d\sigma} \nabla \times \boldsymbol{V}.\hat{\boldsymbol{n}}\, d\sigma \tag{4C.24}$$

For each area element $d\sigma$.

Adding now for all the area elements we get

$$\sum_{\text{all } d\sigma} \oint_{\text{around } d\sigma} \boldsymbol{V}.dr = \iint_{\text{surface } \sigma} \nabla \times \boldsymbol{V}.\hat{\boldsymbol{n}}\, d\sigma \tag{4C.25}$$

From Fig 4C.20 we see that all the interior line integral cancel because along a common curve between two $d\sigma$'s, the two integral, are in opposite direction.

Therefore left hand side of equation (4C.25) becomes simple the line integral around the outside curve bounding the surface. Therefore we can write,

$$\sum_{\text{all } d\sigma} \oint_{\text{around } d\sigma} \boldsymbol{V}.d\boldsymbol{r} = \oint_{\text{curved bounding } \sigma} \boldsymbol{V}.d\boldsymbol{r} = \iint_{\text{open surface } \sigma} \nabla \times \boldsymbol{V}.\hat{\boldsymbol{n}}\, d\sigma \qquad (4C.26)$$

which is Stoke's theorem.

Example of Art 4.31 :

Example 1 : Using Stoke's theorem prove that

$$\oint_C \phi\, d\boldsymbol{r} = \int_S d\boldsymbol{S} \times \nabla\phi = \int_S (\hat{\boldsymbol{n}} \times \nabla\phi)\, dS$$

Where ϕ is a scalar function of $\boldsymbol{r}$, S is a open surface bounded by a closed curve C.

Solution 1 : Let $\boldsymbol{A}$ be a constant vector. Then by Stoke's theorem

$$\oint_C \phi \boldsymbol{A}.\, d\boldsymbol{r} = \iint_S \nabla \times \phi \boldsymbol{A}.\, d\boldsymbol{S}$$

or $$\boldsymbol{A}.\oint_C \phi\, d\boldsymbol{r} = \iint_S \{\phi(\nabla \times \boldsymbol{A}) + \nabla\phi \times \boldsymbol{A}\}.d\boldsymbol{S}$$

since $\boldsymbol{A}$ *is constant vector,* $\nabla \times \boldsymbol{A} = 0$

$$\boldsymbol{A}.\oint_C \phi\, d\boldsymbol{r} = \iint_S \nabla\phi \times \boldsymbol{A}.d\boldsymbol{S} = \iint_S \boldsymbol{A}.d\boldsymbol{S} \times \nabla\phi$$

or, $\boldsymbol{A}.\oint_C \phi\, d\boldsymbol{r} = \boldsymbol{A}.\iint_S d\boldsymbol{S} \times \nabla\phi$; or, $\oint_C \phi\, d\boldsymbol{r} = \iint_S d\boldsymbol{S} \times \nabla\phi$

Example 2 : Evaluate by Stoke's theorem the integral,

$$\oint_C \left(e^x dx + 2y\, dy - dz\right)$$

where C is the curve $x^2 + y^2 = 4$, $z = 2$.

Solution 2 :

$$\oint_C \left(e^x dx + 2y\, dy - dz\right) = \oint_C \boldsymbol{A}.\boldsymbol{dr} = \oint_C A_x dx + A_y dy + A_z\, dz$$

$$\therefore A_x = e^x A_y = 2y \quad A_z = -1$$

$$\text{Now} \quad \nabla \times \boldsymbol{A} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ e^x & 2y & -1 \end{vmatrix} = 0$$

By Stoke's theorem : $\oint_C \boldsymbol{A}.\boldsymbol{dr} = \iint_S \nabla \times \boldsymbol{A}.\hat{\boldsymbol{n}}\, dS$

where S is the surface bounded by the circle.

$x^2 + y^2 = 4$ and $z = 2$. Also $\hat{\boldsymbol{n}} = \boldsymbol{k}$

$$\oint_C \boldsymbol{A}.\boldsymbol{dr} = \oint_C e^x dx + 2y\, dy - dz = 0 \qquad \because \nabla \times \boldsymbol{A} = 0$$

Example 3 : Verify Stoke's theorem for $\boldsymbol{A} = (2x - y)\boldsymbol{i} - yz^2\boldsymbol{j} - y^2z\boldsymbol{k}$, the region of integration being the upper half surface of the sphere $x^2 + y^2 + z^2 = 1$ and C its boundary.

Solution 3 : The boundary is the circle of unit radius in xy – plane in C is a circle of $x^2 + y^2 = 1 = r^2$ $\therefore r = 1$. Substituting $x = r\cos\theta = \cos\theta$ and $y = r\sin\theta = \sin\theta$ and $z = 0$, we get

$$\oint_C A.dr = \oint_C \left[(2x - y)dx - yz^2 dy - y^2 z\, dz\right]$$

or $$\oint_C \boldsymbol{A}.\boldsymbol{dr} = \int_0^{2\pi} \left(2\cos\theta - \sin\theta\right)\left(-\sin\theta\, d\theta\right) \quad \because z = 0$$

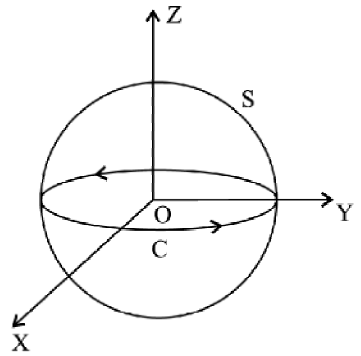


Fig. Example (3)

$$= \int_0^{2\pi} \left(\sin^2\theta - 2\sin\theta\cos\theta\right) d\theta = \int_0^{2\pi} \sin^2\theta\, d\theta - \int_0^{2\pi} \sin 2\theta\, d\theta$$

or, $$\oint_C \boldsymbol{A}.d\boldsymbol{r} = \frac{1}{2}\int_0^{2\pi}(1-\cos 2\theta)\,d\theta - \int_0^{2\pi}\sin 2\theta\,d\theta$$

$$= \frac{1}{2}\int_0^{2\pi} d\theta - \frac{1}{2}\int_0^{2\pi}\cos 2\theta\,d\theta - \int_0^{2\pi}\sin 2\theta\,d\theta$$

$$= \frac{1}{2}.2\pi = \pi \qquad \because \int_0^{2\pi}\cos 2\theta\,d\theta = \int_0^{2\pi}\sin 2\theta\,d\theta = 0$$

Now, $$\nabla\times\boldsymbol{A} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ 2x-y & -yz^2 & -y^2 z \end{vmatrix} = \boldsymbol{k}$$

$$\therefore \iint_S (\nabla\times\boldsymbol{A}).\hat{\boldsymbol{n}}\,ds = \iint_S k.\hat{\boldsymbol{n}}\,ds$$

Now the projection of vector area $\hat{\boldsymbol{n}}\ dS$ on the xy plane whose unit normal is $\boldsymbol{k}$, is given by $\hat{\boldsymbol{n}}\,dS.\boldsymbol{k} = (\hat{\boldsymbol{n}}\,.\boldsymbol{k})\,dS$.

But the projection of dS on xy plane is $dx\ dy$.

Hence $(\hat{\boldsymbol{n}}.\boldsymbol{k})dS = dx\,dy$

$$\therefore \iint_S (\nabla\times\boldsymbol{A})\,.\,\hat{\boldsymbol{n}}\,dS = \iint_S \boldsymbol{k}\,.\,\hat{\boldsymbol{n}}\,dS = \iint_S dx\,dy$$

Now $$\iint dx\,dy = \iint r\,dr\,d\theta = \int_0^1 r\,dr\int_0^{2\pi} d\theta = \frac{1}{2}.2\pi = \pi$$

$\therefore$ Stoke’s theorem is verified.

Exercise of Art 4.31 :

1) Verify Stoke's theorem for $\boldsymbol{A} = (\boldsymbol{y} - z + 2)\boldsymbol{i} + (yz + 4)\boldsymbol{j} - zx\boldsymbol{k}$ over the surface of the cube $x = 0$, $y = 0$, $z = 0$; $x = 2$, $y = 2$, $z = 2$; above the xy plane.

2) Verify Stoke's theorem for the function $\boldsymbol{F} = x^2\boldsymbol{i} - xy\boldsymbol{j}$ integrated round the square in the line $x = 0$, $y = 0$; $x = e$, $y = a$

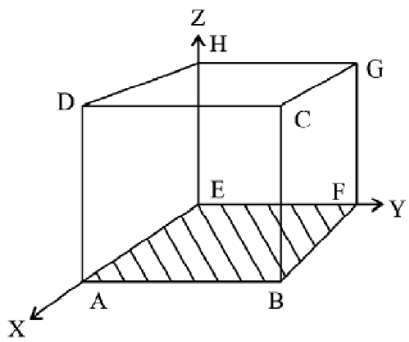


Fig. (Solution 1)

Solution 1 : We have from Stoke's theorem

$$\oint_C \boldsymbol{A}.\boldsymbol{dr} = \iint_S (\nabla \times \boldsymbol{A}) \,.\, \hat{\boldsymbol{n}}\, dS \qquad \text{(i)}$$

We first evalute R.H.S. of equation (1), we have

$$\nabla \times \boldsymbol{A} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ y - z + 2 & yz + 4 & -xz \end{vmatrix} = -y\boldsymbol{i} + (z+1)\boldsymbol{j} - \boldsymbol{k} = \boldsymbol{F}$$

Now we consider the cube

$$\text{Now } \oint_S \boldsymbol{F}.\boldsymbol{dS} = \oint_S \boldsymbol{F}(\boldsymbol{i}dS_x + \boldsymbol{j}dS_y + \boldsymbol{k}dS_z) = \oint_S \boldsymbol{F}.(\boldsymbol{i}\, dy\, dz + \boldsymbol{j}\, dx\, dz + \boldsymbol{k}\, dx\, dy)$$

$$= \iint_{ABCD} \boldsymbol{F}.\boldsymbol{i}\, dy\, dz + \iint_{EFGH} \boldsymbol{F}.-\boldsymbol{i}dy\, dz + \iint_{BFGC} \boldsymbol{F}.\boldsymbol{j}\, dx\, dz + \iint_{AEHD} \boldsymbol{F}.-\boldsymbol{j}\, dx\, dz$$

$$+ \iint_{DCGH} \boldsymbol{F}.\boldsymbol{k}\, dx\, dy + \iint_{ABFE} \boldsymbol{F}.-\boldsymbol{k}\, dx\, dy$$

$$\text{therefore } \iint_S \boldsymbol{F}.\boldsymbol{dS} = \oint_S \boldsymbol{F}.\boldsymbol{dS} - \iint_{ABFE} \boldsymbol{F}.\boldsymbol{k}\, dx\, dy$$

Now $\iint_{ABCD} \boldsymbol{F}.\boldsymbol{i}\, dy\, dz = \iint -y\, dy\, dz = \int_0^2 -y\, dy \int_0^2 dz = \left[\frac{y^2}{2}\right]_0^2 [z]_0^2 = -4$

$$\iint_{EFGH} \boldsymbol{F}.\boldsymbol{i} dy\, dz = \iint +y\, dy\, dz = \int_0^2 +y\, dy \int_0^2 dz = +4$$

$$\iint_{BFGC} \boldsymbol{F}.\boldsymbol{j}\, dx\, dz = \iint_{BFGC} (z-1) dx\, dz = \int_0^2 dx \int_0^2 (z-1) dz = 2\left[\frac{2^2}{2} - 2\right] = 0$$

$$\iint_{AEHD} \boldsymbol{F}.-\boldsymbol{j}\, dx\, dz = -\iint_{AEHD} (z-1)\, dx\, dz = 0$$

$$\iint_{DCGH} \boldsymbol{F}.\boldsymbol{k}\, dx\, dy = \iint_{DCGH} (-1)\, dx\, dy = -\int_0^2 dx \int_0^2 dy = -4$$

$$\therefore \iint \boldsymbol{F}.\boldsymbol{dS} = -4 + 4 + 0 + 0 - 4 = -4$$

Now we evaluate L.H.S. of equation (1)

$$\oint_C \boldsymbol{A}.\boldsymbol{dr} = \oint_C \left[(y - z + 2)dx + (yz + 4)dy - xz\, dz\right]$$

$$= \int_{EA} \boldsymbol{A}.\boldsymbol{dr} + \int_{\boldsymbol{AB}} \boldsymbol{A}.\boldsymbol{dr} + \int_{\boldsymbol{BF}} \boldsymbol{A}.\boldsymbol{dr} + \int_{\boldsymbol{FE}} \boldsymbol{A}.\boldsymbol{dr}$$

Now along EA : $y = 0,\ dy = 0,\ z = 0,\ dz = 0$

$$\therefore \int_{EA} \boldsymbol{A}.\boldsymbol{dr} = \int_0^2 2\, dx = 4$$

Along AB : $x = a,\ dx = 0,\ z = 0,\ dz = 0$

$$\therefore \int_{AB} \boldsymbol{A}.\boldsymbol{dr} = \int_0^2 4\, dy = 8$$

Along BF : $z = 0,\ dz = 0,\ y = 2,\ dy = 0$

$$\int_{BF} \boldsymbol{A}.\boldsymbol{dr} = \int_2^0 4\, dx = -8$$

Along FE : $z = 0,\ dz = 0,\ x = 0,\ dx = 0$

$$\int_{FE} \boldsymbol{A}.\boldsymbol{dr} = \int_2^0 4\, dy = -8$$

$$\therefore \oint_C \boldsymbol{A}.\boldsymbol{dr} = 4+8-8-8=-4$$

$\therefore$ L.H.S of equation (1) is – 4 and R.H.S of equation (1) is also – 4. Therefore Stoke's theorem is verified.

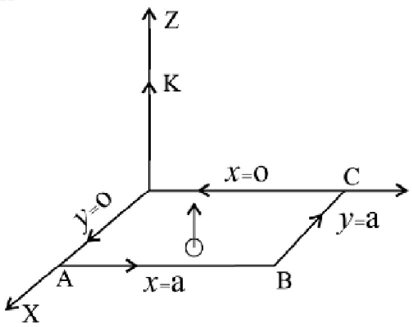


Fig. Solution (2)

Solution 2 : $\boldsymbol{F} = x^2 i - xy\boldsymbol{j} \qquad \therefore \nabla \times \boldsymbol{F} = -y\boldsymbol{k}$

Now, $\iint_S \nabla \times \boldsymbol{F}.\hat{\boldsymbol{n}}\, dS = \iint_S -y\, \boldsymbol{k}.\hat{\boldsymbol{n}}\, dx\, dy$

$$= \iint_S -y\boldsymbol{k}.\boldsymbol{k}\, dx\, dy$$

$$\therefore \iint_S \nabla \times \boldsymbol{F}.\hat{\boldsymbol{n}} ds = \iint_S -y\, dx\, dy = -\int_0^a y\, dy \int_0^a dx$$

$$= -\frac{a^3}{2}$$

Now, $\boldsymbol{F}.\boldsymbol{dr} = (x^2\boldsymbol{i} - xy\boldsymbol{j})\ (\boldsymbol{i}\ dx + \boldsymbol{j}dy) = x^2dx - xy\ dy$

Now $\oint_C \boldsymbol{F}.\boldsymbol{dr} = \int_{OA} \boldsymbol{F}.\boldsymbol{dr} + \int_{AB} \boldsymbol{F}.\boldsymbol{dr} + \int_{BC} \boldsymbol{F}.\boldsymbol{dr} + \int_{CO} \boldsymbol{F}.\boldsymbol{dr}$

Along OA : $y = 0$, $dy = 0$

$$\therefore \int_{OA} \boldsymbol{F}.\boldsymbol{dr} = \int_0^a x^2 dx = \frac{a^3}{3}$$

Along AB : $x = a$, $dx = 0$

$$\therefore \int_{AB} \boldsymbol{F}.\boldsymbol{dr} = \int_0^a -ay\, dy = -\frac{a^3}{2}$$

Along BC : $y = a$, $dy = 0$

$$\therefore \int_{BC} \boldsymbol{F}.\boldsymbol{dr} = \int_0^a x^2 dx = -\frac{a^3}{3}$$

Along CO : $x = 0$, $dx = 0$

$$\therefore \int_{CO} \boldsymbol{F}.\boldsymbol{dr} = 0$$

$$\therefore \oint_C \boldsymbol{F}.\boldsymbol{dr} = \frac{a^3}{3} - \frac{a^3}{2} - \frac{a^3}{3} + 0 = -\frac{a^3}{2}$$

$$\therefore \oint_C \boldsymbol{F}.\boldsymbol{dr} = \iint_S \nabla \times \boldsymbol{F}.\hat{\boldsymbol{n}}\, ds$$

$\therefore$ Stoke's theorem is verified.

Keywords

Double and triple integral, Jacobian, line, surface, volume integral of a vector field, Gauss's divergence theorem, Stoke's theorem and Green's theorem in plane.

4.32 ❒ Summary

- Definitions of double and triple integration as a limit of sums are given. Change of order of integration is exemplified and example of triple integral are given.

- Change of variables of integrands with respective Jacobians are introduced.
- Line, surface, volume integrals of vector field are discussed with examples.
- Elementary proofs of Gauss's divergence theorem, Stoke's theorem and Green's theorem in a plane are given. Also verification of these theorems with examples are provided.

Unit 5 ❒ Orthogonal Curvilinear Co-ordinates

Structure

5.1 ❒ Objectives

Objective of this chapter is to set up an orthogonal curvilinear co-ordinate system and find its unit vectors. The line element, area and volume elements are expressed in terms of orthogonal curvilinear co-ordinate. Now we have derived expressions for gradient, divergence, curl and Laplacian in terms of orthogonal curvilinear co-ordinates and have shown this expressions in some special co-ordinate system like circular cylindrical and spherical polar co-ordinate system.

5.2 ❑ Introduction

In rectangular co-ordinate system the co-ordinate surfaces are planes and they intersect at right angles to each other producing straight co-ordinate axes. In the previous chapters we have defined gradient, Divergence, curl and Laplacian in rectangular co-ordinates *x*, *y*, z. But in solving many physical problems, depending on the symmetry of the problems, we have to express those vector operators in other co-ordinate systems like, cylindrical, spherical etc. in which the surfaces are not all planes and the intersection of the surfaces are curve lines rather than straight lines. Therefore, it has now become necessary to define a co-ordinate system whose co-ordinate surfaces are curved surfaces and the intersections of these curved surfaces produce curved lines as axes of co-ordinate system. This preferred co-ordinate system is called curvilinear co-ordinate system. When the curved surfaces intersect at right angles, we have orthogonal curvilinear co-ordinate system and orthogonal curvilinear co-ordinates are convenient to study the physical problems.

5.3 ❑ Curvilinear co-ordinates

Three curve surfaces u_1 = constant, u_2 = constant and u_3 = constant are taken such that any two surfaces always intersect to produce a curve and all the three surfaces intersect at a point.

For example, (Fig 5.1), surfaces u_1 = constant and u_2 = constant intersect along a curve called u_3 – axis and similarly u_1 – axis and u_2 – axis are defined. We can take these curves of intersections as reference axes to construct a co-ordinate system, called curvilinear co-ordinate system.

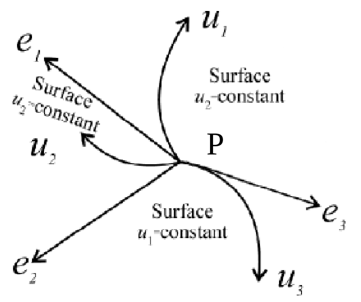


Fig (5.1)

Let the Cartesian co-ordinates and the curvilinear co-ordinates of a same point *P* be (*x*, *y*, *z*) and (u_1, u_2, u_3) respectively. Since there need be point to point correspondence between the co-ordinate systems :

$$\left.\begin{aligned} x &= x(u_1, u_2, u_3) \\ y &= y(u_1, u_2, u_3) \\ z &= z(u_1, u_2, u_3) \end{aligned}\right\} \qquad (5.1)$$

And $$\left.\begin{aligned} u_1 &= u_1(x, y, z) \\ u_2 &= u_2(x, y, z) \\ u_3 &= u_3(x, y, z) \end{aligned}\right\} \qquad (5.2)$$

The functions defined by (5.2) ae continuous having first order continuous derivatives.

If $\boldsymbol{r}$ be the position vector of the point P, then the vectors along tangents to u_1, u_2, u_3 axes will be $\frac{\partial \boldsymbol{r}}{\partial u_1}, \frac{\partial \boldsymbol{r}}{\partial u_2}$ and $\frac{\partial \boldsymbol{r}}{\partial u_3}$ respectively and the unit vectors along these tangents are :

$\boldsymbol{e_1} = \frac{1}{h_1}\frac{\partial \boldsymbol{r}}{\partial u_1}$, $\boldsymbol{e_2} = \frac{1}{h_2}\frac{\partial \boldsymbol{r}}{\partial u_2}$, $\boldsymbol{e_3} = \frac{1}{h_3}\frac{\partial \boldsymbol{r}}{\partial u_3}$, (5.3) respectively,

where $h_i = \left|\frac{\partial \boldsymbol{r}}{\partial u_i}\right|$ (5.4) ; i = 1, 2, 3 and called scale factors which may have dimensions.

Now in order that the co-ordinate surfaces u_1 = constant, u_2 = constant and u_3 = constant uniquely defines a point of intersection, the vectors $\frac{\partial \boldsymbol{r}}{\partial u_1}, \frac{\partial \boldsymbol{r}}{\partial u_2}$ and $\frac{\partial \boldsymbol{r}}{\partial u_3}$ should be non-coplaner and their triple scalar product do not vanish i.e. $\frac{\partial \boldsymbol{r}}{\partial u_1} \cdot \frac{\partial \boldsymbol{r}}{\partial u_2} \times \frac{\partial \boldsymbol{r}}{\partial u_3} \neq 0$ where $\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$

i.e. $$\begin{vmatrix} \frac{\partial x}{\partial u_1} & \frac{\partial y}{\partial u_1} & \frac{\partial z}{\partial u_1} \\ \frac{\partial x}{\partial u_2} & \frac{\partial y}{\partial u_2} & \frac{\partial z}{\partial u_2} \\ \frac{\partial x}{\partial u_3} & \frac{\partial y}{\partial u_3} & \frac{\partial z}{\partial u_3} \end{vmatrix} \neq 0 \qquad (5.5)$$

or, $$\frac{\partial(x, y, z)}{\partial(u_1, u_2, u_3)} \neq 0 \qquad (5.6)$$

where (5.6) gives the Jacobian of transformation : (see article 4C.5).

5.4 ❐ Orthogonal Curvilinear Co-ordinates

The curvilinear co-ordinate system would be orthogonal if the unit vectors given by equations (5.3) along the tangent to the axes are orthogonal.

i.e. $$\boldsymbol{e}_i.\boldsymbol{e}_j = \begin{Bmatrix} 0 & for\ i \neq j \\ 1 & for\ i \neq j \end{Bmatrix} \tag{5.7}$$

and $\boldsymbol{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$; $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$; $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$

5.4.1 : Elements of Arc length, Area and volume

Arc length = since $\boldsymbol{r} = \boldsymbol{r}(u_1, u_2, u_3)$,

$$\boldsymbol{dr} = \frac{\partial \boldsymbol{r}}{\partial u_1}\partial u_1 + \frac{\partial \boldsymbol{r}}{\partial u_2}\partial u_2 + \frac{\partial \boldsymbol{r}}{\partial u_3}\partial u_3 \tag{5.8a}$$

$$= h_1\boldsymbol{e}_1du_1 + h_2\boldsymbol{e}_2du_2 + h_3\boldsymbol{e}_3du_3 \tag{5.8b}$$

using equation (5.3)

Now square of the differential arc length ds is given by :

$$ds^2 = \boldsymbol{dr}.\boldsymbol{dr} = h_1^2\left(du_1\right)^2 + h_2^2\left(du_2\right)^2 + h_3^2\left(du_3\right)^2 \tag{5.9}$$

Using the equation (5.7)

or, $$ds^2 = ds_1^2 + ds_2^2 + ds_3^2 \tag{5.10}$$

where $ds_1 = h_1du_1$, $ds_2 = h_2du_2$ and $ds_3 = h_3du_3$ are the differential arc lengths along u_1,u_2 and u_3 respectively.

Now plane area is a vector given by the cross-product of two vectors. Thus element of area in this case is, $\boldsymbol{ds}_1 \times \boldsymbol{ds}_2 = h_1du_1\boldsymbol{e}_1 \times h_2du_2\boldsymbol{e}_2 = \boldsymbol{e}_1 \times \boldsymbol{e}_2h_1h_2du_1du_2$

Or, $\boldsymbol{ds}_1 \times \boldsymbol{ds}_2 = \boldsymbol{e}_3h_1h_2du_1du_2$ or, $|\boldsymbol{ds}_1 \times \boldsymbol{ds}_2| = h_1h_2du_1du_2$ (5.11)

Similarly, $|\boldsymbol{ds}_2 \times \boldsymbol{ds}_3| = h_2h_3du_2du_3$ (5.12)

And $|\boldsymbol{ds}_3 \times \boldsymbol{ds}_1| = h_3h_1du_3du_1$ (5.13)

Again, volume element having sides $\boldsymbol{ds}_1$, $\boldsymbol{ds}_2$ and $\boldsymbol{ds}_3$ is given by triple scalar product $dV = \boldsymbol{ds}_1.\boldsymbol{ds}_2 \times \boldsymbol{ds}_3 = (\boldsymbol{e}_1.\boldsymbol{e}_2 \times \boldsymbol{e}_3)h_1h_2h_3du_1du_2du_3$

Or, $dV = h_1h_2h_3du_1du_2du_3$ (5.14)

using equation (5.7). Equation (5.14) gives the volume element in orthogonal curvilinear co-ordinate system.

5.4.2 : $\dfrac{\partial \boldsymbol{r}}{\partial \boldsymbol{u_i}}$ and $\nabla \boldsymbol{u_i}$ ($\boldsymbol{i}$ = 1,2,3) forms a reciprocal System of Triads [see article 4A.3.21]

We have $d\boldsymbol{r} = \boldsymbol{i}dx + \boldsymbol{j}dy + \boldsymbol{k}dz$, and $\nabla u_1 = \boldsymbol{i}\dfrac{\partial u_1}{\partial x} + \boldsymbol{j}\dfrac{\partial u_2}{\partial y} + \boldsymbol{k}\dfrac{\partial u_3}{\partial z}$

$$\therefore \nabla u_i.d\boldsymbol{r} = du_i \tag{5.15}$$

Again if we multiply both sides of equation (5.8a) scalarly by ∇u_1, we get

$$\nabla u_1.d\boldsymbol{r} = \left(\nabla u_1.\frac{\partial \boldsymbol{r}}{\partial u_1}\right)\partial u_1 + \left(\nabla u_1.\frac{\partial \boldsymbol{r}}{\partial u_2}\right)\partial u_2 + \left(\nabla u_1.\frac{\partial \boldsymbol{r}}{\partial u_3}\right)\partial u_3 \tag{5.16}$$

Comparing equation (5.15) and (5.16), we get $\nabla u_1.\dfrac{\partial \boldsymbol{r}}{\partial u_1} = 1, \nabla u_1.\dfrac{\partial \boldsymbol{r}}{\partial u_2} = 0,$

$$\nabla u_1.\frac{\partial \boldsymbol{r}}{\partial u_3} = 0$$

Similarly, by multiplying equation (5.8a) scalarly by ∇u_2 and ∇u_3, we get

$$\nabla u_2.\frac{\partial \boldsymbol{r}}{\partial u_1} = 1;\quad \nabla u_2.\frac{\partial \boldsymbol{r}}{\partial u_2} = 1;\quad \nabla u_2.\frac{\partial \boldsymbol{r}}{\partial u_3} = 0$$

And $$\nabla u_3.\frac{\partial \boldsymbol{r}}{\partial u_1} = 0;\quad \nabla u_3.\frac{\partial \boldsymbol{r}}{\partial u_2} = 0;\quad \nabla u_3.\frac{\partial \boldsymbol{r}}{\partial u_3} = 1$$

From these relations we can say that vectors $\dfrac{\partial \boldsymbol{r}}{\partial u_i}$ and ∇u_i (i = 1, 2, 3) form a reciprocal system of triads. Therefore,

$$\nabla u_1 = \frac{\left(\dfrac{\partial r}{\partial u_2} \times \dfrac{\partial r}{\partial u_3}\right)}{\left[\dfrac{\partial r}{\partial u_1}\,\dfrac{\partial r}{\partial u_2}\,\dfrac{\partial r}{\partial u_3}\right]} = \frac{(\boldsymbol{e}_2 h_2 \times \boldsymbol{e}_3 h_3)}{h_1 h_2 h_3 (\boldsymbol{e}_1 \cdot \boldsymbol{e}_2 \times \boldsymbol{e}_3)} = \frac{h_2 h_3 (\boldsymbol{e}_2 \times \boldsymbol{e}_3)}{h_1 h_2 h_3} = \frac{\boldsymbol{e}_1}{h_1}$$

Since $\boldsymbol{e}_1 . \boldsymbol{e}_2 \times \boldsymbol{e}_3 = \mathbf{1}$ & $\boldsymbol{e}_2 \times \boldsymbol{e}_3 = \boldsymbol{e}_1$,

Similarly $\nabla u_2 = \dfrac{\boldsymbol{e}_2}{h_2}$ and $\nabla u_3 = \dfrac{\boldsymbol{e}_3}{h_3}$

Or, we get
$$\left.\begin{aligned} \boldsymbol{e}_1 &= h_1 \nabla u_1 \\ \boldsymbol{e}_2 &= h_2 \nabla u_2 \\ \boldsymbol{e}_3 &= h_3 \nabla u_3 \end{aligned}\right\} \qquad (5.17)$$

5.5 ❒ Gradient in orthogonal curvilinear co-ordinates

We consider a scalar function $\phi(u_1, u_2, u_3)$ where u_1, u_2, u_3 are functions of x, y, z defined by equation 5.2.

Then we have :

$$\left.\begin{aligned} \frac{\partial \phi}{\partial x} &= \frac{\partial \phi}{\partial u_1}\frac{\partial u_1}{\partial x} + \frac{\partial \phi}{\partial u_2}\frac{\partial u_2}{\partial x} + \frac{\partial \phi}{\partial u_3}\frac{\partial u_3}{\partial x} \\ \frac{\partial \phi}{\partial y} &= \frac{\partial \phi}{\partial u_1}\frac{\partial u_1}{\partial y} + \frac{\partial \phi}{\partial u_2}\frac{\partial u_2}{\partial y} + \frac{\partial \phi}{\partial u_3}\frac{\partial u_3}{\partial y} \\ \frac{\partial \phi}{\partial z} &= \frac{\partial \phi}{\partial u_1}\frac{\partial u_1}{\partial z} + \frac{\partial \phi}{\partial u_2}\frac{\partial u_2}{\partial z} + \frac{\partial \phi}{\partial u_3}\frac{\partial u_3}{\partial z} \end{aligned}\right\} \qquad (5.18)$$

Multiplying equation (5.18) by ***i, j, k*** respectively and adding we get :

$$\nabla \phi = \boldsymbol{i}\frac{\partial \phi}{\partial x} + \boldsymbol{j}\frac{\partial \phi}{\partial y} + \boldsymbol{k}\frac{\partial \phi}{\partial z} = \frac{\partial \phi}{\partial u_1}\nabla u_1 + \frac{\partial \phi}{\partial u_2}\nabla u_2 + \frac{\partial \phi}{\partial u_3}\nabla u_3$$

or,
$$\nabla \phi = \frac{\boldsymbol{e}_1}{h_1}\frac{\partial \phi}{\partial u_1} + \frac{\boldsymbol{e}_2}{h_2}\frac{\partial \phi}{\partial u_2} + \frac{\boldsymbol{e}_3}{h_3}\frac{\partial \phi}{\partial u_3} \qquad (5.19)$$

Using (5.17).

5.5.1 : Gradient in circular cylindrical co-ordinates

In this co-ordinate system, the three curvilinear co-ordinates are :

$$u_1 = r, \qquad 0 \le r \le \infty$$

$$u_2 = \theta, \qquad 0 \le \theta \le 2\pi$$

$$u_3 = z, \qquad -\infty \le z \le +\infty$$

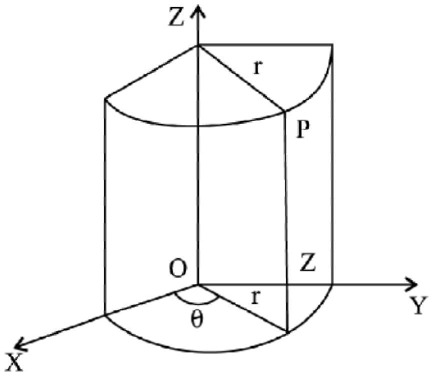


Fig. 5.2

Transformation equations are :

$$\left.\begin{array}{c} x = r\cos\theta \\ y = r\sin\theta \\ z = z \end{array}\right\} \qquad (5.20)$$

Now $\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z = \boldsymbol{i}\, r\cos\theta + \boldsymbol{j}\, r\sin\theta + \boldsymbol{k}\, z$

$$\therefore \ \frac{\partial \boldsymbol{r}}{\partial r} = \boldsymbol{i}\cos\theta + \boldsymbol{j}\sin\theta$$

$$\frac{\partial \boldsymbol{r}}{\partial \theta} = \boldsymbol{i}\, r(-\sin\theta) + \boldsymbol{j}\, r\sin\theta = r\left(-\boldsymbol{i}\sin\theta + \boldsymbol{j}\cos\theta\right)$$

$$\frac{\partial \boldsymbol{r}}{\partial z} = \boldsymbol{k}$$

Therefore unit vectors are :

$$\left.\begin{array}{c} \boldsymbol{e_r} = \dfrac{\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{r}}}{\left|\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{r}}\right|} = \boldsymbol{i}\cos\theta + \boldsymbol{j}\sin\theta \\ \boldsymbol{e}_\theta = \dfrac{\frac{\partial \boldsymbol{r}}{\partial \theta}}{\left|\frac{\partial \boldsymbol{r}}{\partial \theta}\right|} = -\boldsymbol{i}\sin\theta + \boldsymbol{j}\cos\theta \\ \boldsymbol{e_z} = \boldsymbol{k} \end{array}\right\} \qquad (5.21)$$

Scale factors are, $h_{\boldsymbol{r}} = \left|\dfrac{\partial \boldsymbol{r}}{\partial \boldsymbol{r}}\right| = 1;\ h_\theta = \left|\dfrac{\partial \boldsymbol{r}}{\partial \theta}\right| = \boldsymbol{r};\ h_z = \left|\dfrac{\partial \boldsymbol{r}}{\partial z}\right| = 1$

Therefore $\nabla\phi = \frac{\boldsymbol{e}_1}{h_1}\frac{\partial\phi}{\partial u_1} + \frac{\boldsymbol{e}_2}{h_2}\frac{\partial\phi}{\partial u_2} + \frac{\boldsymbol{e}_3}{h_3}\frac{\partial\phi}{\partial u_3}$

$$= \boldsymbol{e_r}\frac{\partial\phi}{\partial \boldsymbol{r}} + \boldsymbol{e_\theta}\frac{1}{\boldsymbol{r}}\frac{\partial\phi}{\partial\theta} + \boldsymbol{k}\frac{\partial\phi}{\partial z} \tag{5.22}$$

5.5.2 : Gradient in spherical polar co-ordinates

In this co-ordinate system, $u_1 = r;\ u_2 = \theta;\ u_3$

Transformation equations are

$$\left.\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned}\right\} \tag{5.23}$$

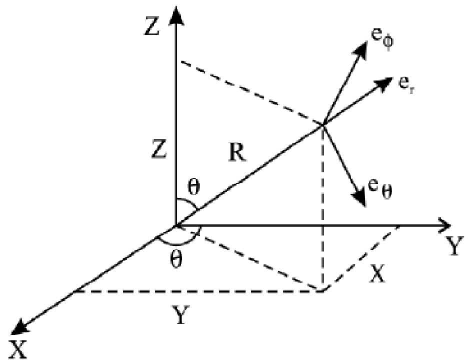


Fig. 5.3

We have $\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$

or,

$$\boldsymbol{r} = \boldsymbol{i}\, r\sin\theta\cos\phi + \boldsymbol{j}\, r\sin\theta\sin\phi + \boldsymbol{k}\, r\cos\theta$$

$$\frac{\partial\boldsymbol{r}}{\partial r} = \boldsymbol{i}\sin\theta\cos\phi + \boldsymbol{j}\sin\theta\sin\phi + \boldsymbol{k}\cos\theta$$

$$\left|\frac{\partial\boldsymbol{r}}{\partial r}\right| = h_r = h_1 = 1$$

$$\frac{\partial\boldsymbol{r}}{\partial\theta} = \boldsymbol{i}\, r\cos\theta\cos\phi + \boldsymbol{j}\, r\cos\theta\sin\phi - \boldsymbol{k}\, r\sin\theta$$

$$\therefore \left|\frac{\partial\boldsymbol{r}}{\partial\theta}\right| = h_\theta = h_2 = r$$

$$\frac{\partial\boldsymbol{r}}{\partial\phi} = \boldsymbol{i}\, r\sin\theta(-\sin\phi) + \boldsymbol{j}\, r\sin\theta\cos\phi$$

$$\therefore \left|\frac{\partial\boldsymbol{r}}{\partial\phi}\right| = h_\phi = h_3 = r\sin\theta$$

$$\left\{\begin{array}{c} \therefore h_1 = h_r = 1 \\ h_2 = h_\theta = r \\ h_3 = h_\phi = r\sin\theta \end{array}\right\} \tag{5.24}$$

Now,

$$\left.\begin{array}{l} \boldsymbol{e}_1 = \boldsymbol{e_r} = \dfrac{\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{r}}}{\left|\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{r}}\right|} = \boldsymbol{i}\sin\theta\cos\phi + \boldsymbol{j}\sin\theta\sin\phi + \boldsymbol{k}\cos\theta \\ \boldsymbol{e}_2 = \boldsymbol{e_\theta} = \dfrac{\frac{\partial \boldsymbol{r}}{\partial \theta}}{\left|\frac{\partial \boldsymbol{r}}{\partial \theta}\right|} = \boldsymbol{i}\cos\theta\cos\phi + \boldsymbol{j}\cos\theta\sin\phi + \boldsymbol{k}\sin\theta \\ \boldsymbol{e}_3 = \boldsymbol{e_\phi} = \dfrac{\frac{\partial r}{\partial \phi}}{\left|\frac{\partial r}{\partial \phi}\right|} = -\boldsymbol{i}\sin\phi + \boldsymbol{j}\cos\theta \end{array}\right\} \tag{5.24a}$$

Therefore, replacing ϕ by ψ ; in (5.19)

$$\nabla\psi = \frac{\boldsymbol{e}_1}{h_1}\frac{\partial\psi}{\partial u_1} + \frac{\boldsymbol{e}_2}{h_2}\frac{\partial\psi}{\partial u_2} + \frac{\boldsymbol{e}_3}{h_3}\frac{\partial\psi}{\partial u_3} = \boldsymbol{e_r}\frac{\partial\psi}{\partial r} + \boldsymbol{e_r}\frac{\partial\psi}{\partial\theta} + \boldsymbol{e}_\phi\frac{1}{r\sin\theta}\frac{\partial\psi}{\partial\phi} \tag{5.25}$$

5.6 ❒ Divergence in orthogonal Curvilinear Co-ordinates

We consider a vector point function $\boldsymbol{A}(u_1,u_2,u_3)$ having components $\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_3$ along the unit vectors $\boldsymbol{e}_1, \boldsymbol{e}_2$, & $\boldsymbol{e}_3$ respectively, such that $\boldsymbol{A} = A_1\boldsymbol{e}_1 + A_2\boldsymbol{e}_2 + A_3\boldsymbol{e}_3 = \boldsymbol{A}_1 + \boldsymbol{A}_2 + \boldsymbol{A}_3$ (5.26)

Since the unit vectors are orthogonal, $\boldsymbol{e}_1 = \boldsymbol{e}_2 \times \boldsymbol{e}_3$; $\boldsymbol{e}_2 = \boldsymbol{e}_3 \times \boldsymbol{e}_1$; $\boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2$ (5.9)

Let us consider the component $\boldsymbol{A}_1$ of the vector point function.

We have, $\boldsymbol{A}_1 = A_1\boldsymbol{e}_1 = A_1\boldsymbol{e}_2 \times \boldsymbol{e}_3 = A_1 h_2 h_3 \nabla u_2 \times \nabla u_3$, using equation (5.17)

Therefore, $\nabla\boldsymbol{A}_1 = \nabla\cdot A_1 h_2 h_3 \nabla u_2 \times \nabla u_3$

$$= \nabla\left(A_1 h_2 h_3\right)\cdot\nabla u_2 \times \nabla u_3 + A_1 h_2 h_3 \nabla\cdot\nabla u_2 \times \nabla u_3$$

But $\nabla \cdot \nabla u_2 \times \nabla u_3 = \nabla u_3 \cdot \nabla \times \nabla u_2 - \nabla u_2 \cdot \nabla \times \nabla u_3 = 0$ (see article 4B.13)

$$\therefore \nabla \cdot \boldsymbol{A}_1 = \nabla\left(A_1 h_2 h_3\right) \cdot \nabla u_2 \times \nabla u_3 \tag{5.27}$$

Now for any function $f(u_1)$, we have

$$\nabla f(u_1) = \boldsymbol{i}\frac{\partial f(u_1)}{\partial x} + \boldsymbol{j}\frac{\partial f(u_1)}{\partial y} + \boldsymbol{k}\frac{\partial f(u_1)}{\partial z} = \boldsymbol{i}\frac{\partial f}{\partial u_1} \cdot \frac{\partial u_1}{\partial x} + \boldsymbol{j}\frac{\partial f}{\partial u_1} \cdot \frac{\partial u_1}{\partial y} + \boldsymbol{k}\frac{\partial f}{\partial u_1} \cdot \frac{\partial u_1}{\partial z}$$

$$= \frac{\partial f}{\partial u_1} \nabla u_1 \tag{5.28}$$

Using the identity (5.28) we get, from equation (5.27),

$$\nabla \cdot \boldsymbol{A}_1 = \frac{\partial}{\partial u_1}\left(A_1 h_2 h_3\right) \nabla u_1 \cdot \nabla u_2 \times \nabla u_3 = \frac{\partial}{\partial u_1}\left(A_1 h_2 h_3\right) \frac{\boldsymbol{e}_1}{h_1} \cdot \frac{\boldsymbol{e}_2}{h_2} \times \frac{\boldsymbol{e}_3}{h_3} \text{ using (5.17)}$$

$$= \frac{\boldsymbol{e_1} \cdot \boldsymbol{e_2} \times \boldsymbol{e_3}}{h_1 h_2 h_3} \frac{\partial}{\partial u_1}\left(A_1 h_2 h_3\right) = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial u_1}\left(A_1 h_2 h_3\right)$$

$$\because \boldsymbol{e}_1 \cdot \boldsymbol{e}_2 \times \boldsymbol{e}_3 = 1$$

Similarly, we can find, $\quad \nabla \cdot \boldsymbol{A}_2 = \dfrac{1}{h_1 h_2 h_3} \dfrac{\partial}{\partial u_2}\left(A_2 h_3 h_1\right)$

$$\nabla \cdot \boldsymbol{A}_3 = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial u_3}\left(A_3 h_1 h_2\right)$$

Thus, we have

$$\nabla \cdot \boldsymbol{A} = \nabla \cdot \boldsymbol{A}_1 + \nabla \cdot \boldsymbol{A}_2 + \nabla \cdot \boldsymbol{A}_3$$

$$= \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u_1}\left(A_1 h_2 h_3\right) + \frac{\partial}{\partial u_2}\left(A_2 h_3 h_1\right) + \frac{\partial}{\partial u_3}\left(A_3 h_1 h_2\right)\right] \tag{5.29}$$

Equation (5.29) gives the divergence of a vector point function in orthogonal curvilinear co-ordinates.

5.6.1. Divergence in circular cylindrical co-ordinates

In this co-ordinates system $u_1 = r$, $u_2 = \theta$, $u_3 = z$; $h_1 = h_r = 1$; $h_2 = h_\theta = r$; $h_3 = h_z = 1$ and $\boldsymbol{e}_1 = \boldsymbol{e}_r$, $\boldsymbol{e}_2 = \boldsymbol{e}_\theta$, $\boldsymbol{e}_3 = \boldsymbol{e}_z = \boldsymbol{k}$ from equation (5.24),

Therefore $\nabla \cdot \boldsymbol{A} = \dfrac{1}{r}\dfrac{\partial}{\partial r}(r\, A_r) + \dfrac{1}{r}\dfrac{\partial A_\theta}{\partial \theta} + \dfrac{\partial A_z}{\partial z}$ (5.30)

where $A_1 = A_r, A_2 = A_\theta, A_3 = A_z$.

5.6.2. Divergence in spherical polar co-ordinate

In this co-ordinate system $u_1 = r, u_2 = \theta, u_3 = \phi$;

$$h_1 = h_r = 1;\ h_2 = h_\theta = r,\ h_3 = h_\phi = r\sin\theta;$$

$$\boldsymbol{e}_1 = \boldsymbol{e}_r,\ \boldsymbol{e}_2 = \boldsymbol{e}_\theta,\ \boldsymbol{e}_3 = \boldsymbol{e}_\phi;\quad A_1 = A_r, A_2 = A_\theta, A_3 = A_\phi$$

Therefore from equation (5.29),

$$\nabla \cdot \boldsymbol{A} = \frac{1}{r^2\sin\theta}\left[\sin\theta\frac{\partial}{\partial r}\left(r^2 A_r\right) + r\frac{\partial}{\partial r}\left(\sin\theta\, A_\theta\right) + r\frac{\partial A_\phi}{\partial \phi}\right]$$

5.7 ❑ Curl in Orthogonal Curvilinear Co-ordinates

We have $\boldsymbol{A} = A_1\boldsymbol{e}_1 + A_2\boldsymbol{e}_2 + A_3\boldsymbol{e}_3$

Now using equation (5.17), we get,

$$A = A_1 h_1 \nabla u_1 + A_2 h_2 \nabla u_2 + A_3 h_3 \nabla u_3$$

$$\therefore \nabla \times \boldsymbol{A} = \nabla \times \left(A_1 h_1 \nabla u_1\right) + \nabla \times \left(A_2 h_2 \nabla u_2\right) + \nabla \times \left(A_3 h_3 \nabla u_3\right) \tag{5.31}$$

Now

$$\nabla \times \left(A_1 h_1 \nabla u_1\right) = \nabla\left(A_1 h_1\right) \times \nabla u_1 + A_1 h_1 \nabla \times \nabla u_1 = \nabla\left(A_1 h_1\right) \times \nabla u_1$$

$$= \frac{\partial}{\partial u_1}\left(A_1 h_1\right)\nabla u_1 \times \nabla u_1 + \frac{\partial}{\partial u_2}\left(A_1 h_1\right)\nabla u_2 \times \nabla u_1 + \frac{\partial}{\partial u_3}\left(A_1 h_1\right)\nabla u_3 \times \nabla u_1 \tag{5.32}$$

Now from equation (5.17)

$$\nabla u_2 \times \nabla u_1 = \frac{\boldsymbol{e}_2 \times \boldsymbol{e}_1}{h_1 h_2} = -\frac{\boldsymbol{e}_3}{h_1 h_2}$$

$$\nabla u_3 \times \nabla u_1 = \frac{\boldsymbol{e}_3 \times \boldsymbol{e}_1}{h_1 h_3} = \frac{\boldsymbol{e}_2}{h_1 h_3}$$

Therefore equations (5.32) becomes

$$\nabla \times (A_1 h_1 \nabla u_1) = \frac{\boldsymbol{e}_2}{h_1 h_3} \frac{\partial}{\partial u_3}(A_1 h_1) - \frac{\boldsymbol{e}_3}{h_1 h_2} \frac{\partial}{\partial u_2}(A_1 h_1)$$

Similarly we get,

$$\nabla \times (A_2 h_2 \nabla u_2) = \frac{\boldsymbol{e}_3}{h_1 h_2} \frac{\partial}{\partial u_1}(A_2 h_2) - \frac{\boldsymbol{e}_1}{h_3 h_2} \frac{\partial}{\partial u_3}(A_2 h_2)$$

$$\nabla \times (A_3 h_3 \nabla u_3) = \frac{\boldsymbol{e}_1}{h_1 h_3} \frac{\partial}{\partial u_2}(A_3 h_3) - \frac{\boldsymbol{e}_2}{h_1 h_3} \frac{\partial}{\partial u_3}(A_3 h_3)$$

Thus

$$\nabla \times \boldsymbol{A} = \frac{\boldsymbol{e}_1}{h_2 h_3}\left[\frac{\partial}{\partial u_2}(A_3 h_3) - \frac{\partial}{\partial u_3}(A_2 h_2)\right] + \frac{\boldsymbol{e}_2}{h_1 h_3}\left[\frac{\partial}{\partial u_3}(A_1 h_1) - \frac{\partial}{\partial u_1}(A_3 h_3)\right]$$

$$+\frac{\boldsymbol{e}_3}{h_1 h_2}\left[\frac{\partial}{\partial u_1}(A_2 h_2) - \frac{\partial}{\partial u_2}(A_1 h_1)\right] \tag{5.33}$$

Equation (5.33) can be written in a determine from

$$\nabla \times \boldsymbol{A} = \frac{1}{h_1 h_2 h_3}\begin{vmatrix} \boldsymbol{e}_1 h_1 & \boldsymbol{e}_2 h_2 & \boldsymbol{e}_3 h_3 \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ A_1 h_1 & A_2 h_2 & A_3 h_3 \end{vmatrix} \tag{5.34}$$

5.7.1. Curl in circular cylindrical co-ordinates

In this co-ordinate system, $h_1 = h_r = 1$; $h_2 = h_\theta = r$; $h_3 = h_z = 1$; $\boldsymbol{e}_1 = \boldsymbol{e}_r$, $\boldsymbol{e}_2 = \boldsymbol{e}_\theta$, $\boldsymbol{e}_3 = \boldsymbol{e}_z = \boldsymbol{k}$; $A_1 = A_r$, $A_2 = A_\theta$, $A_3 = A_z$

From equation (5.34),

$$\nabla\times\boldsymbol{A}=\frac{1}{r}\begin{vmatrix}\boldsymbol{e_r} & r\boldsymbol{e_\theta} & \boldsymbol{k}\\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial\theta} & \dfrac{\partial}{\partial z}\\ A_r & rA_\theta & A_z\end{vmatrix} \tag{5.35}$$

5.7.2. Curl in spherical polar co-ordinates

In this co-ordinate system :

$h_1 = h_r = 1$; $h_2 = h_\theta = r$; $h_3 = h_\phi = r\sin\theta$; $u_1 = r$, $u_2 = \theta$, $u_3 = \phi$; $A_1 = A_r$, $A_2 = A_\theta$, $A_3 = A_\phi$

Therefore from equation (5.34)

$$\nabla\times\boldsymbol{A}=\frac{1}{r^2\sin\theta}\begin{vmatrix}\boldsymbol{e_r} & r\boldsymbol{e_\theta} & r\sin\theta\boldsymbol{e_\phi}\\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial\theta} & \dfrac{\partial}{\partial\phi}\\ A_r & rA_\theta & r\sin\theta A_\phi\end{vmatrix} \tag{5.36}$$

5.8 ❐ Laplacian in Orthogonal Curvilinear Co-ordinate

From equation (5.19) : $\nabla\phi=\dfrac{\boldsymbol{e}_1}{h_1}\dfrac{\partial\phi}{\partial u_1}+\dfrac{\boldsymbol{e}_2}{h_2}\dfrac{\partial\phi}{\partial u_2}+\dfrac{\boldsymbol{e}_3}{h_3}\dfrac{\partial\phi}{\partial u_3}$

let $\boldsymbol{A}=\nabla\phi=A_1\boldsymbol{e}_1+A_2\boldsymbol{e}_2+A_3\boldsymbol{e}_3$ (5.37), comparing (5.19) and (5.37) we get,

$$A_1=\frac{1}{h_1}\frac{\partial\phi}{\partial u_1},\ A_2=\frac{1}{h_2}\frac{\partial\phi}{\partial u_2},\ A_3=\frac{1}{h_3}\frac{\partial\phi}{\partial u_3} \tag{5.38}$$

From equation (5.29) :

$$\nabla\cdot\boldsymbol{A}=\frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial u_1}\left(A_1h_2h_3\right)+\frac{\partial}{\partial u_2}\left(A_2h_3h_1\right)+\frac{\partial}{\partial u_3}\left(A_3h_1h_2\right)\right],$$

substituting A_1, A_2, A_3 from equation (5.38), we get

$$\nabla \cdot \boldsymbol{A} = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial \phi}{\partial u_1}\right) + \frac{\partial}{\partial u_2}\left(\frac{h_3 h_1}{h_2}\frac{\partial \phi}{\partial u_2}\right) + \frac{\partial}{\partial u_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial \phi}{\partial u_3}\right)\right] \qquad (5.39)$$

$$\nabla \cdot \nabla \phi = \nabla^2 \phi = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial \phi}{\partial u_1}\right) + \frac{\partial}{\partial u_2}\left(\frac{h_3 h_1}{h_2}\frac{\partial \phi}{\partial u_2}\right) + \frac{\partial}{\partial u_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial \phi}{\partial u_3}\right)\right] \qquad (5.40)$$

5.8.1. Laplacian in circular cylindrical co-ordinate

In this co-ordinate system : $h_1 = h_r = 1$; $h_2 = h_\theta = r$; $h_3 = h_z = 1$; $u_1 = r$, $u_2 = \theta$, $u_3 = z$; $A_1 = A_r$, $A_2 = A_\theta$, $A_3 = A_z$; $\boldsymbol{e}_1 = \boldsymbol{e}_r$, $\boldsymbol{e}_2 = \boldsymbol{e}_\theta$, $\boldsymbol{e}_3 = \boldsymbol{k}$

Therefore from equation (5.40)

$$\nabla^2 \phi = \frac{1}{r}\left[\frac{\partial}{\partial r}\left(r\frac{\partial \phi}{\partial r}\right) + \frac{\partial}{\partial \theta}\left(\frac{1}{r}\frac{\partial \phi}{\partial \theta}\right) + \frac{\partial}{\partial z}\left(r\frac{\partial \phi}{\partial z}\right)\right]$$

$$= \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \phi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 \phi}{\partial \theta^2} + \frac{\partial^2 \phi}{\partial z^2} \qquad (5.41)$$

5.8.2. Laplacian in spherical co-ordinate system

In this co-ordinate system : $h_1 = h_r = 1$; $h_2 = h_\theta = r$; $h_3 = h_\phi = r \sin \theta$; $u_1 = r$, $u_2 = \theta$, $u_3 = \phi$; $A_1 = A_r$, $A_2 = A_\theta$, $A_3 = A_\phi$; $\boldsymbol{e}_1 = \boldsymbol{e}_r$, $\boldsymbol{e}_2 = \boldsymbol{e}_\theta$, $\boldsymbol{e}_3 = \boldsymbol{e}_\phi$

Therefore from equation (5.40) we get changing ϕ to ψ

$$\nabla^2 \Psi = \frac{1}{r^2 \sin\theta}\left[\sin\theta \frac{\partial}{\partial r}\left(r^2 \frac{\partial \psi}{\partial r}\right) + \frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \psi}{\partial \theta}\right) + \frac{1}{\sin\theta}\frac{\partial^2 \psi}{\partial \phi^2}\right] \qquad (5.42)$$

5.8.3. Exercise

1) Express the vector $\boldsymbol{V} = \boldsymbol{i}x + \boldsymbol{j}2y + \boldsymbol{k}z$ in circular cylindrical co-ordinates.

2) Express the vector $\boldsymbol{V} = \boldsymbol{i}2x - \boldsymbol{j}y + 3\boldsymbol{k}z$ in spherical co-ordinates.

Solution :

Solution 1 : Transformation equations in circular cylindrical co-ordinates are

$$\left.\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}\right\}$$

using equation (5.20).

$$\therefore \boldsymbol{V} = \boldsymbol{i}\, r\cos\theta + \boldsymbol{j}\, 2r\sin\theta + \boldsymbol{k}z$$

Now $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ are given in terms of $\boldsymbol{e}_r$, $\boldsymbol{e}_\theta$ and $\boldsymbol{e}_z$ by solving equation (5.21) and we get

$$\boldsymbol{i} = \boldsymbol{e_r}\cos\theta - \boldsymbol{e_\theta}\sin\theta$$

$$\boldsymbol{j} = \boldsymbol{e_r}\sin\theta + \boldsymbol{e_\theta}\cos\theta$$

$$\boldsymbol{k} = \boldsymbol{e}_z$$

Therefore

$$\boldsymbol{V} = r\cos\theta\left(\boldsymbol{e_r}\cos\theta - \boldsymbol{e_\theta}\sin\theta\right) + 2r\sin\theta\left(\boldsymbol{e_r}\sin\theta + \boldsymbol{e_\theta}\cos\theta\right) + \boldsymbol{e_z}z$$

$$= \boldsymbol{e_r}\left(r\cos^2\theta + 2r\sin^2\theta\right) + \boldsymbol{e_\theta}\left(-r\sin\theta\cos\theta + 2r\sin\theta\cos\theta\right) + \boldsymbol{e_z}z$$

$$= r\,\boldsymbol{e_r}\left(1 + \sin^2\theta\right) + r\,\boldsymbol{e_\theta}\sin\theta\cos\theta + \boldsymbol{e_z}z$$

Solution 2 : Transformation equations for spherical co-ordinates are,

$$\left.\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned}\right\}$$

using equation (5.23).

$$\therefore \boldsymbol{V} = \boldsymbol{i}\, 2r\sin\theta\cos\phi - \boldsymbol{j}\, r\sin\theta\sin\phi + 3\boldsymbol{k}\, r\cos\theta$$

Now $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ are given in terms of $\boldsymbol{e_r}, \boldsymbol{e_\theta}, \boldsymbol{e_\phi}$ by solving equations (5.24a) and we get

$$\boldsymbol{i} = \boldsymbol{e_r}\sin\theta\cos\phi + \boldsymbol{e_\theta}\cos\theta\cos\phi - \boldsymbol{e_\phi}\sin\phi$$

$$\boldsymbol{j} = \boldsymbol{e_r}\sin\theta\sin\phi + \boldsymbol{e_\theta}\cos\theta\sin\phi + \boldsymbol{e_\phi}\sin\phi$$

$$\boldsymbol{k} = \boldsymbol{e_r}\cos\theta - \boldsymbol{e_\theta}\sin\theta$$

Therefore

$$\boldsymbol{V} = 2r\sin\theta\cos\phi\left(\boldsymbol{e_r}\sin\theta\cos\phi + \boldsymbol{e_\theta}\cos\theta\cos\phi - \boldsymbol{e_\phi}\sin\phi\right)$$

$$-r\sin\theta\sin\phi\left(\boldsymbol{e_r}\sin\theta\cos\phi+\boldsymbol{e_\theta}\cos\theta\sin\phi+\boldsymbol{e_\phi}\sin\phi\right)$$

$$+3r\cos\theta\left(\boldsymbol{e_r}\cos\theta-\boldsymbol{e_\theta}\sin\theta\right)$$

$$=r\boldsymbol{e_r}\left(2\sin^2\theta\cos^2\phi-\sin^2\theta\sin^2\phi+3\cos^2\theta\right)$$

$$+r\boldsymbol{e_\theta}\left(2\sin\theta\cos\theta\cos^2\phi-\sin\theta\cos\theta\sin^2\phi-3\sin\theta\cos\theta\right)$$

$$+r\boldsymbol{e_\phi}\left(-2\sin\theta\sin\phi\cos\phi-\sin\theta\sin^2\phi\right)$$

$$=r\boldsymbol{e_r}\left(3\sin^2\theta\cos^2\phi+4\cos^2\theta-1\right)+r\boldsymbol{e_\theta}\sin\theta\cos\theta\left(2\cos^2\phi-\sin^2\phi-3\right)$$

$$-r\boldsymbol{e_\phi}\sin\theta\sin\phi(2\cos\phi-\sin\phi)$$

Keywords

Curvilinear co-ordinate system, orthogonal curvilinear co-ordinate system; gradient, divergence, curl and Laplacian.

5.9 ❐ Summary

- We have defined curvilinear and orthogonal curvilinear systems.
- Expressions for elements of arc length, area and volume have been obtained in orthogonal curvilinear co-ordinate system.
- Expression for gradient, divergence, curl and Laplacian have been obtained in orthogonal curvilinear co-ordinate systems and in circular cylindrical and spherical polar co-ordinate systems.

Unit 6 ❐ Dirac Delta Function

Structure

6.1 ❐ Objectives

The objective of this chapter is to introduce Direc Delta function to the students. It's definition and properties are explained. Also various representations of delta function have been discussed.

6.2 ❐ Introduction

Delta function appears in many physical problems. It was first used by P.A.M Dirac in quantum mechanics and thereafter it became popular among physicists and mathematicians and is popularly known now as Dirac Delta function $\delta(x)$.

The point to be remembered is that $\delta(x)$ is not a function at all in the usual sense. Since its value is not finite at $x = 0$ and it is only treated as if it were a function for certain clearly defined purpose in physics and mathematics.

6.3 ❐ Definition

In one dimension, the Dirac Delta function $\delta(x)$, can be thought of as a function on the real line which is zero everywhere except at the origin where it has such a large value that the integral of the function over an interval containing the point $x = 0$ is equal to unity. Thus

$$\left.\begin{aligned} \delta(x) &= 0, \text{ for } x \neq 0 \\ &= \infty, \text{ for } x = 0 \\ \text{and } \int_{-\infty}^{+\infty} \delta(x)\, dx &= 1 \end{aligned}\right\} \qquad (6.1)$$

δ – function has the unit area under the curve.

When the centre of the delta function is shifted to $x = a$ from the origin equation 6.1 is rewritten as :

$$\left.\begin{aligned} \delta(x-a) &= 0, \text{ for } x-a \neq 0 \\ &= \infty, \text{ for } x-a = 0 \\ \text{and } \int_{-\infty}^{+\infty} \delta(x-a)\, dx &= 1 \end{aligned}\right\} \qquad (6.1a)$$

Equation 6.1a is shown in figure (6.3)

6.4 ❑ Step up/step down function : Unit Impulse Function

We consider a function,

$$\left.\begin{aligned} y_\varepsilon(x) &= \tfrac{1}{2\varepsilon}, \quad -\varepsilon \le x \le +\varepsilon \\ &= 0, \qquad\quad |x| > \varepsilon \end{aligned}\right\} \qquad (6.2)$$

As shown in fig (6.1)

We can make an approximation to δ-function by making a step-up / step-down function shown in fig (6.1).

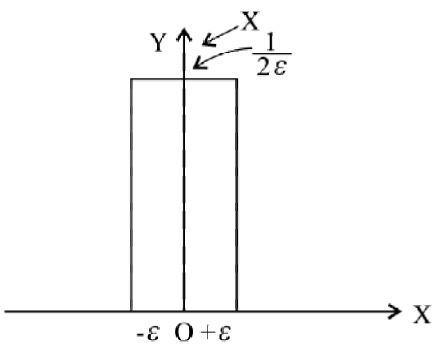


Fig. 6.1

Now let $y = \delta_\varepsilon(x) = u_{-\varepsilon}(x) - u_{+\varepsilon}(x)$

Where $u_{-\varepsilon}(x)$ the step is up function and $u_{+\varepsilon}(x)$ is the step down function. The width of the curve being $\varepsilon - (-\varepsilon) = 2\varepsilon$ and the height is $\frac{1}{2\varepsilon}$ so that area of the curve is $(2\varepsilon)\frac{1}{2\varepsilon} = 1$. Now when $\varepsilon = 0$, we get

$$\delta(x) = \lim_{\varepsilon \to 0} \delta_{\varepsilon}(x) \qquad (6.3)$$

And equation (6.3) is represented by fig (6.2).

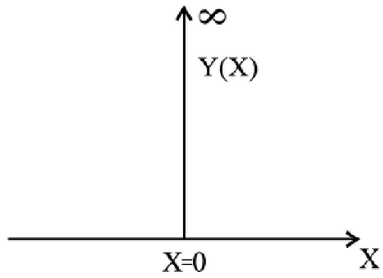


Fig. 6.2

Equation (6.2) also defines a unit impulse function of impulse $F_t(x) \times t = 1$

$$\text{where } F_t(x) = \left.\begin{array}{ll} \dfrac{1}{t}, & -\dfrac{t}{2} \le x \le \dfrac{t}{2} \\ = 0, & |x| > \dfrac{t}{2} \end{array}\right\}$$

A rule for integration of its product with another function $f(x)$ is given by

$$\int_{-\infty}^{+\infty} f(x)\delta(x-0)dx = f(0) \qquad (6.4)$$

when $\delta(x)$ centred at origin. When $\delta(x)$ is centred at $x = a$, we get

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)dx = f(a) \quad (6.5)$$

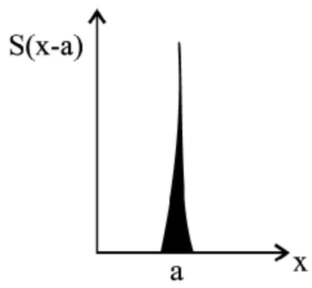


Fig. 6.3

Equation (6.5) is valid for any continuous function $f(x)$, because $\delta(x-a) = 0$ for $x \neq a$ and we can replace the function $f(x)$ by its value at $x = a$ while integrating since,

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)dx = \int_{-\infty}^{+\infty} f(a)\delta(x-a)dx = f(a)\int_{-\infty}^{+\infty} \delta(x-a)dx$$

Now $\int_{-\infty}^{+\infty} \delta(x-a)dx = 1$ by equation (6.1a), and equation (6.5) follows.

The range of integration of equation 6.1 or 6.1a or 6.4 or 6.5 need not be from $-\infty\ to +\infty$. It may be over any region containing the centre of the δ-function where it does not vanish.

It is to be noted that if x has the dimensional length, $\delta(x-a)$ would have the dimension of inverse length.

Similarly if x has the dimension of time, then $\delta(x-a)$ would have the dimension of $(time)^{-1}$.

6.5 ❐ Different representation of the δ-function

a. δ-function as a limiting form of rectangular function :

We suppose :

$$\left.\begin{aligned} y_\varepsilon(x) &= \frac{1}{2\varepsilon}, \quad \text{for} -\varepsilon < x-a < \varepsilon \\ &= 0\,, \qquad \text{for } |x-a| > 0 \end{aligned}\right\} \tag{6.6}$$

We see that as ε decreases, the rectangular distribution becomes narrower as sharper. The integral

$$\int_{-\infty}^{+\infty} y_\varepsilon(x)dx = \frac{1}{2\varepsilon}\int_{a-\varepsilon}^{a+\varepsilon} dx = 1$$

$\varepsilon_1 > \varepsilon_2 > \varepsilon_3$

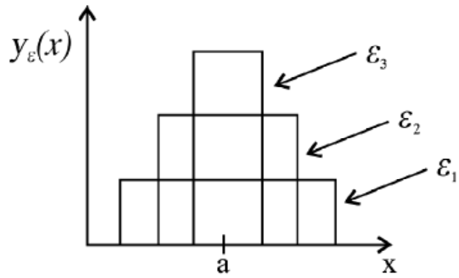


Fig. 6.6

This is true for any value of ε. Thus even in the limit $\varepsilon \to 0$ the structure becomes infinitely peaked, however still retaining the area under the curve as unity so,

$$\lim_{\varepsilon \to 0} \varepsilon\ y_\varepsilon(x) = \delta(x-a)$$

Also

$$\int_{-\infty}^{+\infty} f(x)y_\varepsilon(x)dx = \frac{1}{2\varepsilon}\int_{a-\varepsilon}^{a+\varepsilon} f(x)dx$$

Assuming $f(x)$ to be continuous at $x = a$ and when in the infinitesimal integral $-\varepsilon < x-a < \varepsilon$, $f(x)$ may be assumed to be a constant, we get

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)dx = \lim_{a \to 0}\int_{-\infty}^{+\infty} f(x)\, y_\varepsilon(a)\, dx$$

$$= \lim_{a \to 0}\int_{-\infty}^{+\infty} f(a)\int_{a-\varepsilon}^{a+\varepsilon} dx = f(a)$$

Therefore, the distribution $y_\varepsilon(x)$ in the limit $\varepsilon \to 0$ represent of δ-function.

b. Gaussian representation of the δ-function.

A Gaussian is denoted by

$y_\varepsilon(x) = \frac{1}{\sqrt{2\pi\varepsilon^2}} e^{\frac{(x-a)^2}{2\varepsilon^2}}$; $\varepsilon > 0$, again as ε decreases the Gaussian becomes sharper and in the limit $\varepsilon \to 0$ and will get a δ-function. Also the integral,

$$\int_{-\infty}^{+\infty} y(x)dx = 1$$

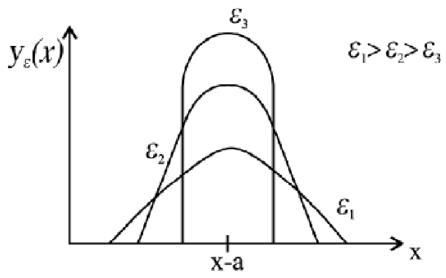


Fig. 6.7

Further it has a width ε and at x = a it has a value $\frac{1}{\sqrt{2\pi\varepsilon^2}}$.

So,

$$\delta(x-a) = \lim_{\varepsilon \to 0} \frac{1}{\sqrt{2\pi\varepsilon^2}} e^{\frac{(x-a)^2}{2\varepsilon^2}}$$

6.5.1 Properties or characteristics of delta function

1. $$\delta(-x) = \delta(x) \tag{6.7}$$

It states that the delta function is an even function of x.

2. $$x\,\delta(x) = 0 \tag{6.8}$$

Since, if we take a continuous function $f(x)$ and find that $\int f(x)x\,\delta(x)\,dx = 0$. This shows that $x\,\delta(x)$ as a factor in the integral is equivalent to zero.

3. $$\delta(x) = \frac{1}{a}\delta(x),\ a > 0 \tag{6.10}$$

We consider $\int_{-\infty}^{+\infty} f(x)\delta(ax)dx = f(0)\int_{-\infty}^{+\infty} \delta(ax)dx.$

Now putting $l = ax, \int_{-\infty}^{+\infty} f(x)\frac{\delta(l)}{a}\delta l = \frac{f(0)}{a}$ (6.11)

But we have $\int_{-\infty}^{+\infty} f(x)\,\delta(ax)dx = f(0)$ (6.4)

Comparing equation (6.11) and (6.4)

$$\delta(ax) = \frac{\delta(l)}{a} = \frac{\delta(ax)}{a}$$

4. $$\delta\left(x^2 - a^2\right) = \frac{1}{2a}\{\delta(x+a) + \delta(x-a)\},\ a > 0 \qquad (6.12)$$

We have $\delta\left(x^2-a^2\right) = \delta[(x-a)\,(x+a)] = \frac{1}{|x+a|}\delta(x-a) + \frac{1}{|x+a|}\delta(x+a)$ using (6.4)

Hence, considering an arbitrary continuous function $f(x)$, we can write,

$$\int f(x)\delta\left(x^2-a^2\right)dx = \int \frac{f(x)}{|x+a|}\delta(x-a)dx + \int \frac{f(x)}{|x-a|}\delta(x+a)dx$$

$$= \frac{1}{2a}[f(a) + f(-a)] \qquad (6.13)$$

using equation (6.10).

The right hand side of the equation (6.13) can be written as

$$\frac{1}{|2a|}\left[\int f(x)\delta(x-a)dx + \int f(x)\delta(x+a)dx\right]$$

Hence we get $\delta\left(x^2-a^2\right) = \frac{1}{2a}[\delta(x-a) + \delta(x+a)]$

5. $$f(x)\delta(x-a) = f(a)\delta(x-a) \qquad (6.14)$$

Since $\delta(x-a) \neq 0$ at x = a but is zero for all other value of x, the product $f(x)\delta(x-a)$ will remain non-zero for x = a and will result in $f(a)\delta(x-a)$.

6.5.2 Delta function in Three Dimension

The three dimensional delta functions are defined as :

$$\left.\begin{aligned} \delta(r) &= 0, \ \text{for } r \neq 0 \\ and \int \delta(r)\, d^3 r &= 1 \end{aligned}\right\} \tag{6.15}$$

Equation (6.4) and (6.5) in three dimensional forms are

$$\int \delta(\boldsymbol{r}) f(\boldsymbol{r}) d^3\boldsymbol{r} = f(0) \tag{6.16}$$

$$\int \delta(\boldsymbol{r} - \boldsymbol{r_0}) f(\boldsymbol{r}) d^3\boldsymbol{r} = f(\boldsymbol{r_0}) \tag{6.17}$$

Key Words

Delta functions, unit impulse function.

6.6 ❐ Summary

- Dirac delta function is defined and explained. Shift of origin considered.
- Rectangular and Gaussian representation discussed.
- Listed the properties of Dirac delta functions.

Unit 7 ❒ Matrices

Structure

7.1 ❐ Objectives

In this chapter we shall discuss various arithmetic operation with matrices covering various terminologies and notation. We shall define a number special matrices which frequently occur in physics and discuss methods of matrix algebra that are useful in solving a system of linear equations in some unknowns.

7.2 ❐ Introduction

Historically study of matrices arose in connection with, successive linear transformtions in vector spaces. The simplest of such transformations are the linear transformtions of components of vectors under rotation of co-ordinate axes as discussed in chapter 4 :

$$A_i' = \sum_{j=1}^{3} a_{ij} \boldsymbol{A_j} \tag{7.1}$$

Where $\boldsymbol{A}_i'$ are the components of vector $\boldsymbol{A}$ in the new co-ordinate system and $\boldsymbol{A}_j$ are those in old co-ordinate system.

where

$$\sum_j a_{ij}^2 = \sum_j a_{ji}^2 = 1\,(i = 1, 2, 3) \tag{7.2}$$

$$\sum_i a_{ij} a_{ik} = \sum_i a_{ji} a_{ki} = 0 \quad (j \neq k) \tag{7.3}$$

Now we consider a further linear transformation of the co-ordinate system in which the same vector has components $\boldsymbol{A}_i''$, which are linearly related to the components $\boldsymbol{A}_i'$ by

$$\boldsymbol{A}_i'' = \sum_{k=1}^{3} b_{ik} \boldsymbol{A}_k' \tag{7.4}$$

It is possible to eliminate the intermediate co-ordinate system and obtain a transformation directly from the components $\boldsymbol{A}_i$ to $\boldsymbol{A}_i''$

$$\boldsymbol{A}_i'' = \sum_{j=1}^{3} \sum_{k=1}^{3} a_{ik} a_{kj} \boldsymbol{A_j} \tag{7.5}$$

$$= \sum_{j=1}^{3}\sum_{k=1}^{3} b_{ik} a_{kj} \boldsymbol{A_j}$$

$$= \sum_{j=1}^{3} c_{ij} \boldsymbol{A_j} \qquad (7.6)$$

Where

$$c_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj} \quad 1 \le i, j \le n$$

In dealing with such transformation it is convenient to introduce the concept of matrices. Now using equation (7.1) and (7.4) we get (7.5), which is the result of two successive linear transformation in the vector space. In fact it is in the study of such successive linear transformations that the branch of matrix algebra historically developed.

For a proper understanding of the basic concept of quantum mechanics, a sound foundation in matrix algebra is essential. Matrices occurs in physics mainly two ways: first in the solution of linear equation and second, in the solution of eigenvalue problems in classical and quantum mechanics.

In this chapter we shall discuss various arithmetic operation with matrices covering various terminologies and notation. We shall define a number special matrices which frequently occur in physics and discuss methods of matrix algebra that are useful in solving a system of linear equations in some unknowns.

7.3 ❒ Definition, Notation and Terminology

A rectangular array of numbers (real or complex) is called a matrix. The array consists of *m* rows and *n* colomns. The individual members of the array are called the elements. Sometime the elements may be functions like $f_1(x)$ etc.

If a matrix has *m* rows and *n* columns, the matrix is of order $m \times n$ (called *m* by *n*). A general *m* by *n* matrix can be written as,

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & & \dots \\ \dots & \dots & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad (7.1)$$

Or $\boldsymbol{A} = \left[a_{ij}\right]_{m\times n}$ (7.8), is shorthand notation.

Terminologies :

1. **Row matrix :** If there be only one row of elements in the matrix, it is called a row matrix. Thus $\boldsymbol{A}$ = [a, b, c, d], is a row matrix of order 1 × 4.

2. **Column matrix :** A matrix having elements in one column only is called a column matrix. Thus

$$\boldsymbol{A} = \begin{bmatrix} p \\ q \\ r \\ s \end{bmatrix} \equiv \{p, q, r, s\}$$, to save space, is a column matrix of order 4 × 1.

3. **Null matrix :** The matrix $\boldsymbol{A}$ of arbitrary order is said to be a null (or zero) matrix if, and only if, every element of $\boldsymbol{A}$ equals zero. We denote a null matrix by **0**. Thus if $\boldsymbol{A}$ = **0**, then $\boldsymbol{A} = \left[a_{ij}\right]_{m\times n} = \mathbf{0}$, thus

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$ is a null matrix of order 2 × 3. It is evident that for any arbitrary matrix $\boldsymbol{A} - (-\boldsymbol{A}) = \mathbf{0}$

4. **Negative matrix :** – $\boldsymbol{A}$ is the negative matrix of $\boldsymbol{A}$, when sign of all the element of $\boldsymbol{A}$ is reversed. If

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ c & -d \end{bmatrix}, \text{ then } -\boldsymbol{A} = \begin{bmatrix} -a & -b \\ -c & +d \end{bmatrix}$$

5. **Transpose of matrix :** If the rows and the columns of a matrix are interchanged, the resulting matrix is called the transpose of the former matrix e.g. if $\boldsymbol{A} = \begin{bmatrix} 2 & 4 & 6 \\ 3 & 5 & 7 \\ 1 & 3 & -5 \end{bmatrix}$,

the transpose of $\boldsymbol{A}$ i.e. $\boldsymbol{A}^{\mathrm{T}}$ or $\tilde{\boldsymbol{A}}$ (called $\boldsymbol{A}$-tilde) is given by, $\boldsymbol{A} = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 5 & 3 \\ 6 & 7 & -5 \end{bmatrix}$.

In notation $\boldsymbol{A} = \left[a_{ij}\right]_{m\times n}$ then $\boldsymbol{A}^{\boldsymbol{T}} = \left[a_{ji}\right]_{\boldsymbol{n}\times\boldsymbol{m}}$

6. **Square matrix :** If the number of rows and columns of a matrix are equal i.e. $m = n$, the matrix is said to be a square matrix. If $\boldsymbol{A}$ is a square matrix of order $n \times n$ we say that $\boldsymbol{A}$ is of order n. Square matrices gives rise to various types of special matrices which frequently occur in physics. For example

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3 & 4 \\ -2 & 5 & 6 \\ 0 & -9 & 0 \end{bmatrix} \text{ is a square matrix of } 3 \times 3.$$

7. **Diagonal matrix :** A square matrix having all its non-diagonal elements as zero is called a diagonal matrix. Let $\boldsymbol{A} = \left[a_{ij}\right]_n$ be a square matrix of order n. The elements a_{11}, a_{22}, a_{33}, ... a_{nn} form the principle diagonal of the matrix. The elements a_{ii} are called the diagonal elements of the equare matrix $\boldsymbol{A}$.

All the remaining elements a_{ij} for $i \neq j$ are called the off-diagonal elements.

Thus in a diagonal matrix $\boldsymbol{A}$, $a_{ii} \neq 0, a_{ij} = 0$, or in short $\left[a_{ij}\right]_n = a_{ij}\,\delta_{ij}$.

For example : $\boldsymbol{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -4 \end{bmatrix}$, is a diagonal matrix of 3×3.

8. **Scalar matrix :** If the elements of a diagonal matrix are all equal, then the matrix is called a scalar matrix. Thus $a_{ii} = x$, $a_{ij} = 0$, $\boldsymbol{A} = \begin{bmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{bmatrix}$, is a scalar matrix of order 3×3.

9. **Unit matrix :** If the elements of a diagonal matrix are all equal to unity, then the matrix is called a unit or identity matrix i.e. $a_{ii} = 1$, $a_{ij} = 0$; $\boldsymbol{A} = \boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, is a unit matrix or order 3 × 3.

10. **Singular matrix** : A square matrix $\boldsymbol{A}$ is a singular matrix, if *det* $\boldsymbol{A}$ = 0. Thus $\boldsymbol{A} = \begin{bmatrix} 2 & 5 & 19 \\ 1 & -2 & -4 \\ -3 & 2 & 0 \end{bmatrix}$**,** is a singular matrix, since *det* $\boldsymbol{A}$ = 0, if $det\ \boldsymbol{A} \neq 0$, then the matrix is called a regular or non-singular matrix.

11. **Determinant of a matrix :** The determinant whose elements are corresponding elements of a square matrix $\boldsymbol{A}$ is called the determinant of matrix $\boldsymbol{A}$ and denoted by det $\boldsymbol{A}$ or $|\boldsymbol{A}|$.

Thus if $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$, then $det\ \boldsymbol{A} = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$. Now we see that *det* $\boldsymbol{A}$ = **0** i.e. $\boldsymbol{A}$ is a singular matrix. Again let $\boldsymbol{A}_1 = \begin{bmatrix} 1 & 4 & 3 \\ 4 & 9 & 6 \\ 5 & 6 & 2 \end{bmatrix}$ and we see $det\ \boldsymbol{A}_1 = \begin{vmatrix} 1 & 4 & 3 \\ 4 & 9 & 6 \\ 5 & 6 & 2 \end{vmatrix} \neq 0$, therefore $\boldsymbol{A}_1$ is a non-singular matrix.

12. **Triangular matrices :** A square matrix in which all the elements below the principle or leading diagonal are zero is called an upper triangular matrix. If however, all the elements above the principle diagonal of a square matrix are zero, then it is called a lower triangular matrix. For example.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -2 & -3 \\ 0 & -4 & 5 & 2 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 7 \end{bmatrix} = \boldsymbol{u}_1$$

Is an upper triangular matrix and $\boldsymbol{B} = \begin{vmatrix} 2 & 0 & 0 \\ -5 & 3 & 0 \\ 2 & 0 & 2 \end{vmatrix} = \boldsymbol{L}_1$, is an lower triangular matrix.

13. **Symmetric matrix :** If a square matrix $\boldsymbol{A} = [a_{ij}]$ such that $\boldsymbol{A^T} = \boldsymbol{A}$ (7.9)

i.e. $[a_{ij}] = [a_{ji}]$ for all i, j. Then the matrix $\boldsymbol{A}$ is called a symmetric matrix. For example

: if $\boldsymbol{A} = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$,

$$\therefore \boldsymbol{A^T} = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \text{ i.e. } \left[a_{ij}\right] = \left[a_{ji}\right] \text{ for all } i,j \ .$$

14. **Skew-symmetric or anti-symmetric matrix :** A square matrix $\boldsymbol{A} = [a_{ij}]$ is called a skew-symmetric or an anti-symmetric matrix if–

i) $a_{ij} = -a_{ji}$ for all values of i, j

ii) $a_{ii} = 0$ i.e. all the leading diagonal elements are zero. Above two properties are satisfied if $\boldsymbol{A^T} = -\mathbf{A}$ (7.10)

for example the matrix : $\begin{bmatrix} 0 & -h & -g \\ h & 0 & -f \\ g & f & 0 \end{bmatrix}$, is a skew-symmetric matrix or anti-symmetric matrix.

Any square matrix can be expressed as the sum of a symmetric and a skew-symmetric matrix in the following manner : $\boldsymbol{A} = \frac{1}{2}\left[\boldsymbol{A} + \boldsymbol{A^T}\right] + \frac{1}{2}\left[\boldsymbol{A} - \boldsymbol{A^T}\right]$ (7.11)

Where first part in R.H.S is a symmetric matrix and the second part is skew-symmetric.

15. **Constant matrix :** If all the non-vanishing elements of a diagonal matrix happen to be equal to each other, it is said to be a constant matrix. The elements of a constant

matrix are thus given by $A=\left[a_{ij}\right]_n = a\delta_{ij} \rightarrow$ (7.12), where = a_{ij} = a = constant for all i, j and a is a scalar when a = 1, we get unit matrix. Thus $\boldsymbol{A} = a\boldsymbol{I}$ (7.13)

This shows that a constant matrix is a constant multiple of the unit-matrix.

16. **Equal matrices :** Two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are said to be equal if and only if a_{ij} = b_{ij} for all values of j. This requires that i) they are of the same order and ii) they have their corresponding element identical. $A = [a_{ij}]$, $B = [b_{ij}]$

7.4 ❒ Complex Matrices

Study of complex matrices with complex elements is useful in quantum mechanics.

17. **Conjugate matrix :** $\boldsymbol{A}$ be a given matrix having complex elements, then the conjugate matrix of $\boldsymbol{A}$, written $\boldsymbol{A}^*$, is the matrix whose corresponding elements are the complex conjugates of the elements of $\boldsymbol{A}$.

That is if $\boldsymbol{A}=\left[a_{ij}\right]_{\boldsymbol{m}\times\boldsymbol{n}}$, then $\boldsymbol{A^*}=\left[a_{ij}{}^*\right]_{\boldsymbol{m}\times\boldsymbol{n}}$. Also if c is any scalar, then,

$$(c\boldsymbol{A})^* = c^*\boldsymbol{A}^* \tag{7.14}$$

For matrix $\boldsymbol{A}$ whose elements are real numbers, the conjugate matrix $\boldsymbol{A}^* = \boldsymbol{A}$.

18. **Hermitian conjugate :** When the two operations of complex conjugation and transposition are carried out one after another on a matrix, the resulting matrix is called the Hermitian conjugate of the original matrix and will be denoted by $\boldsymbol{A}^+$ (called $\boldsymbol{A}$-dagger). The order of the two operation is immaterial, thus $\boldsymbol{A}^+ = (\boldsymbol{A}^*)^{\boldsymbol{T}}$.... (7.15)

For example : $\boldsymbol{A}=\begin{bmatrix} 2+3\boldsymbol{i} & 2-3\boldsymbol{i} \\ 2+\boldsymbol{i} & 3 \end{bmatrix}$

$$\therefore \boldsymbol{A}^+ = \left(\boldsymbol{A}^{\boldsymbol{T}}\right)^* = \left(\boldsymbol{A}^*\right)^T = \begin{bmatrix} 2+3\boldsymbol{i} & 2+\boldsymbol{i} \\ 2-3\boldsymbol{i} & 3 \end{bmatrix}^* = \begin{bmatrix} 2-3\boldsymbol{i} & 2-\boldsymbol{i} \\ 2+3\boldsymbol{i} & 3 \end{bmatrix}$$

19. **Hermitian matrix :** A complex square matrix $\boldsymbol{A} = [a_{ij}]$ is called Hermitian matrix if $(\boldsymbol{A}^*)^T = \boldsymbol{A}$ or $\boldsymbol{A}^+ = \boldsymbol{A}$.

Example : $A=\begin{bmatrix} 0 & -\boldsymbol{i} \\ \boldsymbol{i} & 0 \end{bmatrix}$ $\therefore A^*=\begin{bmatrix} 0 & \boldsymbol{i} \\ -\boldsymbol{i} & 0 \end{bmatrix}$

$$\therefore (\boldsymbol{A}^*)^T = \begin{bmatrix} 0 & -\boldsymbol{i} \\ \boldsymbol{i} & 0 \end{bmatrix} = \boldsymbol{A}$$

Or, $\boldsymbol{A}^+ = \boldsymbol{A}$ i.e. $\boldsymbol{A}$ is Hermitian.

Every diagonal element of a Hermitian matrix must be real ; since $\boldsymbol{A}^+ = \boldsymbol{A}$, or, $[a_{ii}]^+ = [a_{ii}]$

Or, $(a_{ii}{}^*)^T = [a_{ii}]$ or, $(a_{ii})^* = a_{ii}$

$\therefore$ a_{ii} is real.

20. **Symmetric and Hermitian matrix**

We consider equality of $\boldsymbol{A}$, $\boldsymbol{A}^T$ and $\boldsymbol{A}^+$. The equality of $\boldsymbol{A}$ and $\boldsymbol{A}^T$ or of $\boldsymbol{A}$ and $\boldsymbol{A}^+$ will be defined only if $\boldsymbol{A}$ is a square matrix with $m = n$. We then get the following four special matrices e.g. symmetric, anti-symmetric, Hermitian, anti-Hermitian. Symmetric : $\boldsymbol{A} = \boldsymbol{A}^T$; Hermitian : $\boldsymbol{A} = \boldsymbol{A}^+$; Anti-symmetric : $\boldsymbol{A} = -\boldsymbol{A}^T$; Anti-Hermitian (skew-Hermitian) : $\boldsymbol{A} = -\boldsymbol{A}^+$.

21. **Orthogonal matrix** : A square matrix $\boldsymbol{A}$ is called orthgonal if $\boldsymbol{AA}^T = \boldsymbol{I}$, where $\boldsymbol{I}$ is an identity or unit matrix. Since $\boldsymbol{AA}^T = \boldsymbol{I}$, $\boldsymbol{A}^T = +\boldsymbol{A}^{-1}$ i.e. if transpose of $\boldsymbol{A}$ is equal to its inverse.

The orthogonality condition of two square matrices is that the determinant.

$$|\boldsymbol{AA}^T| = \boldsymbol{I}$$

Example : $\boldsymbol{A} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$

A matrix $\boldsymbol{A}$ satisfying the relations $\boldsymbol{AA}^T = \boldsymbol{I}_n$ (7.16a)

$$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}_m \quad (7.16b)$$

Where $\boldsymbol{I}_n$ and $\boldsymbol{I}_m$ are two unit matrices, not necessarily of the same order, is called an orthogonal matrix. Again it can be shown that if $\boldsymbol{A}$ is a finite matrix satisfying both equations (7.16), then $\boldsymbol{A}$ must be square and $\boldsymbol{AA}^T = \boldsymbol{I}$, $\Rightarrow$ $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$.

Let *det* $\boldsymbol{A} = d$. Taking the determinants of both sides of equation 7.16a, we have $d^2 = 1 \Rightarrow d = \pm 1$. This shows that the determinant of an orthogonal matrix can only have values +1 *or* –1. At the same time this shows that $\boldsymbol{A}$ is non-singular, so that $\boldsymbol{A}^{-1}$ exists. Multiplying equation 7.16a by $\boldsymbol{A}^{-1}$ from the right, we have $\boldsymbol{A}^{-1} = \boldsymbol{A}^{T}$.

22. **Unitary matrix :** A complex square matrix $\boldsymbol{A}$ is said to be unitary [not unit or identity matrix], if $\boldsymbol{A}^{-1} = \boldsymbol{A}^{+}$. Therefore $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}\boldsymbol{A}^{+} = \boldsymbol{I}$

So, if the product of the matrix and its Hermitian conjugate is an identity matrix, it is a unitary matrix.

7.5 ❒ Matrix Algebra

Matrix algebra is different from ordinary algebra in as much as vector algebra is different from a scalar algebra. We ordinarily indicate a matrix by a bold face letter like $\boldsymbol{A}$ or $\boldsymbol{B}$ etc. but the later does not have a numerical values, it simply stands for the array.

The various operation of addition, subtraction, multiplication etc. on matrices are called its algebra.

7.5.1 Addition of Matrices : If $\boldsymbol{A}$ and $\boldsymbol{B}$ be two matrices of the same order then their sum $\boldsymbol{A} + \boldsymbol{B}$ is defined as the matrix whose elements are obtained by adding the corresponding elements of $\boldsymbol{A}$ and $\boldsymbol{B}$.

If $\boldsymbol{A} = [a_{ij}]$ and $\boldsymbol{B} = [b_{ij}]$, then $\boldsymbol{A} + \boldsymbol{B} = [a_{ij} + b_{ij}]$. (7.17)

Therefore the sum of two matrices, each of order $m \times n$, is a matrix of the same order $m \times n$ with each element being the sum of the corresponding elements of the given matrices.

Thus it is evident that matrices are useful in case which are added by components, for example, vectors. To explain, suppose,

$$\boldsymbol{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ d_1 & e_1 & f_1 \end{pmatrix} \text{ and } \boldsymbol{B} = \begin{pmatrix} a_2 & b_2 & c_2 \\ d_2 & e_2 & f_2 \end{pmatrix}$$

$$\text{and } \boldsymbol{A} + \boldsymbol{B} = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 & c_1 + c_2 \\ d_1 + d_2 & e_1 + e_2 & f_1 + f_2 \end{pmatrix}$$

Suppose the column in ***A*** and ***B*** represent the displacement of three particles. Then first particle has displacement $ia_1 + jd_1$ in ***A***, the first particle has displacement at a later time $ia_2 + jd_2$ ***B***.

Then the total displacement of the first particle is $i(a_1 + a_2) + j(d_1 + d_2)$ is the first column of matrix ***A*** + ***B***. Similarly the second and third columns represent displacement of the second and third particles.

Properties of matrix addition :

1> Matrix addition is commutative : $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$

2> Matrix addition is associative : $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$

Two matrices must be conformable for addition,; by conformity we mean both the matrices must have the same number of rows and the same number of columns i.e. they are of the same order.

7.5.2 : Scalar multipliction of a matrix :

If a matrix ***A*** is multiplied by a scalar quantity *k*, the each element of ***A*** gets multiplied by *k*. example : if $\boldsymbol{A} = [a_{ij}]$, then $k\boldsymbol{A} = [ka_{ij}]$.

In determinants however only one row or one column is multiplied and not every elements of the determinant.

It is known that the components of a vector $\boldsymbol{A} = a\boldsymbol{i} + b\boldsymbol{j}$ may be conveniently written as elements of a matrix either $\boldsymbol{A} = \begin{pmatrix} a \\ b \end{pmatrix}$, called a column matrix or a column vector or $\boldsymbol{A}^{\boldsymbol{T}} = (a\ b)$ called a row matrix or row vector.

The row matrix $\boldsymbol{A}^{\boldsymbol{T}}$ is the transpose of the matrix ***A***. Now a vector twice the length of ***A*** is $2a\boldsymbol{i} + 2b\boldsymbol{j} = 2\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2a \\ 2b \end{pmatrix}$ & $2\boldsymbol{A}^{\boldsymbol{T}} = (2a\ \ 2b)$

7.5.3 : Matrix Multiplication :

Let $\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *and* $\boldsymbol{B} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$

Then $\boldsymbol{AB} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix} = \boldsymbol{C}$

Thus all the elements of ***C*** may be obtained by using the following simple rule. The element in row *m* and column *n* of the product matrix ***AB*** is equal to row *m* of ***A*** times column *n* of ***B.*** In notation

$$(\boldsymbol{AB})_{mn} = \sum_k \boldsymbol{A_{mk}B_{kn}} \tag{7.18}$$

We consider the elements in a row or a column of a matrix as the components of a vector. Then row times column multiplication for the matrix product ***AB*** corresponds to finding the dot product of the row vector of ***A*** and a column vector of ***B*** e.g. $(a\boldsymbol{i} + b\boldsymbol{j}).(e\boldsymbol{i} + g\boldsymbol{j}) = ae + bg$

The product ***AB*** (maintaining the order) can be found if and only if the number of elements in a row of ***A*** equals the number of elements in a column of ***B***: the matrices ***A,B*** in that order are then called conformable. **The number of rows in *A* and of column in *B* have nothing to do with the question of whether we can find *AB* or not.** Matrix multiplication is not commutative i.e. in general, matrices do not commute under multiplication.

We define the commutator of the matrices ***A*** and ***B*** by $[\boldsymbol{A}, \boldsymbol{B}] = \boldsymbol{AB} - \boldsymbol{BA}$ = commutator of ***A*** and ***B*** (7.19)

If $[\boldsymbol{A}, \boldsymbol{B}] = 0$, then ***A*** and ***B*** commutes (in some special matrices).

Commutators are widely used in classical and quantum mechanics.

Properties of matrix multiplication :

1. Non-commutative : $\boldsymbol{AB} \neq \boldsymbol{BA}$ in general
2. Associative : $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{ABC}$
3. Distributive : $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$ and $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$
4. $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$, where ***I*** is a unit matrix conformable with ***A***.
5. $\boldsymbol{AA}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$ provided $|\boldsymbol{A}| \neq 0$
6. $[\boldsymbol{AB}]^T = \mathbf{B}^\mathrm{T}\mathbf{A}^\mathrm{T}$
7. If $\boldsymbol{A}.\boldsymbol{B} = 0$, it does not mean necessarily that ***A*** or ***B*** is a null matrix.
8. If ***A*** and ***B*** are two square matrix of the same order then $det\,\boldsymbol{AB} = det\boldsymbol{BA}$ = (det ***A***) (*det* ***B***) even when ***A*** and ***B*** do not commute.

7.5.4 : Adjoint and inverse of a matrix : The cofactor of an element in a square matrix ***A*** means exactly the same thing as the cofactor of that element in *det* ***A*** since if

$$\boldsymbol{A} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}, \ |\boldsymbol{A}| = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

The matrix obtained from the cofactor of $|\boldsymbol{A}|$ is given by ***C***, where

$$\boldsymbol{C} = \begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{bmatrix}$$

The transpose of the matrix of the co-factor is

$$\begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix} = \text{adjoint of matrix } \boldsymbol{A} = \text{adj } \boldsymbol{A} = \boldsymbol{C}^T, \qquad (7.20)$$

Now we define the inverse of a matrix ***A*** as the matrix $\boldsymbol{A}^{-1}$ such that $\boldsymbol{A}\boldsymbol{A}^{-1}$ and $\boldsymbol{A}^{-1}\boldsymbol{A}$ are both equal to unit matrix ***I***. It is to be noted that only square matrices can have inverse and acatually some square matrices do not have inverse either.

Now if $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$, then (det $\boldsymbol{A}^{-1}$) (det ***A***) = det ***I*** = 1. If two numbers have product equal to one, then neither of them is zero. Thus det $\boldsymbol{A} \neq 0$ is a requirement for ***A*** to have an inverse.

Thus the condition for a square matrix ***A*** to have an inverse is that ***A*** is non-singular i.e. $|\boldsymbol{A}| \neq 0$.

If a matrix has an inverse, it is called an invertible and if it does not have an inverse, it is called singular.

If **A** be an invertible matrix, then the inverse of matrix ***A*** is given by,

$$\boldsymbol{A}^{-1} = \frac{adj\ \boldsymbol{A}}{|\boldsymbol{A}|} = \frac{\boldsymbol{C}^T}{|\boldsymbol{A}|} \qquad (7.21)$$

7.5.5 : Properties of inverse of a matrix :

1. Inverse of matrix is unique.
2. Every matrix commutes with its inverse, i.e. $\boldsymbol{AA}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$
3. Inverse of the product of a number of matrices (all square and of the same order), none of which is singular, equals the product of the inverses taken in the reverse order i.e. $(\boldsymbol{ABC})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$
4. If **A** be an invertible matrix, then $\boldsymbol{A}^{-1}$ is invertible and $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.
5. Inverse of the transpose of a square matrix is the transpose of its inverse, i.e. $[\boldsymbol{A}^{-1}]^{T} = [\boldsymbol{A}^{T}]^{-1}$

7.5.6 : Properties of orthogonal matrix : [see Art 7.2.2., item 21]

1. Every orthogonal matrix is non-singular i.e. if $\boldsymbol{A}$ is an orthogonal matrix, $|\boldsymbol{A}| \neq 0$
2. Unit matrix is an orthogonal matrix.
3. If $\boldsymbol{A}$ is an orthogonal matrix, then its determinant, $|\boldsymbol{A}| = \pm 1$
4. The product of two orthogonal matrices is also orthogonal.
5. The transpose of a orthogonal matrix is also orthogonal.
6. The inverse of an orthogonal matrix is also orthogonal.

7.5.7 : Properties of unitary matrix : [see 7.4 : item no 22]

A matrix $\boldsymbol{U}$ satisfying the relations, $\boldsymbol{UU}^{+} = \boldsymbol{I}_n \rightarrow$ (7.22a)

$$\boldsymbol{U}^{+}\boldsymbol{U} = \boldsymbol{I}_m \rightarrow \quad (7.22b)$$

is called a unitary matrix. If $\boldsymbol{U}$ is a finite matrix satisfying both equations (7.22), then $\boldsymbol{U}$ must be a square matrix, and $\boldsymbol{UU}^{+} = \boldsymbol{I} \Rightarrow \boldsymbol{U}^{+}\boldsymbol{U} = \boldsymbol{I}$. The elements of a unitary matrix may be complex. In fact it is evident from equations (7.22) that a real unitary matrix is orthogonal.

Let det $\boldsymbol{U} = d$. Taking the determinants of both sides of equation (7.22a) and noting that $\boldsymbol{U}^{+} = \boldsymbol{d}^{*}$, we have $\boldsymbol{dd}^{*} = 1 \Rightarrow |\boldsymbol{d}| = 1$

This shows that the determinant of a unitary matrix can be a complex number of unit magnitude, i.e. a number of the form $e^{i\theta}$, where θ is real. It also shows that a unitary matrix is non-singular and possess an inverse properties :

1) The inverse of a unitary matrix is unitary.
2) The Hermitian conjugate of a unitary matrix is its inverse i.e. $\boldsymbol{U}^{+} = \boldsymbol{U}^{-1}$
3) The product of two unitary matrices is also unitary.
4) A unitary matrix with elements as real numbers is orthogonal.

7.5.8 : Trace of matrix :

The trace of a square matrix $\boldsymbol{A} = [a_{ij}]$ is defined as the sum of its diagonal elements. It is also called spur or the diagonal sum and is denoted by $T_r\boldsymbol{A}$ or $S_p\boldsymbol{A}$. Thus,

$$T_r\boldsymbol{A} = \sum_{i=1}^{n} a_{ii} \tag{7.23}$$

Properties :

1) The trace of sum (or difference) of two matrices is the sum (or difference) of their traces.
2) The trace of the product of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is independent of the order multiplication i.e. $T_r(\boldsymbol{AB}) = T_r(\boldsymbol{BA})$.

   This property is true even when $\boldsymbol{AB} \neq \boldsymbol{BA}$ and the above equation implies that the trace of any commutator $[\boldsymbol{A}, \boldsymbol{B}]$ is zero.
3) The above equation also gives, $T_r(\boldsymbol{ABC}) = T_r(\boldsymbol{BCA}) = T_r(\boldsymbol{CAB})$

   i.e. the trace is the invariant under cyclic permutation of matrices in a product. It is important to note that trace of a number of matrices is not invariant under any permutation, but only under a cyclic permutation of the matrices.

7.5.9 : Rank of a matrix :

An integral number r is said to be the Rank of a matrix $\boldsymbol{A}$; if,

i. There is at least one square sub-matrix of $\boldsymbol{A}$ of order r whose determinant is non-zero.

ii. All the square sub-matrices of $\boldsymbol{A}$ of order $(r + 1)$, have determinants zero.

Generally speaking the rank r of a matrix is the largest order of any non-vanishing minor of the matrix.

Example : Let $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$ be a matrix of 3rd order and $|\boldsymbol{A}| = 0$. Since $|\boldsymbol{A}|$ is zero, the rank of A is not 3. But, there is at least one 2 × 2 sub-matrix $\begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix}$ whose determinant is not zero. In fact none of the minors is zero. So, the rank of $\boldsymbol{A}$ i.e. $r(\boldsymbol{A}) = 2$.

We observe the following :

1) The rank of a non-singular square matrix of order n is n and that of a singular square matrix of order n is less than n.
2) The rank of a null matrix is obviously zero.
3) The rank of the transpose of a matrix $\boldsymbol{A}$ is the same as the rank $\boldsymbol{A}$.
4) The rank of product of two matrices never exceeds the rank of either matrix.
5) The rank of a matrix is not altered, (i) If interchange of rows is made, (ii) If the elements of any row are multiplied by a non-zero number, (iii) If λ times the elements of a row is added to corresponding elements of another row, λ being any number, (iv) If the matrix is pre or post multiplied by a non-singular matrix.

Sub-matrices and rank : Let $\boldsymbol{A}$ be a matrix of order $m \times n$. Any matrix obtained from $\boldsymbol{A}$ by omitting some of its rows or columns is called a sub-matrix $\boldsymbol{A}$. When a matrix is partitioned into a number of blocks, each block is a sub-matrix of the original matrix.

If a matrix $\boldsymbol{A}$ has at least one square non-singular sub-matrix of order r but every square sub-matrix of $\boldsymbol{A}$ of order greater than r is singular, then r is said to be the rank of the matrix $\boldsymbol{A}$. The rank of the matrix $\boldsymbol{A}$ given below is 3 because it has non-singular square sub-matrices of order 3 but not higher than 3.

$$\boldsymbol{A} = \begin{bmatrix} 3 & 5 & 9 & 1 \\ 2 & 0 & -1 & 3 \\ -3 & -6 & 1 & 2 \end{bmatrix}$$

Example : Find the rank of the matrix :

$$\begin{bmatrix} 1 & 0 & 2 & -4 \\ 3 & -1 & 1 & 2 \\ 5 & -1 & 5 & -6 \\ 4 & -1 & 3 & -2 \end{bmatrix}$$

Solution : We note that the determinant of the given matrix is zero. So that its rank is less than 4. Next we consider all possible sub-matrices of order 3 (there are 16 of them) and observe that all of them are also singular, so that rank is less than 3. Finally, we note that the given matrix has non-singular sub-matrices of order 2, so that its rank is 2.

A simple method to find the rank of a matrix is to find the maximum number of linearly independent row vectors or column vectors of the matrix. This maximum number is the rank of the matrix.

7.5.10 : Normal form :

A matrix which commutes with its own Hermitian conjugate is said to be a normal matrix or in normal form. Thus if matrix $\boldsymbol{A}$ is in normal form if and only if

$$[\boldsymbol{A}, \boldsymbol{A}^{+}] = 0 \tag{7.24}$$

It can be easily seen that symmetric, antisymmetric, Hermitian and antihermitian matrices are also normal matrices.

For this reason they often occurs in physics.

Properties :

1) The inner product of the i^{th} and j^{th} rows of the normal matrix equals the inner product of the i^{th} and j^{th} columns.
2) The norm of the i^{th} row of a normal matrix equals that of the i^{th} column.

7.6 ❑ Characteristic Equation of a Square Matrix : Eigenvalues and Eigen vectors of Matries

Let $\boldsymbol{A}$ be a square matrix of order n and $\boldsymbol{X}$ a non-zero column vector. If there exists a scalar λ such that $\boldsymbol{AX} = \lambda\boldsymbol{X}$ (7.25)

Then the vector $\boldsymbol{X}$ is defined as an eigenvector and λ is defined as an eigenvalue corresponding to the eigenvector $\boldsymbol{X}$. Equation (7.25) is called eigenvalue equation and may be written as

$$\boldsymbol{AX} = \lambda \boldsymbol{IX} = (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{X} = 0 \tag{7.26}$$

Characteristic matrix : For a given square matrix $\boldsymbol{A}$, the matrix $(\boldsymbol{A} - \lambda \boldsymbol{I})$ is called the Characteristic matrix of $\boldsymbol{A}$, where λ is a scalar parameter and $\boldsymbol{I}$ the unit matrix of the same order.

Characteristic polynomial : The determinant $|\boldsymbol{A} - \lambda \boldsymbol{I}|$, an expansion will give rise to a polynomial and is known as the characteristic polynomial of matrix $\boldsymbol{A}$.

Characteristic equation : The equation $|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$ is known as the characteristic equation of matrix $\boldsymbol{A}$ determines the eigenvalues of the matrix $\boldsymbol{A}$.

Eigenvectors or characteristic vectors : For each eigenvalues λ, we have a non-zero column vector $\boldsymbol{X}$ that satisfies the equation $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{X} = 0$ (7.27)

The non-zero vector $\boldsymbol{X}$ is known as the eigenvector or the characteristic vector.

Orthogonal vectors : Two vectors $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are said to be orthogonal vectors if the condition $\boldsymbol{X}_1^+ \boldsymbol{X}_2 = 0$ is satisfied. Let $\boldsymbol{X}_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ *and* $\boldsymbol{X}_2 = \begin{bmatrix} x_4 \\ x_5 \\ x_6 \end{bmatrix}$

$$\therefore\ \boldsymbol{X}_1^+ \boldsymbol{X}_2 = \left(x_1 x_2 x_3\right) \begin{pmatrix} x_4 \\ x_5 \\ x_6 \end{pmatrix} = x_1 x_4 + x_2 x_5 + x_3 x_6 = 0$$

Normalised form of vectors : With the condition $\boldsymbol{X}_1^+ \boldsymbol{X}_1 = 1$ we can find out normalised form of vector $\boldsymbol{X}_1$, similarly from the condition $\boldsymbol{X}_2^+ \boldsymbol{X}_2 = 1$, we can find out nromalised form of $\boldsymbol{X}_2$.

7.6.1 : Some theoretical aspects of Eigenvalues and Eigenvectors of matrix

Theorem I :

The determinant of matrix $\boldsymbol{A}$ is the products of its eigenvalues.

Proof : We have $|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$ (7.28)

Equation (7.28) can be expanded as

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & & \dots \\ a_{n1} & a_{2n} & \dots & a_{nn}-\lambda \end{vmatrix} = 0 \tag{7.29}$$

or, $$(-1)^n \left(\lambda^n + \alpha_1 \lambda^{n-1} + \dots + \alpha_n\right) = 0 \tag{7.30}$$

Where $\alpha_1, \alpha_2 \dots \alpha_n$ are constants expressed in terms of the coefficient a_{ij} of the matrix $\boldsymbol{A}$ given by equation (7.7). (see Art 7.2.1)

The identity $$|\boldsymbol{A}-\lambda \boldsymbol{I}| = (-1)^n \left(\lambda^n + \alpha_1 \lambda^{n-1} + \dots + \alpha_n\right) \tag{7.31}$$

If $\lambda_1, \lambda_2 \dots \lambda_n$ are the eigenvalues of $\boldsymbol{A}$, the roots of the right hand side polynomial of (7.31) will be $\lambda_1, \lambda_2 \dots \lambda_n$

Therefore $$|\boldsymbol{A}-\lambda \boldsymbol{I}| = (\lambda_1-\lambda)(\lambda_2-\lambda)\dots(\lambda_n-\lambda) \tag{7.32}$$

Putting $\lambda = 0$ on both sides

$$|\boldsymbol{A}| = \lambda_1 \lambda_2 \dots \lambda_n \tag{7.33}$$

Hence theorem is proved. Also if any of the eigenvalues of $\boldsymbol{A}$ is zero, then $|\boldsymbol{A}| = 0$ i.e. the matrix A is singular.

Theorem II : If $\boldsymbol{A}$ be a square matrix, then its trace is the sum of its eigen values.

Proof : If $\boldsymbol{A}$ be a square matrix of order *n*, then $T_r(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$

Now if $\lambda_1, \lambda_2 \dots \lambda_n$ be eigenvalues of $\boldsymbol{A}$, then we have the identity

$$|\boldsymbol{A}-\lambda \boldsymbol{I}| = (-1)^n \left(\lambda^n + \alpha_1 \lambda^{n-1} + \dots + \alpha_n\right) \tag{7.34}$$

The coefficient of λ^{n-1} in L.H.S is

$$(-1)^{(n-1)}\sum_{i=i}^{n} a_{ii} \tag{7.35}$$

and the coefficient of λ^{n-1} in R.H.S. is

$$(-1)^{(n+1)}\sum_{i=i}^{n} \lambda_i \tag{7.36}$$

Equating these quantities, we have

$$(-1)^{(n-1)}\sum_{i=i}^{n} a_{ii} = (-1)^{(n+1)}\sum_{i=i}^{n} \lambda_i$$

Therefore

$$\sum_{i=1}^{n} a_{ii} = \sum_{i=i}^{n} \lambda_i$$

Thus the theorem is proved.

7.7 ❒ Diagonalization

We have a square matrix ***A*** of order *n* which we have to diagonalize. We now construct a matrix ***P*** of order *n* whose columns are the eigen vectors of the given matrix ***A***. Since the eigenvectors are linearly independent, ***P*** is non-singular. Therefore $\boldsymbol{P}^{-1}$ exists. Now we state that the matrix $\boldsymbol{P}^{-1}\boldsymbol{AP}$ is diagonal whose diagonal element are the eigenvalues of ***A***.

Example 1 : Consider the matrix $\boldsymbol{A} = \begin{pmatrix} 3 & i \\ -i & 3 \end{pmatrix}$. Find the eigenvalues and eigen vectors and hence construct the unitary matrix ***U*** that diagonalize ***A***. Also compute $\boldsymbol{U}^{-1}\boldsymbol{AU}$.

Solution : The characteristic equation of matrix ***A*** is

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} 3-\lambda & i \\ -i & 3-\lambda \end{vmatrix} = 0$$

Or, $(\lambda - 2)(\lambda - 4) = 0$

$\therefore$ the eigen values are $\lambda_1 = 2, \lambda_2 = 4$

for $\lambda_1 = 2$; $\boldsymbol{AX} = 2\boldsymbol{X}, \boldsymbol{X} = \begin{pmatrix} x \\ y \end{pmatrix}$ so that

$$\begin{pmatrix} 3 & i \\ -i & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix} = 2\boldsymbol{I}\begin{pmatrix} x \\ y \end{pmatrix}$$

or, $$\left[\begin{pmatrix} 3 & i \\ -i & 3 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\right]\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

or, $$\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

or, $$\left.\begin{aligned} x + iy &= 0 \\ -ix + y &= 0 \end{aligned}\right\}$$

Both the equations are equivalent. Choosing $x = 1$, the convenient eigenvector $\boldsymbol{X}_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$.

Similarly for $\lambda_2 = 4$, the eigenvector is $\boldsymbol{X}_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$

Normalised eigenvectors are $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$ and $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$

Therefore $\boldsymbol{U} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$

$$\boldsymbol{U}^{+} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

$\therefore$ $\boldsymbol{UU}^{+} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{I}$ thus $\boldsymbol{U}$ is unitary.

now $|\boldsymbol{U}| = -\sqrt{2}i \neq 0$

$$\therefore \boldsymbol{U}^{-1} = -\frac{1}{\sqrt{2i}}\begin{pmatrix} -i & -1 \\ -i & 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ 1 & -i \end{pmatrix}$$

$$\therefore \boldsymbol{D} = \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}\begin{pmatrix} 3 & i \\ -i & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = \begin{pmatrix} 1 & -i \\ 2 & 2i \end{pmatrix}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

Therefore $\boldsymbol{D}$ is a diagonal matrix whose diagonal elements are the eigenvalues of $\boldsymbol{A}$.

Example 2 : A square matrix $\boldsymbol{A} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ is given. Find the eigenvalues and eigenvectors of the matrix $\boldsymbol{A}$. Construct an appropriate matrix which will diagonalize $\boldsymbol{A}$ and find the diagonal matrix.

Solution : Characteristic equation :

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0 \Rightarrow \begin{vmatrix} -\lambda & -i \\ i & -\lambda \end{vmatrix} = 0$$

Therefore eigenvalues are $\lambda = \pm 1$

Using the equation : $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{X} = 0$, we get the eigenvectors $\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

For $\lambda = 1$, $\begin{pmatrix} -1 & -i \\ i & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$

or, $-x_1 - ix_2 = 0,\ ix_1 - x_2 = 0$

These equations are not independent so that x_1 and x_2 are not unique and infinitely many solutions can be obtained. If we take $x_1 = 1$ we get $x_2 = i$

Therefore, eigenvector corresponding to $\lambda = 1$ is $\boldsymbol{X}_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$ and that corresponding to $\lambda = -1$ is $\boldsymbol{X}_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$

Now we construct a 2 × 2 matrix $\boldsymbol{P}$ with the column vectors $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$

Therefore $\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$

Therefore $\boldsymbol{P}^{-1} = \frac{1}{|\boldsymbol{P}|} Adj\ \boldsymbol{P} = \frac{1}{-2i} Adj\ \boldsymbol{P} = -\frac{1}{-2i}\boldsymbol{C}^T$

Where $\boldsymbol{C}$ is the cofactor matrix of $\boldsymbol{A}$ and is given by $\boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$

Where $c_{11} = +(-i); c_{12} = -(i), c_{21} = -(1), c_{22} = +(1)$

$$\therefore \boldsymbol{C} = \begin{pmatrix} -i & -i \\ -1 & 1 \end{pmatrix}$$

Therefore $\boldsymbol{C}^T = \begin{pmatrix} -i & -i \\ -1 & 1 \end{pmatrix}^T = \begin{pmatrix} -i & -1 \\ -i & 1 \end{pmatrix}$

Therefore $\boldsymbol{P}^{-1} = -\frac{1}{2i}\begin{pmatrix} -i & -1 \\ -i & 1 \end{pmatrix}$

Now, $\boldsymbol{AP} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ i & i \end{pmatrix}$

$$\therefore \boldsymbol{P}^{-1}\boldsymbol{AP} = -\frac{1}{2i}\begin{pmatrix} -i & -1 \\ -i & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ i & i \end{pmatrix} = \frac{1}{2i}\begin{pmatrix} -2i & 0 \\ 0 & 2i \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \boldsymbol{D}$$

where $\boldsymbol{D}$ is the diagonal matrix with diagonal elements as the eigen values of $\boldsymbol{A}$.

Thus we can state the following theorem.

If a matrix of order n has n linearly independent eigenvectors, then it is related through a similarity transformation to a diagonal matrix whose diagonal elements are the eigenvalues of the matrix.

However the matrix $\boldsymbol{P}$ is not a unique matrix, because we could arrange the eigenvectors $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ in any order in the construction of $\mathbf{P}$ *i.e.* we could form $\boldsymbol{P}$ in example 2 as

$$\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$$

We can therefore state the general rule that in the process of Diagonalisation $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{D}$, the order of the eigenvalues in $\boldsymbol{D}$ corresponds to the order of the eigenvectors of $\boldsymbol{A}$ in constructing $\boldsymbol{P}$. The matrices $\boldsymbol{A}$ and $\boldsymbol{D}$ are said to be related by a simiarity transformation. The inverse transformation $\boldsymbol{A} = \boldsymbol{PDP}^{-1}$ is also similarity transformations.

7.8 ❒ Solutions of systems of linear homogenous and non-homogenous equations : An application of theory of matrices

7.8.1 : We consider a set of m non-homogeneous linear equations in n unknowns : (m < n)

$$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2$$

$$... \; ... \; ... \; ... \; ... \; ... \; ... \; ... \; ... \; ... \; ...$$

$$a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m$$

Which can be represented as,

$$\begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

or, $\quad \boldsymbol{AX} = \boldsymbol{B}$ (7.37)

and $\quad A_b = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} & b_1 \\ a_{21} & a_{22} & ... & a_{2n} & b_2 \\ ... & ... & ... & ... & \vdots \\ a_{m1} & a_{m2} & ... & a_{mm} & b_m \end{bmatrix}$

$\boldsymbol{A}_b = [\boldsymbol{A}, \boldsymbol{B}]$ is called the augmented matrix of order $n \times (m + 1)$

Now,

i) The equations are consistent, if rank $\boldsymbol{A}$ = rank $\boldsymbol{A}_b$

ii) The solutions are unique, if rank $\boldsymbol{A}$ = rank $\boldsymbol{A}_b = n$, where n= numbers of unknowns.

iii) Infinite solutions, if rank $\boldsymbol{A}$ = rank $\boldsymbol{A}_b = r$, $r < n$

iv) The equations are inconsistent, if rank $\boldsymbol{A} \neq$ rank $\boldsymbol{A}_b$

7.8.2 : Solutions of homogeneous equations : ($m = n$)

For homogeneous system, equatiosn (7.37) can be written as $\boldsymbol{AX} = 0$ (7.38)

If $|\boldsymbol{A}| \neq 0$ then $\boldsymbol{A}^{-1}$ exists

Pre-multiplying both sides by $\boldsymbol{A}^{-1}$, we get

$$\boldsymbol{A}^{-1}\,\boldsymbol{AX} = \boldsymbol{A}^{-1}\mathbf{0} \text{ or, } \boldsymbol{IX} = \mathbf{0}$$

Which shows that $\boldsymbol{X} = \mathbf{0}$ *i.e* $x_1 = x_2 = \ldots = x_n = 0$

This is called the trivial solution.

If however the matrix $\boldsymbol{A}$ is singular i.e. $|\boldsymbol{A}| = 0$, then the equations under considerations will have infinite solutions where some solutions may be non-zero (non-trivial solution).

7.8.3 : Solutions for non-homogeneous system of equations : (m = n)

A system of non-homogeneous equations is represented by : $\boldsymbol{AX} = \boldsymbol{B}$ (7.39)

Now if $|\boldsymbol{A}| \neq 0$, $\boldsymbol{A}^{-1}$ exists. Therefore $\boldsymbol{A}^{-1}\,\boldsymbol{AX} = \boldsymbol{A}^{-1}\boldsymbol{B}$ or, $\boldsymbol{IX} = \boldsymbol{A}^{-1}\,\boldsymbol{B}$

$$\therefore\ \boldsymbol{X} = \boldsymbol{A}^{-1}\,\boldsymbol{B} \tag{7.40}$$

Thus finding the value of $\boldsymbol{A}^{-1}$, we can find out the solution using equation (7.40).

This method of finding out the solution however fails if $\boldsymbol{A}$ is singular. However the solution given by equation (7.40) is unique.

7.9 ❒ Solutions of Coupled Linear Ordinary Differentical Equations in Terms of Eigenvalue Problems

We want to reduce a system of coupled ordinary differential equation to an eigenvalue problems.

We exemplified the process by a specific problems of discussing the vibrations of the two coupled springs shown in fig (7.1)

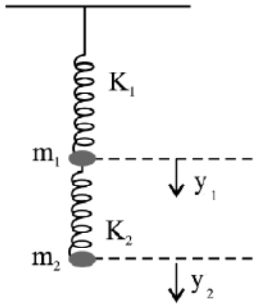


Fig. (7.1)

In the figure y_1, y_2 are the displacements of the two masses.

The equations of motion for the coupled vibrations can be written as

$$\left.\begin{aligned} m_1\ \ddot{y}_1 &= -k_1 y_1 + k_2\left(y_2 - y_1\right) \\ m_2\ \ddot{y}_2 &= -k_2\left(y_2 - y_1\right) \end{aligned}\right\} \tag{7.41}$$

where the dots denote time derivatives.

We now define

$$\left.\begin{aligned} \boldsymbol{y} &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ \textit{and}\ \boldsymbol{A} &= \begin{bmatrix} -\frac{(k_1+k_2)}{m_1} & \frac{k_2}{m_1} \\ \frac{k_2}{m_2} & -\frac{k_2}{m_2} \end{bmatrix} \end{aligned}\right\} \tag{7.42}$$

Equation (7.41) can be written as $\ddot{\boldsymbol{y}} = \boldsymbol{A}\boldsymbol{y}$ (7.43)

With the trial solution, $\boldsymbol{y} = \boldsymbol{x}e^{\omega t} \rightarrow$ (7.44) we get from equation (7.43)

We get $\boldsymbol{x}\omega^2 e^{\omega t} = \boldsymbol{A}\boldsymbol{x}e^{\omega t}$

$$\boldsymbol{A}\boldsymbol{x} = \omega^2 \boldsymbol{x} \tag{7.45}$$

Equation (7.45) is the eigenvalues equation ω^2 being the eigenvalues corresponding to the matrix $\boldsymbol{A}$ with the eigenvector $\boldsymbol{x}$. Thus ω^2 gives the frequency of oscillation and eigenvector x gives the displacement equation of the system.

Now suppose $m_1 = 3$, $m_2 = 2$, $k_1 = 18$, $k_2 = 6$

Therefore $\boldsymbol{A} = \begin{vmatrix} -8 & 2 \\ 3 & -3 \end{vmatrix}$ (7.46)

The eigen values and eigenvectors of this matrix are found to be,

$$\omega^2 = -9,\ x_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix};\quad \omega^2 = -2,\ x_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

The most general solution of the system is therefore,

$$\boldsymbol{y}(\boldsymbol{t}) = b_1 \boldsymbol{x}_1 e^{3it} + b_2 \boldsymbol{x}_1 e^{-3it} + b_3 \boldsymbol{x}_2 e^{\sqrt{2}it} + b_4 \boldsymbol{x}_2 e^{-\sqrt{2}it}$$

or, $$\boldsymbol{y}(\boldsymbol{t}) = a_1 \boldsymbol{x}_1 \cos 3t + a_2 \boldsymbol{x}_1 \sin 3t + a_3 \boldsymbol{x}_2 \cos\sqrt{2}t + a_4 \boldsymbol{x}_2 \sin\sqrt{2}t \qquad (7.47)$$

where a and b are the arbitrary constants.

7.10 ❒ Functions of a Matrix

There are two methods by which a function of a matrix can be evaluated

1) Functions of a diagonalizable matrix
2) Functions of any matrix based on the existence of a minimal polynomial.

7.10.1 : Functions of a diagonalizable matrix

We have $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{D}$, $\boldsymbol{A} = \boldsymbol{PDP}^{-1}$ (7.48)

Where $\boldsymbol{A}$ be a diagonalizable square matrix, $\boldsymbol{P}$ be the diagonalising matrix for $\boldsymbol{A}$ and $\boldsymbol{D}$ is the diagonal matrix containing the eigenvalues of $\boldsymbol{A}$ as its diagonal elements.

Now if f is any function of a matrix, then we have

$f(\boldsymbol{A}) = \boldsymbol{P}\, f\,(\boldsymbol{D})\boldsymbol{P}^{-1}$ (7.49)

Thus, if we can define $f(\boldsymbol{D})$, we can define and evaluate $f(\boldsymbol{A})$.

7.10.2 : Powers of a matrix

We have from second equation of (7.48), we taking k^{th} power

$$\boldsymbol{A}^k = (\boldsymbol{PDP}^{-1})\ (\boldsymbol{PDP}^{-1})\ ....\ (k \text{ times})$$
$$= \boldsymbol{PD}^k\boldsymbol{P}^{-1} \tag{7.50}$$

Similarly, if $m = -k$ is a negative integer and $|\boldsymbol{A}| \neq 0$, then $\boldsymbol{A}^m = \boldsymbol{PD}^m\boldsymbol{P}^{-1} = \boldsymbol{P}(\boldsymbol{D}^{-1})^k$ $\boldsymbol{P}^{-1}$ (7.51)

Example : Find $\boldsymbol{A}^k$, where k is any integer, positive or negative, where $\boldsymbol{A} = \begin{pmatrix} 3 & i \\ -i & 3 \end{pmatrix}$

Solution : The eigenvalues of the matrix $\boldsymbol{A}$ are 2, 4. Eigen vectors are $\begin{pmatrix} 1 \\ i \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -i \end{pmatrix}$

Therefore $\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$, $\boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} = \boldsymbol{D}$

Now $|\boldsymbol{A}| \neq 0$. For any integral k, therefore we have :

$$\boldsymbol{A}^k = \boldsymbol{PD}^k\boldsymbol{P}^{-1} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} 2^k & 0 \\ 0 & 4^k \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2i} \\ \frac{1}{2} & -\frac{1}{2i} \end{pmatrix} = 2^{k-1}\begin{bmatrix} 1+2^k & i(-1+2^k) \\ (1-2^k)i & 1+2^k \end{bmatrix} \tag{7.52}$$

7.10.3 : Roots of a matrix

We consider a diagonal matrix $\boldsymbol{D}$, whose elements are given by $\boldsymbol{D} = \left[d_i\delta_{ij}\right]$. It is evident that $\boldsymbol{D}^k$ is again a diagonal matrix whose diagonal elements are d_i^k, i.e. $\boldsymbol{D}^k = \left[d_i{}^k\delta_{ij}\right]$. Now let $p = \dfrac{1}{k}$ and consider a diagonal matrix $\boldsymbol{D}_0$ whose elements are given by $\boldsymbol{D}_0 = \left[d_i{}^p\delta_{ij}\right]$. Clearly the k^{th} power of $\boldsymbol{D}_0$ will equal $\boldsymbol{D}$ i.e. $\boldsymbol{D}_0^{\,k} = \boldsymbol{D}$. Then consider the matrix $\boldsymbol{B} = \boldsymbol{PD}_0\boldsymbol{P}^{-1} = \boldsymbol{PD}^p\boldsymbol{P}^{-1}$. Taking the k^{th} power of $\boldsymbol{B}$, we find,

$$\boldsymbol{B}^k = (\boldsymbol{PD}^p\boldsymbol{P}^{-1})\ (\boldsymbol{PD}^p\boldsymbol{P}^{-1})\ \ldots\ (\boldsymbol{PD}^p\boldsymbol{P}^{-1})(\boldsymbol{k}^{\textbf{th}} \textbf{ times}) = \boldsymbol{PDP}^{-1} = \boldsymbol{A} \tag{7.53}$$

Thus $\boldsymbol{B} = \boldsymbol{PD}^p\boldsymbol{P}^{-1}$ is a k^{th} root of $\boldsymbol{A}$.

The same result holds good for any fractional power. Thus if q is any fraction, we have

$$\boldsymbol{A}^q = \boldsymbol{PD}^q\boldsymbol{P}^{-1} \tag{7.54}$$

7.10.4 : Series

Suppose u is a series in matrix $\boldsymbol{A}$ with scalar co-efficient a_k as in equation (7.55)

$$u = \sum_{k=0}^{\infty} a_k \boldsymbol{A}^{\boldsymbol{k}} \tag{7.55}$$

Now if and only if every element of the right hand side converges then the series converges. In each case, equation (7.55) can be written as :

$$f(\boldsymbol{A}) = u = \sum_{k=0}^{\infty} a_k \boldsymbol{A}^{\boldsymbol{k}} \tag{7.56}$$

Where $f(\boldsymbol{A})$ is a matrix of the same order as $\boldsymbol{A}$ and whose elements are given by

$$[f(\boldsymbol{A})]_{ij} = \sum_{k=0}^{\infty} a_k \left(\boldsymbol{A}^{\boldsymbol{k}}\right)_{ij} \tag{7.57}$$

Now we may state that a series $f(\boldsymbol{A})$ in a matrix $\boldsymbol{A}$ is convergent if and only if the coresponding algebraic series $f(\lambda)$ is convergent for every eigenvalue λ_i of $\boldsymbol{A}$.

Thus, if

$$f(\lambda) = \sum_{k=0}^{\infty} a_k \boldsymbol{A}^{\boldsymbol{k}} \tag{7.58}$$

exist for $|\lambda| < R$, then,

$$f(\boldsymbol{A}) = \sum_{k=0}^{\infty} a_k \boldsymbol{A}^{\boldsymbol{k}} \tag{7.59}$$

exists if and only if every eigenvalue λ_i of $\boldsymbol{A}$ satisfy $|\lambda_i| < R$. R is called the radius of convergence of the series.

7.11 ❒ Cayley-Hamilton's Theorem

The theorem can be stated as follows,

The square matrix satisfies its own characteristic equation.

Stated mathematically :

If $|\boldsymbol{A}-x\boldsymbol{I}| = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n = 0$

be the characteristic equation of a square matrix $\boldsymbol{A}$, then

$$a_0\boldsymbol{I} + a_1\boldsymbol{A} + a_2\boldsymbol{A}^2 + ... + a_n\boldsymbol{A}^n = 0 \tag{7.60}$$

Where every x is replaced by $\boldsymbol{A}$, and thus

$$a_0 = a_0x^0 = a_0\boldsymbol{A}^0 = a_0\boldsymbol{I}$$

Thus, for any square matrix of order n, there is a matrix polynomial of degree not more than n which equals the null matrix. If $\boldsymbol{A}$ is a square matrix where characteristic polynomial is $|\boldsymbol{A} - x\boldsymbol{I}| = D(x)$, then $D(\boldsymbol{A}) = 0$. If λ_i are the eigenvalues of $\boldsymbol{A}$, equation (7.60) can be written as

$$(\boldsymbol{A}-\lambda_1\boldsymbol{I})(\boldsymbol{A}-\lambda_2\boldsymbol{I})...(\boldsymbol{A}-\lambda_n\boldsymbol{I}) = 0 \tag{7.61}$$

Example : Consider the matrix $\boldsymbol{A} = \begin{bmatrix} -11 & -10 & 5 \\ 5 & 4 & -5 \\ -20 & -20 & 4 \end{bmatrix}$

Show that this matrix satisfies the cayley-Hamilton theorem.

Solution : The eigenvalues of $\boldsymbol{A}$ are 4, –6, –1, we have $(\boldsymbol{A} - 4\boldsymbol{I})(\boldsymbol{A} + 6\boldsymbol{I})(\boldsymbol{A} + \boldsymbol{I})$

$$= \begin{bmatrix} -15 & -10 & 5 \\ 5 & 0 & -5 \\ -20 & -20 & 0 \end{bmatrix} \begin{bmatrix} -5 & -10 & 5 \\ 5 & 10 & -5 \\ -20 & -20 & 10 \end{bmatrix} \begin{bmatrix} -10 & -10 & 5 \\ 5 & 5 & -5 \\ -20 & -20 & 5 \end{bmatrix} = 0$$

7.3.16.1 : Minimal polynomial

In case of some matrices having eigenvalues, with multiplicity greater than unity, there may exist polynomial of degree less than n which equal the zero matrix. The following example follows the statement.

Example 1 : If $\boldsymbol{A} = \begin{bmatrix} q & p & p \\ p & q & p \\ p & p & q \end{bmatrix}$, where p, q are scalar and $p \neq 0$, show that,

$$[\boldsymbol{A}-(q+2p)\boldsymbol{I}][\boldsymbol{A}-(q-p)\boldsymbol{I}]=0$$

Where $q - p$, $q - p$ and $q + 2p$ are the eigenvalues of $\boldsymbol{A}$.

Solution : The eigenvalues of the given matrix are found to be $q + 2p$ with multiplicity unity and $q - p$ with multiplicity two, greater than unity.

From cayley-Hamilton theorem, we get

$$[\boldsymbol{A}-(q+2p)\boldsymbol{I}][\boldsymbol{A}-(q-p)\boldsymbol{I}]^2=0 \tag{7.62}$$

However, now we consider the given matrix polynomial

$$[\boldsymbol{A}-(q+2p)\boldsymbol{I}][\boldsymbol{A}-(q-p)\boldsymbol{I}]=\begin{bmatrix}-2p & p & p\\ p & -2p & p\\ p & p & -2p\end{bmatrix}\begin{bmatrix}p & p & p\\ p & p & p\\ p & p & p\end{bmatrix}$$

$$=\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \tag{7.63}$$

This shows that for the matrix under consideration we have got a polynomial of degree 2 which is equal to zero. i.e.

$$\boldsymbol{A}^2-(2q+p)\boldsymbol{A}+\left(q^2+pq-2p^2\right)\boldsymbol{I}=0 \tag{7.64}$$

From equation (7.61) we see that highest degree of polynomial of $\boldsymbol{A}$ is 3. But from equation (7.64) we see that highest degree under the given condition is 2. Polynomial is equation (7.63) can be termed as minimal polynomial and polynomial in equation (7.61) is called characteristic polynomical of the matrix.

In the given problem the highest degree of the minimal polynomial is evidently 3 when it is equal to the characteristic polynomial. It is also evident that minimal polynomial divides the characteristic polynomial exactly.

7.3.16.2 : Matrix polynomial

An expression of the form

$$f(\boldsymbol{A}) = a_k \boldsymbol{A}^{\boldsymbol{k}} + a_{k-1}\boldsymbol{A}^{\boldsymbol{k}-1} + \ldots + a_1\boldsymbol{A} + a_0\boldsymbol{I} \tag{7.65}$$

Where a_i are scalar co-efficient, is called a matrix polynomial in $\boldsymbol{A}$ of degree K. The existence of a minimal polynomial [it can be shown] for every matrix provides a greater simplification in evaluating matrix polynomials.

7.11.1 Evaluation of functions of any matrix, Diagonalisable or not, using cayley-Hamiltonian theorem

If the degree of the minimal polynomial of a matrix $\boldsymbol{A}$ is m, any functions $f(\boldsymbol{A})$, which is sufficiently differentiable, can be expressed as a linear combination of the m linearly independent matrices

$\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2 \ldots. \boldsymbol{A}^{m-1}$, i.e. $f(\boldsymbol{A}) = s(\boldsymbol{A})$ (7.66), where $s(\boldsymbol{A}) = \alpha_{m-1}\boldsymbol{A}^{\boldsymbol{m}-1} + \alpha_{m-2}\boldsymbol{A}^{\boldsymbol{m}-2}$ $+ \ldots + \alpha_1\boldsymbol{A} + \alpha_0\boldsymbol{I}$ (7.67), where α_i are scalars which can be suitably evaluated by a process.

7.11.2 : Inner product :

In some practical situation in basic science we cannot do without some concept of measurement for distance and angle between two lines etc. The inner product in some way is a generalisation of these everyday concepts.

We discussed in 7.1 the way matrices occurs in physics by virtue of transformation of vectors in vector space and the operation of linear operators on vector space.

A linear transformation from one set to another can be written as

$$\boldsymbol{A}_i'(i=1,2..n) = \sum_{k=1}^{p} b_{ik}\boldsymbol{A}_k\,(i=1,2,3,\ldots,n) \tag{7.68}$$

Or as $\boldsymbol{A}' = \boldsymbol{BA}$ where $\boldsymbol{B} = [b_{ik}]$ is now an $n \times p$ matrix. If now the variables $\boldsymbol{A}'_i$ are further related to the m variables $\boldsymbol{A}''_i$ by a linear transformation of the form

$$\boldsymbol{A}_i''(i=1,2\ldots n) = \sum_{k=1}^{p} c_{ik}\boldsymbol{A}_k'\,(i=1,2,\ldots,m) \tag{7.69}$$

which can be written in the matrix notation as $\boldsymbol{A}'' = \boldsymbol{CA}'$ where $\boldsymbol{A}'' = \{\boldsymbol{A}''_i\}$ is a column vector of order $m \times 1$ and $\boldsymbol{C} = [C_{ij}]$ is a matrix of order $m \times n$.

Therefore the combined transformation can be written as :

$$\boldsymbol{A}'' = \boldsymbol{CA}' = \boldsymbol{CBA} = \boldsymbol{DA} \tag{7.70}$$

The product matrix $\boldsymbol{D} = \boldsymbol{CB}$ will be of order $m \times p$.

Now we define inner product space in matrix notation in line with equation (7.70) which gives a reason of definition of matrix multiplication as we have given earlier equation (7.70).

A vector space $\boldsymbol{V}$ defined over a field $\boldsymbol{F}$, where $\boldsymbol{F}$ is the field of real or complex numbers, becomes an inner product space if with every pair of element. ψ and $\phi \varepsilon V$, there is associated a unique scalar belonging to the field $\boldsymbol{F}$, denoted by (ψ, ϕ) and called the inner product or the scalar product of (ψ, ϕ); for which the following properties hold

$$\left.\begin{aligned} (\boldsymbol{\psi}, \phi) &= (\phi, \boldsymbol{\psi})^* \\ (a\boldsymbol{\psi}, b\phi) &= a^* b(\boldsymbol{\psi}, \phi) \\ (\xi, a\boldsymbol{\psi} + b\phi) &= a^*(\xi, \boldsymbol{\psi}) + b^*(\xi, \phi) \end{aligned}\right\} \tag{7.71}$$

These three equations in equations (7.71) together form the definition of inner product of vectors.

The vector space of n-tuplets of real or complex numbers can be made an inner product space if we define the inner product of two vectors by

$$(\boldsymbol{\psi}, \phi) = \sum_{i=1}^{n} \boldsymbol{\psi}_i^* \phi_i \tag{7.72}$$

Now when we regard vectors as column matrix, their inner product defined in equation (7.72) can also be written in a concise way in the matrix notation.

If $\boldsymbol{\psi} = \{\psi_1, \psi_2, ..., \psi_n\}$ and $\phi = \{\phi_1, \phi_2, ..., \phi_n\}$ are two column vectors,

$$\boldsymbol{\psi}^+ = \left[\psi_1^*, \psi_2^*, ..., \psi_n^*\right] \text{ and } \boldsymbol{\psi}^+\phi = \left[\psi_1^*, \psi_2^*, ..., \psi_n^*\right]\{\phi_1, \phi_2, ..., \phi_n\}$$

$$= \sum_{i=1}^{n} \boldsymbol{\psi}_i^* \phi_i \tag{7.73}$$

Which is the same expression in equation (7.72). Hence, the inner product can be expressed as $(\psi, \phi) = \psi^{+}\phi$ (7.74)

The orthogonality condition is now $\psi^{+}\phi = 0$ (7.73)

Examples :

Example 1 : Show that the matrix $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ is unitary.

Solution : $A^{+} = (A^{*})^{T} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^{T}$

or, $A^{+} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$, Now $A^{+}A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

$\therefore$ A is unitary.

Example 2 : Show that the matrix $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ is orthogonal.

Solution : Let $A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ $\therefore A^{T} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

Now $AA^{T} = \begin{bmatrix} \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \sin\theta\cos\theta \\ -\sin\theta\cos\theta + \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{bmatrix}$

$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$ = unit matrix $\therefore$ A is orthogonal.

Example 3 : If $\boldsymbol{A}=\begin{bmatrix}\cos\theta & i\sin\theta\\ i\sin\theta & \cos\theta\end{bmatrix}$ show that A is unitary.

Solution : $\boldsymbol{A}^{+}=(\boldsymbol{A}^{*})^{\boldsymbol{T}}=\begin{bmatrix}\cos\theta & -i\sin\theta\\ -i\sin\theta & \cos\theta\end{bmatrix}^{T}=\begin{bmatrix}\cos\theta & -i\sin\theta\\ -i\sin\theta & \cos\theta\end{bmatrix}$

$$\therefore \boldsymbol{A}^{+}\boldsymbol{A}=\begin{bmatrix}\cos\theta & -i\sin\theta\\ -i\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\cos\theta & i\sin\theta\\ i\sin\theta & \cos\theta\end{bmatrix}$$

$$=\begin{bmatrix}\cos^2\theta+\sin^2\theta & i\cos\theta\sin\theta-i\sin\theta\cos\theta\\ -i\sin\theta\cos\theta+i\sin\theta\sin\theta & \sin^2\theta+\cos^2\theta\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

$$= \textit{unit matrix}$$

$\therefore$ $\boldsymbol{A}$ is unitary matrix.

Example 4 : Consider the following transformation in three dimension :

$$x'=x\cos\theta+y\sin\theta$$

$$y'=-x\sin\theta+y\cos\theta$$

$$z'=z$$

i) Write down the transformation matrix $\boldsymbol{A}(\boldsymbol{\theta})$.

ii) Show that $\boldsymbol{A}(\theta_1)\boldsymbol{A}(\theta_2)=\boldsymbol{A}\left(\theta_1+\theta_2\right)$

iii) Show that $\boldsymbol{A}(\theta)$ is unitary.

Solution :

i) The transformation in this case is

$$\begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}=\begin{pmatrix}\cos\theta & \sin\theta & 0\\ -\sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}$$

The transformation matrix is $A(\theta)=\begin{pmatrix}\cos\theta & \sin\theta & 0\\ -\sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}$

ii) $$A(\theta_1)A(\theta_2)=\begin{pmatrix}\cos\theta_1 & \sin\theta_1 & 0\\ -\sin\theta_1 & \cos\theta_1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\cos\theta_2 & \sin\theta_2 & 0\\ -\sin\theta_2 & \cos\theta_2 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$=\begin{pmatrix}\cos(\theta_1+\theta_2) & \sin(\theta_1+\theta_2) & 0\\ -\sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) & 0\\ 0 & 0 & 1\end{pmatrix}=A(\theta_1+\theta_2)$$

iii) $A^+(\theta_1)=\left[A^*(\theta)\right]^T=\left[A(\theta)\right]^T$ since $A(\theta)$ is real matrix.

Now $$A^+(\theta)A(\theta)=\left[A(\theta)\right]^T\left[A(\theta)\right]=\begin{pmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\cos\theta & \sin\theta & 0\\ -\sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$=0\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}=I$$

Exercises :

Pauli matrices : The following set of three 2 × 2 matrices is called Pauli spin matrices.

$$\sigma_1=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\sigma_2=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\sigma_3=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

Prove that :

i) $\sigma_i\sigma_j+\sigma_j\sigma_i=2\,\delta_{ij}I_2$ where I_2 is a 2 × 2 unit matrix, $\delta_{ij}=\left.\begin{matrix}0, i\neq j\\ 1, i=j\end{matrix}\right\}$

ii) $\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j = i\boldsymbol{\sigma}_k, i, j, k$ is a cyclic permutation of 1,2,3

iii) $\boldsymbol{\sigma}_i^2 = \boldsymbol{I}_2$

Keywords

Adjoint, inverse, Orthogonal, Hermitian, Trace, Rank, Normal form, matrix functions, Eigen value and Eigen functions.

7.12 ❐ Summary

1. Definitions, notation and terminology of real and complex matrices have been discussed with examples.
2. Properties of inverse, orthogonal, unitay matrices, have been stated.
3. Procedures for Diagonalisation and to find rank of matrix have been discussed.
4. Eigen value equations have been set up and procedure to obtain eigenfunction and eigen values have been indicated. Coupled linear ordinary differential equation have been discussed in terms of eigen value problem.
5. Evaluations of function of any matrix have been incorporated.

Unit 8 ❒ C and C++ Programming Fundamentals

Structure

8.1 ❒ Objectives

This chapter is designed for software programmers with a need to understand the C programming language starting from scratch. This chapter will give you enough understanding on C programming language from where you can take yourself to higher level of expertise.

8.2 ❒ Introduction to Programming

In computing, a program is a specific set of ordered operations for a computer to perform. In the modern computer that John von Neumann outlined in 1945, the program contains a one-at-a-time sequence of instructions that the computer follows. Typically, the program is put into a storage area accessible to the computer. The computer gets one instruction and performs it and then gets the next instruction. The storage area or memory can also contain the data that the instruction operates on. Note that a program is also a special kind of "data" that tells how to operate on "application or user data".

The processs of writing computer programs is known as Programming, and the language of writing programs is known as Programming Language. Example : C & C++.

C is a general-purpose, high-level language that was originally developed by Dennis M. Ritchie to develop the UNIX operating system at Bell Labs. C was originally first implemented on the DEC PDP-11 computer in 1972. In 1978, Brian Kernighan and Dennis Ritchie produced the first publicly available description of C, now known as the K&R (Kernighan and Richie) standard. The UNIX operating system, the C compiler, and essentially all UNIX application programs have been written in C. C has now become a widely used professional language for various reasons : Easy to learn Structured language it produces efficient programs it can handle low-level activities. It can be compiled on a variety of computer platforms.

Keywords : C, C++, auto, break, case, char, const, continue, default, do, double, else, enum, extrn, float, for, goto, if, int, long, register, return, short, signed, sizeof, static, struct, switch, typedef, union, unsigned, void, volatile, while.

Facts about C :

C was invented to write an operating system called UNIX. C is a successor of B language which was introduced around the early 1970s. The language was formalized in 1988 by the American National Standard Institute (ANSI). The UNIX OS was totally written in C. Today C is the most widely used and popular System Programming Language. Most of the state-of-the-art software have been implemented using C. Today's most popular Linux OS and RDBMS MySQL have been written in C.

Why Use C :

C was initially used for system development work, particularly the programs that make-up the operating system. C was adopted as a system development language because it produces code that runs nearly as fast as the code written in assembly language. Some examples of the use of C might be :

- Operating Systems
- OVERVIEW C Programming Language
- Compilers
- Assemblers
- Text Editors
- Print Spoolers
- Network Drivers
- Modern Programs
- Databases
- Language Interpreters
- Utilities

C Programs :

A C program can vary from 3 lines to millions of lines and it should be written into one or more text files with extension ".c"; for example, hello.c. You can use "vi", "vim" or any other text editor to write your C program into a file. This tutorial assumes that you know how to edit a text file and how to write source code inside a program file.

Sample :

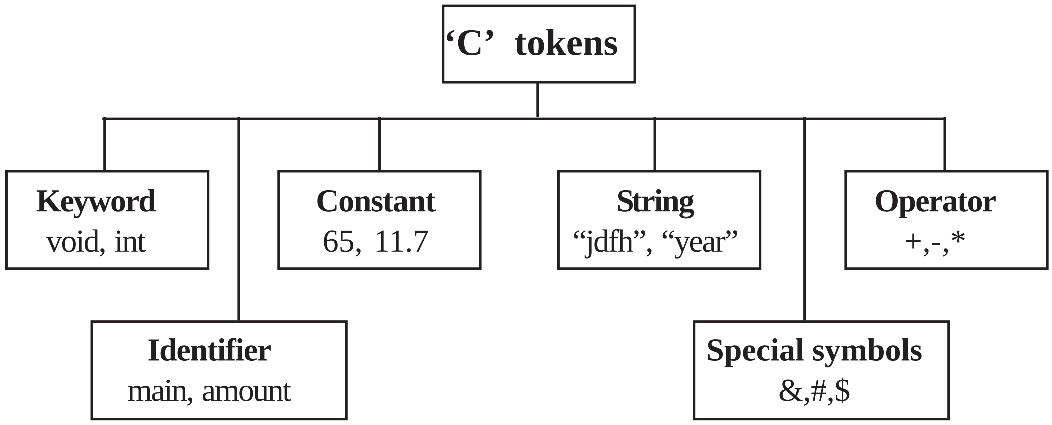
```
# include<stdio.h>          //header file
void main(){                //main Function
        //code
}
```

8.3 ❐ Constants, Variable and Data types

8.3.1 C Tokens :

The smallest individual units and punctuation marks are known as C Tokens. C Programming Language has six types of Tokens.

Keywords : Keywords are special words that are used to give a special meaning to the program and can't be used as variable and constant. They are basically a sequence of characters that have fixed to mean for example break, for, while, do-while, do, if, int, long, char.



Identifiers : Identifiers refer to the variable name, array name, function name. It is user defined and collection of letters and digits but first letter always character.

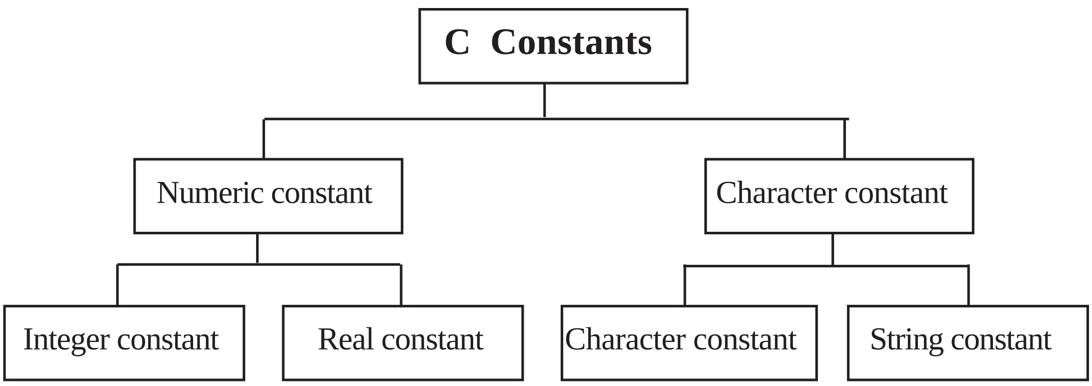
Constants : The quantity which does not change during the execution of a program is known as constant. There are types of constant.

Variables : Variables are used to give the name and allocate memory space. An entity that may vary data during execution. For example, sum, area, a, b, age, city.

String : String is a collection of character. For example, "RAM", "Meerut", "Star" String is represented by a pair of double quotes.

Operators : Operators act as connectors and they indicate what type of operation is being carried out. The values that can be operated by these operators are called operands. They are used to perform basic operations, comparison, manipulation of bits and so on.

8.3.2 C Constants



Integer Constants :

An integer constant is a sequence of digits from 0 to 9 without decimal points or fractional part or any other symbols. There are 3 types of integers namely decimal integer, octal integers and hexadecimal integer.

Decimal Integers consists of a set of digits 0 to 9 preceded by an optional + or - sign. Spaces, commas and non digit characters are not permitted between digits. Example for valid decimal integer constants are

int y=123;//here 123 is a decimal integer constant

Octal Integers constant consists of any combination of digits from 0 through 7 with a O at the beginning. Some examples of octal integers are

Real Constants

Real Constants consists of a fractional part in their representation. Integer constants are inadequate to represent quantities that vary continuously. These quantities are represented by numbers containing fractional parts like 26.082. Example of real constants are

float x = 6.3; //here 6.3 is a double constant.

float y = 6.3f; //here 6.3f is a float constant.

float z = 6.3 e + 2; //here 6.3 e + 2 is a exponential constant.

float s = 6.3L ; //here 6.3L is a long double constant

Real numbers can also be represented by exponential notation. The general form for exponential notation is mantissa exponent. The mantissa is either a real number expressed in decimal notation or an integer. The exponent is an integer number with an optional plus or minus sign.

Single Character Constants

A single character constant represent a single character which is enclosed in a pair of quotation symbols.

Example for character constants are

char p = 'ok' ; // p will hold the value 'o' and k will be omitted

char y = 'u'; // y will hold the value 'u'

char k = '34' ; // k will hold the value '3', and '4' will be omitted

char e = ' '; // e will hold the value ' ', a blank space

chars = '\45'; // s will hold the value ' ', a blank space

All character constants have an equivalent integer value which are called ASCII Values.

String Constants

A string constant is a set of characters enclosed in double quotation marks. The characters in a string constant sequence may be a alphabet, number, special character and blank space. Example of string constants are

"VISHAL" "1234" "God Bless" "!.....?"

Backslash Character Constants [Escape Sequences]

Backslash character constants are special characters used in output functions. Although they contain two characters they represent only one character. Given below is the table of escape sequence and their meanings.

8.3.3 Variables

C variable is an identifier that is used to store a data value and whose value may be changes during the program execution. **C variable** might be belonging to any of the data type like int, float, char etc.

Variable Declaration in C :

All variable which are used in the program should be declared before use. Declaration consists of one or more variable name (that are chosen by programmer) with data type and ending with semicolon.

Example int sum;

Data type **variable name**

Condition for declaring Variable

1. They must begin with letter.
2. Length of a variable should not be more than 31 characters.
3. It should not be Keyword.
4. No white space is allowed.

8.3.4 Data Types

Data types specify a particular kind of data item, as defined by the values variable can take. C language has some predefined set of data types to handle various kinds of data that we can use in our program. These data types have different storage capacities.

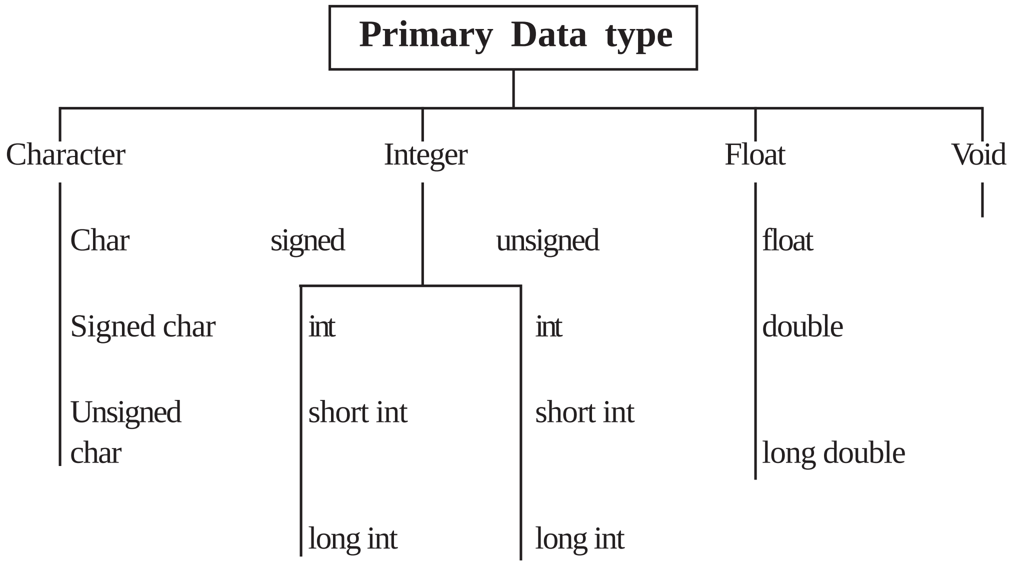
C language supports 2 different types of data types :

1. **Primary data types :**

   These are fundamental data types in C namely integer (int), floating point (float), character (char) and void.

2. **Derived data types :**

   Derived data types are nothing but primary data types but a little twisted or grouped together like **array, structure, union** and **pointer.** These are discussed in details later. Data type determines the type of data a variable will hold. If a variable x is declared an int. It means x can hold only integer values. Every variable which is used in the program must be declared as what data-type it is.



Summary : This has been a lengthy and perhaps disconcerting article. The alphabet of C although of relevance, is not normally a day to day consideration of practicing programmers, so it has been discussed but can now be largely ignore.

Much the same can be said regarding keywords and identifiers, since the topic is not complicated and simply becomes committed to memory. The declaration of variables is rarely a problem, although it is worth re-emphasizing the distinction between a declaration and a definition. If that still remains unclear, you might find of benefit to go back and re-read the description. The standard has substantially affected parts of the language described in this chapter. In particular, the changes to the conversions and the change from 'unsignedness preserving' to 'value preserving' rules of arithmetic may cause some surprises to experienced C programmers. Even they have some real relearning to do.

8.4 ❐ Operators and Expressions

An operator is a symbol that tells the compiler to perform specific mathematical or logical functions. C provides the following types of operators

- Arithmetic Operators
- Relational Operators
- Logical Operators
- Bitwise Operators
- Assignment Operators
- Misc Operators

Arithmetic Operators

Operator	Operation of Operator	Example
+	Adds two operands.	A + B
–	Subtracts second operand from the first.	A – B
*	Multiplies both operands.	A*B
/	Divides numerator by de-numerator	B/A
%	Modulus Operator and remainder of after an integer division.	B % A
++	Increment operator increases the integer value by one	A++
--	Decrement operator decreases the integer value by one.	A--

Relational Operators

Operator	Operation of Operator	Example
= =	Check if the values of two operands are equal or not. If yes, then the condition becomes true.	(A == B)
!=	Check if the values of two operands are equal or not. If the values are not equal, then the condition becomes true	(A != B)
>	Check if the value of left operand is greater than the value of right operand. If yes, then the condition becomes true.	(A > B)

<	Check if the value of left operand is less than the value of right operand. If yes, then the condition becomes true.	(A < B)
>=	Check if the value of left operand is greater than or equal to the value of right operand. If yes, then the condition becomes true.	(A >= B)
<=	Check if the value of left operand is less than or equal to the value of right operand. If yes, then the condition becomes true.	(A <= B)

Logical Operators

Bitwise Operators

Operator	Operation of Operator	Example
&&	Called Logical AND operator. If both the operands are non-zero, then the condition becomes true	(A && B)
||	Called Logical OR Operator. If any of the two operands is non-zero, then the condition becomes true	(A || B)
!	Called Logical NOT Operator. It is used to reverse the logical state of its operand. If a condition is true, then Logical NOT operator will make it false.	! (A && B)

Bitwise operator works on bits and perform bit-by-bit operation. The truth tables for &, I, and ^ is as follows –

A	B	A&B	A | B	A^B
0	0	0	0	0
0	1	0	1	1
1	1	1	1	0
1	0	0	1	1

Assignment Operator

This type of Operators are used to assisgn the result of an expression to an identifier. The most common assignment operator is " = ".

Example : C = A + B will assign the value of A + B to C.

Misc Operators

Operator	Operation of Operator	Example
sizeof()	Returns the size of a variable.	sizeof(a), where a is integer, will return 4.
&	Returns the address of a variable.	& a; returns the actual address of the variable.
*	Pointer to a variable	* a;
? :	Conditional Expression.	If condition is true? then value X : otherwise value Y

C Programming Expresson :

1. In programming, an expression is any legal combination of symbols that represents a value.
2. C Programming provides its own rules of Expression, whether it is legal expression or illegal expession. For example, in the C language × +5 is a legal expression.
3. Every expression consists of at least one operand and can have one or more operators.
4. Operands are values and Operators are symbols that represent particular actions.

Valid C Programming Expression :

C Programming code gets compiled firstly before execution. In the different phases of compiler, C programming expression is checked for its validity.

Expressions	Validity
a + b	Expression is valid since it contain + operator which is binary operator
++ a + b	Invalid Expression

Priority and Expression :

In order to solve any expression we should have knowledge of C Programming Operators and their priorities.

Types of Expression :

In programming, different varities of expressions are given to the compiler. Expressions can be classified on the basis of Position of Operators in an expression –

Type	Explanation	Example
Infix	Expression in which Operator is in between Operands	a + b
Prefix	Expression in which Operator is written before Operands	+ a b
Postfix	Expression in which Operator is written after Operands	ab +

These expressions are solved using the stack.

Example of Expression :

Now we will be looking into some of the C Programming Expressions, Expression can be created by combining the operators and operands

Each of the expression results into the some resultant output value. Consider few expressions in the table below.

Expression Examples, Explanation n1 + n2, this is an expression which is going to add two numbers and we can assign the result of addition to another variable.

x = y. This is an expression which assigns the value of right hand side operand to left side variable.

v = u + a* t, We are multiplying two numbers and result is added to 'u' and total result is assigned to v

x <= y, This expression will return Boolean value because comparison operator will give us output either true or false ++j, This is expression having pre increment operator, it is used to increment the value of j before using it in expression [/table].

Summary : This chapter has described the entire range of control of flow available in C. The only areas that cause even moderate surprise are the way in which cases in a switch statement are not mutually exclusive, and the fact that goto cannot transfer control to any function except the one that is currently active. None of this is intellectually deep and it has never been known to cause problems either to beginners or programmers experienced in other languages. The logical expressions all give integral results. This is perhaps slightly unusual, but once again take very little time to learn.

8.5 ❒ I/O Statements

C programming has several in-built library functions to perform input and output tasks.

Two commonly used functions for I/O (Input/Output) are printf() and scanf().

The scanf() function reads formatted input from standard input (keyboard) whereas the printf() function sends formatted output to the standard output (screen).

Example 1 : C Output

```
#include<stdio.h>// Including header file to run printf() function.
int main()
{
printf("C Programming");// Displays the content inside quotation
return O;
}
```

Output

C Programming

How this program works ?

- All valid C program must contain the main() function. The code execution begins from the start of main () function.
- The printf() is a library function to send formatted output to the screen. The printf() function is declared in "stdio.h" header file.
- Here, stdio.h is a header file (standard input output header file) and #include is a preprocessor directive to paste the code from the header file when necessary. When the compiler encounters printf() function and doesn't find stdio.h header file, compiler shows error.
- The return O; statement is the "Exit status" of the program.

Example 2 : C Integer Output

```
#include<stdio.h>
int main()
{
inttestInteger=5;
printf("Number = %d", testInteger);
```

```
return 0;
}
```

Output

Number = 5

Inside the quotation of printf() function, there is a format string "%d" (for integer). If the format string matches the argument (*testInteger* in this case), it is displayed on the screen.

Example 3 : C Integer Input/Output

```
#include<stdio.h>
int main()
{
inttestInteger;
printf("Enter an integer:");
scanf("%d", &testInteger);
printf("Number = %d", testInteger);
return 0;
}
```

Output

Enter an integer : 4

Number = 4

The scanf() function reads formatted input from the keyboard. When user enters an integer, it is stored in variable *testInteger.*

Note the '&' sign before *testInteger; & testInteger* gets the address of *testInteger* and the value is stored in that address.

Example 4 : C Floats Input/Output

```
#include<stdio.h>
int main()
{
float f;
printf("Enter a number:");
```

```
//%f format string is used in case of floats
scanf("%f",&f);
printf("Value = %f", f);
return 0;
}
```

Output

Enter a number : 23.45

Value = 23.450000

The format string "%f" is used to read and display formatted in case of floats.

Example 5 : C Character I/O

```
#include<stdio.h>
int main()
{
char chr;
printf("Enter a character: ");
scanf("%c",&chr);
printf("You entered %c.",chr);
return 0;
}
```

Output

Enter a character : g

You entered g.

Format string %c is used in case of character types.

Little bit on ASCII code

When a character is entered in the above program, the character itself is not stored. Instead, a numeric value (ASCII value) is stored.

And when we displayed that value using "%c" text format, the entered character is displayed.

Example 6 : ASCII Code

```
#include<stdio.h>
```

```
int main()
{
char chr;
printf("Enter a character: ");
scanf("%c",&chr);
//When %c text format is used, character is displayed in case of character types
printf("You entered %c.\n", chr);
//When %d text format is used, integer is displayed in case of character types
printf("ASCII value of %c is %d.",chr,chr);
return 0;
}
```

Output

```
Enter a character : g
You entered g.
ASCII value of g is 103.
```

The ASCII value of character 'g' is 103. When, 'g' is entered, 103 is stored in variable chr instead of g.

You can display a character if you know ASCII code of that character. This is shown by following example.

Example 7 : C ASCII Code

```
#include<stdio.h>
int main()
{
int chr=69;
printf("Character having ASCII value 69 is %c.", chr);
return 0;
}
```

Output

```
Character having ASCII value 69 is E.
```

More on Input/Output of floats and Integers

Integer and floats can be displayed in different formats in C programming.

Example #7 : I/O of Floats and Integers

```
#include<stdio.h>
int main()
{
int integer = 9876;
float decimal=987.6543;
// Prints the number right justified within 6 columns
printf("4 digit integer right justified to 6 column : %6d\n", integer);
//Tries to print number right justified to 3 digits but the number is not right adjusted because there are only 4 numbers
printf("4 digit integer right justified to 3 column: %3d\n", integer);
//Rounds to two digit places
printf("Floating point number rounded to 2 digits: %.2f\n", decimal);
//Round to 0 digit places
printf("Floating point number rounded to 0 digits: %.f\n", 987.6543);
//Prints the number in exponential notation (scientific notation)
printf("Floating point number in exponential form: %e\n", 987.6543);
Return 0;
}
```

Output

```
4 digit integer right justified to 6 column     :  9876
4 digit integer right justified to 3 column     :  9876
Floating point number rounded to 2 digits       :  987.65
Floating point number rounded to 0 digits       :  988
Floating point number in exponential form       :  9.876543e+02
```

8.6 ❒ Manipulators for data formatting

Formatting output using manipulators

Formatted output is very important in development field for easily read and understand.

C++ offers the several input/output manipulators for formatting, commonly used manipulators are given below .

Manipulator Declaration in

endl	iostream.h
setw	iomanip.h
setprecision	iomanip.h
setf	iomanip.h

endl

endl manipulator is used to Terminate a line and flushes the buffer.

Difference b/w '\n' and endl

When writing output in C++, you can use either std::endl or '\n' to produce a newline, but each has a different effect.

- std::endl sends a newline character '\n' and flushes the output buffer.
- '\n' sends the newline character, but does not flush the output buffer.

The distinction is very important if you're writing debugging messages that you really need to see immediately, you should always use std::endl rather than '\n' to force the flush to take place immediately.

The following is an example of how to use both versions, although you cannot see the flushing occurring in this example.

```
#include <iostream.h>
int main()
{
cout<<"USING'\\n' ...\n";
cout<<"Line 1 \nLine 2 \nLine3 \n";
cout<<"USING end ...."<<endl;
cout<< "Line 1" <<endl<< "Line 2" <<endl << "Line 3" <<endl;
return 0;
}
```

Output

```
USING '\n' ...
Line 1
Line 2
Line 3
USING end ...
Line 1
Line 2
Line 3
```

setw() and setfill() manipulators

setw() manipulator sets the width of the filed assigned for the output.

The field width determines the minimum number of characters to be written in some output representations. If the standard width of the representation is shorter than the field width, the representation is padded with fill characters (using setfill).

setfill() character is used in output insertion operations to fill spaces when results have to be padded to the field width.

Syntax

```
setw([number_of_characters]);
setfill([character]);
```

Consider the example

```
#include <iostream.h>
#include <iomanip.h>
int main()
{
    cout<<"USING setw() ............\n";
    cout<<setw(10) <<11<<"\n";
    cout<<setw(10)<<2222<<"\n";
    cout<<setw(10) <<33333<<"\n";
    cout<<setw(10) <<4<<"\n";

    cout<<"USING setw() &setfill() [type- I]..\n":
```

```
        cout<<setfill('0');
        cout<<setw(10) <<11<<"\n";
        cout<<setw(10)<<2222<<"\n";
        cout<<setw(10) <<33333<<"\n";
        cout<<setw(10) <<4<<"\n";
        cout<<"USING setw() &setfill() [type- II]..\n":
        cout<<setfill('-')<<setw(10) <<11<<"\n";
        cout<<setfill('*')<<setw(10)<<2222<<"\n";
        cout<<setfill('@')<<setw(10) <<33333<<"\n";
        cout<<setfill('#')<<setw(10) <<4<<"\n";
        return 0;
}
```

Output

```
USING setw()............
        11
      2222
     33333
         4
USING setw() &setfill() [type- I]...
0000000011
0000002222
0000033333
0000000004
USING setw() &setfill() [type-II]...
-------  –  11
******2222
@@@@@33333
#########4
```

setf() and setprecision() manipulator

setprecision manipulator sets the total number of digits to be displayed, when floating point numbers are printed.

Syntax

```
setprecision([number_of_digits]);
cout<<setprecision(5)<<1234.537;
//output will be : 1234.5
```

On the default floating-point notation, the precision field specifies the maximum number of meaningful digits to display in total counting both those before and those after the decimal point. Notice that it is not a minimum and therefore it does not pad the displayed number with trailing zeros if the number can be displayed with less digits than the precision.

In both the fixed and scientific notations, the precision field specifies exactly how many digits to display after the decimal point, even if this includes trailing decimal zeros. The number of digits before the decimal point does not matter in this case.

Syntax

```
setf([flag_value], [field bitmask]);
```

field bitmask flag values

adjustfield	left, right or internal
basefield	dec, oct or hex
floatfield	scientific or fixed

Consider the example

```
#include <iostream.h>
#include <iomanip.h>
int main ()
{
    cout<<"USING fixed .................\n";
    cout.setf(ios::floatfield,ios::fixed);
    cout<<setprecision(5)<<1234.537<<endl;
    cout<<"USING scientific .................\n";
    cout.setf(ios::floatfield,ios::scientific);
    cout<<setprecision(5)<<1234.537<<endl;
    return 0;
}
```

Output

```
USING fixed ......................
1234.53700
USING scientific ...............
1234.5
```

Consider the example to illustrate base fields

```
#include <iostream.h>
#include <iomanip.h>
int main()
{
    intnum=10;
    cout<<"Decimal value is :"<<num<<endl;
    cout.setf(ios::basefield,ios::oct);
    cout<<"Octal value is :"<<num<<endl;
    cout.setf(ios::basefield,ios::hex);
    cout<<"Hex value is :"<<num<<endl;
    return 0;
}
```

8.7 ❒ Control System

In any programming language, there is a need to perform different tasks based on the condition and we can control the flow of program in such a way so that it executes certain statements based on the outcome of a condition. In C programming Language we have following decision control statements

1. if statement
2. if-else & else-if statement
3. switch-case statement

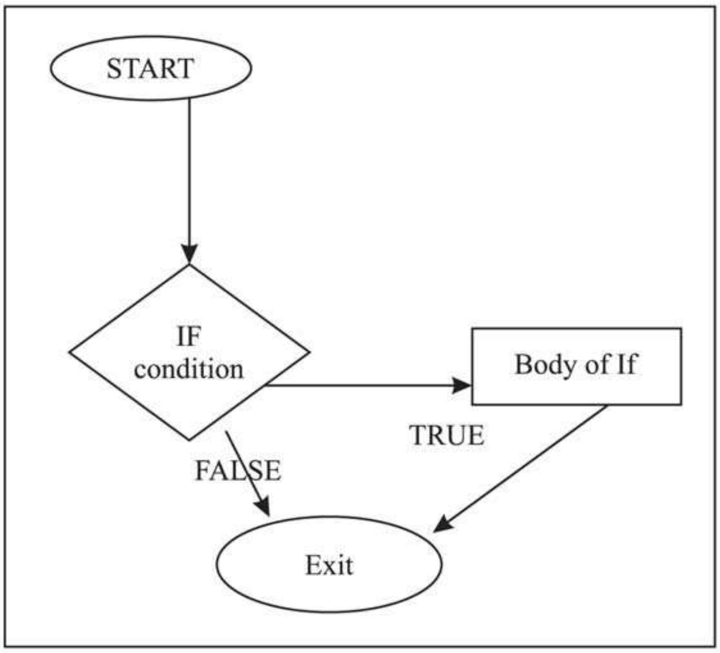
If Statement

It is basically a two way decision statement, used to decide whether a certain statement or block of statements will be executed or not.

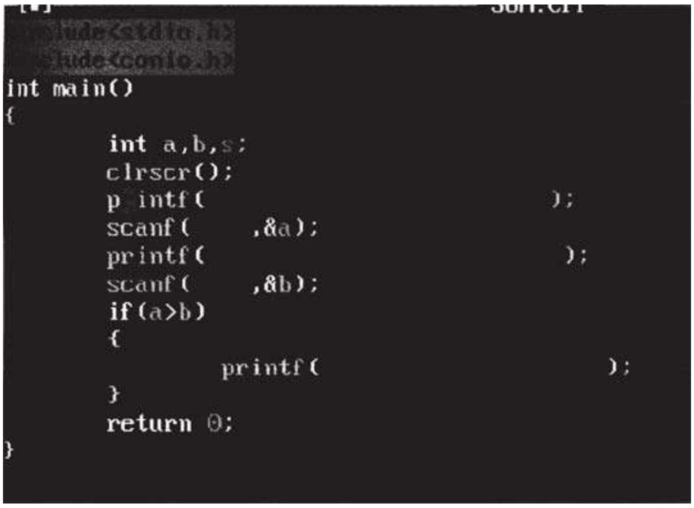
Syntax of if statement

```
If (condition)
{
        Body Of If Statement
}
```

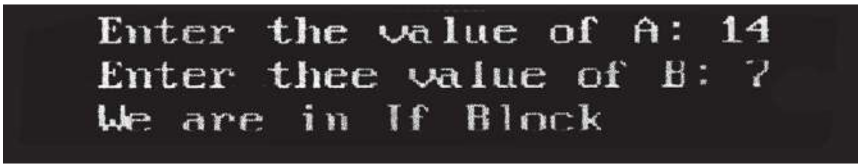
Flow Diagram of if statement



Example of if statement



Output



If-Else Statement

If-Else statement is an extension of if statement. If the condition of if statement is TRUE the If block is executed otherwise Else block is executed.

Syntax of if- Else statement

```
if (Condition)
{
        Body Of If
}
else
{
        Body of Else
}
```

Example of if-else statement

```
#include<stdio.h>
int main()
{
        int a;
        printf(                  );
        scanf(     ,&a);
        if(a%2==0)
                printf(              ,a);
        else
                printf(             ,a);
        return 0;
}
```

Output

```
 enter a value  13
13 is odd
```

The Else If Statement

The else...if statement is useful when you need to check multiple conditions

Syntax Of Else-If Statement

```
if (condition 1)
        Body Of Else If
}
else if (Condition 2)
{
        Body Of Else If
}
else
{
        Body Of Else
}
```

Example Of Else-If Statement

```
#include<stdio.h>
int main()
{
        int a,b,c;
        printf(                                  );
        scanf(     ,&a);
        printf(                                  );
        scanf(     ,&b);
        printf(                                  );
        scanf(     ,&c);
        if(a>b && a>c)
                printf(              );
        else if(b>a && b>c)
                printf(              );
        else
                printf(              );
}
```

Output

```
Enter the Value of A :12

Enter the Value of B :14

Enter the Value of C :16
C is big
```

Nested If ..Else statement

When an if else statement is present inside the body of another “if” or “else” then this is called nested if else.

Example Of Nest If Statement

```
#include<stdio.h>
int main()
{
        int a,b,c;
        printf(                              );
        scanf(      ,&a);
        printf(                              );
        scanf(      ,&b);
        printf(                              );
        scanf(      ,&c);
        if(a>b)
                if(a>c)
                        printf(          );
                else
                        printf(          );
        else if(b>c)
                printf(          );
        else
                printf(          );
}
```

Output

```
Enter the Value of A :12

Enter the Value of B :14

Enter the Value of C :11
B is big
```

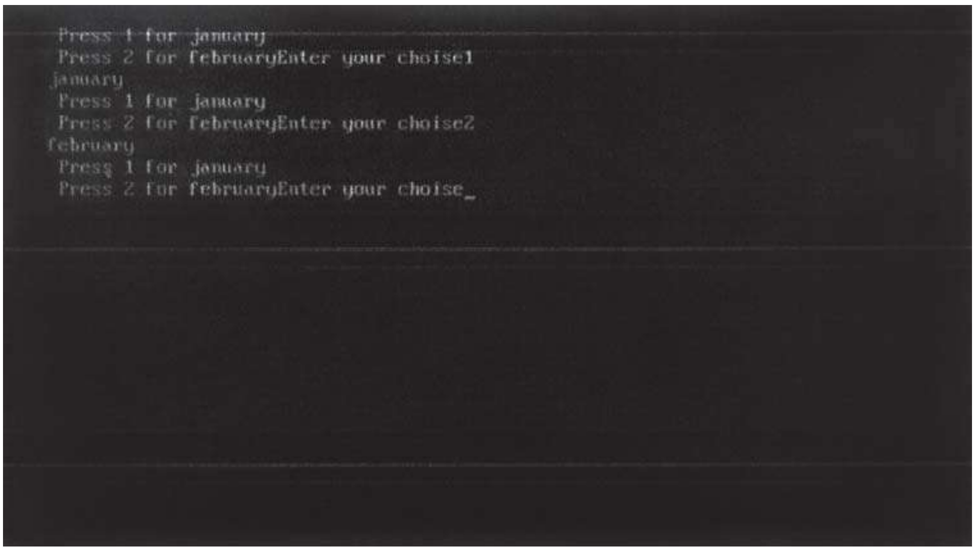
Switch Statement in C/C++

The switch statement is used when we have multiple options and we need to perform a different task for each option.

```
switch (n)
{
      case 1 : //code to be executed if n = 1;
            break ;
      case 2://code to be executed if n = 2;
            break;
      default://code to be executed if n doesn't match any cases
}
```

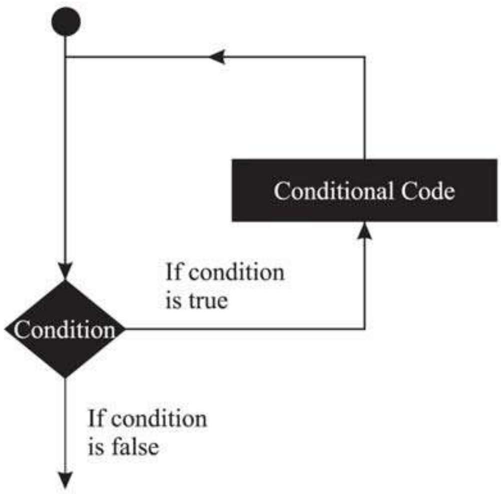
Example of Switch Statement

Output



8.8 ❒ Loop Statement

Loop : Loop is used when we need to repeatedly execute a block of statements according to the condition given in the loop.



C programming language provides the following types of loops to handle looping requirements.

1. for loop
2. while loop
3. do... while loop

For Loop

The syntax of for loop is :

```
for (initialization Statement; testExpression; updateStatement)
{
        // codes
}
```

How for loop works ?

The initialization statement is executed only once.

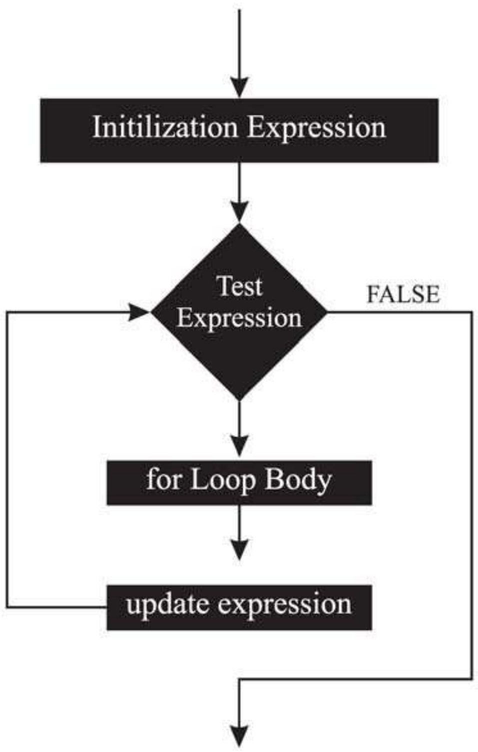
Then, the test expression is evaluated. If the test expression is false (0), for loop is terminated. But if the test expression is true (nonzero), codes inside the body of for loop is executed and the update expression is updated.

This process repeats until the test expression is false.

The *for* loop is commonly used when the number of iterations is known.

To learn more on test expression (when test expression is evaluated to nonzero (true) and 0 (false)), check out relational and logical operators.

For Loop Flowchart



Example for Loop

```
// Program to calculate the sum of first n natural numbers
// Positive integers 1,2,3...n are known as natural numbers
#include<stdio.h>
int main()
{
int num, count, sum =0;
```

```
printf("Enter a positive integer: ");
scanf("%d", &num);
//for loop terminates when n is less than count
for (count = 1; count <=num;++count)
{
        sum += count;
}
printf("Sum = %d", sum);
return 0;
}
```

Output

Enter a positive integer : 10

Sum = 55

While loop

The syntax of while loop is :

```
while (testExpression)
{
//codes
}
```

where, testExpression checks the condition is true or false before each loop.

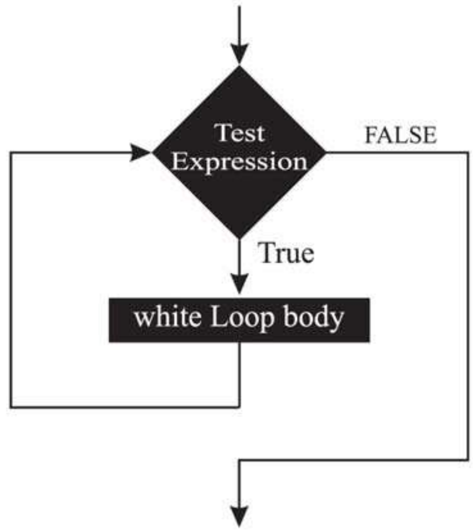
How while loop works?

The while loop evaluates the test expression.

If the test expresion is true (nonzero), codes inside the body of while loop are exectued. The test expression is evaluated again. The process goes on until the test expression is false.

When the test expression is false, the while loop is terminated.

Flowchart of while loop



Example 1 : While Loop

```
// Program to find factorial of a number
// For a positive integer n, factorial = 1*2*3...n
#include<stdio.h>
int main()
{
int number;
long long factorial;
printf("Enter an integer: ");
scanf("%d", &number);
factorial = 1;
//loop terminates when number is less than or equal to 0
while (number >0)
{
factorial *= number://factorial = factorial* number;
--number;
}
printf("Factorial= %lld", factorial);
```

```
return 0;
}
```

Output

```
Enter an integer: 5
Factorial = 120
```

do...while loop

The do..while loop is similar to the while loop with one important difference. The body of do... while loop is executed once, before checking the test expression. Hence, the do...while loop is executed at least once.

do...while loop Syntax

```
do
{
// codes
}
while (testExpression);
```

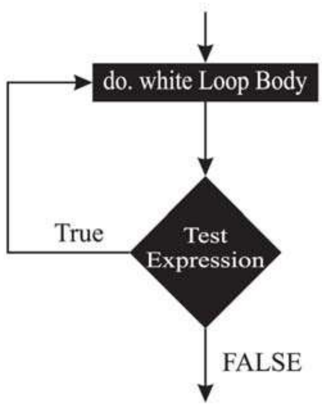
How do...while loop works?

The code block (loop body) inside the braces is executed once.

Then, the test expression is evaluated. If the test expression is true, the loop body is executed again. This process goes on until the test expression is evaluated to 0 (false).

When the test expression is false (nonzero), the do... while loop is terminated.

Flowchart of do... while Loop



Example 2: do...while loop

```
//Program to add numbers until user enters zero
#include<stdio.h>
int main()
{
double number, sum = 0;
//loop body is executed at least once
do
{
printf("Enter a number: ");
scanf("%If".&number);
sum += number;
}
while(number !=0.0);
printf("Sum = % %.2lf", sum);
return 0;
}
```

Output

```
Enter a number  :   1.5
Enter a number  :   2.4
Enter a number  :  –3.4
Enter a number  :   4.2
Enter a number  :     0
Sum = 4.70
```

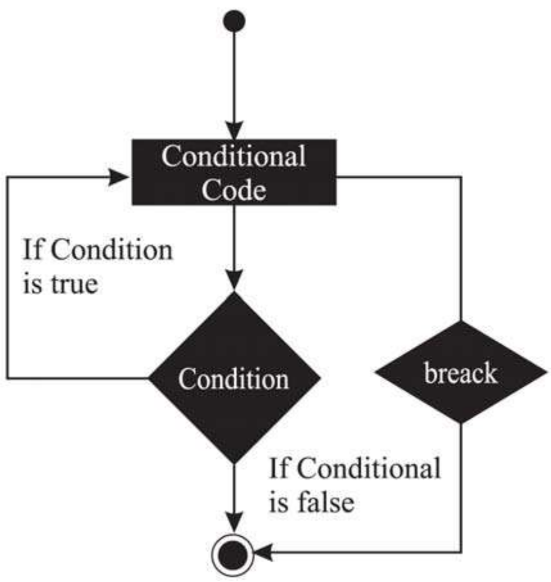
Nested Loop : C programming allows to use one loop inside another loop. Loop inside the another loop is called Nested Loop.

Syntax

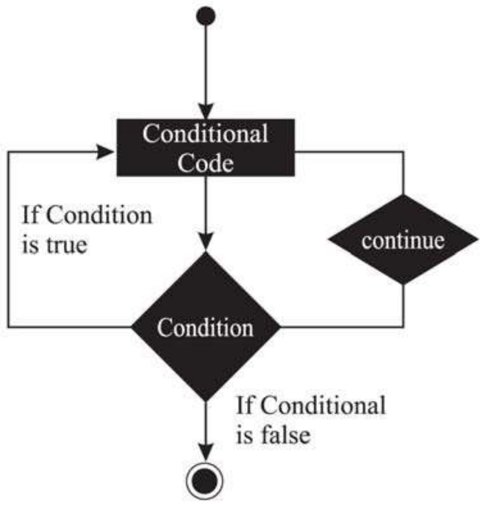
```
while (condition) {

        while (condition) {
              statement (s);
        }
        statement (s);
}
```

Loop Control Statements :

1. **Break** : When a break statement is encountered inside a loop, the loop is immediately terminated and the program control resumes at the next statement following the loop.



2. Continue : When a continue statement is encountered inside a loop, control jumps to the beginning of the loop for next iteration, skipping the execution of statements inside the body of loop for the current iteration.



Theory Questions

1. State whether the following statement are True or False
   (a) Every statement in C program should end with a semicolon.
   (b) Every C program ends with and END word.
   (c) A printf Statement can generate only one output line.
   (d) Like variable constant have a type./
   (e) All static variable are automatically initialized to zero.
   (f) The statement a+ = 10 is valid.
   (g) The expression !(a>=b) is same as a>b.
   (h) = & '==' both are same.
   (i) While loop is an entry control loop.
   (j) An exit controlled loop is executed a minimum of one time.
2. Fill the blanks in the following statements.
   (a) The ____________ Function is used to display the output of the screen.
   (b) The ____________ header file contains mathematical functions.
   (c) A variable can be made constant by using the keyword ____________.
   (d) ___________ is the increment operator.
   (e) ?: Operator known as ___________ operator.

(f) The ___________ statement is used to immediately exit from the loop.

(g) The first part of for loop declaration is ___________ .

(h) The do-while is known as ___________ control loop.

(i) A for loop within a for loop is known as ___________ loop.

(j) while (1) is known as ___________ loop.

3. Programming problems

(a) C "Hello, World!" Program.

(b) C Program to Print an Integer (Entered by the User).

(c) C Program to Add Two Integers.

(d) C Program to Swap Two Numbers.

(e) C Program to find Factorial of a Number.

(f) C Program to Make a Simple Calculator Using switch...case.

(g) C Program to Generate Multiplication Table.

(h) C Program to Check Whether a Number is Prime or Not.

(i) C Program to Check Armstrong Number.

Answer

1.

(a)T (b)F (c)F (d)T (e)T (f)T (g)F (h)F (i)T

2.

(a)printf( )	(b)math.h	(c)constant	(d)++
(e)ternary	(f)break	(g)initialization	(h)exit
(i)nested	(j)infinite		

8.8.1 Jumping out of Loops

While executing any loop, it becomes necessary to skip a part of the loop or to leave the loop as soon as certain condition becomes true, which is called jumping out of loop. C language allows jumping from one statement to another within a loop as well as jumping out of the loop.

There are two keywords in C language to Jumping out or Break Loop.

- break statement
- continue statement

Break Statement

When break statement is encountered inside a loop, the loop is immediately excited and the program continues with the statement immediately following the loop.

```
While (condition check)
    {
    Statement – 1;
    Statement – 2 ;
    if (some condition)
    {
        break;
    }
    Statement – 3 ;
    Statement – 4 ;
    }
        jumps out of the loop no matter how many cycles are left, loop is excited
```

Continue Statement

It causes the control to go directly to the test condition and then continue the loop process. On encountering continue, curson leave the current cycle of loop, and starts with the next cycle.

```
while (condition check)
{
    Statement – 1 ;
    Statement – 2 ;
If (some condition)
{
    Continue ;
jumps to the
next cycle directly.
    Statement – 3 ;
    Statement – 4 ;
    }
```

Not executed for the cycle of loop in which continue is executed.

8.8.2 Goto Statement transfer of control branching within a loop

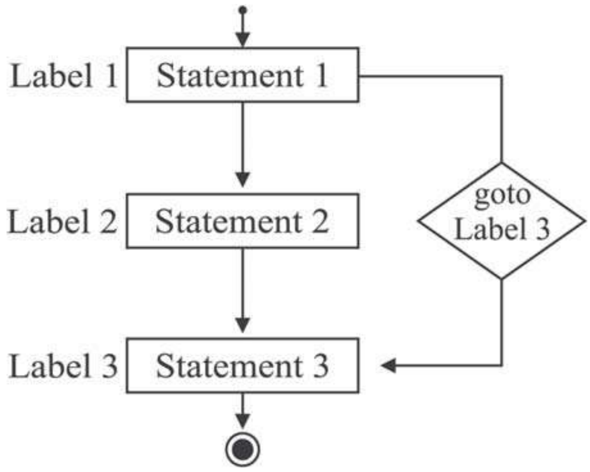
C supports a unique form of a statement that is the goto statement which is used to branch unconditionally within a program from one point to another. Although it is not a good habit to use the goto statement in C, there may be some situations where the use of the goto statement might be desirable.

The goto statement is used by programmers to change the sequence of execution of a C program by shiffing the control to a different part of the same program.

The general form of the goto statement is :

goto level :

A lebel is an identifier required for goto statement to a place where the branch is to be made.



8.9 ❒ Summary

This chapter has introduced many of the basics of the language although informally. Functions in particular, from the basic building block for c. Art 8.4 provides a full descriptions of this fundamental objects, but you should by now understand enough about them to follow their informal use in the intervening material.

8.10 ❒ Exerceises

1. History of C

   Why we use C programming language ?
2. Procedure to create a Program in C Programming Language.
3. What is the importance of header files ?

8.1 ❒ References

1. Mathematical Methods for physicists, 6th edition : Georg B Arfken, Hans J Weber
2. Mathematical methods in the physical sciences (3rd edition) : Marry L Boas
3. Matrices and Tensors in physics (third edition) : A. W Joshi

4. Mathematical physics (first edition) : Dr. Binoy Bhattacharyya
5. Fundamentals of mathematical physics (sixth edition) : Dr. A. B Gupta
6. Vector analysis (second analysis) : Seymour Lipschutz, Dennis Spellman, Murray R. Spiegel
7. Let us C (15th edition) : Yashavant Kanetkar
8. C in depth (third edition) : Dipali Srivastava and S. K. Srivastava
9. Programming with C (second edition) : Brian W Kerninghan and Dennis Ritchie
10. C : The complete reference (fourth edition) : Herbart Schildt
11. C for undergraduates (unpublished) : Debashis Guha
12. Introduction to C (unpublished) : Rabishankar Chottopadhyay

Notes