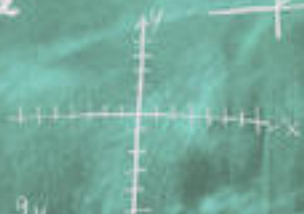


$$X_{1/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

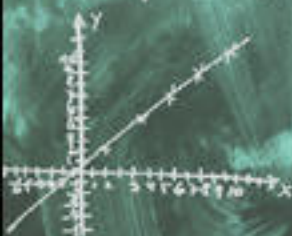


$$X^2 + pX + q = 0$$



$$\begin{aligned} X &= c - 2y \\ X + a &= b \\ f'(x) &= \tan x \end{aligned}$$

$$f(x) = S$$



Fundamentals of Mathematical Physics

Tarasankar Nag

Fundamentals of Mathematical Physics

About the Author



Dr. Tarasankar Nag graduated from the erstwhile Presidency College (now Presidency University) and obtained his Master's degree in Science from Calcutta University. He carried out various researches at Saha Institute of Nuclear Physics and obtained his PhD from Calcutta University. After a brief stay at Saha Institute of Nuclear Physics as a postdoctoral fellow, he joined West Bengal Educational Service in 1983. After rendering his services for more than 25 years in different government colleges, he retired from the erstwhile Presidency College in 2012. Since his retirement, he has been teaching as a guest professor at the Presidency University and Lady Brabourne College. He has carried out various researches in the fields of Plasma Physics, Nonlinear Dynamics & Chaos, and Quantum Mechanics. His current fields of interest include General Relativity and Cosmology. He has 12 research papers to his credit in various national and international journals.

Fundamentals of Mathematical Physics

Tarasankar Nag

(Retd.) Associate Professor, W.B.E.S.,
Presidency College (Erstwhile)



McGraw Hill Education (India) Private Limited
CHENNAI

McGraw Hill Education Offices

Chennai New York St Louis San Francisco Auckland Bogotá Caracas
Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal
San Juan Santiago Singapore Sydney Tokyo Toronto



McGraw Hill Education (India) Private Limited

Published by McGraw Hill Education (India) Private Limited
444/1, Sri Ekambara Naickr Industrial Estate, Alapakkam, Porur, Chennai-600 116

Fundamentals of Mathematical Physics

Copyright © 2017 by McGraw Hill Education (India) Private Limited.

No part of this publication may be reproduced or distributed in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise or stored in a database or retrieval system without the prior written permission of the publishers. The program listings (if any) may be entered, stored and executed in a computer system, but they may not be reproduced for publication.

This edition can be exported from India only by the publishers,
McGraw Hill Education (India) Private Limited.

ISBN (13 digit): 978-93-5260-562-0

ISBN (10 digit): 93-5260-562-4

Managing Director: *Kaushik Bellani*

Director—Science & Engineering Portfolio: *Vibha Mahajan*

Lead—UG Portfolio: *Suman Sen*

Content Development Lead: *Shalini Jha*

Specialist—Product Development: *Amit Chatterjee*

Production Head: *Satinder S Baveja*

Copy Editor: *Taranpreet Kaur*

Assistant Manager—Production: *Anuj K Shrivastava*

General Manager—Production: *Rajender P Ghansela*

Manager—Production: *Reji Kumar*

Information contained in this work has been obtained by McGraw Hill Education (India), from sources believed to be reliable. However, neither McGraw Hill Education (India) nor its authors guarantee the accuracy or completeness of any information published herein, and neither McGraw Hill Education (India) nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that McGraw Hill Education (India) and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

Typeset at Tej Composers, WZ 391, Madipur, New Delhi 110 063 and printed at

Cover Printer:

Visit us at: www.mheducation.co.in

Preface

A sound grasp over the basic principles of Mathematical Physics is essential for pursuing a bachelor course in Physics. There is no dearth of good books that give an adequate coverage of the syllabus. When the esteemed publishing house, McGraw Hill Education (India), approached me for writing a book on the same topic, I felt extremely hesitant. However, after going through the books used by the students, I felt that many popular books lacked plenty of intricate points needed for a comprehensive understanding of the subject. So, when I finally undertook the task of writing this book, I decided that I would try to highlight the finer details of the subject and emphasise the applications to the physical problems as far as possible.

This book is designed to cover the B.Sc. level syllabi of Calcutta University and other major Universities of Eastern Indian. The problems given in the chapters are constructed in a way to enable the students to have an in-depth understanding of the subject. Multiple-choice questions are provided in all the chapters, with a view of making the students well-versed in the basic understanding of the concepts. Appendix given at the end of each chapter is added to supplement the discussions in the main text. The discussion on the Convergence of Integrals at the end of Chapter 5 deserves mention in this regard. The answers to most of the questions are also provided.

Acknowledgements

I am indebted to Prof. Avijit Lahiri and Prof. Ajay Chaudhury for many stimulating discussions, and the reviewers for their useful suggestions. Mention must be made of Ms Prerana Biswas, Mr Sanwar Molla, and my erstwhile students, who provided me with the question papers that enhanced my idea about the standard of questions set

in the examinations. Deepest thanks are due to the management of McGraw Hill Education India, particularly Mr Amit Chatterjee, for rendering me valuable assistance in the preparation of the text. I shall be very delighted to receive feedback from the academic community.

Tarasankar Nag

Publisher's Note:

McGraw Hill Education (India) invites suggestions and comments, all of which can be sent to *info.india@mheducation.com* (kindly mention the title and author name in the subject line).

Piracy-related issues may also be reported.

Contents

Preface

v

1. Functions of Several Variables, Partial Derivatives **1.1–1.36**

- 1.1 Introduction 1.1
- 1.2 Some Useful Relations 1.3
- 1.3 Taylor's Series, Maxima and Minima 1.6
 - Solved Problems* 1.11
- 1.4 Multiple Integrals 1.20
 - Solved Problems* 1.23
 - Summary* 1.24
 - Multiple Choice Questions* 1.25
 - Answers to MCQs* 1.27
 - Exercise* 1.27
 - Hints and Answers to Exercise Questions* 1.29
 - Appendix* 1.32

2. Vector Analysis **2.1–2.42**

- 2.1 Introduction 2.1
- 2.2 Transformation of Axes, Polar Vector, Axial Vector 2.4
- 2.3 Vector Equations of Lines, Curves, Surfaces 2.6
- 2.4 Differentiation of Scalar and Vector Fields 2.7
- 2.5 Integrals of Vectors 2.9
- 2.6 Divergence of a Vector Field; Gauss Divergence Theorem 2.13
- 2.7 Curl of a Vector Field; Stokes Theorem 2.16
- 2.8 Corollaries from Gauss Theorem and Stokes Theorem 2.18
- 2.9 Levi-Civita Symbol and Vector Identities 2.20
 - Solved Problem* 2.21
- 2.10 Vector Relations in Curvilinear Coordinates 2.25
 - Solved Problems* 2.30

<i>Summary</i>	2.34
<i>Multiple Choice Questions</i>	2.34
<i>Answers to MCQs</i>	2.37
<i>Exercise</i>	2.37
<i>Hints and Answer to Exercise Questions</i>	2.40
<i>Appendix</i>	2.41

3. Linear Vector Space and Matrices

3.1–3.34

3.1	Definition of Vector Space	3.1
3.2	Subspace	3.2
3.3	Inner Product; Linear Independence	3.2
3.4	Gram-Schmidt Orthogonalisation Method	3.4
	<i>Solved Problems</i>	3.5
3.5	Linear Operators, Matrices	3.7
3.6	Matrix Representation, Change of Basis	3.8
3.7	Different Types of Representations	3.9
3.8	Cayley-Hamilton's Theorem	3.15
3.9	Rank, Linear Algebraic Equations and Matrices	3.19
3.10	Infinite Dimensional Vector Space	3.20
3.11	Application	3.21
	<i>Solved Problems</i>	3.21
	<i>Summary</i>	3.26
	<i>Multiple Choice Questions</i>	3.27
	<i>Answers to MCQs</i>	3.29
	<i>Exercise</i>	3.29
	<i>Hints and Answers to Exercise Questions</i>	3.31
	<i>Appendix</i>	3.34

4. Tensor Analysis

4.1–4.22

4.1	Introduction	4.1
4.2	Some Important Properties	4.2
4.3	(A) Non-orthogonal Curvilinear Coordinate System	4.4
	(B) Contravariant and Covariant Components	4.5
4.4	Cartesian Tensors, Pseudotensor	4.6
4.5	Covariant Derivative	4.8
4.6	Properties of Christoffel Symbols	4.9
4.7	Tensor Calculus in a General Coordinate System	4.12
	<i>Solved Problems</i>	4.15
	<i>Summary</i>	4.18
	<i>Multiple Choice Questions</i>	4.19
	<i>Answers to MCQs</i>	4.21
	<i>Exercise</i>	4.21
	<i>Hints and Answers to Exercise Questions</i>	4.22

5. Infinite Series**5.1–5.31**

- 5.1 Introduction 5.1
- 5.2 Convergence of a Series of Numbers 5.2
- 5.3 Convergence of a Series of Functions 5.5
- 5.4 Power Series 5.8
- 5.5 Convergence in the Mean 5.9
- 5.6 Summability of a Divergent Series 5.10
- 5.7 Combination of Two Infinite Series 5.11
- 5.8 Asymptotic Series 5.11
- 5.9 Application to Physical Problems 5.12
 - Solved Problems* 5.14
 - Summary* 5.19
 - Multiple Choice Questions* 5.21
 - Answers to MCQs* 5.24
 - Exercise* 5.25
 - Hints and Answers to Exercise Questions* 5.26
 - Appendix* 5.27

6. Complex Variables**6.1–6.26**

- 6.1 Introduction 6.1
- 6.2 Integral Theorems 6.4
- 6.3 Cauchy's Integral Formula 6.5
- 6.4 Laurent's Expansion, Residue 6.8
- 6.5 Theorem of Residues 6.11
 - Solved Problems* 6.12
 - Summary* 6.17
 - Multiple Choice Questions* 6.19
 - Answers to MCQs* 6.21
 - Exercise* 6.21
 - Hints and Answers to Exercise Questions* 6.22
 - Appendix* 6.24

7. Ordinary Differential Equations, Special Functions**7.1–7.55**

- 7.1 Introduction, First Order Ordinary Differential Equations 7.1
- 7.2 Linear Second Order Differential Equation 7.3
- 7.3 Linear Second Order Differential Equations with Constant Coefficients 7.5
- 7.4 Series Solution Method 7.10
- 7.5 Legendre Differential Equation 7.11
- 7.6 Generating Function 7.13
- 7.7 Associated Legendre Polynomial 7.16
- 7.8 Applications of the Legendre Polynomials to Physical Problems 7.17
- 7.9 Hermite Differential Equation 7.18
- 7.10 Bessel Differential Equation 7.20

- 7.11 Equation Having an Irregular Singular Point 7.26
- 7.12 Beta Function, Gamma Function, Dirac Delta Function 7.27
- 7.13 System of Simple Coupled Differential Equations 7.32
 - Solved Problems* 7.33
 - Summary* 7.42
 - Multiple Choice Questions* 7.44
 - Answers to MCQs* 7.48
 - Exercise* 7.48
 - Hints and Answers to Exercise Questions* 7.51
 - Appendix* 7.53

8. Sturm–Liouville Problem and Green’s Function

8.1–8.12

- 8.1 Boundary-Value Problem and Eigenfunctions 8.1
- 8.2 Inhomogeneous Differential Equations, Green’s Function 8.4
 - Solved Problems* 8.6
 - Summary* 8.9
 - Multiple Choice Questions* 8.9
 - Answers to MCQs* 8.10
 - Exercise* 8.10
 - Hints and Answers to Exercise Questions* 8.11

9. Fourier Series and Integral Transforms

9.1–9.21

- 9.1 Introduction, Fourier Expansion, Parseval’s Theorem 9.1
 - Solved Problems* 9.4
- 9.2 Multiple Fourier Series 9.6
- 9.3 Fourier Transform 9.7
- 9.4 Fourier Transform of Derivatives 9.8
- 9.5 Laplace Transform 9.10
 - Solved Problems* 9.13
 - Summary* 9.16
 - Multiple Choice Questions* 9.18
 - Answers to MCQs* 9.19
 - Exercise* 9.19
 - Hints and Answers to Exercise Questions* 9.20

10. Partial Differential Equations

10.1–10.27

- 10.1 Introduction 10.1
- 10.2 First Order Partial Differential Equations 10.1
- 10.3 Second Order Partial Differential Equation 10.3
- 10.4 Selected Partial Differential Equations 10.5
- 10.5 Inhomogeneous Partial Differential Equation 10.20
 - Summary* 10.22
 - Multiple Choice Questions* 10.22
 - Answers to MCQs* 10.24
 - Exercise* 10.24

Hints and Answers to Exercise Questions 10.26
Appendix 10.27

11. Probability and Statistics

11.1–11.26

- 11.1 Introduction 11.1
- 11.2 Joint Probability, Conditional Probability Etcetra 11.2
- 11.3 Random Variables, Distribution Function Etcetra 11.4
- 11.4 Some Standard Distributions 11.7
- 11.5 Application 11.11
- 11.6 Theory of Errors 11.12
 - Solved Problems* 11.16
 - Summary* 11.20
 - Multiple Choice Questions* 11.21
 - Answers to MCQs* 11.23
 - Exercise* 11.23
 - Hints and Answers to Exercise Questions* 11.25
 - Appendix* 11.26

12. Group Theory

12.1–12.25

- 12.1 Introduction 12.1
- 12.2 Multiplication Table, Isomorphism 12.2
- 12.3 Section 12.3
- 12.4 Representation of Groups 12.5
- 12.5 Character Table 12.10
- 12.6 Infinite Groups 12.12
- 12.7 Applications 12.18
 - Solved Problems* 12.20
 - Summary* 12.22
 - Exercise* 12.23
 - Hints and Answers to Exercise Questions* 12.25

13. Linear Integral Equations

13.1–13.11

- 13.1 Introduction 13.1
- 13.2 Degenerate Kernel 13.3
- 13.3 Iterative Solution and other Methods 13.4
- 13.4 Schmidt-Hilbert Theory 13.6
 - Solved Problems* 13.7
 - Summary* 13.9
 - Exercise* 13.9
 - Hints and Answers to Exercise Questions* 13.10

Bibliography

B.1–B.4

Functions of Several Variables, Partial Derivatives

CHAPTER

1

1.1 INTRODUCTION

Functions of several variables are ubiquitous in the everyday world. The objects we encounter in every sphere irrespective of whether it belongs to physical science, chemistry, economics etc. depend on more than one variables for their evolution. We in this chapter discuss various operations on a function of several variables resembling those on the function of a single variable, viz., differentiation, integration, finding out the extrema etc. The chapter ends with a brief discussion of the calculus of variations. The discussions on the exact and the inexact differentials are excluded because of their inclusions in the chapter on differential equations.

Limit and Continuity

The problem of finding out the limit of a function at a point is similar to that of a single variable with the exception that here the point in a multidimensional space can be approached along different paths and hence the single-valuedness of the limiting value has to be ensured for different paths in order that the limit exists at the point. [An identical situation arises in the case of the function of a complex variable which is a particular case of a function of two variables. For a detailed discussion refer to chapter-6]. The ϵ, δ definition of the limit tells us that for the limit to be defined at a point (x_0, y_0) , $|f(x, y) - l| < \epsilon$ for $|x - x_0| < \delta_1$, $|y - y_0| < \delta_2$ (the inequality signs in $|x - x_0|$, $|y - y_0|$ may be replaced by $\leq$ sign in some cases), where δ_1, δ_2 will depend in general on x_0, y_0, ϵ . For example, the functions e^{x+y} , $F(x+y)$ (F being any polynomial in $x+y$) etc. have well-defined limits at any point (x, y) in the two-dimensional plane.

As an exceptional case we may consider $f(x, y) = e^{-\frac{1}{(x^2-y^2)}}$, and its limit at $(0, 0)$. Let us take $y = 0$ and allow x to go to zero. Clearly $f(x, y) \rightarrow 0$ in this case. But if we take on the contrary $x = 0$ and let $y \rightarrow 0$ along the positive y -axis, $f(x, y) \rightarrow \infty$. Similarly, depending on the path of approach the limit may have different values.

In the case of the continuity of a function a similar line of reasoning applies and hence the same test of uniqueness has to be made. The rules regarding the addition, the difference, the product the division of the limits and the continuity of the functions apply equally to the functions of two variables also. These lead to the result that the limit of the sum, difference or the product = sum, difference of the product of the limits etc. provided the other conditions regarding the existence of the limits are satisfied. For example,

$$\lim_{\substack{x \rightarrow a \\ y \rightarrow b}} [f_1(x, y) \pm f_2(x, y) \pm f_3(x, y) \pm \dots] = \lim_{\substack{x \rightarrow a \\ y \rightarrow b}} f_1(x, y) \pm \lim_{\substack{x \rightarrow a \\ y \rightarrow b}} f_2(x, y) \pm \dots,$$

$$\lim_{\substack{x \rightarrow a \\ y \rightarrow b}} f_1(x, y) \cdot f_2(x, y) = \lim_{\substack{x \rightarrow a \\ y \rightarrow b}} f_1(x, y) \cdot \lim_{\substack{x \rightarrow a \\ y \rightarrow b}} f_2(x, y),$$

$$\lim_{\substack{x \rightarrow a \\ y \rightarrow b}} \frac{f_1(x, y)}{f_2(x, y)} = \frac{\lim_{\substack{x \rightarrow a \\ y \rightarrow b}} f_1(x, y)}{\lim_{\substack{x \rightarrow a \\ y \rightarrow b}} f_2(x, y)} \left(\text{if } \lim_{\substack{x \rightarrow a \\ y \rightarrow b}} f_2(x, y) \neq 0 \right),$$

Derivatives—Since the function depends on more than one variable the change in the magnitude of the function as the arguments change by infinitesimally small amounts can be written for the case of two variables as,

$$\Delta f = f(x + \Delta x, y + \Delta y) - f(x, y)$$

$$\text{or,} \quad \Delta f = f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y) + f(x + \Delta x, y) - f(x, y). \quad (1.1)$$

In analogy to the definition of the derivatives of the function of a single variable we define,

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}, \quad \frac{\partial f}{\partial y} = \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}$$

$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ are called the partial derivatives of f with respect to x and y respectively. (These are sometimes

symbolised by f_x, f_y respectively).

Thus, the relation (1.1) can be written as, (in the limits $\Delta x, \Delta y \rightarrow 0$)

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \quad (dx, dy, df \text{ are called the first order differentials})$$

For a function $f(x_1 \dots x_n)$ the total first order differential df can be written as

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n.$$

The second order partial derivatives are defined at (x, y) by

$$\frac{\partial^2 f}{\partial x^2} \text{ (or } f_{xx}) = \lim_{h \rightarrow 0} \frac{f_x(x + h, y) - f_x(x, y)}{h}, \quad \frac{\partial^2 f}{\partial y^2} \text{ (or } f_{yy}) = \lim_{k \rightarrow 0} \frac{f_y(x, y + k) - f_y(x, y)}{k}$$

$$f_{xy} = \frac{\partial^2 f}{\partial x \partial y} = \lim_{h \rightarrow 0} \frac{f_y(x+h, y) - f_y(x, y)}{h}, \quad \frac{\partial^2 f}{\partial y \partial x} = \lim_{k \rightarrow 0} \frac{f_x(x, y+k) - f_x(x, y)}{k} = f_{yx}$$

The last two derivatives merit more discussion. Writing the expressions for f_x, f_y respectively in full we obtain,

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y} &= \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \frac{(f(x+h, y+k) - f(x+h, y)) - [f(x, y+k) - f(x, y)]}{hk} \\ &= \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \frac{[f(x+h, y+k) - f(x+h, y) - f(x, y+k) + f(x, y)]}{hk} \end{aligned}$$

or,

$$\frac{\partial^2 f}{\partial y \partial x} = \lim_{k \rightarrow 0} \lim_{h \rightarrow 0} \frac{[f(x+h, y+k) - f(x+h, y) - f(x, y+k) + f(x, y)]}{hk}.$$

Thus, two derivatives are equal provided the two limits commute or in other words that the second order derivatives are continuous.

As an example let us take $f(x, y) = xy \left(\frac{x^2 - y^2}{x^2 + y^2} \right)$, for $x, y \neq 0$ and $= 0$ for $x, y = 0$

$$f_x(x, y) = \frac{\partial}{\partial x} \left[xy \left(\frac{x^2 - y^2}{x^2 + y^2} \right) \right] = xy \left(\frac{4xy^2}{(x^2 + y^2)^2} \right) + y \left(\frac{x^2 - y^2}{x^2 + y^2} \right)$$

Or $f_x(x, y) = 0$ at $x = y = 0$.

$$f_y(x, y) = xy \left[-\frac{4x^2 y}{(x^2 + y^2)^2} \right] + x \left(\frac{x^2 - y^2}{x^2 + y^2} \right) = 0 \text{ at } x, y = 0$$

$$f_{yx}(0, 0) = \lim_{k \rightarrow 0} \frac{f_x(0, k) - f_x(0, 0)}{k} = \lim_{k \rightarrow 0} -\frac{k}{k} = -1$$

$$f_{xy}(0, 0) = \lim_{h \rightarrow 0} \frac{f_y(h, 0) - f_y(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = 1$$

Thus clearly here $f_{xy}(0, 0) \neq f_{yx}(0, 0)$.

1.2 SOME USEFUL RELATIONS

Let us start with an equation $f(x, y, z) = 0$. We can use this equation to express x and y as functions of y, z and z, x respectively given by $x = x(y, z), y = y(z, x)$,

$$\text{Or } dx = \left(\frac{\partial x}{\partial y} \right)_z dy + \left(\frac{\partial x}{\partial z} \right)_y dz, \quad dy = \left(\frac{\partial y}{\partial x} \right)_z dx + \left(\frac{\partial y}{\partial z} \right)_x dz$$

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

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

Since this relation is valid for arbitrary dx and dz . We put first $dz = 0$ to get $\left(\frac{\partial x}{\partial y} \right)_z \left(\frac{\partial y}{\partial x} \right)_z = 1$ (1.2)

Putting $dx = 0$ we obtain, $\left(\frac{\partial x}{\partial y} \right)_z \left(\frac{\partial y}{\partial z} \right)_x \left(\frac{\partial z}{\partial x} \right)_y = -1$.

These relations are extensively used in thermodynamics. Thus, if $f(P, V, T) = 0$ denotes the equation of state of a gas, we can write for a thermodynamic change $\left(\frac{\partial P}{\partial V} \right)_T \left(\frac{\partial V}{\partial T} \right)_P \left(\frac{\partial T}{\partial P} \right)_V = -1$.

Chain rule—Suppose, $f = f(x, y, z)$, where x, y, z are functions of another variable t say we get,

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt}.$$

This relation can be extended to a function f of n variables.

Similarly if $f = f(x_1, \dots, x_n)$, where $x_1, \dots, x_n$ are functions of $u_1, u_2, \dots, u_n$ say, we obtain,

$$\frac{\partial f}{\partial u_i} = \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial u_i} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial u_i} + \dots + \frac{\partial f}{\partial x_n} \frac{\partial x_n}{\partial u_i}, \text{ for } i = 1, 2, \dots, n$$

Example Let $f = f(x, y)$ and suppose we go over to polar coordinates (r, θ) given by $x = r \cos \theta, y = r \sin \theta$,

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

Thus,

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = \cos \theta \frac{\partial f}{\partial x} + \sin \theta \frac{\partial f}{\partial y}.$$

$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = -r \sin \theta \frac{\partial f}{\partial x} + r \cos \theta \frac{\partial f}{\partial y}.$$

$$\frac{\partial r}{\partial x} = \frac{\partial}{\partial x} (\sqrt{x^2 + y^2}) = \cos \theta, \frac{\partial \theta}{\partial x} = -\frac{\sin \theta}{r},$$

$$\frac{\partial r}{\partial y} = \sin \theta, \frac{\partial \theta}{\partial y} = \frac{\cos \theta}{r}$$

$$\frac{\partial f}{\partial x} = \cos \theta \frac{\partial f}{\partial r} - \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta}, \frac{\partial f}{\partial y} = \sin \theta \frac{\partial f}{\partial r} + \frac{\cos \theta}{r} \frac{\partial f}{\partial \theta}$$

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

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial r} \left(\frac{\cos \theta}{r} \frac{\partial f}{\partial \theta} + \frac{\sin \theta}{r} \frac{\partial f}{\partial r} \right) \sin \theta + \frac{\partial}{\partial \theta} \left(\frac{\cos \theta}{r} \frac{\partial f}{\partial \theta} + \frac{\sin \theta}{r} \frac{\partial f}{\partial r} \right) \left(\frac{\cos \theta}{r} \right).$$

Carrying out the differentiations and adding we get,

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

The derivatives $\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \theta}$ indicate the rates of change of the function f in the direction of r and θ increasing respectively. These are called the directional derivatives of f along those directions.

Similarly, for spherical polar coordinates (r, θ, ϕ) related to (x, y, z) by $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$, $z = r \cos \theta$ respectively we can write,

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2}.$$

Note: Let us take the example of a function $f(x, y) = \frac{xy}{x^2 + y^2}$, $(x, y) \neq (0, 0)$ and $= 0$ otherwise. If we calculate the limit of this function at $(0, 0)$ as we approach the point along a line $y = mx$ we obtain,

$$\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} f(x) = \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{mx^2}{x^2(1+m^2)} = \lim_{x \rightarrow 0} \frac{m}{(1+m^2)} \text{ and clearly depends on } m.$$

Thus, no unique limit of the function exists at $(0, 0)$.

$$\text{But } f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0, f_y(0, 0) = \lim_{k \rightarrow 0} \frac{f(0, k) - f(0, 0)}{k} = 0.$$

Thus, the partial derivatives exist at $(0, 0)$ though the limiting value of the function is not meaningful at $(0, 0)$. This result is characteristic of a function of many variables.

1.2.1 Implicit Differentiation

Suppose we have an equation of the form $f(x, y(x)) = 0$. The equation is such that it is not possible to express y explicitly as a function of x from the equation above. Differentiating with respect to x we get,

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx} = 0, \text{ yielding } \frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}.$$

If we have an equation of the form $F(x, y(x), z(x)) = 0$, we get, $\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{dy}{dx} + \frac{\partial F}{\partial z} \cdot \frac{dz}{dx} = 0$. Thus to find out $\frac{dy}{dx}$ and $\frac{dz}{dx}$ separately we need another relation of the form $\phi(x, y(x), z(x)) = 0$ from which we shall obtain,

$$\frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial y} \frac{dy}{dx} + \frac{\partial \phi}{\partial z} \cdot \frac{dz}{dx} = 0$$

Combining the above two equations we can find out $\frac{dy}{dx}$ and $\frac{dz}{dx}$ at any point x , which are true if

$$\left(\frac{\partial F}{\partial y} \right) \left(\frac{\partial \phi}{\partial z} \right) - \left(\frac{\partial F}{\partial z} \right) \left(\frac{\partial \phi}{\partial y} \right) = \frac{\partial(F, \phi)}{\partial(y, z)} \neq 0.$$

To decide if we can express $(u, v, \dots)$ in term of $(x, y, \dots)$ in a close neighbourhood of the point $(x_0, y_0, \dots)$ we have the implicit function theorem at our disposal [which we will briefly discuss in the **appendix**] when there are two relations of the forms $f(x, y, \dots, u, v, \dots) = 0$ and $\phi(x, y, \dots, u, v, \dots) = 0$ respectively.

1.3 TAYLOR'S SERIES, MAXIMA AND MINIMA

Suppose we are required to find out the local extrema of a function of several variables. This evaluation is important in diverse physical problem particularly in determining the stability of a dynamical system with respect to the perturbation about an equilibrium point. To this end we require the expansion of a function about any point in a series called the Taylor series, the details of which will be provided in a subsequent chapter on infinite series. The Taylor series of the function of a single variable can be written as,

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2!} f''(x_0) + \dots$$

Symbolically, this is written as $f(x) = \left[\exp \left\{ (x - x_0) \frac{d}{dx} \right\} f(x) \right]_{x=x_0}$ which leads to the above series after the formal expansion of the exponential operator.

The Taylor series of a function of two variables gives similarly,

$$f(x, y) = \left[\exp \left\{ (x - x_0) \frac{\partial}{\partial x} + (y - y_0) \frac{\partial}{\partial y} \right\} f(x, y) \right]_{x_0, y_0}.$$

[This can be similarly extended to a function of any number of variables].

Thus we can write for a function of two variables,

$$f(x, y) - f(x_0, y_0) = (x - x_0)f_x + (y - y_0)f_y + \frac{1}{2!} [(x - x_0)^2 f_{xx} + (y - y_0)^2 f_{yy} + 2(x - x_0)(y - y_0)f_{xy}] + \dots,$$

where all the derivatives are assumed to be evaluated at (x_0, y_0) and the relation $f_{xy} = f_{yx}$ is supposed valid. For the point (x_0, y_0) to be a stationary point of $f(x, y)$ (maximum, minimum or saddle point) the first order variation of $f(x, y)$ about the point (x_0, y_0) must vanish. This gives $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$.

Or,

$$\Delta f = f(x, y) - f(x_0, y_0) = \frac{1}{2!} \left[f_{xx} \left(\Delta x + \frac{f_{xy} \Delta y}{f_{xx}} \right)^2 + (\Delta y)^2 \left(f_{yy} - \frac{f_{xy}^2}{f_{xx}} \right) \right], \text{ where } \begin{cases} \Delta x = x - x_0 \\ \Delta y = y - y_0 \end{cases}$$

$$= \frac{1}{2!} \left[f_{yy} \left(\Delta y + \frac{f_{xy} \Delta x}{f_{yy}} \right)^2 + (\Delta x)^2 \left(f_{xx} - \frac{f_{xy}^2}{f_{yy}} \right) \right]$$

If $\Delta f > 0$ (for the point (x_0, y_0) to be a minimum) for arbitrary $\Delta x, \Delta y, f_{xx}, f_{yy} > 0$ and $f_{xx}f_{yy} > f_{xy}^2$ at (x_0, y_0) .

Similarly for (x_0, y_0) to be a maximum, for arbitrary $\Delta x, \Delta y, f_{xx}, f_{yy} < 0$ and $f_{xx}f_{yy} > f_{xy}^2$ at (x_0, y_0) .

For (x_0, y_0) to be a saddle point the condition is that f_{xx}, f_{yy} are of the opposite signs or $f_{xx}f_{yy} < f_{xy}^2$ because then $\Delta f < 0$ for a set of values of $\Delta x, \Delta y$ and > 0 for others.

For a function of n variables, we can write in general with $f_{x_1}^0 f_{x_2}^0 \dots = 0$ at $x_{10}, x_{20}, \dots$

$$\Delta f = \frac{1}{2!} \sum \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right) \Delta x_i \Delta x_j, \text{ the second order derivatives being evaluated at the equilibrium point. } \Delta f$$

can also be expressed as $\Delta f = \sum_{i,j} s_{ij} \Delta x_i \Delta x_j$, where $s_{ij} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_0$. It can be shown that by a proper choice

of axes we can write, $\Delta f = \sum_i \lambda_i (\Delta x_i)^2$, the λ_i 's being the eigenvalues of the matrix S_{ij} (Refer to the chapter on matrices). This if the λ_i 's are all > 0 , $\Delta f > 0$ and hence we have the minima of the function. For λ_i 's all < 0 , $\Delta f < 0$ and the function has a maximum at the equilibrium point. The test fails if all λ_i 's vanish (see **appendix**). This happens for a function of two variables, if f_{xx}, f_{yy}, f_{xy} all vanish simultaneously at the equilibrium point. We have then to go to higher order differentials in order to look for the maxima or the minima of the function.

For $f_{xy}^2 = f_{xx} f_{yy}$, $\Delta f = \frac{1}{2} [\Delta x |f_{xx}|^{1/2} \pm \Delta y |f_{yy}|^{1/2}]^2$ and may be zero for some choice of $\Delta x, \Delta y$ leading to the need for further investigation. The case when the two λ_i 's are of opposite signs corresponds to saddle points. See '**Appendix**' for more detail.

1.3.1 Extrema with Constraints

Often we encounter problems of extremising a function of a certain number of variables, subject to the constraints that a set of functions of the same variables remain constant. For example, the problem of extremisation of the temperature or the potential on a given surface belongs to this category. Below we show how to tackle these types of problems.

Let the function to be extremised be $f(x_1, x_2, \dots, x_n)$ with the conditions that $\phi_1(x_1, \dots, x_n), \dots, \phi_m(x_1, x_2, \dots, x_n)$ are constants.

Thus, we can write $\frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n = 0$ with

$$\sum_{j=1}^n \frac{\partial \phi_i}{\partial x_j} dx_j = 0 \text{ for } i = 1, 2, \dots, m. \quad (1.3)$$

Multiplying the constraints by $\lambda_1, \lambda_2, \dots, \lambda_m$ and adding to the first we obtain, $\sum_{i=1}^n \left[\frac{\partial f}{\partial x_i} + \sum_{l=1}^m \lambda_l \frac{\partial \phi_l}{\partial x_i} \right] dx_i = 0$ for $i = 1, 2, \dots, m$,

We adjust the set of m λ_i 's in such a way that

$$\frac{\partial f}{\partial x_i} + \sum_l \lambda_l \frac{\partial \phi_l}{\partial x_i} = 0 \text{ for } l = 1, 2, \dots, m, \quad (1.4)$$

We can then write, $\sum_{l=m+1}^n \left[\frac{\partial f}{\partial x_l} + \sum_{i=1}^m \lambda_i \frac{\partial \phi_i}{\partial x_l} \right] dx_l = 0$

Since there are m equations of constraint among n variables the set of the remaining $(n-m)$ variables can be chosen to be independent and hence we can write, $\frac{\partial f}{\partial x_l} + \sum_{i=1}^m \lambda_i \frac{\partial \phi_i}{\partial x_l} = 0$. For $l = m+1, \dots, n$. (1.5)

Hence combining (1.4) and (1.5) we can write,

$$\frac{\partial f}{\partial x_l} + \sum_{i=1}^m \lambda_i \frac{\partial \phi_i}{\partial x_l} = 0 \quad \text{for } l = 1, 2, \dots, n. \text{ These equations, } n \text{ in total, combined with the equations of}$$

constraint enable us to find out the m multipliers λ_i (called Lagrange's undetermined multipliers) along with the points $(x_1, \dots, x_n)$ at which the extrema occur. These will be illustrated with examples in the section on 'solved problems'.

Note that instead of introducing the undetermined multipliers we could also use m equations of constraint to eliminate m dx_i 's to obtain ultimately an equation involving $(n-m)$ dx_i 's, which can be considered to be independent of one another, leading to the same results as above. If we have several functions to be extremised with the same set of constraints, we can solve independent variational problems of each of these functions provided they are independent.

1.3.2 Extrema with Respect to Variation of Path (Calculus of Variation)

The problem dealt with here is to find out the extreme value of an integral between two fixed points obtained by variation with respect to the path connecting the two end points.

Let $I = \int_{x_1}^{x_2} f(x, y, y') dx$ be the integral we want to extremise by variation of the path $y(x)$ joining x_1 and x_2 .

Let an infinitesimally varied path be $y(x) + \eta(x)$ where $\eta(x)$, the variation of the path at every x is infinitesimally small and $\eta(x_1) = \eta(x_2) = 0$ (Fig. 1.1).

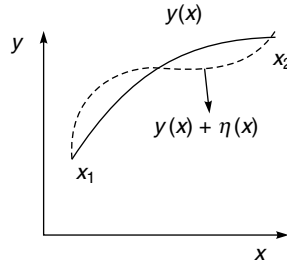


Fig. 1.1 (Dotted line shows the varied path)

$$\begin{aligned} \text{So } I + \delta I &= \int_{x_1}^{x_2} f(x, y + \eta(x), y' + \eta'(x)) dx \\ &= \int_{x_1}^{x_2} \left[f(x, y, y') + \eta(x) \frac{\partial f}{\partial y} + \eta'(x) \frac{\partial f}{\partial y'} \right] dx \end{aligned} \quad (1.6)$$

[Variations of y and y' are made at each fixed x and hence $\delta x = 0$]

$$\text{Now, } \int_{x_1}^{x_2} \eta'(x) \frac{\partial f}{\partial y'} dx = \int_{x_1}^{x_2} \left\{ \frac{d}{dx} \left[\eta(x) \frac{\partial f}{\partial y'} \right] - \eta(x) \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right\} dx$$

$$= \left[\eta(x) \frac{\partial f}{\partial y'} \right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta(x) \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) dx$$

Since $\eta(x) = 0$ at $x = x_1$ and $x = x_2$ we obtain,

$$\int_{x_1}^{x_2} \eta'(x) \frac{\partial f}{\partial y'} dx = - \int_{x_1}^{x_2} \eta(x) \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) dx \text{ thus yielding from (1.6),}$$

$$I + \delta I = I + \int_{x_1}^{x_2} \eta(x) \left[\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right] dx$$

Since for I to be an extremum, the first order variation $\delta I = 0$ and hence η being arbitrary, we obtain,

$$\frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) = \frac{\partial f}{\partial y} \quad (1.7)$$

If I involves the integration of f that depends on y'' as well, we can follow the procedure outlined above obtaining additional terms in Eq. (1.7). (See **Appendix (B)**).

If the function f involves more than one independent functions $y_1(x), y_2(x) \dots$ and their derivatives we obtain by independent variations with respect to each of these functions an equation identical with (1.7) and the number of such variational equations will be the same as that of the independent functions. In classical

mechanics, the trajectory of a system is obtained by extremising the action given by $S = \int_{t_1}^{t_2} L dt$, where L for a

n -degree of freedom system can be written as $L = L(q_1, q_2, \dots, q_n, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_n, t)$. Thus for each q_i , called the generalised coordinate we will end up with an equation similar to (1.7) called the Euler-Lagrange equation

and given by $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = \frac{\partial L}{\partial q_i}$. For a particle moving in three dimensions $L = \frac{1}{2} m \dot{\vec{r}}^2 - V(\vec{r})$ and we obtain,

$m \ddot{\vec{r}} = -\nabla V$ from Eq. (1.7) which clearly agrees with our expectation.

Suppose that in this case also we discuss a constrained variational problem in the form $\int_{x_1}^{x_2} f(x, y, y') dx =$ extremum subject to the condition $\int_{x_1}^{x_2} \phi_i(x, y, y') dx = \text{constant}$ for $i = 1, 2, \dots, m$.

Clearly we have to extremise $\int_{x_1}^{x_2} \left[f + \sum_i \lambda_i \phi_i \right] dx$ or $\int_{x_1}^{x_2} g dx$, where $g = f + \sum_i \lambda_i \phi_i$. The Euler-Lagrange

equations involving the Lagrange's undetermined multipliers are given by,

$$\frac{d}{dx} \left[\frac{\partial g}{\partial y'} \right] = \frac{\partial g}{\partial y} \quad (1.8)$$

Combining the Eq. (1.8) and the relations $\int_{x_1}^{x_2} \phi_i(x, y, y') dx = c_i$

we can find out λ_i 's in terms of the constants c_i 's.

(For example: The problems of determining the path taken by a particle to cover the distance between two fixed points in minimum time or the shape of the surface of minimum area enclosing a given volume belong to these categories).

1.3.3 Differentiation of the Integrals

(i) Let us consider $F(x) = \int_a^b f(x, y) dy$.

We obtain, $F(x) = \phi(x, b) - \phi(x, a)$, where $\phi(x, y) = \int f(x, y) dy$.

$$\text{or, } \frac{dF}{dx} = \frac{\partial \phi(x, b)}{\partial x} - \frac{\partial \phi(x, a)}{\partial x} = \frac{\partial}{\partial x} \left[\int_c^b f(x, z) dz \right] - \frac{\partial}{\partial x} \left[\int_a^c f(x, z) dz \right]$$

[c is any arbitrary constant].

$$\text{or, } \frac{dF}{dx} = \frac{\partial}{\partial x} \left[\int_a^b f(x, z) dz \right] = \int_a^b \frac{\partial f}{\partial x} dz = \int_a^b \frac{\partial}{\partial x} f(x, y) dy$$

(ii) Let $F(x) = \int_{t=\phi(x)}^{t=\psi(x)} f(x, t) dt$

or, $F(x) = g(x, \psi(x)) - g(x, \phi(x))$, where $g(x, t) = \int f(x, t) dt$

$$\begin{aligned} \text{or, } \frac{dF}{dx} &= \frac{d}{dx} [g(x, \psi(x)) - g(x, \phi(x))] \\ &= \frac{\partial}{\partial x} g(x, \psi(x)) - \frac{\partial g(x, \phi(x))}{\partial x} + \frac{\partial g}{\partial \psi} \frac{d\psi}{dx} - \frac{\partial g}{\partial \phi} \frac{d\phi}{dx}. \end{aligned}$$

Clearly $\frac{\partial g}{\partial \psi} = f(x, \psi(x))$, $\frac{\partial g}{\partial \phi} = f(x, \phi(x))$

$$\text{or, } \frac{dF}{dx} = f(x, \psi(x)) \frac{d\psi}{dx} - f(x, \phi(x)) \frac{d\phi}{dx} + \int_{t=\phi(x)}^{t=\psi(x)} \frac{\partial f(x, t)}{\partial t} dt \quad (1.9)$$

We have also there obvious equalities

$$\frac{d}{dx} \int_0^x f(t) dt = f(x), \quad \frac{d}{dx} \int_x^a f(x) dt = -f(x).$$

Note: Additional details regarding the variational principle are given in the appendix.

SOLVED PROBLEMS

Problem 1

(A) Test if the following limits exist

$$(a) \lim_{(x,y) \rightarrow (0,0)} \frac{2y^2}{\sqrt{x^2 + xy}}, \quad (b) \lim_{(x,y) \rightarrow (0,0)} \frac{x^3 + y^3}{x^2 + y^2}$$

$$(c) \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{x^2 + y^2}{2x^2 + 3xy + y^2}, \quad (d) \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{x^2}{x^2 + y^2}$$

(B) Test the continuity of the following functions

$$(a) f(x, y) = \frac{x^3 + y^3}{xy + x^2} \text{ for } x, y \neq 0 \text{ and } = 0 \text{ at } x = \infty; y = 0,$$

$$(b) f(x, y) = \frac{2xy}{x^2 + y^2} \text{ for } x, y \neq 0 \text{ and } = 0 \text{ when } x = y = 0.$$

Solution:

$$(A) (a) \text{ Putting } y = mx, \lim_{(x,y) \rightarrow (0,0)} \frac{2y^2}{\sqrt{x^2 + xy}} = \lim_{(x,y) \rightarrow (0,0)} \frac{2m^2 x}{\sqrt{1+m^2}}. \text{ Clearly the limit is 0 independent of } m.$$

$$(b) \text{ Putting } y = mx, \lim_{(x,y) \rightarrow (0,0)} \frac{x^3 + y^3}{x^2 + y^2} = \lim_{(x,y) \rightarrow (0,0)} \frac{(1+m^3)x}{(1+m^2)} = 0 \text{ irrespective of } m. [\text{when } m \rightarrow \infty, \text{ i.e.,} \\ \text{when we approach along the } y\text{-axis for } x = 0, \text{ the limit is 0. in both these cases}]$$

$$(c) \lim_{(x,y) \rightarrow (0,0)} \frac{x^2 + y^2}{2x^2 + 3xy + y^2} = \frac{1+m^2}{2+3m+m^2} \text{ after putting } y = mx \text{ and is clearly different for different} \\ m. \text{ Hence the limit does not exist at } (0, 0)$$

$$(d) \lim_{(x,y) \rightarrow (0,0)} \frac{x^2}{x^2 + y^2} = \frac{1}{1+m^2}, \text{ putting } y = mx \text{ and if clearly depends on } m. \text{ Thus the limit does not} \\ \text{exist,}$$

$$(B) (a) \text{ Putting } y = mx, \lim_{(x,y) \rightarrow (0,0)} \frac{x^3 + y^3}{xy + x^2} = \left[\frac{x(1+m^3)}{(m+1)} \right]_{(x,y) \rightarrow (0,0)} = 0.$$

 Also $f(0, 0) = 0$ and hence the function is continuous at $(0, 0)$

(b) Putting $y = mx$, $\lim_{(x,y) \rightarrow (0,0)} \frac{2xy}{x^2 + y^2} = \lim_{(x,y) \rightarrow (0,0)} \frac{2m}{1 + m^2}$ and the limit depends clearly on m . Thus

though $f(0, 0) = 0$ the limit of the function does not exist at the point. Hence the function is discontinuous at $(0, 0)$.

Problem 2

The relation between the Cartesian and the spherical polar coordinates is given

by $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$ and $z = r \cos \theta$ respectively. Find out the Jacobian determinant

$$\begin{vmatrix} \frac{\partial r}{\partial x} & \frac{\partial \theta}{\partial x} & \frac{\partial \phi}{\partial x} \\ \frac{\partial r}{\partial y} & \frac{\partial \theta}{\partial y} & \frac{\partial \phi}{\partial y} \\ \frac{\partial r}{\partial z} & \frac{\partial \theta}{\partial z} & \frac{\partial \phi}{\partial z} \end{vmatrix}.$$

Also express $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$ in terms of the derivatives with respect to r, θ, ϕ .

Solution: We have, $r = \sqrt{x^2 + y^2 + z^2}$, $\theta = \cos^{-1} \frac{z}{\sqrt{x^2 + y^2 + z^2}}$, $\phi = \tan^{-1} \frac{y}{x}$

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

$$\frac{\partial \theta}{\partial x} = \frac{xz}{\sqrt{x^2 + y^2} (x^2 + y^2 + z^2)} = \frac{\cos \theta \cos \phi}{r}, \frac{\partial \theta}{\partial y} = \frac{yz}{\sqrt{x^2 + y^2} (x^2 + y^2 + z^2)} = \frac{\cos \theta \sin \phi}{r}$$

$$\frac{\partial \theta}{\partial z} = -\frac{1}{\sqrt{1 - \frac{z^2}{x^2 + y^2 + z^2}}} \left[\frac{1}{\sqrt{x^2 + y^2 + z^2}} - \frac{z^2}{(x^2 + y^2 + z^2)^{3/2}} \right] = -\frac{(x^2 + y^2)^{1/2}}{(x^2 + y^2 + z^2)} = -\frac{\sin \theta}{r}$$

$$\frac{\partial \phi}{\partial x} = \frac{\partial}{\partial x} \left[\tan^{-1} \frac{y}{x} \right] = -\frac{y}{x^2 + y^2} = -\frac{\sin \phi}{r \sin \theta},$$

$$\frac{\partial \phi}{\partial y} = \frac{\partial}{\partial y} \left[\tan^{-1} \frac{y}{x} \right] = \frac{x}{x^2 + y^2} = \frac{\cos \phi}{r \sin \theta}, \frac{\partial \phi}{\partial z} = 0$$

$$\text{Thus } J = \begin{vmatrix} \frac{\partial r}{\partial x} & \frac{\partial \theta}{\partial x} & \frac{\partial \phi}{\partial x} \\ \frac{\partial r}{\partial y} & \frac{\partial \theta}{\partial y} & \frac{\partial \phi}{\partial y} \\ \frac{\partial r}{\partial z} & \frac{\partial \theta}{\partial z} & \frac{\partial \phi}{\partial z} \end{vmatrix} = \begin{vmatrix} \sin \theta \cos \phi & \frac{\cos \theta \cos \phi}{r} & -\frac{\sin \phi}{r \sin \theta} \\ \sin \theta \sin \phi & \frac{\cos \theta \sin \phi}{r} & \frac{\cos \phi}{r \sin \theta} \\ \cos \theta & -\frac{\sin \theta}{r} & 0 \end{vmatrix} = \frac{1}{r^2 \sin \theta}$$

[This proves that at every point except $r = 0$ $J \neq 0$ leading to the result that the (r, θ, ϕ) coordinates are a set of independent coordinates similar to the Cartesian coordinates].

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial f}{\partial r} \cdot \frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial x} + \frac{\partial f}{\partial \phi} \frac{\partial \phi}{\partial x} \\ &= \sin \theta \cos \phi \frac{\partial f}{\partial r} + \frac{\cos \theta \cos \phi}{r} \frac{\partial f}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial f}{\partial \phi}, \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial y} &= \frac{\partial f}{\partial r} \cdot \frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial y} + \frac{\partial f}{\partial \phi} \cdot \frac{\partial \phi}{\partial y} \\ &= \sin \theta \sin \phi \frac{\partial f}{\partial r} + \frac{\cos \theta \sin \phi}{r} \frac{\partial f}{\partial \theta} + \frac{\cos \phi}{r \sin \theta} \frac{\partial f}{\partial \phi}, \quad \frac{\partial f}{\partial z} = -\frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} + \cos \theta \frac{\partial f}{\partial r} \end{aligned}$$

$$\begin{aligned} \text{thus } \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial r} \left[\sin \theta \cos \phi \frac{\partial f}{\partial r} - \frac{\cos \theta \cos \phi}{r} \frac{\partial f}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial f}{\partial \phi} \right] \sin \theta \cos \phi \\ &+ \frac{\partial}{\partial \theta} \left(\sin \theta \cos \phi \frac{\partial f}{\partial r} - \frac{\cos \theta \cos \phi}{r} \frac{\partial f}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial f}{\partial \phi} \right) \frac{\cos \theta \cos \phi}{r} \\ &+ \frac{\partial}{\partial \phi} \left(\sin \theta \cos \phi \frac{\partial f}{\partial r} - \frac{\cos \theta \cos \phi}{r} \frac{\partial f}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial f}{\partial \phi} \right) \frac{\sin \phi}{r \sin \theta} \end{aligned}$$

and after performing the differentiations we get,

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \sin^2 \theta \cos^2 \phi \frac{\partial^2 f}{\partial r^2} + \frac{\cos^2 \theta \cos^2 \phi}{r^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{\sin^2 \phi}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2} \\ &+ \left(\frac{\cos^2 \theta \cos^2 \phi + \sin^2 \phi}{r} \right) \frac{\partial f}{\partial r} - \frac{\sin^2 \phi}{r^2} \cot \theta \frac{\partial f}{\partial \theta} + \frac{2 \sin \phi \cos \phi}{r^2 \sin^2 \theta} \frac{\partial f}{\partial \phi} - \frac{2 \sin \phi \cos \phi}{r} \frac{\partial^2 f}{\partial r^2} \end{aligned}$$

Similarly, we can find out $\frac{\partial^2 f}{\partial y^2}$, $\frac{\partial^2 f}{\partial z^2}$ and after adding we obtain,

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2}.$$

Problem 3

The equations of two surfaces are given by

$$x^2 yz + xy^2 z + xyz^2 = \text{constant and } x^3 + y^3 + z^3 = \text{constant } y = y(x), z = z(x)$$

Find out $\frac{dy}{dx}, \frac{dz}{dx}$ at the points of intersection (x_0, y_0, z_0)

Solution: From the equation of two surfaces we get,

$$\left(2x yz + x^2 z \frac{dy}{dx} + x^2 y \frac{dz}{dx} + y^2 z + 2xy \frac{dy}{dx} z + xy^2 \frac{dz}{dx} + yz^2 + xz^2 \frac{dy}{dx} + 2xyz \frac{dz}{dx} \right) = 0$$

$$\text{or } z(2xy + y^2 + yz) + (x^2 z + 2xyz + xz^2) \frac{dy}{dx} + (x^2 y + xy^2 + 2xyz) \frac{dz}{dx} = 0 \text{ from the first (1.9) (A)}$$

$$\text{and } x^2 + y^2 \frac{dy}{dx} + z^2 \frac{dz}{dx} = 0 \dots (1.9) \text{ (B) from the second.}$$

Solving the two equations we can find out $\frac{dy}{dx}$ and $\frac{dz}{dx}$ at the point of intersection (x_0, y_0, z_0) .

Problem 4

If a function $f(x, y)$ can be written as $f(x, y) = x^n \phi(y/x) = y^n \psi\left(\frac{x}{y}\right)$ where the forms of ϕ and ψ are determined from the form of f , it is called a homogeneous function of x and y of degree n .

Prove that $x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = nf$. Extend it to three variables.

Solution: $f(x, y) = x^n \phi(y/x) = x^n \phi(u)$, where $u = y/x$. Taking $f = x^n \phi(u)$

$$\text{get, } \frac{\partial f}{\partial x} = nx^{n-1} \phi\left(\frac{y}{x}\right) + x^n \phi'(u) \frac{\partial u}{\partial x} = nx^{n-1} \phi(u) - yx^{n-2} \phi'(u)$$

$$\frac{\partial f}{\partial y} = \frac{1}{x} x^n \phi'(u) = x^{n-1} \phi'(u)$$

$$\text{or, } x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = nx^n \phi(u) = nf.$$

For 3 variables the homogeneous function is defined as,

$$f(x, y, z) = x^n \phi\left(\frac{y}{x}, \frac{z}{x}\right) = y^n \psi\left(\frac{x}{y}, \frac{z}{y}\right) = z^n \xi\left(\frac{x}{z}, \frac{y}{z}\right)$$

$$\frac{\partial f}{\partial x} = nx^{n-1} \phi(u, v) + x^n \frac{\partial \phi}{\partial u} \cdot \frac{\partial u}{\partial x} + x^n \frac{\partial \phi}{\partial v} \cdot \frac{\partial v}{\partial x} \quad [\text{Taking } f = x^n \phi(u, v)]$$

where $u = y/x$, $v = z/x$

$$\begin{aligned} \text{or, } \frac{\partial f}{\partial x} &= nx^{n-1}\phi(u, v) - yx^{n-2}\frac{\partial \phi}{\partial u} - zx^{n-2}\frac{\partial \phi}{\partial v} \\ \frac{\partial f}{\partial y} &= x^{n-1}\frac{\partial \phi}{\partial u} + x^{n-1}\frac{\partial \phi}{\partial v} \cdot \frac{\partial v}{\partial y} = x^{n-1}\frac{\partial \phi}{\partial v} \left(\because \frac{\partial v}{\partial y} = 0 \right) \\ \frac{\partial f}{\partial z} &= x^{n-1}\frac{\partial \phi}{\partial v}. \end{aligned}$$

$$\text{Thus } x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} + z \frac{\partial f}{\partial z} = nx^n \phi(u, v) = nf.$$

This can clearly be extended to a homogeneous function of any number of independent variables. This is called Euler's theorem of homogeneous functions.

Problem 5

Find out the stationary points of the function $x^3 + x^2y + xy^2 + y^3 - 4x - 4y$ and state their nature.

Solution: The stationary points are given by $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$, where

$$f(x, y) = x^3 + x^2y + xy^2 + y^3 - 4x - 4y.$$

$$\frac{\partial f}{\partial x} = 3x^2 + 2xy + y^2 - 4 = 0 \text{ and } \frac{\partial f}{\partial y} = 3y^2 + 2xy + x^2 - 4 = 0 \quad (1.10)$$

Solving (1.10) we obtain, $2(x^2 - y^2) = 0$ giving $y = \pm x$.

For $y = x$ we obtain, putting in the equation, $\frac{\partial f}{\partial x} = 0$ we get,

$$x^2 = \frac{2}{3} \text{ or } x = \pm \sqrt{\frac{2}{3}}. \text{ For } y = -x \text{ similarly we get, } x^2 = 2 \text{ given } x = \pm \sqrt{2}$$

Thus the stationary points are $\left(\sqrt{\frac{2}{3}}, \sqrt{\frac{2}{3}}\right), \left(-\sqrt{\frac{2}{3}}, -\sqrt{\frac{2}{3}}\right), (\sqrt{2}, -\sqrt{2}), (-\sqrt{2}, \sqrt{2})$.

$$\text{Now } \frac{\partial^2 f}{\partial x^2} = 6x + 2y = 8x \text{ for } y = x, \frac{\partial^2 f}{\partial y^2} = 6y + 2x = 8x \text{ for } y = x$$

$$\text{and } \frac{\partial^2 f}{\partial x \partial y} = 2(y + x) \text{ for } y = x \text{ leading to } \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2} > 0 \text{ and}$$

$$\frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 = 64x^2 - (4x)^2 > 0 \text{ for } y = x = \sqrt{\frac{2}{3}}$$

For $x = y = \sqrt{\frac{2}{3}}$ we get a minimum,

For $x = y = -\sqrt{\frac{2}{3}}$, $\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial y^2} = 8x < 0$ and $\frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} > \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$

leading to a maximum of the function at these values of x, y .

For $x = -y$, $\frac{\partial^2 f}{\partial x^2} = 4x$, $\frac{\partial^2 f}{\partial y^2} = -4x$, $\frac{\partial^2 f}{\partial x \partial y} = 0$.

Thus the point $(\sqrt{2}, -\sqrt{2})$ or $(-\sqrt{2}, \sqrt{2})$ gives rise to a saddle point.

Problem 6 The temperature T at every point on the circumference of a circle whose equation is given by $x^2 + y^2 = 4$ is, $T = k(x^4 + y^4)$, k being a constant. Find out the position corresponding to the maximum T .

Solution: Introducing Lagrange's undetermined multiplier λ we can say that $k(x^4 + y^4) + \lambda(x^2 + y^2) = \text{extremum}$.

The positions of extrema are given by taking the differential of the above equation to get, $(4kx^3 + 2\lambda x)dx + (4ky^3 + 2\lambda y)dy = 0$

We adjust λ so that $4kx^3 + 2\lambda x = 2x(2kx^2 + \lambda) = 0$ and $2y(2ky^2 + \lambda) = 0$

These yield (i) $x = 0$ or $\lambda = -2kx^2$ and (ii) $y = 0$ or $\lambda = -2ky^2$.

Thus we obtain the stationary points as, (a) $x = 0$, $y^2 = 4$, $\lambda = -8k$ or,

(b) $y = 0$, $x^2 = 4$, $\lambda = -8k$ or,

(c) $x^2 = y^2 = 2$, $\lambda = -2kx^2 = -4k$

of these three sets, (a) and (b) give the stationary points corresponding to the hottest points and $T_{\max} = 16k$ in the region $x^2 + y^2 = 4$.

Problem 7 Show that the path joining two points in a plane where the distance between two points is an extremum is a straight line.

Solution: The distance between two points along a path can be written as

$$S = \int_1^2 \sqrt{dx^2 + dy^2} = \int_1^2 \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx.$$

Thus the equation of the path where the distance is an extremum is given by $\frac{d}{dx} \left[\frac{\partial f}{\partial \left(\frac{dy}{dx} \right)} \right] = \frac{\partial f}{\partial x} = 0$,

where $f = \sqrt{1 + \left(\frac{dy}{dx} \right)^2}$

or, $\frac{d}{dx} \left[\frac{\frac{dy}{dx}}{\left[1 + \left(\frac{dy}{dx} \right)^2 \right]^{\frac{1}{2}}} \right] = 0$ giving $\left(\frac{dy}{dx} \right)^2 = c \left[1 + \left(\frac{dy}{dx} \right)^2 \right]$, c being a constant.

This yields, $\left(\frac{dy}{dx} \right) = \sqrt{\frac{c}{1-c}} = A$, a constant.

Thus the path is a straight line of constant slope A . This path is clearly one of the minimum distance because there can be an infinite number of paths between two points of different shapes and no upper limit exists in respect of the path lengths.

Problem 8(a)

Find the shape of the curve in a plane which has the minimum perimeter for a given area.

Solution: The area of a plane curve can be written as $A = \frac{1}{2} \int (x dy - y dx)$ where the area is in the xy -plane and the limits of the integral are fixed up so as to cover the area. [This can be understood by assuming the area to be divided into small triangles so that it can then be written as $\frac{1}{2} \oint r^2 d\theta$, where the polar coordinates $x = r \cos \theta$ and $y = r \sin \theta$ have been used.]

A better proof can be obtained by using Green's theorem in a plane in vector calculus]. If we parametrise

the curve by a parameter α we can write the area as $\frac{1}{2} \int_{\alpha_1}^{\alpha_2} (x\dot{y} - y\dot{x}) d\alpha = A$. The perimeter of the area is

$P = \int_{\alpha_1}^{\alpha_2} (\dot{x}^2 + \dot{y}^2)^{1/2} d\alpha$, where $\dot{x}$ and $\dot{y}$ are equal to $\frac{dx}{d\alpha}$ and $\frac{dy}{d\alpha}$ respectively. Thus, using Lagrange's

method of undetermined multipliers we can convert the problem to that of the minimisation of $P + \lambda A$,

λ being the multiplier. The variational equation yields, $\frac{d}{d\alpha} \left[\frac{\partial}{\partial \dot{x}} (P' + \lambda A') \right] = \frac{\partial}{\partial x} (P' + \lambda A')$ and

$$\frac{d}{d\alpha} \left[\frac{\partial}{\partial \dot{y}} (P' + \lambda A') \right] = \frac{\partial}{\partial y} (P' + \lambda A') \quad (1.11)$$

where

$$P' = (\dot{x}^2 + \dot{y}^2)^{1/2}, \quad A' = \frac{1}{2} (x\dot{y} - y\dot{x})$$

We obtain, $\frac{d}{d\alpha} \left[\frac{\dot{x}}{(\dot{x}^2 + \dot{y}^2)^{1/2}} - \frac{\lambda}{2} y \right] = \frac{\partial}{\partial x} \left(\frac{\lambda}{2} x \dot{y} \right) = \frac{\lambda}{2} \dot{y}$ and

$$\frac{d}{d\alpha} \left[\frac{\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{1/2}} + \frac{\lambda}{2} x \right] = \frac{\partial}{\partial y} \left(-\frac{\lambda}{2} y \dot{x} \right) = -\frac{\lambda}{2} \dot{x} \quad (1.12)$$

Integrating Eq. (1.12) we obtain,

$$\frac{\dot{x}}{(\dot{x}^2 + \dot{y}^2)^{1/2}} - \lambda y = c_1 \text{ and } \frac{\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{1/2}} + \lambda x = c_2, c_1 \text{ and } c_2 \text{ being constants.}$$

Hence, we obtain, $\frac{dy}{dx} = \frac{c_2 - \lambda x}{c_1 + \lambda y} = \frac{c - x}{c' + y}$, where $c = \frac{c_2}{\lambda}$, $c' = \frac{c_1}{\lambda}$.

Thus, integrating the above equation we get, $\left(c'y + \frac{y^2}{2} \right) = cx - \frac{x^2}{2} + d$, d being any constant.

$$\text{as, } x^2 + y^2 - 2cx + 2c'y = 2d$$

$$\text{or, } (x - c)^2 + (y + c')^2 = R^2, \text{ where } R^2 = c^2 + c'^2 + 2d.$$

which gives the equation of a circle.

Similarly it can be shown that the curve of a given perimeter with the maximum enclosed area is a circle. This is left as an exercise for the students.

Problem 8(b)

(a) Find $\frac{d}{dx} \int_{x^2}^{x^3} \frac{\cos xt}{t} dt$

(b) If $y = \int_a^x f(u) e^{i(x-u)} du$, prove that $y'' + y = f''(x) + if(x)$

Hence find out $y_1'' + y_1$ and $y_2'' + y_2$, where $y_1 = \int_a^x f(x) \cos(x - u) du$

and $y_2 = \int_a^x f(x) \sin(x - u) du$ respectively and a is a constant.

(c) Calculate the appropriate integral and find out $\int_0^\infty x^2 e^{-ax} \sin bx dx$ by differentiating the integral.

Solution: (a) We use the relation.

$$\frac{dF}{dx} = f(x, \psi(x)) \frac{d\psi}{dx} - f(x, \phi(x)) \frac{d\phi}{dx} + \int_{t=\phi(x)}^{t=\psi(x)} \frac{\partial f(x, t)}{\partial x} dt, \text{ where } F(x) = \int_{t=\phi(x)}^{t=\psi(x)} f(x, t) dt$$

$$\begin{aligned} \text{or, } \frac{d}{dx} \int_{x^2}^{x^3} \frac{\cos xt}{t} dt &= \frac{\cos x^4}{x^3} (3x^2) - \frac{\cos x^3}{x^2} (2x) - \int_{x^2}^{x^3} \sin xt \, dt \\ &= \frac{3}{x} \cos x^4 - \frac{2}{x} \cos x^3 + \frac{(\cos x^4 - \cos x^3)}{x} \\ &= \frac{4}{x} \cos x^4 - \frac{3}{x} \cos x^3 \end{aligned}$$

$$(b) \, y = \int_a^x f(x) e^{i(x-u)} du$$

$$\text{or, } y' = f(x) + i \int_a^x f(u) e^{i(x-u)} du = f(x) + iy$$

$$\text{or, } y'' = f'(x) + iy' = f'(x) + if(x) - y$$

$$\text{or, } y'' + y = f'(x) + if(x)$$

Taking the real and the imaginary parts of this equation ($f(x) = \text{real}$).

we obtain, $y_1'' + y_1 = f'(x)$ and $y_2'' + y_2 = f(x)$.

$$(c) \text{ Let } I = \int_0^{\infty} e^{-ax} \sin bx \, dx.$$

$$\begin{aligned} \text{or, } I &= \left[-\frac{e^{-ax} \cos bx}{b} \right]_0^{\infty} - \int_0^{\infty} \frac{ae^{-ax} \cos bx}{b} dx \\ &= \frac{1}{b} - \frac{a}{b} \left[-\frac{e^{-ax} \sin bx}{b} \right]_0^{\infty} + \int_0^{\infty} \frac{ae^{-ax} \sin bx}{b} dx \end{aligned}$$

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

$$\text{or, } I = \frac{b}{a^2 + b^2}.$$

Now the required integral $I_1 = \int_0^{\infty} x^2 e^{-ax} \sin bx \, dx = \frac{d^2}{da^2} I$

$$\text{or, } I_1 = \frac{d^2}{da^2} \left[\frac{b}{a^2 + b^2} \right] = \frac{d}{da} \left[-\frac{2ab}{(a^2 + b^2)^2} \right] = -2b \left[\frac{(a^2 + b^2)^2 - 4a^2(a^2 + b^2)}{(a^2 + b^2)^4} \right]$$

$$\text{or, } I_1 = -\frac{2b(b^2 - 3a^2)}{(a^2 + b^2)^3}.$$

1.4 MULTIPLE INTEGRALS

Similar to the integrals of the function of a single variable, we will discuss in this section the integral with respect to several variables. These occur frequently in almost every discipline of physical science. We will deal in this section with the double and triple integrals of a function of two and three variables and the change of variables in the integration required in many cases to facilitate the calculations.

1. Double integrals

The double integral of a function written as $\int_{y=y_1}^{y_2} \int_{x=x_1}^{x_2} f(x, y) \, dx \, dy$ is defined as $\lim_{\substack{\Delta x_i \rightarrow 0 \\ \Delta y_j \rightarrow 0}} \sum_{i,j} f(x_i, y_j) \Delta x_i \Delta y_j$.

In order that the integral exists the limit should be unique and independent of the order of the variables in the sum. In respect of the integration, the integral should be independent of the order of variables in the integration provided the same region is described in the xy -plane. The limits of integration clearly depend on the shape and the size of the area described. Thus for a rectangle of sides $2a$ and $2b$ respectively in the x and the y -directions with the centre at the origin the integrals are to be performed independently along x and y from $-a$ to a to $-b$ to b respectively. But if the boundary curves in the xy -plane be given by the equations $y_1 = f_1(x)$ and $y_2 = f_2(x)$ respectively while x runs from x_1 to x_2 the integral can be written as

$$\int_{x_1}^{x_2} dx \int_{y_1=f_1(x)}^{y_2=f_2(x)} f(x, y) \, dy, \text{ while if the equations of the boundary curves be represented as } x_1 = \phi_1(y) \text{ and } x_2 = \phi_2(y)$$

respectively, where y varies from y_1 to y_2 we can write the same integral as $\int_{y_1}^{y_2} dy \int_{\phi_1(y)}^{\phi_2(y)} f(x, y) \, dx$. Clearly, the

two integrals must be identical for the double integral to exist. For example, in the case of the region being a

circle whose equation is given by $x^2 + y^2 = a^2$ we can write the integral as $\int_{-a}^a dy \int_{-\sqrt{a^2-y^2}}^{\sqrt{a^2-y^2}} f(x, y) \, dx$ or alterna-

tively as $\int_{-a}^a dx \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} f(x, y) \, dy$. For the case of a region bounded by an ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ we can write

the integral as $\int_{-a}^a dx \int_{-\frac{b}{a}\sqrt{a^2-x^2}}^{\frac{b}{a}\sqrt{a^2-x^2}} f(x, y) dy$ or as $\int_{-b}^b dy \int_{-\frac{a}{b}\sqrt{b^2-y^2}}^{\frac{a}{b}\sqrt{b^2-y^2}} f(x, y) dx$.

Here, the region defined above is such that any line parallel to the x or y -axis cuts the curve in two distinct points. But if the region be simply connected but of such a shape that lines parallel to the x or the y -axis intersect the bounding curves in more than two points along one or both the directions the entire region is split up into parts each of which is such that the lines parallel to the x or the y -axis cut the bounding curve in two points only. The previous analysis is then applied and the required double integral is given by the sum of the separately calculated double integrals over these regions.

2. Triple integrals

The triple integral of a function written as $\int_{z=z_1}^{z_2} \int_{y=y_1}^{y_2} \int_{x=x_1}^{x_2} f(x, y, z) dx dy dz$ is defined as $\lim_{\substack{\Delta x_i \rightarrow 0 \\ \Delta y_j \rightarrow 0 \\ \Delta z_k \rightarrow 0}} \sum_{ijk} f(x_i, y_j, z_k)$

$\Delta x_i, \Delta y_j, \Delta z_k$. The uniqueness of the limit and its independence of the order of the variables must be ensured in order that the integral exists. In terms of the integral it can be said that the result of integration should be independent of the order of integration variables if the region covered in space remains the same. The statements made in the case of double integrals apply equally well in this case also. For example if the region be of such a shape that the lines parallel to the axes cut the bounding surface in more than two points the region is to be divided into suitable parts and the required integral is the sum of the separate integrals over each of the parts.

3. Change of variables

A particular set of coordinates may be best suited for a selected type of calculations. For example, the use of plane polar coordinates will be convenient for calculations of the integral over a region bounded by a circle while the spherical polar coordinates serve the purpose best for calculations in problems involving spherical symmetry. The pertinent question in this respect is that of conversion from the original set of coordinates to one that is more suitable in regard to the computation of multiple integrals. Let the original set of coordinates in the computation of double integral be (u_1, u_2) . An infinitesimally small area in (u_1, u_2) coordinates may be represented by a parallelogram and represented vectorially by $d\vec{u}_1 \times d\vec{u}_2$. Let us change over to a more

suitable pair of coordinates (v_1, v_2) . An infinitesimally small change du_1 gives rise to a change $\frac{\partial v_1}{\partial u_1} du_1$ and $\frac{\partial v_2}{\partial u_1} du_1$ along the v_1 and the v_2 directions respectively. Similarly for the small change du_2 the corresponding changes along the v_1 and the v_2 directions are $\frac{\partial v_1}{\partial u_2} du_2$ and $\frac{\partial v_2}{\partial u_2} du_2$ respectively. Thus the elementary area in

(v_1, v_2) coordinates is given by $|d\vec{v}_1 \times d\vec{v}_2| = \left| \left(\vec{e}_1 \frac{\partial v_1}{\partial u_1} du_1 + \vec{e}_2 \frac{\partial v_1}{\partial u_2} du_2 \right) \times \left(\vec{e}_1 \frac{\partial v_2}{\partial u_1} du_1 + \vec{e}_2 \frac{\partial v_2}{\partial u_2} du_2 \right) \right|$

$$= \begin{vmatrix} \frac{\partial v_1}{\partial u_1} & \frac{\partial v_1}{\partial u_2} \\ \frac{\partial v_2}{\partial u_1} & \frac{\partial v_2}{\partial u_2} \end{vmatrix} |(\bar{e}_1 \times \bar{e}_2) du_1 du_2|$$

$$= \left| J \begin{pmatrix} v_1 & v_2 \\ u_1 & u_2 \end{pmatrix} \right| |d\bar{u}_1 \times d\bar{u}_2|, \text{ where } \left| J \begin{pmatrix} v_1 & v_2 \\ u_1 & u_2 \end{pmatrix} \right| = \text{modulus of the Jacobian determinant.}$$

From plane polar to Cartesian coordinates, we get,

$$dx dy = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} dr d\theta = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} dr d\theta = r dr d\theta$$

Similarly, in three dimensions an infinitesimally small volume can be approximated as a parallelopiped. We can write an analogous result for transition from (u_1, u_2, u_3) to (v_1, v_2, v_3) coordinates. Thus, we obtain for

$$\text{the elementary volume in } (v_1, v_2, v_3) \text{ triad as } |d\bar{v}_1, (d\bar{v}_2 \times d\bar{v}_3)| = \begin{vmatrix} \frac{\partial v_1}{\partial u_1} & \frac{\partial v_1}{\partial u_2} & \frac{\partial v_1}{\partial u_3} \\ \frac{\partial v_2}{\partial u_1} & \frac{\partial v_2}{\partial u_2} & \frac{\partial v_2}{\partial u_3} \\ \frac{\partial v_3}{\partial u_1} & \frac{\partial v_3}{\partial u_2} & \frac{\partial v_3}{\partial u_3} \end{vmatrix} |(\bar{e}_1 (\bar{e}_2 \cdot \bar{e}_3)) du_1 du_2 du_3|$$

$$= \left| J \begin{pmatrix} v_1 & v_2 & v_3 \\ u_1 & u_2 & u_3 \end{pmatrix} \right| |d\bar{u}_i \cdot (d\bar{u}_2 \times d\bar{u}_3)|.$$

$$\text{From spherical polar to Cartesian coordinates we have, } dx dy dz = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix} dr d\theta d\phi.$$

Putting $x = r \sin \theta \cos \phi, y = r \sin \theta \sin \phi, z = r \cos \theta$, we obtain, $dx dy dz = r^2 dr \sin \theta d\theta d\phi$. For a general

$$\text{transformation from the } (u_1, u_2, \dots, u_n) \text{ to the } (v_1, v_2, \dots, v_n) \text{ coordinates we have } dU = \left| J \begin{pmatrix} u_1 & u_2 & \dots & u_n \\ v_1 & v_2 & \dots & v_n \end{pmatrix} \right| dV,$$

where dU and dV represent the elementary volumes in $(u_1, u_2, \dots, u_n)$ and $(v_1, v_2, \dots, v_n)$ coordinates respectively.

This result carries over to the multiple integral of a function over a set of coordinates (as is clear from Riemann's definition of integrals). We can write in general,

$$\int f(x_1 \dots x_n) dx_1 \dots dx_n = \int \left| J \begin{pmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \end{pmatrix} \right| \phi(y_1, \dots, y_n) dy_1 \dots dy_n$$

where the two sets of coordinates cover the same volume. For change from the Cartesian to spherical polar coordinates we have, $\int f(x, y, z) dx dy dz = \int \phi(r, \theta, \phi) r^2 dr \sin \theta d\theta d\phi$. [An elementary knowledge of vector algebra on the part of the students has been assumed. (For example, an infinitesimal element of volume in u_1, u_2, u_3 coordinates = $d\bar{u}_1 \cdot (d\bar{u}_2 \times d\bar{u}_3)$ etc.). Students will have a detailed exposure to these in Chapter 2].

SOLVED PROBLEMS

Problem 9

Find

(a) $\int_0^\infty \int_0^\infty e^{-(x^2+y^2)^2} dy dx$

(b) $\int \int x e^{y^3} dx dy$ over the region bounded by the lines $x = 0, y = 1$ and $y = x$

(c) $\int_1^5 \int_0^2 \int_{y^{1/2}}^2 z^2 (\sin x^3) dx dy dz$

Solution: (a) We transform to polar coordinates by $x = \rho \cos \theta, y = \rho \sin \theta$

$$\text{Thus } I = \int_{\theta=0}^{2\pi} \int_{\rho=0}^{\infty} e^{-\rho^4} \rho d\rho d\theta, \text{ since } dx dy \rightarrow \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} \end{vmatrix} d\rho d\theta = \rho d\rho d\theta$$

$$\text{Or } I = 2\pi \int_{\rho=0}^{\infty} \rho e^{-\rho^4} d\rho = \frac{\pi}{2} \int_0^\infty x e^{-1/2-x} dx, \text{ where } x = \rho^4, d\rho = \frac{dx}{4x^{3/4}}$$

$$\text{Or } I = \frac{\pi}{2} \sqrt{\pi} = \frac{1}{2} \pi^{3/2}$$

$$(b) \quad I = \int \int x e^{y^3} dx dy = \int_0^1 e^{y^3} dy \int_0^y x dx = \frac{1}{2} \int_0^1 y^2 e^{y^3} dy$$

$$\text{or, } I = \frac{1}{6} \int_0^1 e^z dz = \frac{1}{6} (e - 1)$$

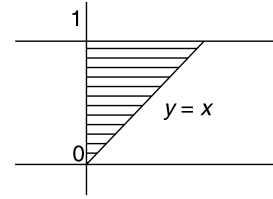


Fig. 1.2

(The region of integration is shown in Fig. 1.2)

$$(c) \quad \text{Let us evaluate } \int_0^2 \int_{y^{1/2}}^2 \sin x^3 dx dy = \int_0^2 \sin x^3 \left[\int_0^{x^2} dy \right] dx$$

$$= \int_0^2 x^2 \sin x^3 dx = \frac{1}{3} \int_0^8 \sin t dt = \frac{1}{3} [1 - \cos 8]$$

$$\text{Or,} \quad \text{The given integral} = \int_1^5 \int_0^2 \int_{y^{1/2}}^2 z^2 \sin x^3 dx dy dz = \int_1^5 \frac{1}{3} (1 - \cos 8) z^2 dz$$

$$= \frac{1}{3} \left[\frac{124}{3} \right] (1 - \cos 8) = \frac{124}{9} (1 - \cos 8)$$

SUMMARY

- The limit, continuity and the differentiability of a function of several variables have been discussed in the text. The calculation becomes much more involved in this case relative to that in the case of the function of a single variable because the uniqueness of the relevant limits has to be ensured in all the cases irrespective of the path of approach to the point. The chain rule of differentiation along with the principle of implicit differentiation (discussed in a greater detail in the appendix). have been dealt with.
- The Taylor series of a function of two variables has been discussed with the problem of existence of the maxima, minima and the saddle point. The limitation of this method has been stressed. The problem of the constrained extrema (i.e. a function $f(x, y)$ needed to be extremised subject to some other functions of the variables ($\phi(x, y)$, $\psi(x, y)$, say, remaining constants) has been dwelt upon. The efficacy of the method outlined here can of course be generalised to a function of n -variables. The calculus of varia-

tions has been discussed, where an integral of a function $\int_{x_1}^{x_2} f(x, y, y_x \dots) dx$ between two fixed end

points has to be extremised in the unrestricted case as well as in the constrained case when some other integrals between the same two points have to be kept fixed. This has been further elaborated in the appendix, which also discusses the problem of extrema of a many-variable function.

- The problem of multiple integrals (double and triple integrals in this chapter) has been discussed along with the technique of the change of variables used in the evaluation.

MULTIPLE CHOICE QUESTIONS

1. If $x^3 + y^3 + 2x^2y = 0$, $\frac{dy}{dx}$ is
- (a) 0 (b) $-\left(\frac{4xy + 3x^2}{3y^2 + 2x^2}\right)$ (c) $-\frac{x^2}{y^2}$ (d) $\left(\frac{4xy + 3x^2}{3y^2 + 2x^2}\right)$

2. If $y - z^2(x, y) - xz(x, y) = 0$, $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are respectively.

- (a) $\frac{1}{2z+x}, \frac{1}{x+2z}$ (b) $\frac{1}{2z+x}, -\frac{z}{x+2z}$
 (c) $-\frac{z}{x+2z}, \frac{1}{2z+x}$ (d) $-\frac{z}{x+2z}, -\frac{1}{2z+x}$

(Hint: Differentiate w.r.t. x and y)

3. If $x^2 + y^2(x) + 2z^2(x) = 0$ and $3x^2 - 4y^2(x) + 5z^2(x) = 0$, $\frac{dy}{dx}$ and $\frac{dz}{dx}$ are respectively

- (a) $\frac{x}{13y}, -\frac{7x}{13z}$ (b) $\frac{x}{13y}, \frac{7x}{13z}$ (c) $-\frac{x}{13y}, \frac{-7x}{13z}$ (d) None of these

4. If $x = r^2 + r\theta, y = 2r^2 - 3r\theta$, $\frac{\partial r}{\partial x}$ and $\frac{\partial r}{\partial y}$ are respectively

- (a) $\frac{1}{10r}, \frac{3}{10r}$ (b) $-\frac{1}{10r}, -\frac{3}{10r}$ (c) $\frac{3}{10r}, \frac{1}{10r}$ (d) $-\frac{3}{10r}, -\frac{1}{10r}$

5. If $f(x, y) = x^2 + 2y^2 + 3xy$ and $x = r \cos \theta, y = r \sin \theta$ then $\frac{\partial f}{\partial r}$ and $\frac{\partial f}{\partial \theta}$ are respectively

- (a) $(2x + 3y) \cos \theta + (4y + 3x) \sin \theta, (2x + 3y) \sin \theta + (4y + 3x) \cos \theta$
 (b) $(2x + 3y) \cos \theta + (4y + 3x) \sin \theta, (2x + 3y) \cos \theta - (4y + 3x) \sin \theta$
 (c) $(2x + 3y) \cos \theta + (4y + 3x) \sin \theta, r[(4y + 3x) \cos \theta - (2x + 3y) \sin \theta]$
 (d) $r[(4y + 3x) \cos \theta + (2x + 3y) \sin \theta], (2x + 3y) \cos \theta + (4y + 3x) \sin \theta$

6. If $f(x, y, z) = x^3 + 3x^2y + z^2x + xyz$ then $x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z}$ is

- (a) $3f$ (b) $4f$
 (c) $2f$ (d) not simply related to f

(Hint: Use the theorem on homogeneous functions)

7. If $f(x, y) = \sin(x^2 - y^2) g(x, y) = e^{\frac{1}{x-y^2}}$

- (a) $f(x, y)$ and $g(x, y)$ have well-defined limits everywhere
 (b) Neither of these has a well-defined limit at $x = y = 0$
 (c) $f(x, y)$ has a well-defined limit everywhere but $g(x, y)$ has no well-defined limit at $x = y = 0$
 (d) $f(x, y)$ has no well-defined limit at $x = y = 0$ but $g(x, y)$ has a well-defined limit at $x = y = 0$

8. If $f(x, y, z)$ and $\phi(x, y, z)$ are homogeneous functions of degree n
- (a) $\frac{f}{\phi}$ is a homogeneous function of n th. degree in x, y, z
 - (b) $\frac{f}{\phi}$ is a function of $y/x, z/x$
 - (c) $\frac{f}{\phi}$ is a homogeneous function of degree $n - 1$
 - (d) None of these
9. If $f(x, y) = \exp\left(-\frac{1}{x^2 - y^2}\right)$, which of the statements is true?
- (a) $f(x, y)$ can not be expanded about $x = y = 0$
 - (b) $f(x, y)$ can be expanded about every point in the xy -plane
 - (c) $f(x, y)$ can be expanded for all x about $y = 0$
 - (d) $f(x, y)$ can be expanded for all y about $x = 0$
10. $f(x, y) = x(xy - 1)$ has
- (a) a maximum at $x = y = 0$
 - (b) a minimum at $x = y = 0$
 - (c) no stationary point anywhere
 - (d) a saddle point at $x = y = 0$
11. $f(x, y) = 2x^2 + 3y^2 - 16$ has
- (a) a local maximum at $x = y = 0$
 - (b) a local minimum at $x = y = 0$
 - (c) a saddle point at $x = y = 0$
 - (d) a global minimum at $x = y = 0$
12. $f(x, y) = x^2 - y^2 - 4$ has
- (a) a maximum at $x = y = 0$
 - (b) no stationary point at $x = y = 0$
 - (c) a minimum at $x = y = 0$
 - (d) a saddle point at $x = y = 0$
13. The temperature $T(x, y) = (x^2 + y^2)k$ with $xy = 1$ ($k = \text{constant}$). Then T has
- (a) a minimum at $x = y \pm 1$ and $x = -y = \pm 1$
 - (b) a maximum at the values of x, y stated above
 - (c) a minimum at $x = y = 1$ and a maximum at $x = -y = 1$
 - (d) a maximum at $x = y = -1$ and a minimum at $x = y = 1$

(Hint: Use the method of extremum with constraints)

14. If $I(x) = \int_{\psi(x)}^{\phi(x)} e^{xt} dt$ then $\frac{dI}{dx}$ is

(a) $e^{x\phi(x)}\phi'(x) - e^{x\psi(x)}\psi'(x)$

(b) $\int_{\psi(x)}^{\phi(x)} te^{xt} dt$

- (c) $e^{x\phi(x)} - e^{x\psi(x)}$ (d) $e^{x\phi(x)}\phi'(x) - e^{x\psi(x)}\psi'(x) + \int_{\psi(x)}^{\phi(x)} te^{-xt} dt$
15. $\int_{x_1}^{x_2} \left[\frac{1}{2}m\dot{x}^2 - \frac{1}{2}\alpha x - \beta x^3 \right] dx$ has an extremum for fixed x_1, x_2 . Then the path is given by
 (a) $m\ddot{x} = -\alpha - 3\beta x^2$ (b) $m\ddot{x} = 0$
 (c) $m\ddot{x} = \alpha - 3\beta x^2$ (d) no well-defined path
16. The density of the material of a square lamina with the vertices at $(0, 0)$, $(0, a)$, $(a, 0)$ and (a, a) respectively is αxy ; α being a constant. The c.g. of the lamina is at
 (a) $(0, 0)$ (b) $\left(\frac{a}{2}, \frac{a}{2}\right)$
 (c) $\frac{a}{2}, \frac{2a}{3}$ (d) $\frac{2a}{3}, \frac{2a}{3}$
17. $\iint x^2 y^2 dx dy$ over the region bounded by $y = 0, y = a, y = 2x, y = 2(x - a)$ is
 (a) $\frac{a^6}{a}$ (b) $\frac{8a^6}{a}$ (c) $\frac{103a^6}{360}$ (d) None of these
18. $\iiint x^2 y^2 z^2 dx dy dz$ over the volume of a sphere of radius a is,
 (a) $4\pi a^9$ (b) $\frac{a^9}{27}$ (c) $\frac{2\pi a^9}{945}$ (d) $\frac{4\pi a^9}{945}$

(Hint: Use spherical polar coordinates and transform the integral into one over r, θ, ϕ)

Answers to MCQs

1. (b)	2. (c)	3. (a)	4. (c)	5. (c)	6. (a)	7. (c)
8. (b)	9. (a)	10. (c)	11. (d)	12. (d)	13. (a)	14. (d)
15. (a)	16. (d)	17. (c)	18. (d)			

EXERCISE

- If $u = x^2 + y^2, v = 2xy$ find out the first order and the second order partial derivatives of u with respect to x and y . Also find out, $\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v}, \frac{\partial y}{\partial u}, \frac{\partial y}{\partial v}$.
- Express $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$ in cylindrical coordinates (ρ, θ, z) ($x = \rho \cos \theta, y = \rho \sin \theta$) and spherical polar coordinates (r, θ, ϕ) ($x = r \sin \theta \cos \phi, y = r \sin \theta \sin \phi, z = r \cos \theta$) respectively.
- Given the equation $Ax^4 + By^4 + Cx^2y + Dxy^2 = \text{constant}$, find out $\frac{d^2 y}{dx^2}$ and $\frac{d^2 r}{d\theta^2}$, where $x = r \cos \theta, y = r \sin \theta$. [$y = y(x), r = r(\theta)$].

4. Given the equations $u + 3v - 2x^2 + y^2 = 0$ and $2u + v + 3xy = 0$, find out $\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v}, \frac{\partial y}{\partial u}, \frac{\partial y}{\partial v}$ with the condition for their existence.
5. Find out the extreme and their nature in respect of the following functions.
 - (i) $f(x, y) = Ax^2 + By^2 + Cxy$.
 - (ii) $A \sin x + B \cos y + C$, where A, B, C are constants.
 - (iii) $(x^2 + 2y^2) \exp [-(x^2 + y^2)]$.
 - (iv) xy
6. Find out the extremum of the distance from the origin to the plane $2x + 3y + 4z = 6$ and its nature.
7. Determine the stationary points of (a) $x^4 + y^4 + z^4$ with the condition that $x^2 + y^2 + z^2 = 4$, (b) $A \sin x + B \cos y$ with the constraint that $C \sin x + D \cos y = \text{constant}$, (c) $3x^2 + 4y^2 + 5z^2$ with the constraints that (i) $x^2 + y^2 + z^2 = 1$, (ii) $2x^2 + y^2 + 2z^2 = 3$. (A, B, C, D are constants)
8. Light ray travels in the xz -plane from $(0, 0)$ to $(0, z_0)$ where the refractive index μ varies as $\alpha\beta z$, α and β being constants. Find out the equation of the path of the ray. What happens if μ suddenly changes from μ_1 to μ_2 as the path crosses the line $z = \frac{h}{2}$. (Euler's equation (Eq. 1.7) is not applicable here because of discontinuity at $z = \frac{h}{2}$. Directly find out the extremum by calculating the optical path length and its extremum with respect to variation in x).
9. Find out the equation of a curve for which the area enclosed between the curve and the x -axis with a given length between two given points $x = x_1$ and $x = x_2$ is maximum.
10. (a) Show that the extremum of the electrostatic energy $\int (\nabla \phi)^2 dv$ gives the equation $\nabla^2 \phi = 0$
 (b) Show that the extremum of $\int \psi^* \left(-\frac{\hbar^2}{2m} \nabla^2 \psi + V \phi \right) dv$ with $\int \psi^* \psi dv$ constant leads to $\nabla^2 \psi + \frac{2m}{\hbar^2} (E - V) \psi = 0$, E being a constant. (This is Schrodinger equation of quantum mechanics, where the operator $-\frac{\hbar^2}{2m} \nabla^2 + V$ refers to the Hamiltonian operator, ψ is the state function and $\int \phi^* \psi dv$ the total probability of having the particle in the volume).
11. Find the equation of the curve that represents the shape of a uniform chain hanging freely between two fixed points under a constant gravitational field.
[Hint: Write the potential energy and minimise it]
12. Deduce the equations obtained by the variational principle applied to the following.
 - (a) $\int_{x_1}^{x_2} \left(\frac{1}{2} m \dot{x}^2 - \frac{1}{2} K_1 x^2 - K_2 x^4 \right) dx$,
 - (b) $\int_{x_1}^{x_2} \left[f(x) \left(\frac{dy}{dx} \right)^2 + g(x) y^2 \right] dx$
 - (c) $\frac{1}{2} \int dt \int dv [\dot{\phi}^2 - m^2 \phi^2 - (\nabla \phi)^2]$
 - (d) $\int \left[\frac{\hbar}{2\pi i} (\psi^* \dot{\psi} - \dot{\psi}^* \psi) + V \psi^* \psi + \frac{\hbar^2}{8\pi m} \bar{\nabla} \psi \cdot \bar{\nabla} \dot{\psi} \right] dv dt, \left[\dot{\psi} = \frac{d\psi}{dt} \right]$

- (e) $\frac{1}{2} \int [A\dot{\phi}^2 - (\bar{\nabla}\phi)^2] dv dt$ [for (c), (d), (e) refer to the appendix B(c)].
13. (a) If $F(x, y) = A \sin^{-3} x + B \cos^3 y + C \sin^2 x \cos^2 y$, where $x = a + bt$, $y = ct^2$, find out $\frac{d^2 F}{dt^2}$
- (b) Find out $\frac{d}{d(\cos x)} \int_0^{x^2} \sin(xt) dt$
- (c) Calculate $\frac{dF}{dx}$ directly, where $F(x) = \int_{x^3}^{x^4} e^{-x^2 t} \sin xt dt$ and verify the relation
- $$\frac{dF}{dx} = f(x, x^4) \frac{d}{dx}(x^4) - f(x, x^3) \frac{d}{dx}(x^3) + \int_{x^3}^{x^4} \frac{\partial f(x, t)}{\partial x} dt, \text{ where } f(x, t) = e^{-x^2 t} \sin xt$$
14. (a) Find out the centre of gravity of a cylinder with the origin at the centre of the base circle and the density of the material of the cylinder $f = \alpha z$, α being a constant.
- (b) Find out the centre of gravity of the hemisphere with the centre at the origin and the density of the material $f = Ar^2 \cos \theta$, A being a constant and the other symbols having their usual meanings.
- (c) Determine the centre of gravity of the region bounded by $x = 0$, $y = 0$, $z = 0$ and $2x + 3y + z = 2$ with the density ρ given by $\rho = kxyz$, k being a constant.
- (d) Find out the moment of inertia of a figure bounded by the parabola $x^2 = ay$, the straight line $y = a$ relative to the line $x = a$, the density of the material being given by $\sigma = k|x|y$, k being a constant.
15. (a) Find out the volume of the slice cut out from the sphere $x^2 + y^2 + z^2 = R^2$ by the plane $z = h$.
- (b) Find out the volume of the portion of a right circular cone between the vertex and the section with the plane $z = z_0$, where the z -axis is taken as the axis of the cone.
- (c) Find out the largest surface area of the cylinder that can be inscribed in a sphere of radius r .

Hints and Answers to Exercise Questions

1. $u_x = 2x$, $u_y = 2y$, $u_{xx} = 2$, $u_{yy} = 2$, $u_{xy} = u_{yx} = 0$, $v_x = 2y$, $v_y = 2x$, $v_{xx} = 0$, $v_{yy} = 0$, $v_{xy} = v_{yx} = 2$.

There will be 16 possible pairs of values of x and y arising from the sets

(a) $x = \pm 1/2(\sqrt{u+v} + \sqrt{u-v})$, $y = \pm 1/2(\sqrt{u+v} - \sqrt{u-v})$

(b) $y = \pm 1/2(\sqrt{u+v} - \sqrt{u-v})$, $x = \pm 1/2(\sqrt{u+v} + \sqrt{u-v})$.

One representative set is, for example,

(i) $x = \frac{1}{2}(\sqrt{u+v} + \sqrt{u-v})$, $y = \frac{1}{2}(\sqrt{u+v} - \sqrt{u-v})$

$$\frac{\partial x}{\partial u} = \frac{1}{4}[(u+v)^{-1/2} + (u-v)^{-1/2}], \quad \frac{\partial x}{\partial v} = \frac{1}{4}[(u+v)^{-1/2} - (u-v)^{-1/2}]$$

$$\frac{\partial y}{\partial u} = \frac{1}{4}[(u+v)^{-1/2} - (u-v)^{-1/2}], \quad \frac{\partial y}{\partial v} = \frac{1}{4}[(u+v)^{-1/2} + (u-v)^{-1/2}].$$

The results from the other sets follow similarly.

$$2. \quad \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial f}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}, \quad \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2}$$

(Hint: Express $\frac{\partial f}{\partial x_i}$ as $\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial \rho} \cdot \frac{\partial \rho}{\partial x_i} + \frac{\partial f}{\partial \theta} \cdot \frac{\partial \theta}{\partial x_i} + \frac{\partial f}{\partial z} \cdot \frac{\partial z}{\partial x_i}$ or, as,

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial r} \cdot \frac{\partial r}{\partial x_i} + \frac{\partial f}{\partial \theta} \cdot \frac{\partial \theta}{\partial x_i} + \frac{\partial f}{\partial \phi} \cdot \frac{\partial \phi}{\partial x_i} \text{ in cylindrical and spherical polar coordinates respectively,}$$

where $i = 1, 2, 3$ imply the x, y and z coordinates respectively. Find out ρ, θ and r, θ, ϕ respectively as functions of x, y, z . Calculate the derivatives and of the first and then of the second order of f with respect to x, y and z respectively. The calculation is a bit tedious but it is instructive to do these because of immense applications of the derivatives in various fields).

$$3. \quad \frac{dy}{dx} = -\frac{(4Ax^3 + 2Cxy + Dy^2)}{(4By^3 + 2Dxy + Cx^2)} \text{ and}$$

$$\frac{dr}{d\theta} = -\frac{[(4r^4 \sin \theta \cos \theta (A \cos^2 \theta - B \sin^2 \theta) + Cr^3 \cos \theta (2 \sin^2 \theta - \cos^2 \theta) + Dr^3 \sin \theta (2 \cos^2 \theta - \sin^2 \theta)]}{[4r^3 (A \cos^4 \theta + B \sin^4 \theta) + 3r^2 \sin \theta \cos \theta (C \cos \theta + D \sin \theta)]}.$$

4. Now proceed to calculate $\frac{d^2 y}{dx^2}, \frac{d^2 r}{d\theta^2}$. The final result is cumbersome and not given here.

$$\frac{\partial x}{\partial u} = \frac{5x-4y}{12x^2+6y^2}, \quad \frac{\partial y}{\partial u} = -\frac{(8x+3y)}{(12x^2+6y^2)}, \quad \frac{\partial x}{\partial v} = \frac{9x-2y}{12x^2+6y^2}, \quad \frac{\partial y}{\partial v} = -\frac{2y(4x+9y)}{3(4x^2+2y^2)}$$

5. (i) The extrema arc at $x = y = 0$ provided $e^2 \neq 4AB$. If A, B be both > 0 , and $4AB > c^2$, we have a minimum at $x = y = 0$. If A, B be both < 0 and $4AB > c^2$, we have a maximum at $x = y = 0$. If A and

B be of opposite signs we have a saddle point. There exist extrema at $y = -\frac{2Ax}{c}$ but $c^2 - 4AB = 0$ there leading to failure of the test at this level.

(ii) The extrema are at $x = (2n+1)\pi/2, n = 0, \pm 1, \dots$ and $y = m\pi, m = 0, \pm 1, \dots$ respectively.

The nature of the extrema will depend on the signs of A and B . For example the points $x = \pm \frac{\pi}{2}, \pm \frac{5\pi}{2}, \dots; y = 0, \pm 2\pi, \pm 4\pi, \dots$ give maxima for $A, B > 0$ but minima for $A, B < 0$ and saddle points for A and B of the opposite signs. For the extrema having the other sets of values ($x =$

$\pm \frac{3\pi}{2}, \pm \frac{7\pi}{2}, \dots; y = \pm \pi, \pm 3\pi, \dots$) we will obtain different results for the same A, B .

(iii) $x = 0, y = 0$ gives minimum while $x = 0, y = \pm 1$ and $x = \pm 1, y = 0$ yield the maximum and the saddle point respectively.

(iv) There are no extrema

6. There are minima at $x = \frac{12}{29}, y = \frac{18}{29}, z = \frac{24}{29}$.

7. (a) The minimum value of $x^4 + y^4 + z^4 = \frac{16}{3}$ at $x^2 = y^2 = z^2 = 4/3$ and the maximum value of $x^4 + y^4 +$

$z^4 = 8$ at $x = 0, y^2 = z^2 = 2$ (or $y = 0$, or $x^2 = z^2 = 2$, or $z = 0, x^2 y^2 = 2$)

There are other stationary points at $x = y = 0, z^2 = 2$ (or $x = z = 0, y^2 = 2$ or $y = z = 0, x^2 = 2$). The derivatives of second order are of different signs indicating the possibility of saddle points here.

(b) If $C \sin x + D \cos y = K$, K being a constant, the extrema are at $\cos x = 0, \cos y = \frac{K \mp C}{D}$ and $\sin y = 0, \sin x = \frac{K \mp D}{C}$.

The extrema are seen to be saddle points. The magnitudes of $A \sin x + B \cos y$ at the saddle points are $\pm A + \frac{B(K \mp C)}{D}$ and $\pm B + \frac{A(K \mp D)}{C}$ respectively.

8. The path of the light ray in the xz -plane is given by $(\alpha - \beta z) \frac{d^2 z}{dx^2} = -\beta \left[1 + \left(\frac{dz}{dx} \right)^2 + \left(\frac{dz}{dx} \right)^4 \right]$. If the

light ray travels from a medium with refractive index μ_1 , to another one with refractive index μ_2 the

change of the direction of the path is given by $\frac{\sin \theta_1}{\sin \theta_2} = \frac{\mu_2}{\mu_1}$, where θ_1 and θ_2 are the angles made by the

incident and the refracted ray respectively with the normal at the point of incidence to the boundary.

9. The curve is a circular arc passing through the two points $x = x_1$ and $x = x_2$.

(Hint: $\int_{x_1}^{x_2} f(x) dx = \text{maximum}$ with the constraint that $\int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx} \right)^2} dx = \text{fixed}$).

11. The equation of the curve is, $z = K + a^{-1} \cosh(ax + b)$, where a, b , and K are constants.

12. (a) $m\ddot{x} = -K_1 x - 4K_2 x^3$, (b) $f(x) \frac{d^2 y}{dx^2} + f'(x) \frac{dy}{dx} = g(y(\alpha))$

(c) $\ddot{\phi} - \nabla^2 \phi = -m^2 \phi$ (d) $i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + v\psi$, (e) $A\ddot{\phi} - \nabla^2 \phi = 0$

13. (a) (Hint: Find out $\frac{dF}{dt}$ from $\frac{dF}{dt} = \frac{dF}{dx} \cdot \frac{dx}{dt} + \frac{dF}{dy} \cdot \frac{dy}{dt}$ and therefrom $\frac{d^2 F}{dt^2}$).

- (b) $\left[\left(\frac{1 - \cos x^3}{x^2} \right) - 3x \sin x^3 \right] \cdot \frac{1}{\sin x}$
14. (a) $\left(0, 0, \frac{2h}{3} \right)$, the z -coordinate being measured from the base.
- (b) $\bar{x} = \bar{y} = 0, \bar{z} = \frac{5a}{9}$ (a = radius of the hemisphere).
- (c) $\bar{x} = \frac{2}{7}, \bar{y} = \frac{4}{21}, \bar{z} = \frac{2}{7}$
- (d) $\frac{11ma^2}{8}$, m being the mass.
15. (a) $\pi h \left(R^2 - \frac{h^2}{3} \right)$,
- (b) $\frac{z_0^3}{3} \pi \tan^2 \theta$, θ being the semivertical angle
- (c) $\pi r^2 (1 + \sqrt{2})$.

APPENDIX

(A) Implicit Function Theorem

Often the dependent variables are known not as explicit functions of the independent variables but there is an implicit relation among them. The question is, "Can we utilise these relations to express the dependent variables as explicit functions of independent variables at least in some neighbourhood of the initially chosen set of independent variables?" The answer to this question is furnished by the implicit function theorem. In the following section we give some examples of the method instead of stating the theorem formally with the hope that these will provide a better understanding of the main problem and its solution.

- (a) Let us start with a relation of the form $f(x, y(x)) = 0$. Clearly we obtain, $\frac{dy}{dx} = - \frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$ (Section 1.2 (c)).

So, provided $\frac{\partial f}{\partial y} \neq 0$ at the point (x_0, y_0) say we can perform an expansion about $x = x_0, y = y_0$ and in a small region about the point y can be written a function of x (of course, if the function $y(x)$ be a rapidly varying function of x , the region of validity of the approximation will be highly limited).

- (b) Similarly, if we have relations of the form $f(x, y(x), z(x)) = 0$ and $g(x, y(x), z(x)) = 0$ respectively we obtain,

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dx} + \frac{\partial f}{\partial z} \cdot \frac{dz}{dx} = 0 \quad \text{and} \quad \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y} \cdot \frac{dy}{dx} + \frac{\partial g}{\partial z} \cdot \frac{dz}{dx} = 0$$

Thus we obtain,
$$\frac{dz}{dx} = - \frac{\left(\frac{\partial f}{\partial y} \frac{\partial g}{\partial v} - \frac{\partial f}{\partial v} \frac{\partial g}{\partial y} \right)}{\left(\frac{\partial f}{\partial y} \frac{\partial g}{\partial z} - \frac{\partial f}{\partial z} \frac{\partial g}{\partial y} \right)} = - \frac{\left[\frac{\partial(f, g)}{\partial(y, x)} \right]}{\left[\frac{\partial(f, g)}{\partial(y, z)} \right]}$$

Similarly,
$$\frac{dy}{dx} = - \frac{\left[\frac{\partial(f, g)}{\partial(x, z)} \right]}{\left[\frac{\partial(f, g)}{\partial(y, z)} \right]},$$
 where $\frac{\partial(f, g)}{\partial(y, x)}$ is the Jacobian determinant given by $\frac{\partial f}{\partial y} \frac{\partial g}{\partial x} - \frac{\partial f}{\partial x} \frac{\partial g}{\partial y}$.

Thus similar to the case (a) above we can find out y and z as function of x in some neighbourhood of a chosen point (x_0, y_0, z_0) provided the denominator determinant does not vanish there.

(c) If we have two functional relations of the form $f(x, y, p, q) = 0$ and $g(x, y, p, q) = 0$ respectively we can

proceed similarly to obtain,
$$\frac{\partial x}{\partial p} = - \frac{\left[\frac{\partial(f, g)}{\partial(p, y)} \right]}{\left[\frac{\partial(f, g)}{\partial(x, y)} \right]}, \quad \frac{\partial x}{\partial q} = - \frac{\left[\frac{\partial(f, g)}{\partial(q, y)} \right]}{\left[\frac{\partial(f, g)}{\partial(x, y)} \right]}, \quad \frac{\partial y}{\partial p} = - \frac{\left[\frac{\partial(f, g)}{\partial(p, x)} \right]}{\left[\frac{\partial(f, g)}{\partial(x, y)} \right]}, \quad \text{and} \quad \frac{\partial y}{\partial q} = - \frac{\left[\frac{\partial(f, g)}{\partial(q, x)} \right]}{\left[\frac{\partial(f, g)}{\partial(x, y)} \right]}.$$

Provided the determinant appearing in the denominator does not vanish we can express x, y in terms of (p, q) at least locally in some neighbourhood of the points (x_0, y_0, p_0, q_0) . The extension to a larger number of variables is straightforward. The formal statement of the theorem is thus, "If the determinant

$\left(\frac{\partial(f_1 \dots f_n)}{\partial(y_1 \dots y_n)} \right) \neq 0$ we can from an implicit relation of the form $f_i(x_1 \dots x_n, y_1 \dots y_n) = 0$ for $i = 1, 2, \dots, n$ find out $y_1, y_2, \dots, y_n$ locally as functions of $x_1 \dots x_n$.

(B) Extension of the Variational Principle

(a) Previously in section 1.3(d) we started with $f = f(x, y, y_x)$ and found out the equation of the path by

equating $\delta \int_{x_1}^{x_2} f(x, y, y_x) dx$ to zero where the δ -variation implies a variation with respect to the path

connecting two fixed end-points. In some cases f may contain higher order derivatives of y as well. In those cases we have,

$$\delta \int_{x_1}^{x_2} f(x, y, y_x, y_{xx} \dots) dx = 0. \quad \text{Let us consider } f = f(x, y, y_x, y_{xx}).$$

$$\text{or, } \delta \int_{x_1}^{x_2} f(x, y, y_x, y_{xx}) dx = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial y_x} \delta y_x + \frac{\partial f}{\partial y_{xx}} \delta y_{xx} \right) dx = 0$$

$$\text{Now } \int_{x_1}^{x_2} \frac{\partial f}{\partial y_{xx}} \delta y_{xx} dx = \int_{x_1}^{x_2} \frac{\partial f}{\partial y_{xx}} \frac{d}{dx} (\delta y_x) dx = \int_{x_1}^{x_2} \left[\frac{d}{dx} \left(\frac{\partial f}{\partial y_{xx}} \delta y_x \right) - \frac{d}{dx} \left(\frac{\partial f}{\partial y_{xx}} \right) \delta y_x \right] dx$$

Since $\delta y, \delta y_x \dots = 0$ at the end-points we obtain,

$$\begin{aligned} \int_{x_1}^{x_2} \frac{\partial f}{\partial y_{xx}} \delta y_{xx} dx &= - \int_{x_1}^{x_2} \left[\frac{d}{dx} \left(\frac{\partial f}{\partial y_{xx}} \right) \delta y_x \right] dx \\ &= - \int_{x_1}^{x_2} \left[\frac{d}{dx} \left\{ \left(\frac{d}{dx} \frac{\partial f}{\partial y_{xx}} \right) \delta y \right\} - \frac{d^2}{dx^2} \left(\frac{\partial f}{\partial y_{xx}} \right) \delta y \right] dx. \end{aligned}$$

$$\text{Again since the end points are fixed } \int_{x_1}^{x_2} \frac{\partial f}{\partial y_{xx}} \delta y_{xx} dx = \int_{x_1}^{x_2} \frac{d^2}{dx^2} \left(\frac{\partial f}{\partial y_{xx}} \right) \delta y dx$$

[Since $\delta y = 0$ at $x = x_1, x_2$]

$$\text{Thus we get combining the terms } \delta \int_{x_1}^{x_2} f(x, y, y_x, y_{xx}) dx = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y_x} \right) + \frac{d^2}{dx^2} \left(\frac{\partial f}{\partial y_{xx}} \right) \right] \delta y dx$$

Since $\delta \int_{x_1}^{x_2} f(x, y, y_x, y_{xx}) dx = 0$ for arbitrary δy and arbitrary limits x_1, x_2

$$\text{we obtain, } \frac{\partial f}{\partial y} = \frac{d}{dx} \left(\frac{\partial f}{\partial y_x} \right) - \frac{d^2}{dx^2} \left(\frac{\partial f}{\partial y_{xx}} \right).$$

For f depending in general on $y, y_x^{(1)}, \dots, y_x^{(n-\lambda)}$ we get

$$\frac{\partial f}{\partial y} + \sum_{i=1}^n (-1)^i \frac{d^i}{dx^i} \left(\frac{\partial f}{\partial y_x^{(i)}} \right) = 0 \dots (B) \quad (1)$$

(b) If $f = f(x, y_1, y_{1x}, \dots, y_2, y_{2x}, \dots)$ we can write an equation separately for each y_i thus obtaining the

$$\text{equations similar to (B) (1) given by } \frac{\partial f}{\partial y_i} + \sum_{j=1}^n (-1)^j \frac{d^j}{dx^j} \left(\frac{\partial f}{\partial y_{i,x}^{(j)}} \right) = 0 \text{ for } i = 1, 2, \dots, n.$$

(c) If the number of independent variables be more than one, such as, $f = f(q, q_x, q_y, q_z, q_t)$, say, we have to

find out the extremum of $\int f(q, q_x, q_y, q_z, q_t) dx dy dz dt$ with the end point fixed. Thus $\delta \int_1^2 f(q, q_x, q_y, q_z, q_t) dx dy dz dt = 0$.

[This occurs in the case of continuous systems with the Lagrangian density playing the role of f].

The procedure followed in section 1.3(d) leads to the equation $\sum_i \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial \left(\frac{\partial q}{\partial x_i} \right)} \right) = \frac{\partial f}{\partial q}$, for $i = 1, 2, 3, 4$ in this case.

(C) Extension of the Extremum Problem to a n -variable Function

The following outlines the discussion of section 1.3 for $f(x_1 \dots x_n)$, $n \geq 2$.

For a function $f(x_1 \dots x_n)$ we have stationary points when $\frac{\partial f}{\partial x_i} = 0$, for $i = 1, \dots, n$. Further if the determinant $\left| \frac{\partial^2 f}{\partial x_i \partial x_j} \right|_0 \neq 0$ where $i, j = 1, \dots, n$ we can have maxima, minima or saddle points. If $\det \left| \frac{\partial^2 f}{\partial x_i \partial x_j} \right| = 0$ (which corresponds to the eigenvalues λ 's referred in section 1.3 being zero) the conclusion is unpredictable. Also it can be shown that (i) if the determinants $d_1 \dots d_n > 0$, at the equilibrium point ($d_1 \dots d_n$ are given below) we have minima, (ii) if $d_1 \dots d_n$ are alternately $<$ and > 0 we have maxima.

$$\text{Here } d_1 = \frac{\partial^2 f}{\partial x_i^2}, i = 1, \dots, n, d_2 = \begin{vmatrix} \frac{\partial^2 f}{\partial x_i^2} & \frac{\partial^2 f}{\partial x_i \partial x_j} \\ \frac{\partial^2 f}{\partial x_i \partial x_j} & \frac{\partial^2 f}{\partial x_j^2} \end{vmatrix}, (i, j = 1, 2, \dots, n)$$

$$d_3 = \begin{vmatrix} \frac{\partial^2 f}{\partial x_i^2} & \frac{\partial^2 f}{\partial x_i \partial x_j} & \frac{\partial^2 f}{\partial x_i \partial x_k} \\ \frac{\partial^2 f}{\partial x_i \partial x_j} & \frac{\partial^2 f}{\partial x_j^2} & \frac{\partial^2 f}{\partial x_j \partial x_k} \\ \frac{\partial^2 f}{\partial x_i \partial x_k} & \frac{\partial^2 f}{\partial x_k \partial x_j} & \frac{\partial^2 f}{\partial x_k^2} \end{vmatrix}, i, j, k = 1, \dots, n$$

and so on.

(These statements are not proved here and students may see the advanced texts on calculus for proofs).

For example, we take the function $f(x, y, z) = 3x^2 + 5y^2 + 3z^2 - 3xy + 2xz + 6z$

The stationary points are given by $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0$, which yield, $6x - 3y = 0$, $10y - 3x = 0$, $6z + 2x + 6 = 0$.

Solving we get, $x = y = 0$, $z = -1$.

We will have 3 determinants d_1, d_2, d_3 in this case.

They are all positive in this case as can be easily shown.

Hence the stationary point $(0, 0, 1)$ is a point of minimum of the function.

These conditions satisfied by $d_1, d_2, d_3, \dots$ yield those given in the text (see section 1.3 (A)).

Also, it is to be noted that these extrema refer to the local extrema, where $f(x_{10}, x_{20}, \dots) > \text{or} < f(x_1, x_2, \dots)$ for $(x_1, x_2, \dots)$ lying in a close neighbourhood of $x_{10}, x_{20}, \dots$. To find out the global extrema where $f(x_{10}, x_{20}, \dots) > \text{or} < f(x_1, x_2, \dots)$ for all $(x_1, x_2, \dots)$ belonging to the domain of definition of the function, we have to find out first the local extrema and then to ascertain the global maximum or minimum from them. In some cases, the form of the function enables us to determine the global extrema very easily. For example $f(x, y, z) = A - (x^2 + y^2 + z^2)$, A being a positive constant has a global maximum at $x = y = z = 0$.

Vector Analysis

CHAPTER

2

2.1 INTRODUCTION

Vectors occur in almost every branch of physical science, classical mechanics, electrodynamics, relativity to mention a few of them. From your school days you know that a vector is a quantity possessing both magnitude and direction while a scalar has magnitude alone. Thus the velocity, momentum, force etc. belong to the world of vectors while temperature, potential etc. to that of scalars. You have also learnt that a vector is represented geometrically by a directed line segment in space. In this chapter we will address the question of presenting a more rigorous definition of vectors, of introducing a proper representation of vectors and the various operations on vectors. We defer the problem of defining vectors in the most general and abstract way to the Chapter 3.

To represent mathematically a vector and the results of various operations on it we need a coordinate system. The most simple coordinate system is the Cartesian system, where the axes are fixed in space and mutually perpendicular to one another. We then write a vector $\vec{A}$ in terms of the unit vectors $\vec{i}$, $\vec{j}$ and $\vec{k}$ as $\vec{A} = \vec{i}A_x + \vec{j}A_y + \vec{k}A_z$ (Fig. 2.1), where A_x , A_y and A_z are called the components of $\vec{A}$ along the x , y and the z -axes respectively. Of course, there are other coordinate systems in which the axes are not fixed in direction in space but may be mutually perpendicular. We will come across such coordinate systems in the concluding section of this chapter. It is to be remembered, however, that a vector exists independently of any coordinate system and any

physical equation involving vectors (or scalars), Newton's law $\vec{F} = \frac{d\vec{P}}{dt}$ being one of the simplest examples in this context, remains meaningful irrespective of any coordinate system. It is only when we try to represent mathematically the equation that the question of a coordinate system becomes relevant.

Thus the sum of two vectors $\vec{A}$ and $\vec{B}$, where $\vec{B} = \vec{i}B_x + \vec{j}B_y + \vec{k}B_z$ is another vector $\vec{C}$, where $\vec{C} = \vec{i}C_x + \vec{j}C_y + \vec{k}C_z$, $C_x = A_x + B_x$, $C_y = A_y + B_y$, $C_z = A_z + B_z$. $\vec{C}$ is a vector lying in the same three-dimensional space being represented according to the parallelogram law by the diagonal of the parallelogram drawn through the point where the vectors intersect (Fig. 2.2). (Any linear combination of two vectors $\vec{A}, \vec{B}$ given by $\vec{C} = m\vec{A} + n\vec{B}$, m and n being constants, also belongs to the same space). The vector addition is commutative, i.e., $\vec{A} + \vec{B} = \vec{B} + \vec{A}$. The difference of two vectors $\vec{A} - \vec{B}$, given by $\vec{A} - \vec{B} = \vec{A} + (-1)\vec{B}$ is obtained by adding $\vec{A}$ to a vector equal and opposite to $\vec{B}$ according to the parallelogram law. The vector addition is also associative in the sense that $(\vec{A} + \vec{B}) + \vec{C} = \vec{A} + (\vec{B} + \vec{C})$ (Fig. 2.3)

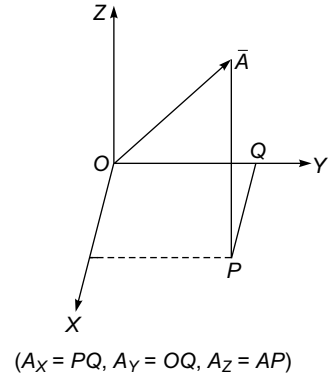


Fig. 2.1

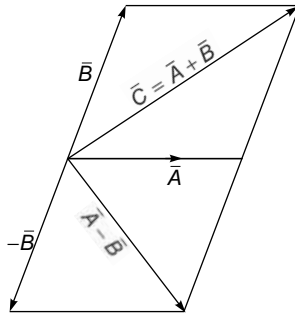


Fig. 2.2 (Parallelogram law)

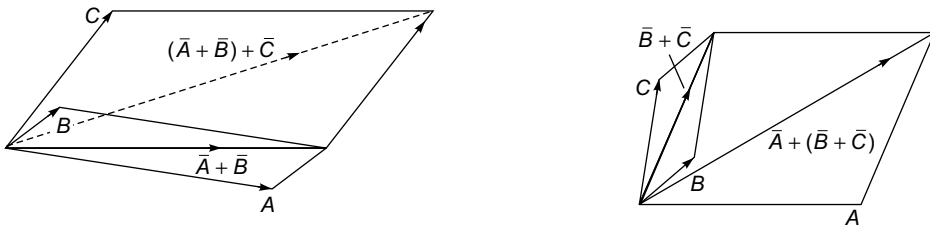


Fig. 2.3 Demonstration of the associative law

2.1.1 Product of Two Vectors

There can be three different types of product of vectors

- Scalar or Dot Product** - The scalar (or dot) product of two vectors defined as $\vec{A} \cdot \vec{B} = AB \cos \theta$ is a scalar, where $\vec{A}$ and $\vec{B}$ are the two vectors and θ is the angle between the two.

Example The work done in causing a displacement given by $\int_A^B \vec{F} \cdot d\vec{r}$, or the kinetic energy of a moving body given by $\int_1^2 \vec{F} \cdot \vec{v} dt$, $\vec{F}$ being the force and $\vec{v}$ the velocity of the body (assuming that the kinetic energy at $t = t_1$ is zero), are examples of the scalar product.

- (b) **Vector (or cross) product** - The vector (or cross) product of two vectors $\vec{A}, \vec{B}$ defined by $\vec{A} \times \vec{B} = |\vec{A}\vec{B} \sin \theta| \vec{n}$ is a vector whose magnitude is $|\vec{A}\vec{B} \sin \theta|$, direction specified by the unit vector $\vec{n}$ is perpendicular to the plane containing $\vec{A}$ and $\vec{B}$ and the sense is such that if we rotate a right handed screw in the anticlockwise direction from $\vec{A}$ to $\vec{B}$, the tip of the screw moves in the positive direction of $\vec{n}$ according to the convention of the right handed system of coordinates. For the left handed system the direction will be exactly opposite (Fig. 2.4).

Example The angular momentum $\vec{L}$ of a particle $\vec{r}$ given by $\vec{L} = \vec{r} \times \vec{p}$ ($\vec{r}$, the radius vector of the particle with respect to an origin 0, $\vec{p}$, its linear momentum and $\vec{L}$ is the angular momentum with respect to 0), the moment of a force $\vec{F}$ about the point 0 given by $\vec{N} = \vec{r} \times \vec{F}$ are examples of the vector product.

- (c) **Direct product** - The direct product of two vectors $\vec{A}, \vec{B}$ written as $\vec{A} \vec{B}$ is an object consisting of nine components ($A_x B_x, A_x B_y, A_x B_z, A_y B_x, A_y B_y, A_y B_z, A_z B_x, A_z B_y, A_z B_z$) behaves as a tensor of second rank (described in detail in Chapter 4). It is to be noted that a scalar and a vector behave as tensors of rank zero and one respectively. The examples of tensors are given in Chapter 4 and hence not discussed here.

Note: In three dimensions a vector $\vec{A}$ is written as $\vec{A} = \vec{i}A_x + \vec{j}A_y + \vec{k}A_z$. The vectors $\vec{i}, \vec{j}$ and $\vec{k}$ are said to form a basis in terms of which $\vec{A}$ is expressed. In order to form a basis the vectors must be linearly independent (explained in detail in Chapter 3). If $\vec{A}_1, \vec{A}_2, \dots, \vec{A}_n$ are linearly independent vectors in n -dimensions, the equation $C_1 \vec{A}_1 + C_2 \vec{A}_2 + \dots + C_n \vec{A}_n = 0$ has the only solution $C_1 = C_2 = \dots = C_n = 0$. That the unit vectors $\vec{i}, \vec{j}$ and $\vec{k}$ are linearly independent can be proved by taking the dot products of $\vec{i}, \vec{j}, \vec{k}$ successively with the equation $C_1 \vec{i} + C_2 \vec{j} + C_3 \vec{k} = 0$ to give $C_1 = C_2 = C_3 = 0$. Of course for linear independence in three dimensions the three vectors need be non-coplanar, not necessarily perpendicular (Sec 2.1.2).

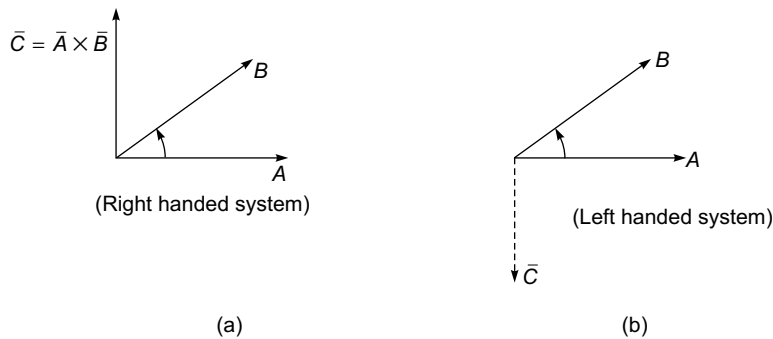


Fig. 2.4

2.1.2 Triple Product

- (a) *Scalar triple product* - The scalar triple product $\vec{A} \cdot (\vec{B} \times \vec{C})$ is equal to $\vec{A} \cdot [BC \sin \theta \vec{n}] = (BC \sin \theta)(\vec{A} \cdot \vec{n})$. Now $BC \sin \theta$ is the area of the parallelogram formed by $\vec{B}$ and $\vec{C}$, the sides of the base respectively and $\vec{n}$ the unit normal to the area (Fig. 2.5).

Clearly, $\vec{A} \cdot \vec{n}$ is the height of the parallelepiped formed by the sides $\vec{A}$, $\vec{B}$ and $\vec{C}$ respectively. Thus, $\vec{A} \cdot (\vec{B} \times \vec{C})$ is the volume of the parallelepiped. It can easily be seen that

$$\vec{A} \cdot (\vec{B} \times \vec{C}) = \vec{B} \cdot (\vec{C} \times \vec{A}) = \vec{C} \cdot (\vec{A} \times \vec{B})$$

- (b) *Vector triple product* - The vector triple product is defined by $\vec{A} \times (\vec{B} \times \vec{C})$ (Note that here the order of $\vec{A}$, $\vec{B}$ and $\vec{C}$ are important).

The vector $\vec{A} \times (\vec{B} \times \vec{C})$ lies in the plane defined by $\vec{B}$ and $\vec{C}$ hence can be written as $m\vec{B} + n\vec{C}$. Since this holds good for any choice of $\vec{A}$, $\vec{B}$ and $\vec{C}$ we take $\vec{A} = \vec{i} A_x + \vec{j} A_y + \vec{k} A_z$, $\vec{B} = \vec{i} B_x + \vec{j} B_y$ and $\vec{C} = \vec{k} C_z$ for our convenience. Thus, $[\vec{i} A_x + \vec{j} A_y + \vec{k} A_z] \times [\vec{i} B_x C_z - \vec{j} B_y C_z] = m[\vec{i} B_x + \vec{j} B_y] + n\vec{k} C_z$.

Since $\vec{i}$, $\vec{j}$, $\vec{k}$ are linearly independent we equate the coefficients of $\vec{i}$, $\vec{j}$ and $\vec{k}$ respectively to obtain, $m = A_z C_z = \vec{A} \cdot \vec{C}$, $n = -(A_x B_x + A_y B_y) = -\vec{A} \cdot \vec{B}$. Since this is a vector relation this is valid equally for all choices of the coordinate system. Hence we can write in general,

$$\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B})$$

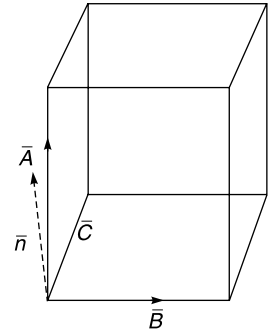


Fig. 2.5

Note: The condition that the three vectors $\vec{A}$, $\vec{B}$ and $\vec{C}$ be linearly independent is that $\vec{A} \cdot (\vec{B} \times \vec{C}) \neq 0$, i.e., that the three vectors are non-coplanar. This can also be seen in the following way.

Let any fourth vector $\vec{D}$ be written as a linear combination of $\vec{A}$, $\vec{B}$ and $\vec{C}$. Thus we write, $\vec{D} = C_1 \vec{A} + C_2 \vec{B} + C_3 \vec{C}$. Expanding in terms of the components we can write, $D_x = C_1 A_x + C_2 B_x + C_3 C_x$, $D_y = C_1 A_y + C_2 B_y + C_3 C_y$ and $D_z = C_1 A_z + C_2 B_z + C_3 C_z$. Solving the three inhomogeneous equations we see that C_1 , C_2 and

C_3 have got unique values provided the determinant $\begin{vmatrix} A_x & B_x & C_x \\ A_y & B_y & C_y \\ A_z & B_z & C_z \end{vmatrix}$ is non-zero and hence the scalar triple product $\vec{A} \cdot (\vec{B} \times \vec{C}) \neq 0$. (2.1)

Students can also check this by expanding $\vec{A}$, $\vec{B}$, $\vec{C}$ in terms of $\vec{i}$, $\vec{j}$, $\vec{k}$ and applying the condition of linear independence.

2.2 TRANSFORMATION OF AXES, POLAR VECTOR, AXIAL VECTOR

Let us consider a transformation from a set of Cartesian axes to another identical set. The transformation is, in general, a linear one and can be represented as,

$x' = a_{11}x + a_{12}y + a_{13}z$, $y' = a_{21}x + a_{22}y + a_{23}z$ and $z' = a_{31}x + a_{32}y + a_{33}z$, where (x, y, z) and (x', y', z') refer to the components of the position vector of a point relative to the two sets respectively, the origins of two sets being assumed to coincide. For rotation about the z -axis (say) of the coordinate system, through an angle θ in the anticlockwise direction, $a_{11} = \cos \theta$, $a_{12} = \sin \theta$, $a_{13} = 0$, $a_{21} = -\sin \theta$, $a_{22} = \cos \theta$, $a_{23} = 0$, $a_{31} = a_{32} = 0$, $a_{33} = 1$ (Fig. 2.6).

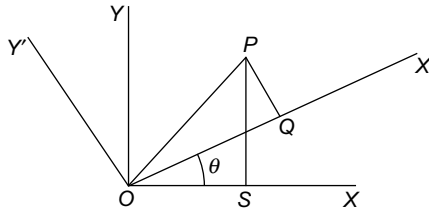


Fig. 2.6

$x = OS, y = PS, x' = OQ, y' = PQ$. By simple geometry it can be proved that $x' = x \cos \theta + y \sin \theta, y' = -x \sin \theta + y \cos \theta, z' = z$.

The above rule of transformation is valid equally for all types of vectors under rotation of axes, viz., for a vector $\bar{A}$ we can write thus,

$$A'_x = A_x a_{11} + A_y a_{12} + A_z a_{13}, A'_y = A_x a_{21} + A_y a_{22} + A_z a_{23}, A'_z = A_x a_{31} + A_y a_{32} + A_z a_{33} \quad (2.1)$$

For inversion of an axis the corresponding component of the vector changes sign. Now, let us consider the transformation of the components of the vector $\bar{C} = \bar{A} \times \bar{B}$. Clearly, for the inversions of all the three axes all the components of the vectors $\bar{A}$ and $\bar{B}$ change sign whereas the component of the vector $\bar{C}$ does not change sign. Similarly for an inversion of only one axis, the corresponding component of $\bar{A}$ or $\bar{B}$ changes sign leaving the other components unaffected while just the opposite thing happens for $\bar{C}$. (Thus for an inversion of the x -axis, A_x and B_x alter signs leaving A_y, A_z, B_y, B_z the same while $C_y (\bar{C} = \bar{A} \times \bar{B})$ and C_z change signs leaving C_x unchanged). For a pure rotation however it can be seen that the transformation of the components of $\bar{C}$ follows the same rule as that obeyed by $\bar{A}$ or $\bar{B}$. (see Eq. 2.2). Take, for simplicity a rotation about the z -axis, say, where

$$\begin{aligned} a_{11} &= \cos \theta, \\ a_{12} &= \sin \theta, \\ a_{13} &= 0, a_{21} = -\sin \theta, \\ a_{22} &= \cos \theta, a_{23} = 0 \\ a_{31} &= a_{32} = 0 \\ a_{33} &= 1 \end{aligned}$$

Using the formula of vector product we can easily show that $\bar{C} = \bar{A} \times \bar{B}$, transforms in the same way as $\bar{A}$ or $\bar{B}$. The vector $\bar{A}$ or $\bar{B}$ is called the polar vector (for example, position vector, velocity vector, the electric field vector etc.) while the vector $\bar{C}$ formed by taking the cross product of two polar vectors is called the axial vector (for example, angular momentum, moment of a force about a point, the vector $\bar{B}$ etc.). This difference in the rule of transformation is best expressed in the following way.

$A'_i = a_{ij} A_j$ (Summation over j is implied), $B'_i = a_{ij} B_j$ while $C'_i = a_{ij} C_j |\det a_{ij}|$. Thus $|\det a_{ij}|$, the determinant of the matrix formed by the coefficient a_{ij} , equals 1 for pure rotation or for an even number of reflections of axes (for $x \rightarrow -x, y \rightarrow -y, z \rightarrow z$). For an odd number of reflections where $x \rightarrow -x, y \rightarrow -y, z$

$\rightarrow -z$, say the matrix is given by $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ with determinant -1 . The transformations of the polar and

the axial vectors are hence identical in the first case and different in the second. The transformation matrix for rotation through θ about the z -axis in anticlockwise direction (described in detail in the next chapter) is given

by $\begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and for an even number of reflections $x \rightarrow -x, y \rightarrow -y, z \rightarrow z$, say, by $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

with determinant 1 in either case.

(Note that in the latter case the handedness of the coordinate system changes). It can be easily shown that $\vec{A} \cdot (\vec{B} \times \vec{C})$ changes sign on an odd number of inversions of axes and is hence called a pseudo scalar.

The characteristic relations expressing the transformations of the components of the polar and the axial vectors as discussed above are valid, in general, for an arbitrary number of dimensions and regarded often as the criteria for a set of quantities to behave as the components of a vector (See Chapter 4) for more details). In the Chapter 3 a more general and abstract definition of vectors is given.

2.3 VECTOR EQUATIONS OF LINES, CURVES, SURFACES

(1) Line

- (a) *Straight line* - Since a straight line maintains a constant direction the equation of a straight line is given by $(\vec{r} - \vec{r}_0) \times \vec{e} = 0$, where $\vec{r}, \vec{r}_0$ are the radii vectors of a moving point and a fixed point on the line with respect to an origin O respectively and $\vec{e}$ is the unit vector along the line. The equation is also written as $\vec{r} = \vec{r}_0 + \vec{i}u + \vec{j}v + \vec{k}w$, where $\vec{i}, \vec{j}, \vec{k}$ are the unit vectors along the x, y and the z -axes respectively (Fig. 2.7).

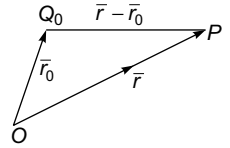


Fig. 2.7

- (b) *Circle* - The equation is $\vec{r} = (\vec{i} \cos \theta + \vec{j} \sin \theta) a$, $\vec{i}, \vec{j}$ being the unit vectors along the two perpendicular axes through the centre, a the radius and θ the angle made by the radius vector with the x -axis (here the origin is the centre of the circle).
- (c) *Ellipse* - The vector equation of an ellipse is $\vec{r} = (\vec{i} a \cos \phi + \vec{j} b \sin \phi)$; the centre of the ellipse being the origin and $(a \cos \phi, b \sin \phi)$ the parametric representation of any point on the ellipse
- $$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$
- (d) The equation of a circular helix is, $\vec{r} = \vec{i} a \cos \theta + \vec{j} a \sin \theta + \vec{k} b \theta$, the xy -plane being the plane of the circle and $2\pi b$ the pitch of the helix.
- (e) The equation of a plane through a point with radius vector $\vec{r}$ relative to the origin O is $\vec{r} \cdot \vec{n} = p$, where $\vec{n}$ is the unit vector perpendicular to the plane and p the length of the normal from O to the plane. For a plane passing through three non-collinear points the equation is $(\vec{r} - \vec{r}_1) \cdot [(\vec{r} - \vec{r}_2) \times (\vec{r} - \vec{r}_3)] = 0$. ($\vec{r}_1, \vec{r}_2, \vec{r}_3$ are the radius vectors of the three points).
- (f) The equation of a surface whose equation is given by $z = f(x, y)$ is $\vec{r} = \vec{i}x + \vec{j}y + \vec{k} f(x, y)$. For a sphere of radius a the equation is, $\vec{r} = \vec{i}x + \vec{j}y + \vec{k} (a^2 - x^2 - y^2)^{1/2}$. It can also be written more simply as $|\vec{r}| = a$, the centre being taken as the origin.

Thus, in general a point of the surface is denoted by the radius vector $\vec{r}$ which is a function of two independent parameters p, q and s is a parameter specifying the surface so that $\vec{r} = \vec{r}(s) = \vec{r}(p(s), q(s))$. The

infinitesimal element of area, which is approximately a parallelogram with sides $\frac{\partial \bar{r}}{\partial p} dp$ and $\frac{\partial \bar{r}}{\partial q} dq$ respectively is equal to $\left(\frac{\partial \bar{r}}{\partial p} \times \frac{\partial \bar{r}}{\partial q} \right) dp dq$ (Fig. 2.8).

Thus the total area of the surface bounded by the range of the variables

$$p \text{ and } q = \int \left| \frac{\partial \bar{r}}{\partial p} \times \frac{\partial \bar{r}}{\partial q} \right| dp dq.$$

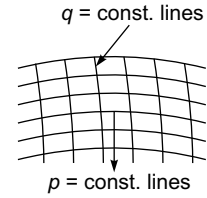


Fig. 2.8

The infinitesimal element of length tangential to the surface is given by

$$ds = |d\bar{r} \cdot d\bar{r}|^{1/2} = |A dp^2 + B dq^2 + C dp dq|^{1/2}, \text{ where } A = \frac{\partial \bar{r}}{\partial p} \cdot \frac{\partial \bar{r}}{\partial p}, B = \frac{\partial \bar{r}}{\partial q} \cdot \frac{\partial \bar{r}}{\partial q}, C = 2 \frac{\partial \bar{r}}{\partial p} \cdot \frac{\partial \bar{r}}{\partial q}.$$

Thus, a spherical surface is specified by parameters θ, ϕ . The infinitesimal element of area tangential to the surface is $\left| \frac{\partial \bar{r}}{\partial \theta} \times \frac{\partial \bar{r}}{\partial \phi} \right| d\theta d\phi$. Now

$\bar{r} = \bar{i} R \sin \theta \cos \phi + \bar{j} R \sin \theta \sin \phi + \bar{k} R \cos \theta$, R being the radius so that the element of area $d\bar{s} = R^2 \sin \theta [\sin \theta \cos \phi \bar{i} + \sin \theta \sin \phi \bar{j} + \cos \theta \bar{k}] d\theta d\phi$ [Assume this expression of $\bar{r}$ for the time being. This will be discussed further in the context of spherical polar coordinates] (Fig. 2.9).

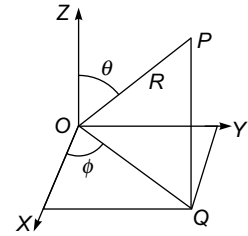


Fig. 2.9

Thus the magnitude of the surface area of the entire sphere.

$$= \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} R^2 \sin \theta d\theta d\phi = 4\pi R^2 [\bar{i} \sin \theta \cos \phi + \bar{j} \sin \theta \sin \phi + \bar{k} \cos \theta \text{ is the unit normal to the surface}].$$

The square of the infinitesimal line element tangential to the sphere is given by

$$ds^2 = \frac{\partial \bar{r}}{\partial \theta} \cdot \frac{\partial \bar{r}}{\partial \theta} d\theta^2 + \frac{\partial \bar{r}}{\partial \phi} \cdot \frac{\partial \bar{r}}{\partial \phi} d\phi^2 + 2 \frac{\partial \bar{r}}{\partial \theta} \cdot \frac{\partial \bar{r}}{\partial \phi} d\theta d\phi$$

or, ds^2 by explicit calculation $= R^2 (d\theta^2 + \sin^2 \theta d\phi^2)$.

2.4 DIFFERENTIATION OF SCALAR AND VECTOR FIELDS

A scalar and a vector field are specified by specifying the scalar and the vector (at a particular time) at every point in a certain region of space.

For example the specification of the temperature or the potential at every point constitutes a scalar field while the specification of the velocity or the force at every point constitutes a vector field.

Thus if $\bar{r}$ and $\bar{r} + \Delta \bar{r}$ be two closely separated points, the change $\Delta \phi$ of a scalar ϕ in going from one point to the other is given by $\Delta \phi = \phi(\bar{r} + \Delta \bar{r}) - \phi(\bar{r}) = \phi(x + \Delta x, y + \Delta y, z + \Delta z) - \phi(x, y, z)$ in terms of the coordinates of two points.

Thus $\Delta\phi = \frac{\partial\phi}{\partial x} \Delta x + \frac{\partial\phi}{\partial y} \Delta y + \frac{\partial\phi}{\partial z} \Delta z$ (neglecting higher order terms in $\Delta x, \Delta y, \Delta z$) or, the rate of change of ϕ with respect to a displacement along a particular direction represented by $\bar{r}$ is given by

$$\frac{d\phi}{ds} = \lim_{\Delta s \rightarrow 0} \frac{\Delta\phi}{\Delta s} = \frac{\partial\phi}{\partial x} \frac{dx}{ds} + \frac{\partial\phi}{\partial y} \frac{dy}{ds} + \frac{\partial\phi}{\partial z} \frac{dz}{ds}, \text{ where } \frac{dx}{ds}, \frac{dy}{ds} \text{ and } \frac{dz}{ds} \text{ are the } x, y, z \text{ components of the unit vector } \bar{n}, \text{ respectively} \quad (2.2)$$

Similarly for a vector function $\bar{A}(x, y, z)$ the derivative $\frac{d\bar{A}}{ds}$ is given by

$$\frac{d\bar{A}}{ds} = \lim_{\Delta s \rightarrow 0} \frac{\Delta\bar{A}}{\Delta s} = \frac{\partial\bar{A}}{\partial x} \frac{dx}{ds} + \frac{\partial\bar{A}}{\partial y} \frac{dy}{ds} + \frac{\partial\bar{A}}{\partial z} \frac{dz}{ds}.$$

Thus writing in Cartesian components $\bar{A} = \bar{i} A_x + \bar{j} A_y + \bar{k} A_z$ we obtain,

$$\begin{aligned} \frac{d\bar{A}}{ds} &= \left(\bar{i} \frac{\partial A_x}{\partial x} + \bar{j} \frac{\partial A_y}{\partial x} + \bar{k} \frac{\partial A_z}{\partial x} \right) \frac{dx}{ds} + \left(\bar{i} \frac{\partial A_x}{\partial y} + \bar{j} \frac{\partial A_y}{\partial y} + \bar{k} \frac{\partial A_z}{\partial y} \right) \frac{dy}{ds} \\ &\quad + \left(\bar{i} \frac{\partial A_x}{\partial z} + \bar{j} \frac{\partial A_y}{\partial z} + \bar{k} \frac{\partial A_z}{\partial z} \right) \frac{dz}{ds} \\ &= \bar{i} \frac{dA_x}{ds} + \bar{j} \frac{dA_y}{ds} + \bar{k} \frac{dA_z}{ds} \left[\because \frac{dA_l}{ds} = \frac{\partial A_l}{\partial x} \frac{dx}{ds} + \frac{\partial A_l}{\partial y} \frac{dy}{ds} + \frac{\partial A_l}{\partial z} \frac{dz}{ds}, l = x, y, z \right] \end{aligned} \quad (2.3)$$

Thus defining a vector $\nabla\bar{\phi}$, called the gradient of ϕ whose components in Cartesian coordinates are $\frac{\partial\phi}{\partial x}$, $\frac{\partial\phi}{\partial y}$ and $\frac{\partial\phi}{\partial z}$ respectively we can write from (2.2), $\frac{d\phi}{ds} = \nabla\bar{\phi} \cdot \bar{n}$. (2.4)

Equation (2.4) being a scalar equation we can interpret it as defining $\nabla\bar{\phi}$ in a coordinate independent way as a vector whose scalar product with a unit vector $\bar{n}$ gives the spatial rate of change of ϕ along the direction $\bar{n}$. Clearly $\frac{d\phi}{ds}$ is maximum when $\nabla\bar{\phi} \parallel \bar{n}$. Thus the direction of $\nabla\bar{\phi}$ denotes that of the maximum spatial rate of change of ϕ . If we consider two surfaces specified by two constant values of ϕ , the maximum rate of change of ϕ occurs at a point along the normal drawn at that point to the $\phi = \text{constant}$ surface.

Thus, $\nabla\bar{\phi}$ at a point is along the normal drawn at that point to the $\phi = \text{constant}$ surface passing through it. Hence, if we are given a surface $\phi = \text{constant}$, the unit normal to the surface is given by $\frac{\nabla\bar{\phi}}{|\nabla\bar{\phi}|}$. The gradient

thus behaves as a vector operator that acting on a scalar gives a vector. The components $\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}$

of $\nabla\bar{\phi}$ transform like those of a vector as can be seen by applying the rule $\frac{\partial}{\partial x'_i} = \sum_j \frac{\partial}{\partial x_j} \frac{\partial x_j}{\partial x'_i}$ and $\phi = \phi'$, ϕ being a scalar. $\left(\frac{\partial x_j}{\partial x'_i} = a_{ij} \text{ as mentioned in Section 2.2} \right)$

2.5 INTEGRALS OF VECTORS

1. Line Integral

The line integral of a vector field along a given curve in space joining the points 1 and 2 given by $\int_1^2 \bar{A} \cdot d\bar{r}$ is equal to $\lim_{\Delta r_i \rightarrow 0} \sum_i A_i \Delta r_i \cos \theta_i \Delta r_i$, θ_i being the angle between the vector $\bar{A}$ and the line element $\Delta \bar{r}_i$ at the i th point of division of the curve. In terms of the components it can also be represented as $\int_1^2 (A_x dx + A_y dy + A_z dz)$, where each of the components of $\bar{A}$ is a function of the coordinates of the point. If $\int_1^2 \bar{A} \cdot d\bar{r}$ is independent of the curve joining 1 and 2, then $\oint \bar{A} \cdot d\bar{r}$ around any closed curve in a simply connected region, i.e., a region with no holes (See Appendix) is clearly zero, because

$$\begin{aligned} \oint \bar{A} \cdot d\bar{r} &= \int_{c_1 1}^2 \bar{A} \cdot d\bar{r} + \int_{c_2 2}^1 \bar{A} \cdot d\bar{r} \text{ (Fig. 2.10)} \\ &= \int_{c_1}^2 \bar{A} \cdot d\bar{r} - \int_{c_2}^2 \bar{A} \cdot d\bar{r} \\ &= 0. \end{aligned}$$

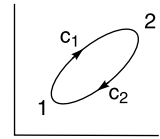


Fig. 2.10

If $\bar{A} = \nabla\bar{\phi}$, where ϕ is a single-valued function of position, then $\oint \bar{A} \cdot d\bar{r} = \oint d\phi = 0$. Conversely, if $\oint \bar{A} \cdot d\bar{r} = 0$ along any simple closed curve (i.e., non self-intersecting) in the region, we can write $\bar{A} \cdot d\bar{r} = d\phi$, ϕ being a single-valued function of the coordinates. Since the above equality is valid for arbitrary dx , dy and dz , $A_x = \frac{\partial \phi}{\partial x}$, $A_y = \frac{\partial \phi}{\partial y}$ and $A_z = \frac{\partial \phi}{\partial z}$, allowing us to find out $\phi(x, y, z)$ uniquely except for a pure additive constant. Thus, $\bar{A} = \nabla\bar{\phi}$ is the necessary and sufficient condition for $\oint \bar{A} \cdot d\bar{r} = 0$ over any closed path.

Note:

- (1) The condition of single-valuedness is important here. For example, if we take $\bar{A} = -\bar{i} \frac{y}{x^2 + y^2} + \bar{j} \frac{x}{x^2 + y^2}$. Putting $y = r \sin \phi$, $x = r \cos \phi$

$$\oint \bar{A} \cdot d\bar{r} \text{ along a circle} = \int_0^{2\pi} (\cos^2 \phi + \sin^2 \phi) d\phi = 2\pi.$$

Putting $A_x = -\frac{y}{x^2 + y^2} = \frac{\partial \phi}{\partial x}$, $A_y = \frac{x}{x^2 + y^2} = \frac{\partial \phi}{\partial y}$, we get,

$\phi = \tan^{-1} \frac{y}{x} + \text{constant}$ after integration. Clearly here through $\bar{A} = \nabla \bar{\phi}$, ϕ is no longer a single-valued function of position because of the multi-valuedness of $\tan^{-1} \frac{y}{x}$. Hence $\oint \bar{A} \cdot d\bar{r} \neq 0$ here. The components A_x, A_y are not defined at the point $(0, 0)$ and hence an infinitesimally small region around $(0, 0)$ has to be deleted in defining the integral. The region thus becomes annular shaped and hence is multiply connected. If we convert the region to a simply connected one the integral becomes zero. (Refer to appendix).

- (2) Suppose the integration is performed along a curve which includes regions where the single-valuedness is absent in the relation between the variables. For example, in the integration around a circle $y = \sqrt{a^2 - x^2}$ for θ lying between 0 and π and $y = -\sqrt{a^2 - x^2}$ for θ between π and 2π . In such a case the integration region has to be divided into parts in each of which y is a single-valued function of x and the total integral is the sum of all such integrals, over the respective regions.

Vector fields for which $\oint_C \bar{F} \cdot d\bar{r} = 0$ over any closed path C in space are said to be conservative. This terminology arises from analogy with mechanical problems. For such a field $\int_A^B \bar{F} \cdot d\bar{r}$ is independent of the path connecting A and B and depends on the coordinates of the end points alone. Thus if $\bar{F}$ denotes the force acting on the particle then $\int_A^B \bar{F} \cdot d\bar{r}$ denotes the work done in moving it from A to B . Thus writing

$$\bar{F} = m \frac{d\bar{V}}{dt} \text{ we can write, } \int_A^B \bar{F} \cdot d\bar{r} = \int_A^B m \frac{d\bar{V}}{dt} \cdot \bar{v} dt = d\left(\frac{1}{2} m V^2\right) \Big|_A^B = T_B - T_A, \text{ where } T_B \text{ and } T_A \text{ are the}$$

kinetic energies of the particle at B and A respectively. According to our earlier discussions $\bar{F}$ can then be written as $\bar{F} = -\nabla \bar{\phi}$, where ϕ is the potential energy at that point ($-ve$ sign is given because $\bar{F}$ acts from regions of high to low potential). Thus we can write $\int_A^B \bar{F} \cdot d\bar{r} = \int_A^B -\nabla \bar{\phi} \cdot d\bar{r} = -d\phi \Big|_A^B = \phi_A - \phi_B$.

Thus equating $T_B - T_A$ to $\phi_A - \phi_B$ we obtain that $T_A + \phi_A = T_B + \phi_B$ and hence the mechanical energy of the particle is conserved.

Example Let us check if $\bar{F} = \bar{i}(y^2 + z^2) + \bar{j}(2xy + y^2) + \bar{k}(2xz)$ is conservative. For a conservative field we can find out a ϕ so that $\bar{F} = \nabla \bar{\phi}$. Putting $\bar{F} = \nabla \bar{\phi}$ we get, $\frac{\partial \phi}{\partial x} = y^2 + z^2$, $\frac{\partial \phi}{\partial y} = 2xy + y^2$, $\frac{\partial \phi}{\partial z} = 2xz$.

Integrating thus, we obtain, $\phi = x(y^2 + z^2) + \psi(y, z)$, where $\psi(y, z)$ is any arbitrary function of y, z .

$$\text{Now, } \frac{\partial \phi}{\partial y} = 2xy + \frac{\partial \psi}{\partial y} = 2xy + y^2, \text{ which gives } \psi = \frac{y^3}{3} + \eta(z).$$

The last condition gives, $2xz + \frac{d\eta}{dz} = 2xz$, yielding $\eta = \text{constant}$. Hence $\phi = x(y^2 + z^2) + \frac{y^3}{3} + \text{constant}$ proving thereby that the vector field is conservative.

2. Surface integral

The surface integral of a vector over a surface $\oint_s \vec{A} \cdot d\vec{s}$ is defined in analogy with Riemann's definition of the definite integral as $\lim_{\Delta s_i \rightarrow 0} \sum_i \vec{A}_i \cdot \Delta \vec{s}_i$, where the entire surface s is assumed to be divided into infinitesimal surfaces. If the

surface be closed bounding a volume then the positive direction of the vector $\Delta \vec{s}_i$ is taken to be pointing outwards from the volume. If the surface be open then the positive direction of the area vector is found out by the movement of the tip of the right handed screw when it is moved in an anticlockwise direction along the boundary of the area (Figs. 2.11(a) and 2.11(b)).

If the equation of the surface be specified by two parameters u, v the surface integral over a surface where u, v vary in a certain range is given by $\int f(x(u, v), y(u, v)) F(u, v) du dv$.

In certain problems the calculation of the surface integral is simplified if the surface is projected on some coordinate plane, say, the xy -plane. Thus $ds |\cos \theta| = dA$, where θ is the angle between the positive direction of the normal $\vec{n}$ to ds and the z -axis. Thus $ds |\vec{n} \cdot \vec{k}| = dA$. The surface integral $\int f(x, y, z) ds$ becomes $\int f(x, y, z) \frac{dA}{|\vec{n} \cdot \vec{k}|}$, (dA being $dx dy$, the projection of ds on the xy -plane). If the equation of the surface be given by $\phi(x, y, z) = \text{constant}$, $\vec{n} = \frac{\nabla \phi}{|\nabla \phi|}$, so that the surface integral = $\int f(x, y, z) \frac{dA}{|\nabla \phi \cdot \vec{k}|} |\nabla \phi|$ (Fig. 2.12).

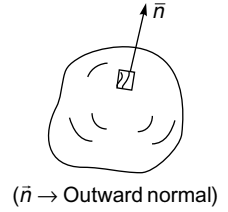


Fig. 2.11(a)

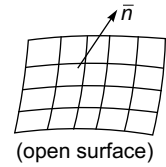


Fig. 2.11(b)

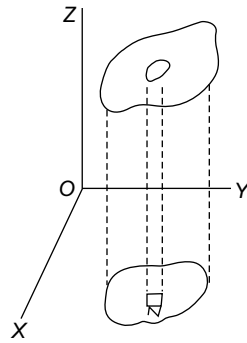


Fig. 2.12

Example

- (1) Surface area of a hemisphere of radius a ,

$$S = a \int \frac{dx dy}{z} \frac{|\nabla \phi|}{|\nabla \phi|} = \frac{\vec{r}}{a} [\vec{r} = \vec{i}x + \vec{j}y + \vec{k}z]$$

$$= a \int \frac{dx dy}{\sqrt{a^2 - x^2 - y^2}} = a \int_{-a}^a dx \int_{y=-\sqrt{a^2-x^2}}^{y=\sqrt{a^2-x^2}} \frac{dy}{\sqrt{a^2 - x^2 - y^2}}$$

or, $S = 2\pi a^2$ after carrying out the integration.

- (2) Let us calculate $\int \bar{A} \cdot d\bar{s}$ over the surface of a hemisphere $x^2 + y^2 + z^2 = R^2$ having its centre at the origin, where $\bar{A} = \bar{j}y^3$.

Following the previous prescription $\int_s \bar{A} \cdot d\bar{s} = a \int y^3 \frac{\bar{j} \cdot \bar{r}}{az} dx dy$

$$\text{or, } \int_s \bar{A} \cdot d\bar{s} = \iint \frac{y^4}{\sqrt{a^2 - x^2 - y^2}} dx dy \quad [y \text{ from } -\sqrt{a^2 - x^2} \text{ to } \sqrt{a^2 - x^2} \text{ and } x \text{ from } -a \text{ to } a]$$

$$= \frac{3\pi}{8} \int_{-a}^a (a^2 - x^2)^2 dx = \frac{2\pi a^5}{5}$$

3. Volume Integral

The volume integral of a scalar function $f(x, y, z)$ over a volume is given by $\int_V f(x, y, z) dv = \lim_{\Delta v_i \rightarrow 0} \sum_i f(x_i, y_i, z_i) \Delta v_i$ according to the usual definition of Riemann integration. An identical definition follows in the case of the integration of a vector over a volume.

Example The volume of a hemisphere with the equation of the surface given by $x^2 + y^2 + z^2 = a^2$ ($z > 0$) is,

$$V = \iiint dx dy dz = \iint dx dy \int_0^{\sqrt{a^2 - x^2 - y^2}} dz = \iint \sqrt{a^2 - x^2 - y^2} dx dy$$

where x extends from $-a$ to a and for a fixed x , y runs from $-\sqrt{a^2 - x^2}$ to $\sqrt{a^2 - x^2}$.

$$\text{Thus } V = \int_{-a}^a dx \int_{-p}^p \sqrt{p^2 - y^2} dy, \text{ where } p = \sqrt{a^2 - x^2}$$

$$\text{or, } V = \int_{-a}^a \frac{(a^2 - x^2)}{2} dx = \frac{2}{3} \pi a^3$$

Suppose we want to find out the centroid of such a uniform hemisphere. The coordinates of the centroid are

$$\bar{x} = \frac{\iiint x dx dy dz}{V} = 0, \bar{y} = \frac{\iiint y dx dy dz}{V} = 0$$

[because x and y are odd functions of x and y respectively].

$$\bar{z} = \frac{\iiint_V z \, dx \, dy \, dz}{V} = \frac{2}{3V} \int_{-a}^a (a^2 - x^2)^{3/2} dx \text{ after performing the } z \text{ and } y \text{ integrations respectively.}$$

$$\text{or, } \bar{z} = \frac{3}{8} a.$$

[The integral could be carried out more simply by using spherical polar coordinates which we will discuss in a subsequent section].

2.6 DIVERGENCE OF A VECTOR FIELD; GAUSS DIVERGENCE THEOREM

The surface integral $\int_S \bar{A} \cdot d\bar{s}$ is called the flux of the vector $\bar{A}$ over the surface S . If $\bar{A}$ represents the velocity vector of flow of a fluid, $\int_S \bar{A} \cdot d\bar{s}$ represent the volume of the fluid flowing through the boundary of the surface per unit time. For the closed surface enclosing a volume v the flux, following our convention of sign for the positive direction of the normal to the surface, is clearly positive if there is a net outflow through the area. The flux is a pseudoscalar if $\bar{A}$ be a polar vector since the area vector is an axial vector. For no source or sink of the fluid the volume of fluid flowing out of the volume through the bounding surface is equal to the loss of the liquid from the volume.

$$\text{If } \rho \text{ be the density of the fluid clearly } \frac{\partial}{\partial t} \int_V \rho dV = - \int_S \rho \bar{V} \cdot d\bar{s} \quad (2.5)$$

Let us calculate the flux $\int_S \bar{A} \cdot d\bar{s}$ across a surface S surrounding a point. Since an extremely small surface can be looked upon as almost flat, we assume an infinitesimally small volume surrounding the point and bounded by the surface, that is small and hence essentially flat.

Let the rectangular parallelepiped be centred at (x, y, z) with sides Δx , Δy and Δz respectively (Fig. 2.13). Let the vector $\bar{A}$ have components A_x, A_y, A_z at the centre. The flux of $\bar{A}$ across the area $ABCD$ is equal to $\left(A_x + \frac{\partial A_x}{\partial x} \cdot \frac{\Delta x}{2} \right) \Delta y \Delta z$

$\Delta y \Delta z$ [The areas being very small we take the magnitude of A_x over the entire area $ABCD$ as being almost equal to that of A_x at the centre of the area]. Similarly the flux of $\bar{A}$ through the opposite surface $EFGH$ can be taken to be equal to $-\left(A_x - \frac{\partial A_x}{\partial x} \cdot \frac{\Delta x}{2} \right) \Delta y \Delta z$, the negative sign preceding the bracket arising from the fact that the direction of the outward normal to $EFGH$ is along the negative x -axis. Thus summing over the contributions from all the faces we obtain, $\oint_S \bar{A} \cdot d\bar{s}$ over the surfaces

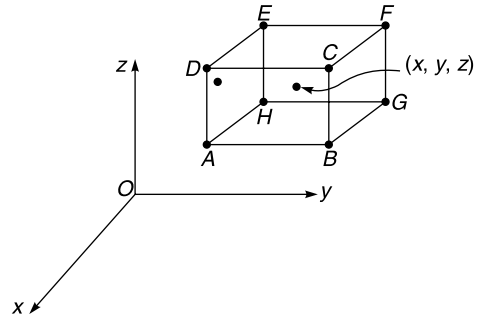


Fig. 2.13

$$\begin{aligned}
 &= \left(A_x + \frac{\partial A_x}{\partial x} \cdot \frac{\Delta x}{z} \right) \Delta y \Delta z - \left(A_x - \frac{\partial A_x}{\partial x} \cdot \frac{\Delta x}{z} \right) \Delta y \Delta z \\
 &+ \left(A_y + \frac{\partial A_y}{\partial y} \cdot \frac{\Delta y}{z} \right) \Delta x \Delta z - \left(A_y - \frac{\partial A_y}{\partial y} \cdot \frac{\Delta y}{z} \right) \Delta x \Delta z \\
 &+ \left(A_z + \frac{\partial A_z}{\partial z} \cdot \frac{\Delta z}{z} \right) \Delta x \Delta y - \left(A_z - \frac{\partial A_z}{\partial z} \cdot \frac{\Delta z}{z} \right) \Delta x \Delta y \\
 &= \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) \Delta x \Delta y \Delta z.
 \end{aligned}$$

The bracketed expression $\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$ defines the divergence of the vector $\vec{A}$ in Cartesian coordinates and is denoted by $\vec{\nabla} \cdot \vec{A}$. Thus we can write the coordinate independent definition of $\vec{\nabla} \cdot \vec{A}$ as,

$$\vec{\nabla} \cdot \vec{A} = \lim_{\Delta V \rightarrow 0} \frac{\oint_{\Delta s} \vec{A} \cdot d\vec{s}}{\Delta V} \quad (2.6)$$

[Divergence of $\vec{A}$ is also denoted by $\text{div } \vec{A}$]. The sign $\oint$ implies that the integration is taken over a closed surface. Thus contrary to the gradient operator which acting on a scalar transforms it into a vector, the divergence operator operating on a vector changes it into a scalar. That $\text{div } \vec{A}$ is a scalar can be proved from its expression in the Cartesian coordinates by utilising the transformation property of the components of a vector under the change of coordinates.

(This is left as an exercise for the students).

Similarly by integrating ϕ over the surface enclosing the infinitesimal volume we obtain,

$$\begin{aligned}
 \int_{\Delta s} \phi d\vec{s} &= \left(\phi + \frac{\partial \phi}{\partial x} \frac{\Delta x}{z} \right) \Delta y \Delta z \vec{i} \\
 &- \left(\phi - \frac{\partial \phi}{\partial x} \frac{\Delta x}{z} \right) \Delta y \Delta z \vec{i} + \left(\phi + \frac{\partial \phi}{\partial y} \frac{\Delta y}{z} \right) \Delta z \Delta x \vec{j} \\
 &- \left(\phi - \frac{\partial \phi}{\partial y} \frac{\Delta y}{z} \right) \Delta z \Delta x \vec{j} + \left(\phi + \frac{\partial \phi}{\partial z} \frac{\Delta z}{z} \right) \Delta x \Delta y \vec{k} \\
 &- \left(\phi - \frac{\partial \phi}{\partial z} \frac{\Delta z}{z} \right) \Delta x \Delta y \vec{k}, \text{ where } \phi \text{ is the value of } \phi(x, y, z) \text{ at the centre of the volume.}
 \end{aligned}$$

Or,

$$\int_{\Delta s} \phi d\vec{s} = \left(\vec{i} \frac{\partial \phi}{\partial x} + \vec{j} \frac{\partial \phi}{\partial y} + \vec{k} \frac{\partial \phi}{\partial z} \right) \Delta x \Delta y \Delta z$$

Or, $\vec{\nabla} \phi$ can be defined in a coordinate independent way as, $\vec{\nabla} \phi = \lim_{\Delta V \rightarrow 0} \frac{\int_{\Delta s} \phi d\vec{s}}{\Delta V}$ (2.7)

(Recall another coordinate independent alternative definition of $\bar{\nabla}\phi$ in Eq. (2.4))

Suppose we consider a closed surface enclosing a finite volume.

The entire closed surface S is supposed to be divided into two parts by a cut. One part consists of the bounding surface S_a and the cut surface S_{ab} . The other part comprises the surface S_b and the cut surface S_{ab} .

$$\text{Thus, } \int_1 \bar{A} \cdot d\bar{s} + \int_2 \bar{A} \cdot d\bar{s} = \int_{S_a} \bar{A} \cdot d\bar{s} + \int_{S_{ab}} \bar{A} \cdot d\bar{s} + \int_{S_b} \bar{A} \cdot d\bar{s} + \int_{S_{ab}} \bar{A} \cdot d\bar{s}$$

The cut surface S_{ab} belongs to both the surfaces 1 and 2 with the difference that the positive directions of the normal directed outwards from the enclosed volume point in the opposite directions in the two cases. (shown by $\bar{n}_1$ and $\bar{n}_2$ in Fig. 2.14, where $\bar{n}_2 = -\bar{n}_1$). Thus the integrals along S_{ab} cancel to lead to the result

$$\text{that } \int_1 \bar{A} \cdot d\bar{s} + \int_2 \bar{A} \cdot d\bar{s} = \int_{S_a} \bar{A} \cdot d\bar{s} + \int_{S_b} \bar{A} \cdot d\bar{s}.$$

We can go on subdividing the surfaces obtained by cutting the original surface further and further till in the limit of an infinitely large number of such surfaces, each one becomes infinitely small and can be said to enclose an infinitesimally small rectangular volume. Since the contributions of the cut surfaces common to two adjacent surfaces cancel out because of the opposite directions of the outwardly drawn normals, we obtain,

$$\int_S \bar{A} \cdot d\bar{s} = \text{the sum of the surface integrals of } \bar{A} \text{ over the original closed surfaces}$$

$$= \lim_{i \rightarrow \infty} \sum_i \int \bar{A}_i \cdot d\bar{s}_i$$

Now, in the infinitesimally small limit we can use the result in the Eq. (2.6) to obtain,

$$\int_S \bar{A} \cdot d\bar{s} = \lim_{\substack{i \rightarrow \infty \\ \Delta V_i \rightarrow 0}} \sum_i (\bar{\nabla} \cdot \bar{A})_i \Delta V_i = \int_V (\bar{\nabla} \cdot \bar{A}) dV \text{ according to, Riemann's definition of the volume integral.}$$

Thus we can state that if $\bar{A}$ be a vector field with the components having partial derivatives that are finite and continuous in a region,

$$\int_i \bar{A} \cdot d\bar{s} = \int_V (\bar{\nabla} \cdot \bar{A}) dV, \text{ provided the two integrals exist.} \quad (2.8)$$

In the limit of an infinitely large number of divisions there will be some remaining parts grazing the edge which may not enclose rectangular volumes. But the contributions of these residual parts are negligibly small compared to the original integral and hence the result (2.8) follows in general. The relation (2.8) is called Gauss divergence theorem.

On applying (2.8) the equation of the flow of a fluid becomes

$$\frac{\partial}{\partial t} \int_V \rho dV = - \int_V \bar{\nabla} \cdot (\rho \bar{v}) dV.$$

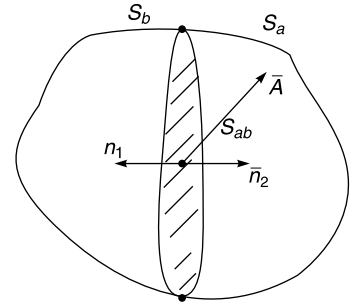


Fig. 2.14

Since this is valid for an arbitrary volume we can write,

$$\frac{\partial \rho}{\partial t} + \bar{\nabla} \cdot (\rho \bar{V}) = 0, \text{ the equation of continuity.}$$

For an incompressible fluid $\rho = \text{constant}$ and hence $\bar{\nabla} \cdot \bar{V} = 0$. The velocity vector is thus a divergenceless vector and called solenoidal. Another popular example is the vector $\bar{B}$ (magnetic induction) for which $\bar{\nabla} \cdot \bar{B} = 0$.

As another application of the divergence theorem we recall the Gauss theorem of electrostatics that states that $\int_S \bar{E} \cdot d\bar{s} = \frac{Q}{\epsilon_0}$, where Q is the charge enclosed in the volume bounded by closed surface S . Applying the divergence theorem we obtain,

$$\int_{\tau} \bar{\nabla} \cdot \bar{E} d\tau = \int_{\tau} \frac{\rho}{\epsilon_0} d\tau, \text{ yielding } \bar{\nabla} \cdot \bar{E} = \frac{\rho}{\epsilon_0}, \rho \text{ being the charge density at the point where } \bar{E} \text{ is determined.}$$

2.7 CURL OF A VECTOR FIELD; STOKES THEOREM

Let us consider the line integral of a vector field around an infinitesimally small rectangle in the xy -plane in the anticlockwise direction (Fig. 2.15).

$$\text{The line integral } \oint \bar{A} \cdot d\bar{l} = \int_A^B \bar{A} \cdot d\bar{l} + \int_B^C \bar{A} \cdot d\bar{l} + \int_C^D \bar{A} \cdot d\bar{l} + \int_D^A \bar{A} \cdot d\bar{l}$$

Let the vector $\bar{A}$ possess components A_x, A_y, A_z at the centre of the rectangular loop. The rectangle being infinitesimally small we can take the magnitudes of the components on the sides as approximately equal to those at the centres of the sides in carrying out the integrations above.

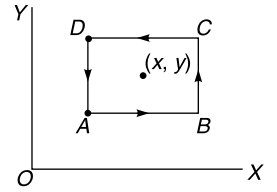


Fig. 2.15

$$\begin{aligned} \text{Thus } \oint \bar{A} \cdot d\bar{l} = & \left(A_x - \frac{\partial A_x}{\partial y} \frac{\Delta y}{z} \right) \Delta x + \left(A_y + \frac{\partial A_y}{\partial x} \frac{\Delta x}{z} \right) \Delta y - \left(A_x + \frac{\partial A_x}{\partial y} \frac{\Delta y}{z} \right) \Delta x \\ & - \left(A_y - \frac{\partial A_y}{\partial x} \frac{\Delta x}{z} \right) \Delta y \end{aligned}$$

(The negative signs in the third and the fourth terms arise from the sides being along the negative directions of the x and the y -axes respectively). Thus $\oint \bar{A} \cdot d\bar{l} = \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \Delta x \Delta y$ (2.9)

Now, we define a vector $\text{curl } \bar{A}$ whose components along the axes are $\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}$, $\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}$ and $\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}$ respectively in Cartesian coordinates. Thus, we can write Eq. (2.9) as, $\oint \bar{A} \cdot d\bar{l} = (\text{Curl } \bar{A})_z \Delta x \Delta y$ in the limit of an infinitesimally small area. Clearly this relation is valid for an arbitrarily oriented rectangle

because it can be rotated to be aligned in the xy -plane and the left hand side being a scalar remains unchanged.

Hence in general, we can write, $(\text{Curl } \bar{A})_n = \lim_{\Delta S \rightarrow 0} \frac{\oint \bar{A} \cdot d\bar{l}}{\Delta S}$, where $(\text{Curl } \bar{A})_n$ denotes the component of curl $\bar{A}$ normal to the area, the positive direction of the normal to the area being the direction of the movement of the head of a right handed screw that is moved in an anticlockwise direction along the boundary curve of the area. (2.10)

This gives clearly a coordinate-independent definition of curl $\bar{A}$. [$\text{Curl } \bar{A}$ written also as $\bar{\nabla} \times \bar{A}$ is

expressed in Cartesian coordinates as
$$\begin{vmatrix} \bar{i} & \bar{j} & \bar{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}.$$

Any arbitrary open surface may be supposed to be divided into small rectangular areas by lines drawn parallel to the coordinate axes (Fig. 2.16(a)).

If we consider two adjacent rectangles $ABCD$ and $BEFC$, for example, and calculate $\oint_{BADC} \bar{A} \cdot d\bar{l} + \oint_{CFEB} \bar{A} \cdot d\bar{l}$, the integrals along the common side are

carried out along CB in the first but along BC in the second, and hence cancel out in the sum (Fig. 2.16(b)). Thus, the sum goes over into the integral along $EADF$. If we calculate the sum of the infinitely large number of integrals $\oint \bar{A} \cdot d\bar{l}$ (all the loops are traversed in the anticlockwise direction) along the boundaries, the integrals along the side common to two adjacent loops cancel out leaving us the integral along the outer boundary C enclosing the area.

Thus, we can write, $\oint_c \bar{A} \cdot d\bar{l} = \lim_{\Delta s_i \rightarrow 0} \sum_i \oint \bar{A}_i \cdot \Delta \bar{l}_i$

Again, by virtue of Eq. (2.10), $\lim_{\Delta s_i \rightarrow 0} \oint \bar{A}_i \cdot \Delta \bar{l}_i = (\text{Curl } \bar{A})_{n_i} \Delta s_i |_{\Delta s_i \rightarrow 0}$

Thus we can write following the definition of the surface integral,

$$\oint_c \bar{A} \cdot d\bar{l} = \lim_{\Delta s_i \rightarrow 0} \sum_i (\text{Curl } \bar{A})_{n_i} \Delta s_i = \int_s (\text{Curl } \bar{A}) \cdot d\bar{s} \quad (2.11)$$

where S is an open area bounded by the contour.

The necessary conditions for the validity of the above relation Eq. (2.11) known as Stokes' theorem, are:

- The function and its partial derivatives should exist everywhere in the region bounded by the contour.
- The open surface should be two-sided (or orientable) so that a unique normal exists at every point.
- The boundary curve should be continuous or piecewise continuous. Also, it should be noted that all

open surfaces having a common boundary loop will give the same value of the integral $\int_s \text{Curl } \bar{A} \cdot d\bar{s}$ as can easily be seen from Eq. (2.11).

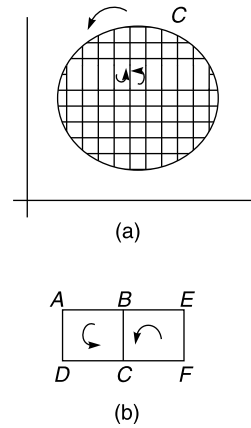


Fig. 2.16

For a closed surface, the boundary loop vanishes and, therefore, $\oint_c \vec{A} \cdot d\vec{l} = 0$ thus yielding $\int_s \text{Curl } \vec{A} \cdot d\vec{s} = 0$, where S is now a closed surface. Again, applying Gauss divergence theorem $\int_s \text{Curl } \vec{A} \cdot d\vec{s} = \int_V \text{div} (\text{Curl } \vec{A}) dV = 0$, V being the volume enclosed by the closed surface. Since this is valid for any arbitrary volume V , $\text{div} (\text{Curl } \vec{A}) = 0$ everywhere within the volume.

(Students are instructed to prove this vector identity by other means).

Similarly $\int_s \text{Curl} (\text{grad } \vec{\phi}) \cdot d\vec{s} = \oint_c \text{grad } \vec{\phi} \cdot d\vec{l} = \oint_c d\phi = 0$ (S being a closed surface the boundary loop disappears).

Since this relation is valid for an arbitrary surface $\text{Curl} (\text{grad } \vec{\phi}) = 0$.

Thus $\nabla \vec{\phi}$ is called an irrotational vector. A vector which can be represented as the Curl of another vector is divergenceless and called solenoidal.

Hence the total flux of such a vector over any closed surface is zero. The lines of these vectors form closed loops. An example is the lines of the magnetic induction vector $\vec{B}$ which is divergenceless and can be expressed as $\text{Curl } \vec{A}$, $\vec{A}$ being called the vector potential.

Note: The region bounded by the contour in Stokes theorem must be simply connected, i.e. must not contain any holes. For example, the annular region between two concentric circles is not simply connected. It is said to be multiply connected. To emphasise this point let us take $\vec{F} = -\vec{i} \frac{y}{x^2 + y^2} + \vec{j} \frac{x}{x^2 + y^2}$ as in Section (2.5).

Clearly, $\frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) = \frac{\partial}{\partial y} \left(-\frac{y}{x^2 + y^2} \right)$ yielding $\vec{\nabla} \times \vec{F} = 0$.

Thus $\int_s (\vec{\nabla} \times \vec{F}) \cdot d\vec{s} = 0$. If we write $\vec{F} = \nabla \phi$, $\phi = \tan^{-1} \frac{y}{x}$ and is multivalued.

Since $\vec{F}$ is not defined at $(0, 0)$ we delete a small region surrounding $(0, 0)$. The region is now multiply connected.

$$\oint \vec{F} \cdot d\vec{l} = \oint \left[-\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy \right].$$

Putting $x = \cos \theta$, $y = \sin \theta$ we get $\oint \vec{F} \cdot d\vec{l} = \int_0^{2\pi} d\theta = 2\pi$ leading to a departure from Stokes law. (Refer to Appendix for further clarification)

2.8 COROLLARIES FROM GAUSS THEOREM AND STOKES THEOREM

2.8.1(a) Green's Theorem in a Plane

Applying Stokes theorem to a simply connected region in the xy -plane we can write from

$$\oint_s \vec{A} \cdot d\vec{l} = \int_s (\vec{\nabla} \times \vec{A}) \cdot d\vec{s} \quad (\text{Fig. 2.17(a)})$$

or,

$$\oint_c (A_x dx + A_y dy) = \iint_R \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) dx dy$$

Writing $A_x = P$, $A_y = Q$, where P and Q are well-behaved functions of x, y we get,

$$\oint_c (P dx + Q dy) = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy \quad (2.12)$$

Since Stoke's theorem involves scalars on both the sides, the relation (2.12), called Green's theorem in a plane is valid for any region in a plane that can be made to lie in the xy - plane by properly orienting the coordinate axes.

If the region consists of two parts connected by the bridge AB the sum of the integrals along the boundaries of two pieces becomes equal to the integral along the original boundary, the integrals along the by joining line AB cancelling each other. Thus the above equality (Eq. 2.12) is valid here also (Fig. 2.17(b)). Also if the region be multiply connected consisting of the original boundary c_0 and a few holes c_1, c_2 (say), it can be converted to a simply connected region by connecting bridges AB, CD

so that the total integral equals $\int_{c_0} (P dx + Q dy) - \int_{c_1} (P dx + Q dy) - \int_{c_2} (P dx + Q dy)$,

the contributions along AB and CD being zero due to two traversals along them in the opposite directions. (Fig. 2.17(c))

Putting $Q = x$, $P = -y$ we can obtain, $\frac{1}{2} \oint_c (x dy - y dx) = \iint_R dx dy = \text{Area } A$

Similarly, writing $x = r \cos \theta$, $y = r \sin \theta$, the area can be written also as,

$$A = \frac{1}{2} \oint r^2 d\theta.$$

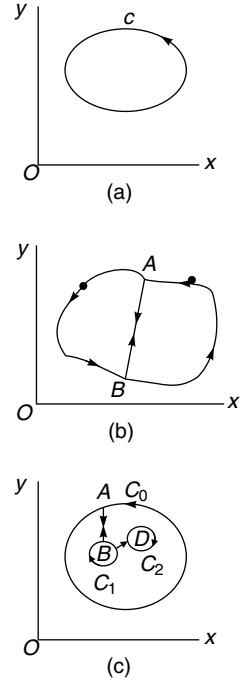


Fig. 2.17

(b) Let us consider a vector $\vec{A} = \phi \nabla \psi - \psi \nabla \phi$. Applying Gauss divergence theorem we obtain,

$$\int_s (\phi \nabla \psi - \psi \nabla \phi) \cdot d\vec{s} = \int_V \nabla \cdot (\phi \nabla \psi - \psi \nabla \phi) dV$$

where S is any closed surface enclosing a volume V .

We thus get, $\int_s (\phi \nabla \psi - \psi \nabla \phi) \cdot d\vec{s} = \int_V (\phi \nabla^2 \psi - \psi \nabla^2 \phi) dV \quad (2.13)$

This is called Green's theorem.

(c) Let ϕ_1 and ϕ_2 be two solutions of the equation $\nabla^2 \phi = f(\vec{r})$

Putting $\psi = \phi_1 - \phi_2$ we get, $\nabla^2 \psi = 0$.

Using Gauss divergence theorem we get, $\int_S \psi \bar{\nabla} \psi \cdot d\bar{s} = \int_V \bar{\nabla} \cdot (\psi \bar{\nabla} \psi) dV$

$$\text{Or,} \quad \int_S \psi \bar{\nabla} \psi \cdot d\bar{s} = \int_V [\psi \nabla^2 \psi + [\bar{\nabla} \psi]^2] dV$$

$$\text{Or,} \quad \int_S \psi \bar{\nabla} \psi \cdot d\bar{s} = \int_V [\nabla \psi]^2 dV \quad (2.14)$$

Now, the left hand side can be made to vanish by (a) taking the surface to an infinite distance so that ψ and $\bar{\nabla} \psi$ behave with distance in such a way that the integral vanishes there (for example for $\psi \propto \frac{1}{r^m}$, $\psi \nabla \psi \propto \frac{1}{r^{2m+1}}$ and since the surface area $\propto r^2$ the integral will tend to zero if $2m + 1 > 2$ or $m > \frac{1}{2}$) or (b) by imposing the boundary conditions on a closed surface at a finite distance so that ψ or $\nabla \psi$ becomes zero on the surface. We then obtain from Eq. (2.14) that $\nabla \psi = 0$. Hence ϕ_1 and ϕ_2 can differ atmost by a constant. Thus if $\psi = 0$ at any point in the volume of our choice $\psi = 0$ everywhere and hence we obtain a unique solution of the equation $\nabla^2 \phi = f(\bar{r})$. This uniqueness theorem is applied in problems of electrostatics, where ϕ is the electrostatic potential at a point and $f(\bar{r})$ is the charge density there.

2.9 LEVI-CIVITA SYMBOL AND VECTOR IDENTITIES

We have already discussed in the section 2.2 that the vector $\bar{C} = \bar{A} \times \bar{B}$ behaves under transformation of axes in a way different from that of $\bar{A}$ or $\bar{B}$. To enlighten this further we introduce the Levi-Civita symbol ϵ_{ijk} which has the value 1 for i, j, k a cyclic permutation of 1, 2, 3, and -1 otherwise, and which is zero when at least two indices are the same.

In terms of this symbol we can write $C_i = \epsilon_{ijk} A_j B_k$ (the sum over j and k is implied). The presence of this factor ϵ_{ijk} accounts for the transformation property of C mentioned. in Section 2.2 in a way explained in chapter 4. The symbol ϵ_{ijk} satisfies the relation $\epsilon_{ijk} \epsilon_{ilm} = \delta_{jl} \delta_{km} - \delta_{jm} \delta_{kl}$ ($\delta_{ij} = 1$ if $i = j$ and $= 0$ if $i \neq j$ (Kronecker delta)) as can easily be verified. This particular identity helps us establish the vector identities comparatively easily.

Thus

$$(1) \bar{\nabla} \times (\bar{A} \times \bar{B})|_i = \epsilon_{ijk} \frac{\partial}{\partial x_j} (\epsilon_{klm} A_l B_m) \text{ [Sum over repeated indices is implied]}$$

$$= \epsilon_{kij} \epsilon_{klm} \left(A_l \frac{\partial B_m}{\partial x_j} + B_m \frac{\partial A_l}{\partial x_j} \right)$$

$$= (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \left(A_l \frac{\partial B_m}{\partial x_j} + B_m \frac{\partial A_l}{\partial x_j} \right)$$

$$= A_i (\bar{\nabla} \cdot \bar{B}) - B_i (\bar{\nabla} \cdot \bar{A}) + (\bar{B} \cdot \bar{\nabla}) A_i - (\bar{A} \cdot \bar{\nabla}) B_i$$

$$\text{Or, } \bar{\nabla} \times (\bar{A} \times \bar{B}) = \bar{A} (\bar{\nabla} \cdot \bar{B}) - \bar{B} (\bar{\nabla} \cdot \bar{A}) + (\bar{B} \cdot \bar{\nabla}) \bar{A} - (\bar{A} \cdot \bar{\nabla}) \bar{B}$$

$$\begin{aligned}
(2) \quad \bar{\nabla} \times (\bar{\nabla} \times \bar{A})|_i &= \epsilon_{ijk} \frac{\partial}{\partial x_j} \left(\epsilon_{klm} \frac{\partial A_m}{\partial x_l} \right) \\
&= \epsilon_{kij} \epsilon_{klm} \frac{\partial}{\partial x_j} \left(\frac{\partial A_m}{\partial x_l} \right) \\
&= (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \frac{\partial}{\partial x_j} \left(\frac{\partial A_m}{\partial x_l} \right) \\
&= \frac{\partial}{\partial x_i} \left(\frac{\partial A_j}{\partial x_j} \right) - \frac{\partial}{\partial x_j} \left(\frac{\partial A_i}{\partial x_j} \right) \\
&= [\nabla(\bar{\nabla} \cdot \bar{A}) - \nabla^2 \bar{A}]_i
\end{aligned}$$

$$\text{or, } \bar{\nabla} \times (\bar{\nabla} \times \bar{A}) = \bar{\nabla}(\bar{\nabla} \cdot \bar{A}) - \nabla^2 \bar{A}$$

$$(3) \quad \bar{\nabla} \cdot (\bar{A} \times \bar{B}) = \frac{\partial}{\partial x_i} (\epsilon_{ijk} A_j B_k)$$

$$\text{or, } \bar{\nabla} \cdot (\bar{A} \times \bar{B}) = \epsilon_{ijk} \left(A_j \frac{\partial B_k}{\partial x_i} + B_k \frac{\partial A_j}{\partial x_i} \right)$$

$$= \bar{B} \cdot (\bar{\nabla} \times \bar{A}) - \bar{A} \cdot (\bar{\nabla} \times \bar{B}) \quad [\text{Using } \epsilon_{ijk} = \epsilon_{kij} = -\epsilon_{jik}]. \quad (2.15)$$

In this way several other vector identities can be proved using this symbol.

SOLVED PROBLEM

Problem 1

Find the unit vector along the line of intersection of the plane,

$$a_1x + b_1y + c_1z = d_1 \text{ and } a_2x + b_2y + c_2z = d_2$$

Solution: The unit vector normal to the respective planes are $\bar{n}_1 = \frac{ia_1 + j\bar{b}_1 + k\bar{c}_1}{\sqrt{a_1^2 + b_1^2 + c_1^2}}$ and $\bar{n}_2 = \frac{ia_2 + j\bar{b}_2 + k\bar{c}_2}{\sqrt{a_2^2 + b_2^2 + c_2^2}}$ respectively. [Write the equation of the plane in the form $\phi(x, y, z) = \text{constant}$. The unit normal to the plane = $\frac{\bar{\nabla}\phi}{|\bar{\nabla}\phi|}$].

The vector along the line of intersection of the planes must be perpendicular to both these vector. The unit vector along the line of intersection of the planes is therefore $\bar{n}_1 \times \bar{n}_2$, where $\bar{n}_1$ and $\bar{n}_2$ are given above.

Problem 2

Find out the angles between the straight lines joining the origin with the points

(x_1, y_1, z_1) and (x_2, y_2, z_2) respectively and also the unit vector perpendicular to both.

Solution: The angle between the vectors $\vec{A}$ and $\vec{B} = \sin^{-1} \left| \frac{|\vec{A} \times \vec{B}|}{(\vec{A}) \times (\vec{B})} \right|$. Thus the angle between the two straight lines above is given by

$$\begin{aligned} &= \sin^{-1} \left[\frac{|(\vec{i}x_1 + \vec{j}y_1 + \vec{k}z_1) \times (\vec{i}x_2 + \vec{j}y_2 + \vec{k}z_2)|}{|(\vec{i}x_1 + \vec{j}y_1 + \vec{k}z_1)| \times |(\vec{i}x_2 + \vec{j}y_2 + \vec{k}z_2)|} \right] \\ &= \sin^{-1} \left[\frac{\{(y_1z_2 - z_1y_2)^2 + (x_2z_1 - x_1z_2)^2 + (x_1y_2 - x_2y_1)^2\}^{1/2}}{\{\sqrt{x_1^2 + y_1^2 + z_1^2} \sqrt{x_2^2 + y_2^2 + z_2^2}\}} \right] \end{aligned}$$

If the unit vector perpendicular to both the lines has direction gives l, m, n then $lx_1 + my_1 + nz_1 = lx_2 + my_2 + nz_2 = 0$

or,

$$m(x_2y_1 - x_1y_2) + n(x_2z_1 - x_1z_2) = 0$$

or,

$$n = \frac{(x_1y_2 - x_2y_1)}{(x_2z_1 - x_1z_2)} m$$

Again from $lx_1 + my_1 + nz_1 = 0$ we obtain

$$lx_1 + my_1 + m \frac{(x_1y_2 - x_2y_1)z_1}{(x_2z_1 - x_1z_2)} = 0$$

or,

$$lx_1 + m \frac{(x_1y_2z_1 - x_1y_1z_1)}{(x_2z_1 - x_1z_2)} = 0$$

or,

$$l = m \frac{(y_1z_2 - y_2z_1)}{(x_2z_1 - x_1z_2)}$$

Again from $l^2 + m^2 + n^2 = 1$ we can find out m (actually two values of m will be obtained corresponding to two directions of the perpendicular) and hence l and n from the previous equation.

Thus the unit normal vector is $\vec{i}l + \vec{j}m + \vec{k}n$.

Problem 3

(a) Suppose that for two fixed vectors $\vec{A}$ and $\vec{B}$ we have a relation $\lambda \vec{A} + \vec{B} \times \vec{C} =$

0. Is $\vec{C}$ uniquely determined?

(b) Suppose we have a relation $\frac{d\vec{A}}{dt} = \lambda \vec{A} \times \vec{B}$, $\vec{B}$ being a constant vector. Determine the characteristics of the vector $\vec{A}$ that remain unchanged with time.

(c) Show that for a given vector $\vec{A}$ in a space of n -dimensions with given basis vectors (not necessarily perpendicular) the components are uniquely fixed.

Solution:

- (a) $\bar{C}$ is not uniquely determined because we can add $v\bar{B}$ to $\bar{C}$ keeping the relation unaltered where v is any arbitrary constant. Thus there can be an infinite number of choices of $\bar{C}$.
- (b) We take the dot product of the relation with $\bar{A}$. We obtain, $\frac{d}{dt}(A^2) = 0$ leading to the result that the norm of $\bar{A}$ is fixed in time. Similarly, taking the dot product with $\bar{B}$ we get, $\frac{d}{dt}(\bar{B} \cdot \bar{A}) = 0$. Since $\bar{B}$ is constant, the angle between $\bar{A}$ and $\bar{B}$ remains constant. The common example is the equation of motion of a charged particle in a uniform magnetic field given by $\frac{d\bar{v}}{dt} = \frac{e}{m} \bar{v} \times \bar{B}$. (Lorentz-force, the symbols having their usual meanings). Taking the scalar product with $\bar{v}$ and $\bar{B}$ respectively we obtain the result that the magnitude of $\bar{v}$ and the angle $\bar{v}$ makes with $\bar{B}$ remain unchanged.
- (c) Let $\bar{A}$ have different components to start with. Thus we write,

$$\bar{A} = \bar{e}_1 A_1 + \bar{e}_2 A_2 + \dots + \bar{e}_n A_n = \bar{e}_1 A'_1 + \bar{e}_2 A'_2 + \dots + \bar{e}_n A'_n$$

($\bar{e}_1 \dots \bar{e}_n$ being the basis vector). Thus we get,

$$\bar{e}_1(\bar{A}_1 - A'_1) + \bar{e}_2(\bar{A}_2 - A'_2) + \dots + \bar{e}_n(\bar{A}_n - A'_n) = 0$$

Since $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n$ are linearly independent, $A_1 - A'_1 = A_2 - A'_2 = \dots = A_n - A'_n = 0$
Hence the components are uniquely determined.

Problem 4

- (a) Find the angle between the surfaces $\phi_1(x, y, z) = c_1$ and $\phi_2(x, y, z) = c_2$, where c_1, c_2 are constants at the point (x_0, y_0, z_0) .
- (b) Find a unit vector normal to the surface $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ at the point $(1, 2, 3)$ and the equation of the tangent plane at the point.
- (c) Find the directional derivative of $\phi = xy^2 + yz^2 + zx^2$ at $(1, 3, 4)$ along $3\bar{i} + 4\bar{j} + 7\bar{k}$.

Solution: (a) The angle between the surfaces = angle between the normals to the surfaces at the point (x_0, y_0, z_0) ,

The unit normals at (x_0, y_0, z_0) , $\bar{n}_1$ and $\bar{n}_2$ to the surfaces $\phi_1 = c_1$ and $\phi_2 = c_2$ are given by

$$\frac{\bar{\nabla}\phi_1}{|\bar{\nabla}\phi_1|} \text{ and } \frac{\bar{\nabla}\phi_2}{|\bar{\nabla}\phi_2|} \text{ respectively at } (x_0, y_0, z_0).$$

$$\text{Thus the angle } \theta \text{ is given by } \cos^{-1} \left[\frac{\bar{\nabla}\phi_1 \cdot \bar{\nabla}\phi_2}{|\bar{\nabla}\phi_1| |\bar{\nabla}\phi_2|} \right].$$

- (b) The unit vector normal to the surface $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ at the point (1, 2, 3) is given by $\frac{\left(\frac{\bar{i}x}{a^2} + \frac{\bar{j}y}{b^2} + \frac{\bar{k}z}{c^2}\right)}{\sqrt{\frac{x^2}{a^4} + \frac{y^2}{b^4} + \frac{z^2}{c^4}}}$ at (1, 2, 3), and is represented, say, by $\bar{n}$.

The equation of the tangent plane at (1, 2, 3) is given by

$$(\bar{r} - \bar{r}_o)\bar{n} = 0, \text{ where } \bar{n} = \left(\frac{\bar{i}}{a^2} + \frac{\bar{j}}{b^2} + \frac{\bar{k}}{c^2}\right) \times \frac{1}{\sqrt{\frac{1}{a^4} + \frac{4}{b^4} + \frac{9}{c^4}}}$$

or, $(x-1)n_x + (y-2)n_y + (z-3)n_z = 0$, where n_x , n_y and n_z are the x , y and z -components of $\bar{n}$.

- (c) Directional derivative of $\phi = xy^2 + yz^2 + zx^2$ at (1, 3, 4) along

$$3\bar{i} + 4\bar{j} + 7\bar{k} = \frac{\bar{i}(y^2 + 2xz) + \bar{j}(2xy + z^2) + \bar{k}(2yz + x^2)}{[(y^2 + 2xz)^2 + (2xy + z^2)^2 + (2yz + x^2)^2]^{1/2}} \left(\frac{3\bar{i} + 4\bar{j} + 7\bar{k}}{\sqrt{3^2 + 4^2 + 7^2}} \right)$$

[$\therefore$ Directional derivative = $\bar{\nabla}\phi \cdot \bar{n}$]

at $x = 1, y = 2$ and $z = 3$.

Problem 5

- (a) Calculate $\oint \bar{A} \cdot d\bar{t}$ around a square of side $2a$ in the anticlockwise direction, and centred at (x_0, y_0) where

$$\bar{A}(x, y) = \frac{1}{(x^2 + y^2)}(-y\bar{i} + x\bar{j}) \text{ Fig. 2.18(a)}$$

- (b) Find out $\oint (x^2 + y^2) dx + x^3 dy$, where c is a right angled triangle with vertices at (0, 0), (1, 0) and (1, 2) respectively. Hence prove Green's theorem in a plane in the two cases Fig. 2.18(b).

Solution: 5(a)

$$\begin{aligned} \oint \bar{A} \cdot d\bar{t} &= \int_{x_0-a}^{x_0+a} -\frac{y}{x^2 + y^2} dx + \int_{y_0-a}^{y_0+a} -\frac{x}{x^2 + y^2} dy - \int_{x_0+a}^{x_0-a} \frac{y}{x^2 + y^2} dx + \int_{y_0+a}^{y_0-a} \frac{x}{x^2 + y^2} dy \\ &= -\left[\tan^{-1} \frac{x}{y_0-a} \right]_{x_0-a}^{x_0+a} + \left[\tan^{-1} \frac{y}{x_0+a} \right]_{y_0-a}^{y_0+a} - \left[\tan^{-1} \frac{x}{y_0+a} \right]_{x_0+a}^{x_0-a} + \left[\tan^{-1} \frac{y}{x_0-a} \right]_{y_0+a}^{y_0-a} \\ &= \tan^{-1} \left[\frac{x_0-a}{y_0-a} \right] - \tan^{-1} \left[\frac{x_0+a}{y_0-a} \right] + \tan^{-1} \left[\frac{y_0+a}{x_0+a} \right] - \tan^{-1} \left[\frac{y_0-a}{x_0+a} \right] \end{aligned}$$

$$- \tan^{-1} \left[\frac{x_0 - a}{y_0 + a} \right] + \tan^{-1} \left[\frac{x_0 + a}{y_0 + a} \right] + \tan^{-1} \left[\frac{y_0 - a}{x_0 - a} \right] - \tan^{-1} \left[\frac{y_0 + a}{x_0 - a} \right]$$

$$= 0 \text{ [Applying the formula } \tan^{-1} x + \cot^{-1} x = \pi/2]$$

$$\oint \left[\frac{\partial}{\partial y} \left(-\frac{y}{x^2 + y^2} \right) - \frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) \right] dx dy$$

$$= 0.$$

Thus Green's theorem is proved.

5(b) The equation of the line AC is, $y = 2x$

$$\oint [(x^2 + y^2) dx + x^3 dy] = \int_0^1 x^2 dx + \int_0^2 dy + \int_1^0 (5x^2 + 2x^3) dx$$

[Putting $y = 2x$ for the line AC]

$$= \frac{1}{3} + 2 - \left(\frac{5}{3} + \frac{1}{2} \right) = \frac{1}{6}$$

$$\oint \left[\frac{\partial}{\partial x} (x^3) - \frac{\partial}{\partial x} (x^2 + y^2) \right] dx dy = \int_{x=0}^1 \int_{y=0}^{2x} (3x^2 - 2y) dx dy = \int_0^1 (6x^3 - 2x^2) dx$$

$$= \frac{3}{2} - \frac{4}{3} = \frac{1}{6}.$$

Hence, Green's theorem in a plane is demonstrated.

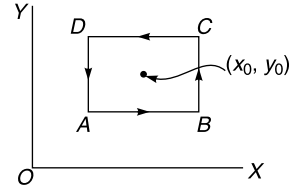


Fig. 2.18(a)

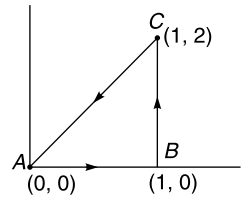


Fig. 2.18(b)

2.10 VECTOR RELATIONS IN CURVILINEAR COORDINATES

Hitherto we have used Cartesian coordinates to specify the position of a point in space and deduced the vector relations in respect of these coordinates. But apart from the Cartesian coordinates there are other coordinate systems which may prove to be more convenient in specific cases. In the following section, we will first give a general description of these coordinate systems and later discuss a few specific system.

In three dimensions the position of a point is fixed by specifying three independent coordinates denoted by q_1, q_2, q_3 say.

Thus, we can write $\vec{r} = \vec{r}(q_1, q_2, q_3)$. The infinitesimal displacement $d\vec{r}$ is given by $d\vec{r} = \sum_i \frac{\partial \vec{r}}{\partial q_i} dq_i$. The square of the infinitesimal line element ds^2 is thus $= \frac{\partial \vec{r}}{\partial q_i} \cdot \frac{\partial \vec{r}}{\partial q_j} dq_i \cdot dq_j$ [Sum over the repeated indices is implied]. Putting $g_{ij} = \frac{\partial \vec{r}}{\partial q_i} \cdot \frac{\partial \vec{r}}{\partial q_j}$, we obtain that $ds^2 = g_{ij} dq_i dq_j$, g_{ij} being the metric of the space.

The unit tangent vector $\bar{e}_i$ along the direction of q_i is $\bar{e}_i = \frac{1}{h_i} \frac{\partial \bar{r}}{\partial q_i}$, h_i being called the scale factor corresponding to q_i . Thus the infinitesimal area in the $q_i q_j$ surface is $h_i h_j (\bar{e}_i \times \bar{e}_j) dq_i dq_j$. The infinitesimal volume element is $h_i h_j h_k [\bar{e}_i (\bar{e}_j \times \bar{e}_k)] dq_i dq_j dq_k$. The system is orthogonal if the unit vectors satisfy the relation $\bar{e}_i \cdot \bar{e}_j = \delta_{ij}$.

Example 1 *Spherical polar coordinate system* – Here $q_1 = r$, $q_2 = \theta$ and $q_3 = \phi$, where r represents the distance of the point from the origin, θ the colatitude and ϕ the longitude.

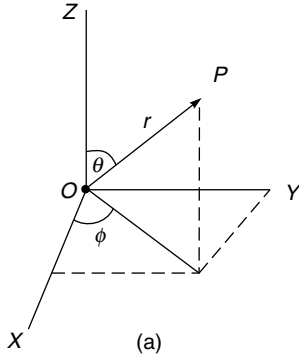


Fig. 2.19(a)

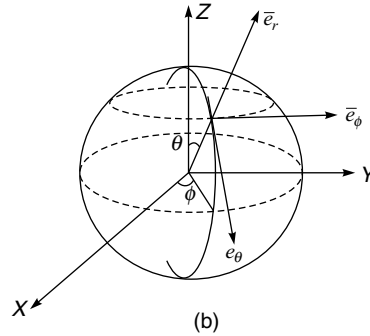


Fig. 2.19(b)

Spherical polar coordinates

The relations between the Cartesian and the spherical polar coordinates are given by $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$ and $z = r \cos \theta$ respectively (Fig. 2.19(a)). The $r = \text{constant}$ surface represents a sphere with centre 0, the $\theta = \text{constant}$ surface a cone with the semi-vertical angle θ whose intersection with the sphere is a circle and the $\phi = \text{constant}$ surface a plane containing the z -axis and the line $y/x = \text{constant}$. The unit vectors $\bar{e}_r, \bar{e}_\theta, \bar{e}_\phi$ shown in the Fig. 2.19(b) intersect orthogonally at each point. This can also be shown as follows.

Here $h_r = 1$, $h_\theta = r$ and $h_\phi = r \sin \theta$ so that

$$\bar{e}_r = \frac{\partial \bar{r}}{\partial r} = \bar{i} \frac{\partial x}{\partial r} + \bar{j} \frac{\partial y}{\partial r} + \bar{k} \frac{\partial z}{\partial r} = \bar{i} \sin \theta \cos \phi + \bar{j} \sin \theta \sin \phi + \bar{k} \cos \theta$$

$$\bar{e}_\theta = \frac{1}{r} \frac{\partial \bar{r}}{\partial \theta} = \frac{1}{r} \frac{\partial}{\partial \theta} (\bar{i}x + \bar{j}y + \bar{k}z) = \bar{i} \cos \theta \cos \phi + \bar{j} \cos \theta \sin \phi - \bar{k} \sin \theta$$

$$\bar{e}_\phi = \frac{1}{r \sin \theta} \frac{\partial \bar{r}}{\partial \phi} = \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} (\bar{i}x + \bar{j}y + \bar{k}z) = -\bar{i} \sin \phi + \bar{j} \cos \phi$$

Clearly the unit vectors are mutually orthogonal.

The coordinate surface elements are given by $r^2 \sin^2 \theta d\phi \bar{e}_r$, $r \sin \theta dr d\theta \bar{e}_\theta$ and $r d\theta dr \bar{e}_\phi$ respectively. The volume element is given by $r^2 \sin \theta dr d\theta d\phi$. Here $g_{rr} = 1$, $g_{\theta\theta} = r^2$, $g_{\phi\phi} = r^2 \sin^2 \theta$, $g_{ij} (i \neq j) = 0$.

Example 2 *Cylindrical coordinates* – Here $q_1 = \rho$, $q_2 = \theta$, $q_3 = z$. Here z represents the distance of the point along the polar axis, ρ the radius of the circle in the xy -plane through the point and θ the angle the unitvector along $\bar{\rho}$ makes with the x -axis. Thus the relations between the Cartesian coordinates and the cylindrical coordinates are $x = \rho \cos \theta$, $y = \rho \sin \theta$, $z = z$ Fig. 2.20(a).

The unit vector $\bar{e}_\rho$ points along the radius to the circle in the plane perpendicular to z -axis, $\bar{e}_\theta$ along the tangent at ρ to the circle and $\bar{e}_z$ along the z -axis (Fig. 2.20(b)).

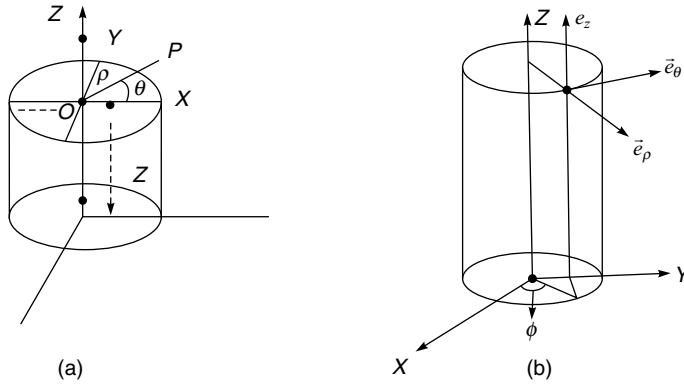


Fig. 2.20 Cylindrical coordinate system

Thus $\bar{e}_\rho, \bar{e}_\theta, \bar{e}_z$ form a right handed triad.

$$\bar{e}_\rho = \bar{i} \cos \theta + \bar{j} \sin \theta, \quad \bar{e}_\theta = -\bar{i} \sin \theta + \bar{j} \cos \theta.$$

Here also the unit vectors are seen to be mutually orthogonal making the coordinate system an orthogonal coordinate system. Here $h_\rho = 1$, $h_\theta = r$, $h_z = 1$.

The surface elements are $\rho d\theta dz \bar{e}_\rho$, $\rho d\theta d\rho \bar{e}_z$ and $\bar{e}_\theta d\rho dz$. The volume element is, $\rho d\theta d\rho dz$.

The components $g_{\rho\rho} = 1$, $g_{\theta\theta} = \rho^2$, $g_{zz} = 1$.

The square of the infinitesimal line element $ds^2 = d\rho^2 + \rho^2 d\theta^2 + dz^2$, while in spherical polar coordinates it is, $ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2$.

2.10.1 Vector Relations in Coordinates of Specific Systems

(1) Gradient – The components of the gradient are given by

$$(\nabla\Psi)_1 = \lim_{\Delta q_1 \rightarrow 0} \frac{\Psi(q_1 + \Delta q_1, q_2, q_3) - \Psi(q_1, q_2, q_3)}{h_1 \Delta q_1} = \frac{1}{h_1} \frac{\partial \Psi}{\partial q_1}$$

$$(\nabla\Psi)_2 = \lim_{\Delta q_2 \rightarrow 0} \frac{\Psi(q_1, q_2 + \Delta q_2, q_3) - \Psi(q_1, q_2, q_3)}{h_2 \Delta q_2} = \frac{1}{h_2} \frac{\partial \Psi}{\partial q_2}$$

$$(\nabla\Psi)_3 = \lim_{\Delta q_3 \rightarrow 0} \frac{\Psi(q_1, q_2, q_3 + \Delta q_3) - \Psi(q_1, q_2, q_3)}{h_3 \Delta q_3} = \frac{1}{h_3} \frac{\partial \Psi}{\partial q_3}$$

Thus for the spherical polar coordinates,

$$(\nabla\Psi)_r = \frac{\partial \Psi}{\partial r}, (\nabla\Psi)_\theta = \frac{1}{r} \frac{\partial \Psi}{\partial \theta}, (\nabla\Psi)_\phi = \frac{1}{r \sin \theta} \frac{\partial \Psi}{\partial \phi}$$

For the cylindrical coordinates,

$$(\nabla\Psi)_\rho = \frac{\partial \Psi}{\partial \rho}, (\nabla\Psi)_\theta = \frac{1}{\rho} \frac{\partial \Psi}{\partial \theta}, (\nabla\Psi)_z = \frac{\partial \Psi}{\partial z}.$$

(2) Divergence: The divergence of a vector $\bar{A}$ is expressed in the coordinate independent form as

$$\text{div } \bar{A} = \lim_{\Delta v \rightarrow 0} \frac{\int \bar{A} \cdot d\bar{s}}{\Delta v}.$$

Let the vector $\bar{A}$ be expressed in terms of the unit vector as $\bar{A} = \bar{e}_1 A_1 + \bar{e}_2 A_2 + \bar{e}_3 A_3$. The sides of the parallelepiped in the orthogonal curvilinear coordinate system are $h_1 \Delta q_1 (= AB)$, $h_2 \Delta q_2 (= DH)$ and $h_3 \Delta q_3 (= DA)$ respectively (Fig. 2.21). The flux through the face $ADHE = -A_1 h_2 h_3 \Delta q_2 \Delta q_3$, the negative sign arising from the fact that the positive direction of the normal to this face is in the negative direction of the q_1 -axis. Similarly, the flux through the opposite

$$\text{face } BCGF = \left[A_1 h_2 h_3 + \frac{\partial}{\partial q_1} (A_1 h_2 h_3) dq_1 \right] dq_2 dq_3.$$

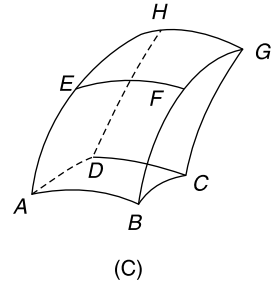


Fig. 2.21

[The sides are assumed to be infinitesimally small so that the magnitudes of the components are assumed to be the same throughout any surface area]

The sum of two contributions $= \frac{\partial}{\partial q_1} (A_1 h_2 h_3) dq_1 dq_2 dq_3$. Similarly the sums of the contributions from the other faces are $\frac{\partial}{\partial q_2} (A_2 h_3 h_1) dq_1 dq_2 dq_3$ and $\frac{\partial}{\partial q_3} (A_3 h_1 h_2) dq_1 dq_2 dq_3$ respectively.

Thus $\int_{\Delta s} \bar{A} \cdot d\bar{s}$, where Δs is the total surface area covering all the six faces

$$= \left[\frac{\partial}{\partial q_1} (A_1 h_2 h_3) + \frac{\partial}{\partial q_2} (A_2 h_3 h_1) + \frac{\partial}{\partial q_3} (A_3 h_1 h_2) \right] dq_1 dq_2 dq_3.$$

Hence divergence of the vector $\bar{A}$ evaluated at the centre of the infinitesimal volume element is given by

$$\text{div } \bar{A} = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} (A_1 h_2 h_3) + \frac{\partial}{\partial q_2} (A_2 h_3 h_1) + \frac{\partial}{\partial q_3} (A_3 h_1 h_2) \right] \quad (2.16)$$

$$\text{Thus in spherical polar system } \nabla \cdot \bar{A} = \frac{1}{r^2 \sin \theta} \left[\frac{\partial}{\partial r} (r^2 \sin \theta A_r) + \frac{\partial}{\partial \theta} (r^2 \sin \theta A_\theta) + \frac{\partial}{\partial \phi} (r A_\phi) \right] \quad (2.17(A))$$

$$\text{and in cylindrical coordinates, } \nabla \cdot \bar{A} = \frac{1}{\rho} \left[\frac{\partial}{\partial \rho} (\rho A_\rho) + \frac{\partial}{\partial \theta} (A_\theta) + \frac{\partial}{\partial z} (\rho A_z) \right] \quad (2.17(B))$$

(3) Curl – We start as in the previous case with the coordinate independent definition of $\text{Curl } \bar{A}$ as,

$(\text{Curl } \bar{A})_n = \lim_{\Delta s \rightarrow 0} \frac{\oint \bar{A} \cdot d\bar{l}}{\Delta s}$, where the contour c is described in an anticlockwise sense so that the open surface of which C is the boundary lies to the left, and $(\text{Curl } \bar{A})_n$ denotes the component of $\text{Curl } \bar{A}$ normal to the area.

Suppose we calculate the circulation of the vector A , i.e., $\oint \bar{A} \cdot d\bar{l}$ around the boundary of the curved rectangle shown in the direction $ABCD$. Let AB and AD be along the q_1 and q_2 directions respectively (Fig. 2.22).

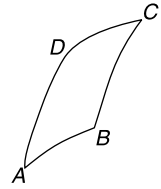


Fig. 2.22

$$\begin{aligned} \text{Thus } \oint \bar{A} \cdot d\bar{l} &= A_1 h_1 \Delta q_1 + \left[A_2 h_2 + \frac{\partial}{\partial q_1} (A_2 h_2) \Delta q_1 \right] \Delta q_2 - \left[A_1 h_1 + \frac{\partial}{\partial q_2} (A_1 h_1) \Delta q_2 \right] \Delta q_1 - [A_2 h_2 \Delta q_2] \\ &= \left[\frac{\partial}{\partial q_1} (A_2 h_2) - \frac{\partial}{\partial q_2} (A_1 h_1) \right] \Delta q_1 \Delta q_2. \end{aligned}$$

Thus $(\text{Curl } \bar{A})_3 = \frac{1}{h_1 h_2} \left[\frac{\partial}{\partial q_1} (A_2 h_2) - \frac{\partial}{\partial q_2} (A_1 h_1) \right]$, where the subscript 3 represents the component along the q_3 direction which is the positive normal to the area.

Thus writing the other components in a similar way we can express $\text{Curl } \bar{A}$ as,

$$\begin{aligned} \text{Curl } \bar{A} &= \bar{e}_1 \left[\frac{1}{h_2 h_3} \left\{ \frac{\partial}{\partial q_2} (A_3 h_3) - \frac{\partial}{\partial q_3} (A_2 h_2) \right\} \right] \\ &+ \bar{e}_2 \left[\frac{1}{h_3 h_1} \left\{ \frac{\partial}{\partial q_3} (A_1 h_1) - \frac{\partial}{\partial q_1} (A_3 h_3) \right\} \right] + \bar{e}_3 \left[\frac{1}{h_1 h_2} \left\{ \frac{\partial}{\partial q_1} (A_2 h_2) - \frac{\partial}{\partial q_2} (A_1 h_1) \right\} \right] \end{aligned} \quad (2.18)$$

Thus in spherical polar coordinates,

$$\begin{aligned} \text{Curl } \bar{A} &= \bar{e}_r \left[\frac{1}{r^2 \sin \theta} \left\{ \frac{\partial}{\partial \theta} (A_\phi \cdot r \sin \theta) - \frac{\partial}{\partial \phi} (r A_\theta) \right\} \right] \\ &+ \bar{e}_\theta \left[\frac{1}{r^2 \sin \theta} \left\{ \frac{\partial}{\partial \phi} (A_r) - \frac{\partial}{\partial r} (r \sin \theta A_\phi) \right\} \right] + \bar{e}_\phi \left[\frac{1}{r} \left\{ \frac{\partial}{\partial r} (A_\theta r) - \frac{\partial}{\partial \theta} (A_r) \right\} \right] \end{aligned} \quad (2.19(A))$$

and in cylindrical coordinates,

$$\text{Curl } \bar{A} = \frac{\bar{e}_\rho}{\rho} \left[\frac{\partial A_z}{\partial \theta} - \frac{\partial}{\partial z} (r A_\theta) \right] + \bar{e}_\theta \left[\frac{\partial}{\partial z} (A_\rho) - \frac{\partial}{\partial \rho} A_z \right] + \frac{\bar{e}_z}{\rho} \left[\frac{\partial}{\partial \rho} (\rho A_\theta) - \frac{\partial}{\partial \theta} (A_\rho) \right] \quad (2.19(B))$$

The Laplacian ∇^2 operator can be expressed as (from the expressions of the divergence and the gradient operators, as

$$\nabla^2 = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left(\frac{h_3 h_1}{h_2} \frac{\partial}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial}{\partial q_3} \right) \right] \quad (2.20)$$

In Chapter 4 a short discussion on non-orthogonal curvilinear coordinate system will be given.

SOLVED PROBLEMS

Problem 6

Prove that a vector is uniquely determined in a region if at every point in the region

$\bar{\nabla} \times \bar{A}$, $\bar{\nabla} \cdot \bar{A}$ are given and the component of $\bar{A}$ normal to the boundary is given.

Solution: Let $\bar{A}_1$ and $\bar{A}_2$ be two vectors satisfying the boundary conditions stated above so that $\bar{\nabla} \times \bar{A}_1 = \bar{\nabla} \times \bar{A}_2$ and $\bar{\nabla} \cdot \bar{A}_1 = \bar{\nabla} \cdot \bar{A}_2$. Thus $\bar{\nabla} \times (\bar{A}_1 - \bar{A}_2) = 0$ so that $\bar{A}_1 - \bar{A}_2 = \bar{\nabla} \phi$, where ϕ is a scalar function of the coordinates. (The region is assumed to be simply connected). From $\bar{\nabla} \cdot (\bar{A}_1 - \bar{A}_2) = 0$ we obtain that $\nabla^2 \phi = 0$. Applying Gauss divergence theorem we get, $\int_S \phi \bar{\nabla} \phi \cdot d\bar{s} = \int_V \bar{\nabla} \cdot (\phi \bar{\nabla} \phi) dv$ where S is the closed boundary enclosing the volume V . Since $(\bar{A}_1 - \bar{A}_2) \cdot \bar{n} = 0$, at every point of the boundary the surface integral on the left hand side vanishes leading to $\int_V \bar{\nabla} \cdot (\phi \bar{\nabla} \phi) dv = 0$ so that $\int_V (\nabla \phi)^2 dv = 0$, and hence $\bar{\nabla} \phi = 0$ at every point in the region. Since $\bar{A}_1 - \bar{A}_2 = \bar{\nabla} \phi$ the uniqueness is established.

Note: This is in electrostatics since $\bar{\nabla} \times \bar{E} = 0$ at every point and if the charge density be specified everywhere $\bar{\nabla} \cdot \bar{E}$ is fixed up. Hence we obtain a unique solution of the electric field if the component of $\bar{E}$ normal to the boundary be given uniquely.

Problem 7

If ψ be a scalar and $\bar{F}$ a vector

(a) $\oint_{C'} \Psi dt = \int_S [\bar{n} \times \nabla \Psi] \cdot d\bar{s}$, (b) $\oint_{C'} d\bar{r} \times \bar{F} = \int_S [(\bar{n} \times \bar{\nabla}) \times \bar{F}] \cdot d\bar{s}$, the symbols having their usual meanings.

Solution: We take $\bar{c}$ to be any constant vector.

or, $\oint_{C'} (\bar{c} \cdot \phi dt) = \oint_{C'} \phi \bar{c} \cdot d\bar{r} = \int_S [\bar{\nabla} \times (\phi \bar{c})] \cdot d\bar{s}$, applying Stokes theorem to an open surface S bounded by a closed curve C'

(2.21)

Now $\bar{\nabla} \times (\Psi \bar{c}) = \Psi(\bar{\nabla} \times \bar{c}) + \bar{\nabla}\Psi \times \bar{c} = \bar{\nabla}\Psi \times \bar{c}$ [$\bar{c} = \text{constant}$, $\bar{\nabla} \times \bar{c} = 0$] (2.22)

Using Eq. (2.22) in Eq. (2.21) we get,

$$\bar{c} \cdot \oint_{c'} \Psi d\bar{r} = \int_s (\bar{\nabla}\Psi \times \bar{c}) \cdot d\bar{s} = \bar{c} \cdot \int_s (\bar{n} \times \bar{\nabla}\Psi) \cdot d\bar{s} \quad [d\bar{s} = \bar{n} ds].$$

Since this is valid for any arbitrary vector c' we obtain,

$$\oint_{c'} \Psi d\bar{r} = \int_{s'} [\bar{n} \times \bar{\nabla}\Psi] ds$$

(b) Using again a constant vector $\bar{V}$ as above we get

$$\oint_{c'} \bar{V} \cdot (d\bar{r} \times \bar{F}) = \oint_{c'} d\bar{r} \cdot (\bar{F} \times \bar{V}) = \int_{s'} [\bar{\nabla} \times (\bar{F} \times \bar{V})] \cdot d\bar{s}$$

or, $\oint_{c'} \bar{V} \cdot (d\bar{r} \times \bar{F}) = \int_s [(\bar{V} \cdot \bar{\nabla}) \bar{F} - (\bar{\nabla} \cdot \bar{F}) \bar{V}] \cdot d\bar{s}$, the other terms vanishing since $\bar{F}$ is a constant vector.

$$\text{or, } \oint_{c'} \bar{V} \cdot (d\bar{r} \times \bar{F}) = \int_s [\bar{n} \cdot (\bar{V} \cdot \bar{\nabla}) \bar{F} - (\bar{n} \cdot \bar{V}) \bar{\nabla} \cdot \bar{F}] \cdot d\bar{s} \quad (2.23)$$

$$\text{Now, } \bar{V} \cdot [(\bar{n} \times \bar{\nabla}) \times \bar{F}] = -(\bar{n} \cdot \bar{V}) (\bar{\nabla} \cdot \bar{F}) + \bar{n} \cdot [(\bar{V} \cdot \bar{\nabla}) \bar{F}]$$

$$\text{Or, from (2.23) } \oint_{c'} \bar{V} \cdot (d\bar{r} \times \bar{F}) = \int_s \bar{V} \cdot [(\bar{n} \times \bar{\nabla}) \times \bar{F}] ds \quad \dots (2.24)$$

Since this holds good for any arbitrary $\bar{V}$ we get,

$$\oint_{c'} \bar{V} \cdot (d\bar{r} \times \bar{F}) = \int_s [(\bar{n} \times \bar{\nabla}) \times \bar{F}] ds$$

Problem 8

(a) Calculate the work done by a force

$\bar{F} = \frac{i}{3}(3x^2y + y^3) + \frac{j}{3}(x^3 + 3xy^2)$ in moving a particle from (0, 0, 0) to (1, 1, 0) along the straight line joining the two points.

(b) Will the work done be different along any other path? Explain.

Solution:

$$(a) \text{ Work done by the force } W = \int_{(0,0,0)}^{(1,1,0)} \bar{F} \cdot d\bar{r}$$

$$\text{or, } W = \frac{1}{3} \int_{(0,0,0)}^{(1,1,0)} [(3x^2y + y^3)dx + (x^3 + 3xy^2)dy]$$

The equation of the straight line is $y = x$.

$$\text{or, } W = \frac{2}{3} \int_0^1 (3x^3 + x^3) dx = \frac{8}{3} \int_0^1 x^3 dx \text{ on putting the relation } y = x \text{ in the expression for } W.$$

$$\text{or, } W = \frac{2}{3}.$$

- (b) If we can write $\vec{F} = \nabla\phi$, where ϕ is a single-valued function of the coordinates, $\int_1^2 \vec{F} \cdot d\vec{r}$ can be written as $\phi_2 - \phi_1$ and hence it is independent of the path connecting 1 and 2. For this to be possible in this case, $\frac{\partial\phi}{\partial x} = \frac{1}{3}(3x^2y + y^3)$ and $\frac{\partial\phi}{\partial y} = \frac{1}{3}(x^3 + 3xy^2)$. Solving the two equations, we obtain, $\phi = (x^3y + xy^3) \frac{1}{3}$ (setting the trivial addition constant = 0).

Since we can express $\vec{F} = \nabla\phi$ in this case we get,

$$\text{or, } \int_1^2 \vec{F} \cdot d\vec{r} = \frac{1}{3} [x^3y + xy^3]_{(0,0,0)}^{(1,1,0)} = \frac{2}{3} \text{ for any path connecting the two points.}$$

Problem 9

Find $\oint_c \vec{F} \cdot d\vec{l}$, where $\vec{F} = x^2y\vec{i} + xy^2\vec{j} + (x^3 + y^3)\vec{k}$ along the curve c , where c is given by the intersection of the sphere $x^2 + y^2 + z^2 = 16$ and the plane $z = 2$. Verify Stokes' theorem in this case. (Circulation is anticlockwise) taking the surface to be the area bounded by the curve.

Solution: $\oint_c \vec{F} \cdot d\vec{l} = \int [x^2y dx + xy^2 dy]$ where c is the circle $x^2 + y^2 = 12$. Putting $x = \sqrt{12}\cos\theta$, $y = \sqrt{12}\sin\theta$, we get,

$$\oint_c \vec{F} \cdot d\vec{l} = \int_0^{2\pi} [24\sqrt{3}\cos^2\theta\sin\theta(-\sqrt{12}\sin\theta) + 24\sqrt{3}\cos\theta\sin^2\theta(\sqrt{12}\cos\theta)] d\theta = 0$$

$$\vec{\nabla} \times \vec{F} = 3\vec{i}y^2 - 3\vec{j}x^2 + \vec{k}(y^2 - x^2).$$

or, $\int_{\delta} (\vec{\nabla} \times \vec{F}) \cdot d\vec{s} = \iint (y^2 - x^2) dx dy$. Since the area of intersection is a circle,

$$\int_{\delta} (\vec{\nabla} \times \vec{F}) \cdot d\vec{s} = \int_{r=0}^{\sqrt{12}} \int_{\theta=0}^{2\pi} r^3 (\sin^2\theta - \cos^2\theta) d\theta dr = 0$$

Thus Stokes' law is verified. Clearly $\int_{\delta} (\vec{\nabla} \times \vec{F}) \cdot d\vec{s}$ will be zero if the area of integration be any area bounded by the same curve.

Problem 10

Calculate $\iint_S \vec{F} \cdot \vec{n} \, ds$, where $\vec{F} = 4x^2\vec{i} + 2y\vec{j} + z^3\vec{k}$ and S is the surface bounded by the region $x^2 + y^2 = 9$, $z = 0$, $z = 2$.

Solution: $\iint_S \vec{F} \cdot \vec{n} \, ds = \int_S z^3 \, dx \, dy$, where the surface S is the plane surface, the equations being given by $x^2 + y^2 = 9$, $z = 0$ and $x^2 + y^2 = 9$, $z = 2$ respectively.

Thus the total surface integral over the plane surfaces

$$= 0 + 8 \iint dx \, dy = 8 \times \pi \times 9 = 72\pi$$

The surface integral over the curved surface $S_1 = \int_{S_1} \vec{F} \cdot \vec{n} \, ds$

$$= \int (4x^2\vec{i} + 2y\vec{j} + z^3\vec{k}) \cdot \vec{n} \, ds \quad [\vec{n} = \text{unit normal to } S_1 \text{ at any point on } S_1]$$

$$= \int [4(\vec{e}_r \cos \theta - \vec{e}_\theta \sin \theta)r^2 \cos^2 \theta + 2(\vec{e}_r \sin \theta + \vec{e}_\theta \cos \theta)r \sin \theta + z^3\vec{k}] \cdot \vec{e}_r \, r \, d\theta \, dz$$

$$= 18 \times 2 \times \pi = 36\pi$$

The total surface integral as $\iint_S \vec{F} \cdot \vec{n} \, ds = 36\pi + 72\pi = 108\pi$

$$\text{Again } \int_V (\vec{\nabla} \cdot \vec{F}) \, dv = \int (8x + 2 + 3z^2) \, dx \, dy \, dz$$

$$= 2 \times \pi \times 9 \times 2 + 8 \times \pi \times 9 = 108\pi.$$

Thus, Gauss divergence theorem is verified.

Problem 11

Calculate $\int (\vec{\nabla} \times \vec{F}) \cdot \vec{n} \, ds$ where $\vec{F} = \vec{i}xy + \vec{j}y^2z + \vec{k}xyz$ and S is a rectangle with vertices at $(0, 0, 0)$, $(0, 2, 0)$, $(1, 2, 0)$ and $(1, 0, 0)$ verify Stokes theorem.

$$\text{Solution: } \vec{\nabla} \times \vec{F} = \vec{i}(xy - y^2) - \vec{j}yz - \vec{k}x$$

$$\text{or, } \int_S (\vec{\nabla} \cdot \vec{F}) \, d\vec{s} = -\int x \, dx \, dy = -1.$$

$$\begin{aligned} \oint \vec{F} \cdot d\vec{l} &= \int_1^0 xy \, dx \quad [\text{because all other line integrals vanish}] \\ &= 2 \int_1^0 x \, dx = -1. \end{aligned}$$

Hence Stokes' law is verified.

Problem 12

Calculate $\int_S \vec{V} \cdot d\vec{s}$, where $\vec{V} = yz^2 \vec{i} + (y^2 + \sin(xz))\vec{j} + yz\vec{k}$ and S is the boundary of the region comprised of the parabolic cylinder $z = 1 - y^2$, $z = 0$, $x = 0$, $x = 4$, taking the help of Gauss divergence theorem.

Solution: $\int_S \vec{V} \cdot d\vec{s} = \int_V \nabla \cdot \vec{V} \cdot d\tau$ from Gauss theorem

$$\text{Now, } \nabla \cdot \vec{V} = 3y^2 + y$$

Or, $\int_S \vec{V} \cdot d\vec{s} = \iiint_V (3y^2 + y) dx dy dz$, where x runs from 0 to 4, z from 0 to $1 - y^2$ and y from 0 to 1.

$$\text{Or, } \int_S \vec{V} \cdot d\vec{s} = 4 \iint (3y^2 + y) dy dz = 4 \int_0^1 (3y^2 + y)(1 - y^2) dy$$

$$= 4 \left[1 - \frac{3}{5} + \frac{1}{2} - \frac{1}{4} \right] = \frac{13}{5}$$

SUMMARY

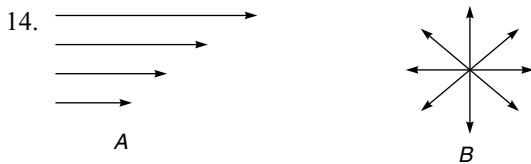
- A vector which is diagrammatically shown by a directed line segment in space requires for its mathematical representation a basis, which comprises linearly independent vectors whose number is equal to the dimension of the space. The concepts of linear independence, different types of products (viz., scalar product, vector product, direct product) have been discussed in the text.
- The change of the components of a vector under a transformation of the axes has been discussed (through analyzed more fully in Chapter 4) and the different behaviours of two types of vectors, polar and axial respectively, have been highlighted.
- The differentiation and the integrations of the vector fields (line integral, surface integral, volume integral) have been discussed with examples. The relevant theorems of vector calculus, such as, Stokes' theorem, Gauss divergence theorem, Green's theorem etc., have been dealt with keeping in mind their diverse applications to physical problems.
- The curvilinear coordinate systems have been discussed and the different operators, viz., gradient, divergence, curl have been evaluated in orthogonal curvilinear coordinate systems. The general curvilinear coordinate systems have been briefly described in the last section with a greater detail left for the Chapter 4.

MULTIPLE CHOICE QUESTIONS

- Let there be three types of terms $\vec{A} \times \vec{B}$, $\vec{A} \cdot \vec{B}$, $\vec{A} \cdot (\vec{B} \times \vec{C})$ which of the following statements is true? ($\vec{A}$, $\vec{B}$, $\vec{C}$ are polar vectors)
 - The first one is a polar vector, the second one a proper scalar and the third a pseudoscalar
 - The first one is an axial vector, the second a proper scalar and the third a pseudoscalar
 - The first one is an axial vector while the other two are proper scalars
 - The first one is an axial vector while the other two are pseudoscalars

2. The equations $\vec{r} \cdot \vec{n} = k$, $\frac{\vec{r} \cdot \vec{n}}{r} = l$, $\vec{r}$ and $\vec{n}$ being the position vector and the unit vector along a fixed direction while k and l are constants represent
- two planes perpendicular to $\vec{n}$
 - a plane perpendicular to $\vec{n}$ for the first and a circle for the second
 - a plane perpendicular to $\vec{n}$ for the first and a cone with the axis along $\vec{n}$
 - a cone for the first and a cylinder for the second.
3. You are given two vectors $\vec{A}, \vec{B}$ and three other constant vectors $\vec{C}, \vec{D}$ and $\vec{E}$ respectively. $\vec{A}$ is perpendicular to $\vec{C}$ while $\vec{B}$ is perpendicular to $\vec{D}$ and $\vec{E}$ both. (All are three dimensional vectors).
- $\vec{A}$ and $\vec{B}$ are fixed in magnitude and direction
 - $\vec{A}$ is not fixed in direction but $\vec{B}$ is fixed in magnitude and direction
 - $\vec{A}$ is fixed neither in magnitude nor in direction while $\vec{B}$ is fixed in direction only
 - $\vec{A}$ and $\vec{B}$ are fixed neither in magnitude nor in direction
4. The component of $\vec{A}$ perpendicular to $\vec{B}$ is
- $\vec{A} - \left(\frac{\vec{A} \cdot \vec{B}}{B} \right) \frac{\vec{A}}{A}$
 - $\vec{A} - \frac{(\vec{A} \cdot \vec{B})\vec{B}}{B^2}$
 - $\frac{\vec{A} \times \vec{B}}{B}$
 - $\vec{A} - \frac{(\vec{A} \times \vec{B})\vec{B}}{B^2}$
5. If the vectors $\vec{A}, \vec{B}$, and $\vec{C}$ are coplanar then $\vec{B} + \lambda\vec{C}, \vec{C} + \mu\vec{A}, \vec{A} + \nu\vec{B}$ (λ, μ, ν are constants) are
- always coplanar
 - coplanar definitely if $\lambda\mu\nu \neq -1$
 - never coplanar
 - none of these
6. The unit vector perpendicular to $\vec{A} = \vec{i} + \vec{j} + \vec{k}$ and $\vec{B} = 2\vec{i} + 3\vec{j} + 4\vec{k}$ is
- $\frac{2\vec{i} - \vec{j} + \vec{k}}{\sqrt{87}}$
 - $\frac{\vec{i} - 2\vec{j} + \vec{k}}{\sqrt{87}}$
 - $\frac{\vec{i} + 2\vec{j} + 3\vec{k}}{\sqrt{3}}$
 - $\frac{\vec{i} - \vec{j} + 2\vec{k}}{\sqrt{87}}$
7. If the acceleration of a particle is towards a fixed centre the area described per unit time at the centre
- is constant
 - increases linearly with time
 - is a quadratically increasing function of time
 - increases upto a time and decreases continuously
8. The rate of change of $\phi = xy + yz + zx$ along the direction of $\vec{i} + \vec{j} + \vec{k}$ at $(1, 2, 3)$ is
- $\sqrt{3}$
 - $4\sqrt{3}$
 - $2\sqrt{3}$
 - $3 + \sqrt{3}$
9. The angle between the surfaces $xy + yz + zx = 3$ and $x^2 + y^2 - z^2 = 1$ at $(1, 1, 1)$ is
- $\cos^{-1} 1/3$
 - $\cos^{-1} 2/3$
 - $\pi/2$
 - $\cos^{-1} 1/\sqrt{3}$
10. The vector is $\nabla [\ln |\vec{r}|]$ is
- along $\vec{r}$ and increasing with r
 - along $\vec{r}$ and decreasing with r
 - perpendicular to $\vec{r}$
 - a constant vector along $\vec{r}$

11. The temperature T given by $T = x^2 + y^2 + z^2$ increases at the maximum rate at $(1, 1, 1)$ along a direction that makes
 - (a) equal angles with the axes
 - (b) angles 30° , 60° and 90° with the axes
 - (c) 30° , 120° and 90° with the axes
 - (d) 60° , 60° and 45° with the axes
12. The vector $5\vec{i} + \vec{j}$ is obtained from the vector $2\vec{i} + 3\vec{j}$ by
 - (a) a rotation about the z -axis and a scaling
 - (b) a rotation about the z -axis combined with a translation
 - (c) a translation only
 - (d) a rotation about the x -axis
13. There are given two vectors $\vec{i}xy + \vec{j}yz + \vec{k}zx$ and $\vec{i}xy^2 + \vec{j}yz^2 + \vec{k}zx^2$ respectively. $\int_1^2 \vec{F} \cdot d\vec{r}$ ($\vec{F}$ represents the vector) is
 - (a) dependent on the path for both
 - (b) independent of the path for both
 - (c) independent of the path for the first and depends on the path for the second
 - (d) dependent on the path for both



The vectors $\vec{A}$ and $\vec{B}$ in the above figure have

- (a) vanishing curl and divergence respectively
 - (b) nonvanishing curl and divergence respectively
 - (c) nonvanishing curl and vanishing divergence respectively
 - (d) vanishing curl and nonvanishing divergence respectively
15. The condition that the vector $\vec{v} = a\vec{i}(y+z) + b\vec{j}(z+x) + c\vec{k}(x+y)$ is irrotational is
 - (a) $a = b = c$
 - (b) $a + b + c = 0$
 - (c) $a = b = -c$
 - (d) $a = -b = c$
 16. $\int_s d\vec{s}$ over a closed spherical surface of radius r is
 - (a) zero
 - (b) $4\pi r^2 \vec{n}$ ($\vec{n}$ is a unit vector along $\vec{r}$)
 - (c) $\pi r^2 \vec{n}$
 - (d) $\frac{4}{3}\pi r^3$
 17. $\int_s \vec{r} \cdot d\vec{s}$ over the surface of a hemisphere of radius r is
 - (a) $4\pi r^3$
 - (b) zero
 - (c) $2\pi r^3$
 - (d) $\frac{4}{3}\pi r^3$
 18. A particle travels along a circle with constant angular velocity $\vec{\omega}$. Its linear velocity has
 - (a) zero divergence and zero curl

- (b) zero divergence but nonzero curl
 (c) non zero divergence but zero curl
 (d) non zero divergence and nonzero curl.
19. Which of the statements is true?
 A vector in three dimensions can
 (a) always be written as the sum of three vectors
 (b) be written as the sum of three vectors only if they are non coplanar
 (c) be written as the sum of three vectors only if they are coplanar
 (d) never be written as the sum of three vectors
20. The equation of the tangent plane to the surface $xy + yz + zx = 3$ at $(1, 1, 1)$ is
 (a) $x + y + z = 6$ (b) $x + y + z = 3$
 (c) $2x + 3y + 4z = 6$ (d) $x + y + z = 0$
21. $\int_{\delta} (\bar{\nabla} \times \bar{A}) \cdot d\bar{s}$ over a closed surface is
 (a) dependent on $\bar{A}$
 (b) dependent on $\bar{A}$ and the nature of the surface
 (c) zero always
 (d) depends only on the nature of the surface
22. The surface integral of $\bar{r}$ over a cube of side unity is
 (a) zero
 (b) 3
 (c) dependent on the choice of the origin
 (d) same as that over a sphere of unit radius
23. The condition that the vector $\bar{A} = \bar{i}(2x + 2ay + z) + \bar{j}[(a + 2)x + 5y + z] + \bar{k}(x + (b - 2)y + z)$ is irrotational is that
 (a) $a = 2, b = 1$ (b) $a = 2, b = 3$
 (c) $a = 1, b = 2$ (d) a and b can have any set of values

Answers to MCQs

1. (b)	2. (c)	3. (c)	4. (b)	5. (b)	6. (b)	7. (a)
8. (b)	9. (a)	10. (b)	11. (a)	12. (a)	13. (c)	14. (b)
15. (a)	16. (a)	17. (c)	18. (b)	19. (b)	20. (b)	21. (c)
22. (b)	23. (b)					

EXERCISE

- Prove that a vector equation of the form $\bar{A} = \bar{B}$ remains the same in all coordinate systems. Should $\bar{A}$ and $\bar{B}$ be vectors of identical type? (Type means polar or axial).
- Can $\bar{C}$ be uniquely determined if for given $\bar{A}$ and $\bar{B}$ the equation $\bar{A} \cdot \bar{C} = \bar{B}$ or $\bar{A} \times \bar{C} = \bar{B}$ is satisfied.

3. Calculate $\bar{A} \times (\bar{B} \times \bar{C})$, $(\bar{A} \times \bar{B}) \times (\bar{C} \times \bar{D})$ by using Levi-civita symbol ϵ_{ijk} [i.e. $(\bar{B} \times \bar{C})_i = \epsilon_{ijk} B_j C_k$ and use $\epsilon_{ijk} \epsilon_{ilm} = \delta_{jl} \delta_{km} - \delta_{jm} \delta_{kl}$].

Prove by using the same symbol that $\bar{A} \cdot (\bar{B} \times \bar{C}) = 0$.

4. (a) Prove that the planes $x + y + z = 1$ and $5x - 7y + 2z = 8$ are orthogonal (Hint:- The planes are orthogonal if their normals at any point are perpendicular).
 (b) Find out the equation of a line normal to the plane $3x + 4y + 2z = 0$ at the origin.
 (c) Find out the smaller angle between the planes $2x + 3y + 5z = 6$ and $3x - y + 2z = 2$ (Hint:- Angle between the planes is equal to that between the normal to the planes).
 5. If the temperature at a point (x, y, z) be given by $T(x, y, z) = ax + by + cz$, a, b, c being constants, find out the spatial rate of change of T along a direction given by $\bar{n} = (i + 2j + 2k)\frac{1}{3}$. In what direction is the rate maximum?

6. (a) Establish, $\bar{\nabla} \phi = \lim_{\Delta v \rightarrow 0} \frac{\Delta \phi}{\Delta v}$, $\oint_S d\bar{s} = 0$ from Gauss theorem.

(b) In the text the following formula were deduced.

$$\oint_c \Psi d\bar{r} = \int_S [\bar{n} \times \bar{\nabla} \Psi] ds, \oint_S d\bar{r} \times \bar{F} = \int_S [(\bar{n} \times \bar{\nabla}) \times \bar{F}] ds.$$

Use an alternative method to deduce the above.

[Hint: Proceed in a method as in the deduction of Stokes law]

7. Utilizing the Levi-civita symbol calculate the following: $\nabla[\bar{a} \cdot (\bar{b} \times \bar{r})]$, $\nabla(\bar{a} \cdot \bar{r})$, $\bar{\nabla} \times [(\bar{a} \times \bar{r})]$ ($\bar{a}, \bar{b}$ are constant vectors).
 8. Calculate the work done by a force $\bar{F} = \bar{i}(y + z) + \bar{j}(x + z) + \bar{k}(x + y)$ in displacing a point from $(0, 0, 0)$ to $(1, 1, 1)$ by first moving along a straight line from $(0, 0, 0)$ to $(1, 1, 0)$ and then along the z -axis to $(1, 1, 1)$. Will the work done be different along other alternative routes? Explain.
 9. Calculate $\oint_c \bar{F} \cdot d\bar{l}$ along the bounding curve C which is the curve of intersection of the hemisphere $x^2 + y^2 + z^2 = 4$, $z \geq 0$ with the plane $z = 0$ and $\bar{F}$ is the vector $(2x^2 + 3y^2)\bar{i} + xy^2 z^2 \bar{j} + yz^3 \bar{k}$. Hence verify Stokes' law.
 10. Verify if the following force are conservative and find out the potentials in the respective cases.
 (a) $\bar{F} = i(y^2 + z^2) + j(x^2 + z^2) + k(x^2 + y^2)$.
 (b) $\bar{F} = iyz + jzx + kxy$.
 (c) $\bar{F} = \bar{r} \sin r$ ($\bar{r} = \bar{i}x + \bar{j}y + \bar{k}z$).
 (d) $\bar{F} = \bar{i}(2xy + y^3 + 3) + \bar{j}(x^2 + 3xy^2 + y^6)$.
 11. Deduce the condition that the coordinate system (u, v, w) given by $x = a_1 u + b_1 v + c_1 w$, $y = a_2 u + b_2 v + c_2 w$, $z = a_3 u + b_3 v + c_3 w$ is orthogonal. Deduce the expression of the infinitesimal line element and volume element in this system.

[Hint: For an orthogonal system $\bar{e}_j \cdot \bar{e}_c = \delta_{ij}$, where $\bar{e}_i = \frac{1}{h_i} \frac{\partial \bar{r}}{\partial q_i}$]

12. (a) Calculate $\int_S \vec{V} \cdot d\vec{s}$, where $\vec{V} = (3x^2 - y^2z^2)\vec{i} + (z^2 + y^2)\vec{j} + (2yz + x^2)\vec{k}$ and S is the surface $(x-2)^2 + (y-1)^2 + z^2 = 16$. Also verify the divergence theorem.
- (b) Do the same for $\vec{V} = xy\vec{i} + (x^2y + xy^2)\vec{j} + 2xyz\vec{k}$ and S is the surface of the cylinder $x^2 + y^2 = 16$ bounded by the planes $z = 1$ and $z = 4$.
13. A particle describes a motion on a curve in two dimensions. Find out the components of velocity and acceleration in plane polar coordinates.
[Hint: $\vec{v} = \vec{i}v_x + \vec{j}v_y$ and express $\vec{i}, \vec{j}, v_x, v_y$ in terms of $\vec{e}_r, \vec{e}_\theta, v_r, v_\theta$ respectively]
14. A particle describes a motion in three dimensions with the force on it being given by $\left(A - B\frac{dz}{dt}\right)\vec{e}_z$, A and B being constants. Find out the displacement at any instant of time.
15. The relativistic equation of motion of a charged particle in an electromagnetic field is given by $\frac{d\xi}{dt} = \vec{F} \cdot \vec{v}$, where ξ is the energy given by $\xi = mc^2$, m being the relativistic mass and $\vec{F}$ the force on it being given by $\vec{F} = e(\vec{E} + \vec{v} \times \vec{B})$. Find out the trajectory of the particle when (a) $\vec{E} = 0$ and the particle has an initial component of velocity parallel to a uniform $\vec{B}$ in the z -direction, (b) $\vec{B} = 0$ and the initial velocity $\vec{v}_0$ is perpendicular to the constant electric field $\vec{E}$.
16. Verify Stokes' theorem for a vector $\vec{F} = \vec{i}(2x^2y + xy^2) + \vec{j}(yz^2 + xyz) + \vec{k}(xz^2 + x^2y)$ with the surface of integration comprising six faces of a cube given by $x = 0, x = 1, y = 0, y = 1, z = 0, z = 1$ the last face being kept open.
17. Calculate $\int_S (\vec{V} \times \vec{F}) \cdot \vec{n} ds$, where $\vec{F} = 2xz\vec{i} + 3y\vec{j} - 2x^2y\vec{k}$ and S is the surface of the region bounded by $x = 0, y = 0, z = 0$ and $x + y + z = 5$. Hence verify Stokes law.
[Hint: Express $\vec{n} ds$ as $\left(\frac{\vec{i} + \vec{j} + \vec{k}}{\sqrt{3}}\right) \frac{dx dy}{(\vec{n} \cdot \vec{k})}$ or as $(\vec{i} + \vec{j} + \vec{k}) dx dy$.]
18. Find out which of the following line integrals are meaningful and deduce their expressions.
- (i) $\int_{(0,0,0)}^{(1,1,1)} [\vec{i}yz + \vec{j}xz + \vec{k}(xy + z^2)] \cdot d\vec{r}$
- (ii) $\int_{(1,1,1)}^{(2,2,2)} [3\vec{i}x^2y + \vec{j}(x^3 + y^3) + \vec{k}(x^2z^2 + yz^3)] \cdot d\vec{r}$
- (iii) $\int_{(a,b,c)}^{(2a,2b,2c)} [2\vec{i}y \sin x + \vec{j}(\cos x + z^2) + \vec{k}(2yz + z^2)] \cdot d\vec{r}$
19. Calculate $\int [(y^2 + x^2)dx + (xy + 1)dy]$ along the x -axis from $(0, 0)$ to $(1, 0)$, then along a straight line from $(1, 0)$ to $(1, 1)$ and then back to $(0, 0)$ along the y -axis. Hence verify Green's theorem in a plane.

[**Hint:** In calculating $\iint f(x, y) dx dy$ over the region vary y from 0 to $1 - x$ and x from 0 to 1 or x from 0 to $1 - y$ and y from 0 to 1]

20. Show that a vector $\vec{A}$ can be decomposed into two parts, one being solenoidal (divergence is zero) and the other irrotational ($\text{Curl} = 0$) (Hint: Let $\vec{B} = \vec{\nabla} \times \vec{A}$ and $\Psi = \vec{\nabla} \cdot \vec{A}$. Introduce $\vec{C}(\vec{r}) = \int_{\vec{r}'} \frac{\vec{B}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}'$. Show that $\vec{\nabla} \cdot \vec{C} = 0$ and $\nabla^2 \vec{C} = -\vec{\nabla} \times \vec{A}$. Now use the relation $\vec{\nabla} \times (\vec{\nabla} \times \vec{C}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{C}) - \nabla^2 \vec{C}$ to obtain $\vec{A} = \vec{\nabla} \times \vec{C} + \vec{\nabla} \phi$ where $\nabla^2 \phi = \Psi$)

Hints and Answers to Exercise Questions

2. $\vec{C}$ is non-unique in either of the two cases.
 [**Hint:** Assume that $\vec{C}'$ is another solution satisfying the above equations. Then find out the relation between $\vec{C}$ and $\vec{C}'$]
4. (b) The equation of the line is $\frac{x}{3} = \frac{y}{4} = \frac{z}{2}$.
 (c) The angle $\theta = \cos^{-1} \left(\frac{13}{2\sqrt{133}} \right)$.
5. $\frac{a+2b+2c}{3}$; Along $\vec{\nabla} T$.
7. 0, $\vec{a}$, $2\vec{a}$
8. 3; No. [**Hint:** Prove that $\vec{F} = \vec{\nabla} \phi$ where ϕ is a scalar]
9. 0
10. (a) $\phi = x^2y + xy^3 + 3x + \frac{y^7}{7}$
 (b) $\phi = xyz$
 (c) $\phi = \sin r - r \cos r$
 (d) $x^2y + xy^3 + 3x + \frac{y^7}{7}$ [The arbitrary constant has been set equal to zero]
11. $\sum_{i=1}^3 a_i b_i = \sum_{i=1}^3 a_i c_i = \sum_{i=1}^3 b_i c_i = 0$; $ds^2 = \left(\sum_i a_i^2 \right) du^2 + \left(\sum_i b_i^2 \right) dv^2 + \left(\sum_i c_i^2 \right) dw^2$,
 and $dV = \left[\left(\sum_i a_i^2 \right) \left(\sum_i b_i^2 \right) \left(\sum_i c_i^2 \right) \right]^{1/2} du dv dw$
 provided the coordinate system is chosen to be orthogonal. For nonorthogonal system use $ds^2 = g_{ij} dq^i dq^j$, where $g_{ij} = \frac{\partial \vec{r}}{\partial q^i} \cdot \frac{\partial \vec{r}}{\partial q^j}$ and $q^1 = u$, $q^2 = v$, $q^3 = w$, and similar such relations for volume and surface elements.

12. (a) $\frac{4^6\pi}{3}$ (b) 192π .
13. $v_r = \frac{dr}{dt}, v_\theta = r \frac{d\theta}{dt}, a_r = \frac{d^2r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2, a_\theta = \frac{1}{r} \frac{d}{dt} \left(r^2 \frac{d\theta}{dt} \right).$
14. $x = x_0 + u_x t, y = y_0 + u_y t, z = \frac{A}{B} t + C e^{-\frac{B}{m}t} + z_0 - \left(\frac{A}{B} + C \right),$ where x_0, y_0, z_0 are the initial values of x, y, z and C is an arbitrary constant.
15. (a) Helix
(b) The orbit equation is complicated. Due to the change of energy because of the presence of the electric field, the mass will vary with time. Solve the equations $= \frac{d\xi}{dt} = e\bar{E} \cdot \bar{u}$ and $\frac{d\bar{p}}{dt} = e\bar{E}, \xi$ being the energy.
18. (i) Meaningful; 4/3 (ii) and (iii) integrals are not meaningful
[Hint: Test if the vectors can be written as $\nabla\phi$]
19. $-\frac{1}{6}$

APPENDIX

Simply Connected and Multiply Connected Regions

A region is simply connected if any closed curve in it can be continuously contracted to a point without leaving the region. If it cannot be done it is multiply connected. For example, the region within a closed curve is simply connected while the region within the curve with holes is multiply connected. (Fig. 2.23(a) and 2.23(b) respectively) while that shown in Fig. 2.23(c) is again simply connected (also see Fig. 2.24).

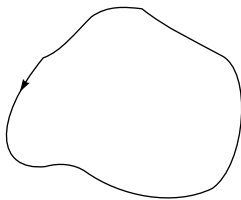


Fig. 2.23(a) Simply connected region

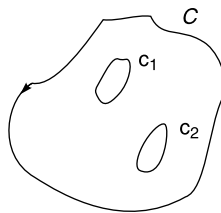


Fig. 2.23(b) Multiply connected

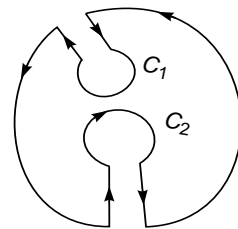


Fig. 2.23(c) Multiply connected region converted to a simply connected region.

Thus, referring to Section (2.5) or (2.7) (see Fig. 2.24) we obtain integrating along C consisting of C_1, C_2, AB and DC ,

$$\begin{aligned}\oint_C \bar{A} \cdot d\bar{r} &= \oint \left[\frac{x}{x^2 + y^2} dy - \frac{y}{x^2 + y^2} dx \right] \\ &= \oint_{C_1} \left[\frac{x}{x^2 + y^2} dy - \frac{y}{x^2 + y^2} dx \right] - \oint_{C_2} \left[\frac{x}{x^2 + y^2} dy - \frac{y}{x^2 + y^2} dx \right]\end{aligned}$$

= 0 because each of the integrals gives 2π and the integrals along the bridge's AB and DC cancel.

$$\left[\oint_{C_1} \left[\frac{x}{x^2 + y^2} dy - \frac{y}{x^2 + y^2} dx \right] = \int_0^{2\pi} d\theta = 2\pi, \text{ on putting } x = r \cos \theta, y = r \sin \theta. \right]$$

Since $\oint_C \bar{A} \cdot d\bar{r} = \int_s \text{Curl } \bar{A} \cdot d\bar{s} = 0$ after the region becomes simply connected, Stokes' law is verified.

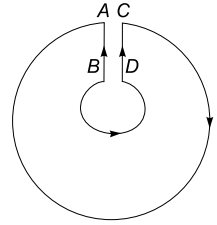


Fig. 2.24

Linear Vector Space and Matrices

CHAPTER

3

3.1 DEFINITION OF VECTOR SPACE

In this chapter we will regard vectors not simply as the directed line segments but as the members of a linear vector space with the properties enlisted below.

A linear vector space comprises a set of elements $a, b, \dots$ where the operations of addition and multiplication by scalars is defined and which satisfy the following conditions.

- (i) $a + b = b + a$, i.e., the addition is commutative and the sum is a member of the space,
- (ii) The addition is associative, i.e., $a + (b + c) = (a + b) + c$.
- (iii) The space is closed under multiplication by a scalar, i.e., if α be a scalar then αa is also a member of the same space.
The product follows the rules below:-
- (iv) $\alpha(a + b + c) = \alpha a + \alpha b + \alpha c$, (3.1)
- (v) $(\alpha + \beta)a = \alpha a + \beta a$, (vi) $\alpha(\beta a) = (\alpha\beta)a$
- (vi) There exists an inverse $-a$ corresponding to each a such that $a - a = a + (-1)a = 0$, where 0 is called the null element of the space.
- (vii) $a + 0 = 0 + a = a$ for every a .

Examples

- (i) The set of ordinary vectors in a finite dimensional space. The null element is the null vector, the inverse is the vector with the same magnitude as but the direction opposite to that of the original vector and the addition obeys the parallelogram law.
- (ii) The set of finite dimensional square or rectangular matrices, the null element being the null matrix and the inverse, the matrix with all the elements changed in sign.
- (iii) The set of all possible solutions of a homogeneous differential equation satisfying proper boundary conditions.

The vector space is thus a generalization of the concept of the space of ordinary vectors. The space is finite or infinite dimensional according as it contains a finite or an infinite number of elements. Thus the example (i) above represents a finite dimensional vector space while the examples (ii) and (iii) represent infinite dimensional spaces. Examples of a few vector spaces with certain additional features are briefly introduced in the appendix.

3.2 SUBSPACE

A part of the original vector space that satisfies the same axioms as the original vector space under the same rules of addition and multiplication by scalars exhibited in Eq. (3.1) is called a subspace of the original vector space. Thus (i) a set of integers forms a vector space of which the set of even integers including zero is a subspace both being infinite dimensional, (ii) the vectors lying in a lower dimensional projection of a finite dimensional space form a subspace.

3.3 INNER PRODUCT; LINEAR INDEPENDENCE

The concept of the inner product of two vectors of a vector space is important in the context of determining the magnitude (or norm) of a vector, ascertaining the linear independence of the vectors, etc. The inner product of two vectors a, b is a scalar denoted by (a, b) that satisfies the following properties:

- (i) $(a, b) = (b, a)^*$, where a, b are vectors in a complex vector space and the symbol $*$ denotes complex conjugation,
- (ii) $(a, b + c) = (a, b) + (a, c)$,
- (iii) $(a, \alpha b) = \alpha (a, b)$,
- (iv) $(\alpha a, b) = \alpha^* (a, b)$, α being a complex scalar.

In the case of ordinary finite dimensional vectors the inner product, also called the scalar or the dot product, of two vectors $\vec{a}$ and $\vec{b}$ is given by $(a, b) = \vec{a} \cdot \vec{b} = ab \cos \theta$, θ being the angle between two vectors while the norm of the vector $\vec{a}$ is $|\vec{a}| = (a, a)^{1/2}$. The work done by a force in covering a displacement

given by $\int_A^B \vec{F} \cdot d\vec{r}$ is an example of the inner product of two vectors. If the vector space be spanned by the

complex functions $f(x), g(x)$ etc., the variable x being defined in an interval (x_1, x_2) the inner product (f, g)

is defined by $\int_{x_1}^{x_2} f^*(x) g(x) w(x) dx$, $w(x)$ being the weight factor not necessarily equal to unity. Two vectors

a, b are orthogonal if $(a, b) = 0$.

Let us consider a vector $\gamma = \alpha + \lambda (\beta, \alpha) \beta$, where α and β are any two vectors and λ is a real constant.

Since $(\gamma, \gamma) \geq 0$ we obtain, $(\alpha, \alpha) + \lambda^2 (\beta, \alpha) (\alpha, \beta) (\beta, \beta) + 2\lambda (\alpha, \beta) (\beta, \alpha) \geq 0$

Since λ is real, $(\alpha, \beta)^2 (\beta, \alpha)^2 \leq (\alpha, \alpha) (\beta, \beta) (\alpha, \beta) (\beta, \alpha)$.

Or, $(\alpha, \alpha) (\beta, \beta) \geq |(\alpha, \beta)|^2$, Since $(\beta, \alpha) = (\alpha, \beta)^*$ (3.2)

This is called Canchy-Schwarz inequality and is used extensively in physics, particularly in quantum mechanics in demonstrating the uncertainty relation and determining the minimum uncertainty product wave function. In ordinary vector analysis Canchy-Schwarz inequality reduces to $(a, b)^2 = a^2 b^2 \cos^2 \theta \leq a^2 \cdot b^2$, [θ is the angle between two vectors] an inequality which is almost trivially satisfied.

3.3.1 Linear Independence of Vectors

A set of vectors $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n$ in a n -dimensional vector space is said to be linearly independent if the relation $C_1 \bar{A}_1 + C_2 \bar{A}_2 + \dots + C_n \bar{A}_n = 0$ has the only solution $C_1 = C_2 = \dots = C_n = 0$.

Taking the scalar products of the above relation with $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n$ successively we get,

$$\begin{aligned} C_1 \bar{A}_1 \cdot \bar{A}_1 + C_2 \bar{A}_1 \cdot \bar{A}_2 + \dots + C_n \bar{A}_1 \cdot \bar{A}_n &= 0 \\ C_1 \bar{A}_2 \cdot \bar{A}_1 + C_2 \bar{A}_2 \cdot \bar{A}_2 + \dots + C_n \bar{A}_2 \cdot \bar{A}_n &= 0 \\ C_1 \bar{A}_n \cdot \bar{A}_1 + C_2 \bar{A}_n \cdot \bar{A}_2 + \dots + C_n \bar{A}_n \cdot \bar{A}_n &= 0 \end{aligned} \quad (3.3)$$

The necessary condition for linear dependence is thus, ($C_1, \dots, C_n$ not all zero)

$$\begin{vmatrix} \bar{A}_1 \cdot \bar{A}_1 & \bar{A}_1 \cdot \bar{A}_2 & \dots & \bar{A}_1 \cdot \bar{A}_n \\ \bar{A}_2 \cdot \bar{A}_1 & \bar{A}_2 \cdot \bar{A}_2 & \dots & \bar{A}_2 \cdot \bar{A}_n \\ \vdots & \vdots & \ddots & \vdots \\ \bar{A}_n \cdot \bar{A}_1 & \bar{A}_n \cdot \bar{A}_2 & \dots & \bar{A}_n \cdot \bar{A}_n \end{vmatrix} = 0 \quad (3.4)$$

The determinant is called Gram's determinant and hence the condition that the vectors are linearly independent is given by the Gram's determinant being non-vanishing. Thus, a set of orthogonal vectors is seen to be clearly linearly independent. A set of linearly independent vectors is, for example represented by

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \dots \text{ where } \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ means a vector with components } x_1, x_2, \dots, x_n \text{ respectively.}$$

For a set of n functions ($f_1(x), f_2(x) \dots f_n(x)$) covering a function space we obtain by multiplying by $f_1, \dots, f_n$ successively and integrating over the entire interval of x an analogous condition,

$$\begin{vmatrix} \int_a^b f_1^2 dx & \int_a^b f_1 f_2 dx & \dots & \int_a^b f_1 f_n dx \\ \int_a^b f_1 f_2 dx & \int_a^b f_2^2 dx & \dots & \int_a^b f_2 f_n dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_a^b f_n f_1 dx & \int_a^b f_n f_2 dx & \dots & \int_a^b f_n^2 dx \end{vmatrix} = 0 \quad (3.5)$$

for the functions to be linearly dependent.

An alternative condition exists for testing if the functions spanning the space in an interval are linearly independent or not. By differentiating the relation $c_1 f_1 + \dots + c_n f_n = 0$ consecutively 1, 2, ..., $n-1$ times we obtain the set of relations.

$$\begin{aligned} c_1 f_1 + c_2 f_2 + \dots + c_n f_n &= 0 \\ c_1 f_1' + c_2 f_2' + \dots + c_n f_n' &= 0 \\ \vdots & \\ c_1 f_1^{n-1} + c_2 f_2^{n-1} + \dots + c_n f_n^{n-1} &= 0 \end{aligned}$$

The condition for linear independence is that

$$\begin{vmatrix} f_1 & f_2 & \dots & f_n \\ f_1' & f_2' & \dots & f_n' \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{n-1} & f_2^{n-1} & \dots & f_n^{n-1} \end{vmatrix} \neq 0, \quad (3.6)$$

From (3.6), clearly the functions $1, x, x^2, \dots, x^{n-1}$ can be shown to form a set of n linearly independent functions in any range of values of x .

The linearly independent vectors (or the functions in the space of functions) form the basis of the vector space. The maximum number of linearly independent vectors in a vector space is called the dimension of the vector space. Thus, in a n -dimensional vector space, the maximum number of linearly independent vectors is n . Of course, the choice of the basis vectors in a vector space is not unique but the dimension of the vector space being fixed, the number of basis vectors is the same for any choice. Thus for a three-dimensional vector space, the unit vectors $\bar{i}, \bar{j}, \bar{k}$ or any three linearly independent combinations of the vectors $\bar{i}, \bar{j}, \bar{k}$ can be chosen as the basis vectors.

The n linearly independent vectors in a vector space of n -dimensions, make a complete finite dimensional vector space. Any vector in the space can then be expanded as a linear combination of the independent vectors. In the case of an infinite dimensional vector space, the completeness condition is more involved because of the problem of convergence of the infinite set.

3.4 GRAM-SCHMIDT ORTHOGONALISATION METHOD

We have seen in the last section that a set of orthogonal vectors is linearly independent but the converse is not necessarily true. For example, in the simple case of a two dimensional vector space, any two non-collinear vectors can form a basis. It is frequently convenient to work with a set of mutually orthogonal vectors primarily for simplicity of calculation. The method we discuss below provides us with a method of constructing a set of orthogonal vectors out of a set of linearly independent vectors.

Let $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n$ be the set of n linearly independent vectors. Let $\bar{B}_1, \bar{B}_2, \dots, \bar{B}_n$ be the desired orthogonal set. Let us choose $\bar{B}_1 = \bar{A}_1$.

Let $\bar{B}_2 = \bar{A}_2 + C_1 \bar{A}_1$, the constant C_1 being chosen so that $\bar{B}_2 \cdot \bar{B}_1 = 0$ or, $\bar{A}_2 \cdot \bar{B}_1 = -C_1 \bar{A}_1 \cdot \bar{B}_1 = -C_1 |\bar{B}_1|^2$ giving $\bar{B}_2 = \bar{A}_2 - \frac{\bar{A}_2 \cdot \bar{B}_1}{|\bar{B}_1|^2} \bar{B}_1$.

Thus, $\bar{B}_2$ is obtained by subtracting from $\bar{A}_2$ the projection of $\bar{A}_2$ on $\bar{B}_1$ so that $\bar{B}_2$ becomes perpendicular to $\bar{B}_1$ (or $\bar{A}_1$). Next, let $\bar{B}_3 = \bar{A}_3 + d_1 \bar{A}_2 + d_2 \bar{A}_1$, the constants d_1 and d_2 being chosen so that $\bar{B}_3 \cdot \bar{B}_1 = \bar{B}_3 \cdot \bar{B}_2 = 0$

$$\text{or,} \quad (\bar{A}_3 + d_1 \bar{A}_2 + d_2 \bar{A}_1) \cdot \bar{B}_1 = \bar{A}_3 \cdot \bar{B}_1 + d_1 \bar{A}_2 \cdot \bar{B}_1 + d_2 \bar{A}_1 \cdot \bar{B}_1 = 0$$

$$\text{and} \quad (\bar{A}_3 + d_1 \bar{A}_2 + d_2 \bar{A}_1) \cdot \bar{B}_2 = \bar{A}_3 \cdot \bar{B}_2 + d_1 \bar{A}_2 \cdot \bar{B}_2 + d_2 \bar{A}_1 \cdot \bar{B}_2 = 0$$

These two equations enable us to determine d_1 and d_2 and hence $\bar{B}_3$. $\bar{B}_3$ can be seen to be $\bar{A}_3$ minus its projection on a plane containing $\bar{B}_1$ and $\bar{B}_2$ so that $\bar{B}_3$ becomes automatically perpendicular to $\bar{B}_1$ and $\bar{B}_2$.

Thus continuing the process we can obtain a set of n orthogonal vectors $\bar{B}_1, \bar{B}_2, \dots, \bar{B}_n$. If we want the set of vectors to be normalized, we have to divide each of the vectors $\bar{B}_1, \bar{B}_2, \dots, \bar{B}_n$ by its magnitude.

We can proceed in a similar way, in the case of a set of functions. Let us start with a set of linearly independent functions $1, x, x^2, \dots$ and try to construct a set of orthonormal functions $f_1(x), f_2(x), \dots$ in the interval $(-1, 1)$, the orthonormalisation condition being given by $\int_{-1}^1 f_l f_m dx = \frac{2}{2l+1} \delta_{lm}$.

Let us take $f_1 = 1$. This satisfies the normalisation condition above with $l = 0$.

Let $f_2 = c_1 f_1 + c_2 x$, the constants $c_{1,2}$ being so adjusted that $\int_{-1}^1 f_2 f_1 dx = 0$ and $\int_{-1}^1 f_2^2 dx = \frac{2}{3}$. On calculation, this gives $c_1 = 0$ and $c_2 = \pm 1$.

Taking the positive square root we get, $f_2 = x$.

Next, let $f_3 = d_1 f_1 + d_2 f_2 + d_3 x^2$, d_1, d_2 and d_3 being so chosen that $\int_{-1}^1 f_3 f_1 dx = \int_{-1}^1 f_3 f_2 dx = 0$

and $\int_{-1}^1 f_3^2 dx = \frac{2}{5} \cdot (l = 2)$.

We get, $d_2 = 0, d_1 = -\frac{1}{2}, d_3 = \frac{3}{2}$ so that $f_3 = \frac{1}{2}(3x^2 - 1)$,

Proceeding in this way we will obtain the complete set of orthonormalised functions which are simply Legendre polynomials of different orders. By starting with the same set of linearly independent functions $1, x, x^2, \dots$ and changing the interval of x and normalisation condition, we obtain other types of orthogonal polynomials such as Hermite, Laguerre polynomials etc.

Note: In the case of a complex vector space the orthogonalisation condition $\bar{B}_2 \cdot \bar{B}_1 = 0$ becomes, $(A_2 + C_1 A_1, A_1) = 0$ which gives $(A_2, A_1) + C_1^* (A_1, A_1) = 0$ and hence $C_1 = -\frac{\overline{(A_2, A_1)}}{(A_1, A_1)}$, where $[\overline{A_2, A_1}]$ denotes the complex conjugate of (A_2, A_1) $\left[(A_2, A_1) = \sum_i A_{2i}^* A_{1i} \right]$.

Thus for two vectors $(-1, 1)$ and $(1, -i)$, $C_1 = -\frac{\overline{(-1+i)}}{2} = \left(\frac{1+i}{2} \right)$

or, $\bar{B}_2 = (1, -i) + \left(\frac{1+i}{2} \right) (-1, i) = \left(\frac{1}{2}(1-i), \frac{1}{2}(1-i) \right)$

This can be continued to a space of any finite dimension.

SOLVED PROBLEMS

Problem 1

If $\begin{pmatrix} x \\ y \end{pmatrix}$ and $\begin{pmatrix} u \\ v \end{pmatrix}$ are linearly independent write $\begin{pmatrix} a \\ b \end{pmatrix}$ in terms of $\begin{pmatrix} x \\ y \end{pmatrix}$ and $\begin{pmatrix} u \\ v \end{pmatrix}$.

Solution: Let $\begin{pmatrix} a \\ b \end{pmatrix} = c_1 \begin{pmatrix} x \\ y \end{pmatrix} + c_2 \begin{pmatrix} u \\ v \end{pmatrix}$

or, $a = c_1 x + c_2 u$, $b = c_1 y + c_2 v$ yielding the result

$$c_1 = \frac{av - bu}{xv - yu}, \quad c_2 = \frac{bx - ay}{xv - yu}$$

The method applies to the problem of expanding the state of a spin $\frac{1}{2}$ particle in terms of the spin up and the spin down eigenstates. Thus $\begin{pmatrix} a \\ b \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, where $c_1 = a$, $c_2 = b$.

Problem 2

Find out the dimension of the vector space spanned by the columns of

$$\begin{pmatrix} 2 & 3 & -1 \\ -1 & 4 & -16 \\ 0 & -8 & 24 \end{pmatrix}$$

Solution: Let us consider the equation $c_1 \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ 4 \\ -8 \end{pmatrix} + c_3 \begin{pmatrix} -1 \\ -16 \\ 24 \end{pmatrix} = 0$

$$\text{Thus we obtain, } 2c_1 + 3c_2 - c_3 = 0 \quad \dots(1)$$

$$-c_1 + 4c_2 - 16c_3 = 0 \quad \dots(2)$$

$$-8c_2 + 24c_3 = 0 \quad \dots(3)$$

These give, $c_2 = 3c_3$ and $c_1 = -4c_3$. Thus putting these in the first equation we can express any column of the matrix as a linear combination of the other two and hence the dimension of the vector space which is the maximum number of linearly independent vectors belonging to the space is 2.

Problem 3

Construct a set of orthogonal basis vectors lying in the plane $2x + y + 2z = 7$.

Solution: We choose any vector with components satisfying the above relation. For example, we can take $x = 2$, $y = 1$, $z = 1$. From the condition of orthogonality with the second vector we have $2x_1 + y_1 + z_1 = 0$, x_1 , y_1 and z_1 being the components of the second one. Again, since $2x_1 + y_1 + 2z_1 = 7$,

we get, $z_1 = 7$. Thus, the sets of orthogonal vectors are $\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} x_1 \\ -7 - 2x_1 \\ 7 \end{pmatrix}$. Of course, depending on

the components of the first vector we can have an infinite number of choices of the second one.

CONCEPT CHECK

- (a) Prove that a given vector can be expanded uniquely in terms of a given set of basis vectors.
(b) Prove that every non-empty subset of a finite dimensional vector space is linearly independent.
[Hint: Assume non-unique expansion and use the definition of linear independence in (a)]
- Suppose we want to construct a vector space of polynomials of degree n . How many basis vectors should we have? Is the choice of basis vectors unique?
- Are the vectors $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ linearly independent? If they are not orthogonal, construct an orthogonal set of vectors out of them.
- Let $|a\rangle = |e\rangle - |f\rangle$, $|b\rangle = |f\rangle - |g\rangle$ and $|c\rangle = |e\rangle - |g\rangle$ so that $|a\rangle + |b\rangle = |c\rangle$. Applying Schwarz inequality prove that $d(e, g) \leq d(e, f) + d(f, g)$, where $d(e, g) = \sqrt{\langle e | c \rangle}$, $d(e, f) = \sqrt{\langle a | a \rangle}$, $d(f, g) = \sqrt{\langle f | g \rangle}$.
(The symbol $|a\rangle$ stands for the vector a in a basis-independent representation. Later in Section 3.11 a brief discussion will be given about the application of such a representation. $\langle a | b \rangle$ represents the inner product of $|a\rangle$ and $|b\rangle$).
- Prove that the intersection of two subspaces is a subspace.
[Hint: A vector lying in the intersection of two subspaces belongs to both].
- Start with the set of linearly independent functions $(1, x, x^2, \dots)$. Construct a set of orthonormal functions out of these with the orthonormalisation condition being given by $\int_{-\infty}^{\infty} f_n f_m e^{-x^2} dx = 2^n n! \sqrt{\pi} \delta_{nm}$, f_n being the n th. function.

3.5 LINEAR OPERATORS, MATRICES

Apart from the elements of the vector space, the other relevant objects are operators acting on the vector space. The operators acting on a finite vector space can transform it into (i) another one of a different dimension or (ii) one of the same dimensions. The operators are assumed to be linear operators with the following properties.

A linear operator A satisfies the following properties

- $A(c_1 a + c_2 b + \dots) = c_1 Aa + c_2 Ab + \dots$, where $c_1, c_2 \dots$ are constants and a, b , the elements of the vector space.
- $(AB)a = A(Ba)$, A and B being two operators. Clearly, $(AB)a$ is defined if A acts on the set of vectors spanned by $B(a)$. If A acts on a vector space that excludes $B(a)$, $(AB)a$ is undefined. In this chapter we will focus on vector spaces of finite dimension so that the elements are vectors having a finite number of components and the operators are matrices which are rectangular or square arrays of various dimensions. Thus, a rectangular matrix of dimension $(m \times n)$ acting on a space of dimension $(n \times 1)$, which represents a n -dimensional vector space comprising vectors with n components, transforms it into a m -dimensional vector space. If $m = n$, the matrix is a square matrix which keeps the vector space invariant, i.e., transforms the vectors of the same space among themselves. Of course, in some cases the elements of the square matrix A of dimension $(n \times n)$ may be such that acting on a vector space of n -dimensions A changes it into one of effectively lower dimension. For example if the only non-vanishing elements of the matrix A be A_{ij} for $i = 1, 2$ and $j = 1, 2$ the vector space of dimension n is reduced effectively to one of dimension 2. Some elementary definitions regarding matrices are as follows:

- (i) **The inverse of the matrix** A is defined as A^{-1} such that $A^{-1}A = AA^{-1} = I$, I being the identity matrix with the diagonal elements all unity and the non-diagonal elements zero. It can be determined by the formula $A^{-1} = \frac{\text{adj}(A)}{\det(A)}$, where $\text{adj}(A)$ denotes the matrix adjoint to A with the elements being the cofactors of A and $\det(A)$ is the determinant of the matrix A . The product of two matrices AB is defined to be such that $(AB)_{ij} = \sum_k A_{ik} B_{kj}$. There is another type of product of matrices known as the direct product which is defined briefly in the appendix.
- (ii) $A^\dagger$, **the hermitian adjoint of the matrix** A , is that formed by taking the complex conjugate of the transposed matrix.
Thus $A^\dagger = (A^T)^*$ [A^T , the transpose of the matrix A is formed by interchanging the rows and the columns of A]. Clearly $(AB)^\dagger = B^\dagger A^\dagger$
- (iii) **The trace of the matrix** A is defined by $\text{Tr } A = \sum_i A_{ii}$. For finite dimensional matrices $\text{Tr}(ABC) = \text{Tr}(CAB) = \text{Tr}(BCA)$. For infinite dimensional matrices this relation is not, generally valid, as is illustrated by the commutation relation $xp - px = i\hbar$, where the trace of the L.H.S is not zero obviously.
- (iv) **Function of a matrix** $f(A)$: Analogous to $f(x)$ we can have similar expressions for $f(A)$. We can have similar power series expansions given by $e^A = I + A + \frac{A^2}{2!} + \dots$, $\sin A = A - \frac{A^3}{3!} + \dots$, for example, where the equality means that the two operators on the two sides of the equation acting on a vector space give identical results. Similarly $B = A^{1/2}$ means that B operating twice on a vector space gives the same results as that of A .

3.6 MATRIX REPRESENTATION, CHANGE OF BASIS

In order to obtain a mathematical representation of the matrix operating on a vector space we have to choose a set of linearly independent basis vectors spanning the vector space. For example, if $|i\rangle$, $|j\rangle$ and $|k\rangle$ be a triad of orthonormal basis vectors covering a three dimensional vector space and a matrix operator A be such that $A|i\rangle = \cos \theta |i\rangle + \sin \theta |j\rangle$, $A|j\rangle = -\sin \theta |i\rangle + \cos \theta |j\rangle$, $A|k\rangle = |k\rangle$, the matrix elements of A are given by $A_{ii} = \langle i|A|i\rangle = \cos \theta$, $A_{ij} = \langle j|A|i\rangle = \sin \theta$.

$$A_{ik} = \langle k|A|i\rangle = 0, A_{ji} = \langle i|A|j\rangle = -\sin \theta, A_{jj} = \langle j|A|j\rangle = \cos \theta, \\ A_{ki} = \langle i|A|k\rangle = 0, A_{kj} = \langle j|A|k\rangle = 0, A_{kk} = \langle k|A|k\rangle = 1.$$

Thus the matrix A takes the form
$$\begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The matrix A is clearly a rotation matrix rotating a 3-dimensional vector about the z-axis.

In general, if $|1\rangle$, $|2\rangle$, ..., $|n\rangle$ be a set of orthonormalised vectors in n dimensions and A be the matrix operator acting on the n -dimensional space, the matrix representation of A is,

$$\begin{vmatrix} \langle 1|A|1\rangle, & \langle 2|A|1\rangle, & \dots & \langle n|A|1\rangle \\ \langle 1|A|2\rangle, & \langle 2|A|2\rangle, & \dots & \langle n|A|2\rangle \\ \dots & \dots & \dots & \dots \\ \langle 1|A|n\rangle, & \langle 2|A|n\rangle, & \dots & \langle n|A|n\rangle \end{vmatrix}.$$

Change of basis – Since the representation of a matrix depends on the set of basis vectors spanning the vector space, the change of basis will alter the representation. Let $|A\rangle$ and $|B\rangle$ be two vectors in the vector space so that $|A\rangle = \sum_i |e_i\rangle A_i$ and $|B\rangle = \sum_i |e_i\rangle B_i$, the e_i 's being the basis vectors. Let the coordinate axes be changed by the action of the operator R so that the basis vector $|e_i\rangle$ transforms to $|e_i'\rangle$ given by $|e_i'\rangle = R |e_i\rangle$. We can write, $|A\rangle = \sum_i |e_i\rangle A_i = \sum_i |e_i'\rangle A_i'$

$$\text{Or, } \sum_i |e_i\rangle A_i = R \sum_i |e_i'\rangle A_i' \text{ so that } A_i' = \sum_j A_j \langle e_i' | e_j \rangle = \sum_j (R^{-1})_{ij} A_j$$

Thus $|A'\rangle = R^{-1} |A\rangle$, $|B'\rangle = R^{-1} |B\rangle$.

If $|B\rangle = S |A\rangle$, S being the appropriate matrix relating $|A\rangle$ and $|B\rangle$, we obtain. $|B'\rangle = R^{-1} |B\rangle = R^{-1} S |A\rangle = R^{-1} S R |A'\rangle$

Thus $|B'\rangle = S' |A'\rangle$ where $S' = R^{-1} S R$.

The matrix $R^{-1} S R$ is defined as the similarity transform of S by the matrix R .

$\det S' = \det (R^{-1} S R) = \det S$, Trace $S' = \text{Tr} (R^{-1} S R) = \text{Tr} (R R^{-1} S)$ giving $\text{Tr} S' = \text{Tr} S$. Thus the determinant and the trace of a matrix S remain unaltered by the similarity transformation. This type of transformation, where the basis is changed, is called a passive transformation. In the active transformation where the vector itself is rotated, keeping the basis unaltered S' will be equal to $R S R^{-1}$.

3.7 DIFFERENT TYPES OF REPRESENTATIONS

1. Reducible and Irreducible Representations

If by a change of basis a matrix A can be brought to the form $A = \begin{vmatrix} \boxed{A_1} & & & \circ \\ & \boxed{A_2} & & \\ & & \ddots & \\ \circ & & & \boxed{A_n} \end{vmatrix}$, it is said to

be in the block-diagonal or the completely reducible form. (The circles at the right uppermost and the left lowermost ends indicate that the terms apart from those in the blocks are zero). Each block A_i acts on a subspace of dimension lower than the original space, as can be seen from the form of A . The i th subspace is left invariant by the action of A , i.e., the elements of the subspace are transformed among themselves by the action of A . The entire matrix A can be written as $A = A_1 \oplus A_2 \oplus A_3 \oplus \dots \oplus A_n$, where the sum $A_1 \oplus A_2$ denotes the direct sum of A_1 and A_2 , each operating on a subspace non-overlapping with the other. If no change of basis transforms the matrix A to such a form the matrix A is said to have the irreducible representation,

Example

(1) Let us consider a matrix M of the form

$$\begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{vmatrix}$$

Let the matrix M operate on a vector $\begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{vmatrix}$ to give $\begin{vmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{vmatrix}$.

We can therefore write $\begin{vmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{vmatrix}$.

Carrying out the calculations we can write,

$$\begin{vmatrix} x'_1 \\ x'_2 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} \text{ and } \begin{vmatrix} x'_3 \\ x'_4 \end{vmatrix} = \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix} \begin{vmatrix} x_3 \\ x_4 \end{vmatrix}.$$

Thus the result of the operation of the entire 4×4 matrix M can be viewed as the sum of the actions of two (2×2) blocks, the first one being $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ and the second one $\begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix}$ acting on two dimensional vector spaces $\begin{vmatrix} x_1 \\ x_2 \end{vmatrix}$ and $\begin{vmatrix} x_3 \\ x_4 \end{vmatrix}$ respectively.

The two two-dimensional subspaces non-overlapping with each other, are left invariant under the action of the respective blocks. Thus the entire matrix M is in a reducible representation while each lower dimensional block of dimension 2 here has an irreducible representation if by no change of basis it can be reduced any further. Below are given two examples from physical problem.

- (a) If by a change of axes the direction of the relative velocity between two inertial frames be taken to coincide with the x -axis, the operation of the 4×4 Lorent-z transformation matrix M can be viewed as the sum of three blocks (i) of dimension 2 acting on a 2- D subspace consisting of x and t and (ii) the other two of dimension 1 each acting on a one-dimensional subspace containing y and z respectively. Thus we can write,

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v \\ -\gamma v & \gamma \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}, y' = y, z' = z \cdot \begin{pmatrix} \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \end{pmatrix}.$$

- (b) The state formed by the coupling of the angular momenta J_1 and J_2 will have J values given by $|J_1 - J_2| \leq J \leq |J_1 + J_2|$.

The rotation matrix D^J acts on a vector space which can be decomposed into subspaces corresponding to the values of J equal to $|J_1 - J_2|, |J_1 - J_2| + 1, \dots, J_1 + J_2$. The entire matrix D^J

can be written as $D^J = D^{|J_1-J_2|} \oplus D \dots \oplus D^{J_1+J_2}$, where each submatrix $D^{|J_i|}$ acts on a vector space of dimension $2J_i + 1$ and keeps it invariant.

2. Diagonal Representation

A matrix can in many cases be transformed by a suitable choice of axes to the diagonal form. The diagonal elements are the eigenvalues of the matrix and we can construct an eigenvalue equation $AX = \lambda X$, X being the eigenvector corresponding to the eigenvalue λ . Thus, if D be the diagonal matrix given

$$\text{by } D = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}, \text{ to which the matrix } A \text{ has been converted and } \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix}, \begin{pmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{pmatrix} \dots \text{ are the}$$

eigenvectors for the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ respectively, we can start from the eigenvalue equation and construct the matrix equation $AS = SD$, where S is the matrix given by

$$S = \begin{pmatrix} x_{11} & \dots & x_{n1} \\ x_{12} & \dots & x_{n2} \\ \vdots & \dots & \vdots \\ x_{1n} & \dots & x_{nn} \end{pmatrix}. \text{ The eigenvalue equation is given by } Ax = \lambda x \text{ so that we get a set of } n \text{ homogeneous}$$

linear algebraic equations connecting the components of x . The condition of non-trivial solutions of the components of x gives $\det(A - \lambda I) = 0$, where I is the identity matrix. By solving this equation we obtain the magnitudes of the eigenvalues and the corresponding eigenvector by solving $Ax = \lambda x$ (See

“Solved Problems”). The new basis vectors are given by $S \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{bmatrix}, S \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{bmatrix}$ and so on.

Thus, $S^{-1}AS = D$ and the matrix has thus been converted by a similarity transformation to the diagonal form. The new basis vectors are the eigenvectors of the matrix A , while the matrix S formed by the basis vectors is the diagonalising matrix.

Let us now investigate the condition of diagonalisability of a matrix. We consider two following cases:

(a) *The eigenvalues are all different.*

Let $X_1, X_2, \dots, X_n$ be n eigenvectors corresponding to n distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. The matrix can clearly be diagonalised if $X_1, \dots, X_n$ be linearly independent of one another because only then we obtain a matrix S whose inverse exists.

To prove this we consider an equation of the form $\alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n = 0$, $\alpha_1 \dots \alpha_n$ being constants. Operating $A, A^2, \dots$ on the equation (1.6) we get,

$$\begin{aligned} \alpha_1 \lambda_1 X_1 + \alpha_2 \lambda_2 X_2 + \dots + \alpha_n \lambda_n X_n &= 0 \\ \alpha_1 \lambda_1^2 X_1 + \alpha_2 \lambda_2^2 X_2 + \dots + \alpha_n \lambda_n^2 X_n &= 0 \\ \alpha_1 \lambda_1^{n-1} X_1 + \alpha_2 \lambda_2^{n-1} X_2 + \dots + \alpha_n \lambda_n^{n-1} X_n &= 0 \end{aligned} \quad (3.7)$$

This equation can be written in the form

$$(\alpha_1 X_1 \dots \alpha_n X_n) \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \dots & \lambda_n^{n-1} \end{pmatrix} = 0$$

Since $\lambda_1 \dots \lambda_n$ are all distinct the determinant

$$\begin{vmatrix} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \dots & \lambda_n^{n-1} \end{vmatrix} \neq 0$$

Hence, the only solution of the homogeneous equations (Eq. 3.7) is that $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$. Thus the vectors $(X_1 \dots X_n)$ are linearly independent and hence the matrix can be diagonalised.

Note: If the eigenvalues are not all distinct, i.e., some of them are degenerate, the above argument falls through and the matrix is not necessarily diagonalisable. One of the examples is the matrix $\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$. The eigenvalue a is doubly degenerate. This matrix is not diagonalisable because there is only one independent eigenvector $\begin{pmatrix} x \\ 0 \end{pmatrix}$ for both the eigenvalues ($x = \text{arbitrary}$).

One indeed needs to know if there are conditions which ensure the diagonalisability of a matrix independently of the nature of the eigenvalues.

To this end it is essential to look into the condition of simultaneous diagonalisability of two operators.

- (i) Let A and B be two operators which have simultaneous eigenstates $|\psi\rangle$.

Thus we get, $A|\psi\rangle = \lambda|\psi\rangle$, $B|\psi\rangle = \mu|\psi\rangle$, λ and μ being the respective eigenvalues.

or, $(AB - BA)|\psi\rangle = (\mu\lambda - \lambda\mu)|\psi\rangle = 0$.

Since this is valid for a complete set of eigenstates $|\psi\rangle$, we get, $AB - BA = 0$.

Hence, the operators A and B commute.

- (ii) Let us assume that two operators A and B commute. Let $|\psi\rangle$ be an eigenstate of the operator A corresponding to the eigenvalue λ so that $A|\psi\rangle = \lambda|\psi\rangle$

Since $AB\psi = BA\psi = \lambda B\psi$, we can say that $B\psi$ is an eigenstate of A with the eigenvalue λ .

If the eigenvalue be non-degenerate, $B\psi$ can at best be a multiple of ψ (one eigenvalue implies one eigenstate)

or, $B\psi = \mu\psi$, μ being any constant.

Hence the commutivity of the operators A and B in this case implies simultaneous diagonalisability.

- (b) *Degenerate case*

Suppose now that the eigenvalue is degenerate and the degeneracy is 2-fold implying thereby that there are two eigenstates ψ_1 and ψ_2 of the operator A corresponding to the same eigenvalue λ . Then $A(c_1\psi_1 + c_2\psi_2) = \lambda(c_1\psi_1 + c_2\psi_2)$

Let us now see what is the condition of the state $c_1\psi_1 + c_2\psi_2$ to be an eigenstate of B .

Suppose, $B[c_1 |\psi_1\rangle + c_2 |\psi_2\rangle] = \mu [c_1 |\psi_1\rangle + c_2 |\psi_2\rangle]$, μ being the eigenvalue.

Taking the inner product with $|\psi_1\rangle$ and $|\psi_2\rangle$ respectively we obtain,

$$c_1 B_{11} + c_2 B_{12} = c_1 \mu, \text{ where } B_{11} = \langle \psi_1 | B | \psi_1 \rangle,$$

$$B_{12} = \langle \psi_1 | B | \psi_2 \rangle \text{ and } |\psi_1\rangle, |\psi_2\rangle \text{ being mutually orthogonal, } \langle \psi_1 | \psi_2 \rangle = 0 \quad (3.8) (A)$$

$$\text{and} \quad c_1 B_{21} + c_2 B_{22} = c_2 \mu \quad (3.8) (B)$$

Equation (3.8) (A) and (3.8) (B) can be rewritten as,

$$c_1 (B_{11} - \mu) + c_2 B_{12} = 0 \text{ and}$$

$$c_1 B_{21} + c_2 (B_{22} - \mu) = 0$$

$$\text{For non-trivial solutions, } \begin{vmatrix} B_{11} - \mu & B_{12} \\ B_{21} & B_{22} - \mu \end{vmatrix} = 0$$

We obtain a quadratic equation in μ which yields, in general, two distinct values of μ . Thus we have two simultaneous eigenstates $|\lambda, \mu_1\rangle$ and $|\lambda, \mu_2\rangle$. The general state $|\psi\rangle$ can be written as, $|\psi\rangle = \sum c_{\lambda\mu} |\lambda, \mu\rangle$

It can be easily generalized to the case when the degeneracy is higher than two-fold.

If the two roots of the above quadratic equation be identical, the degeneracy is not lifted. We then have to search for other operators. $C, D \dots$ which commute with A and B and go on finding out simultaneous eigenstates of the operators till the degeneracy is lifted. This leads to the idea of a complete set of commuting observables whose eigenvalues enable us to define the state uniquely.

Examples

(i) *One-dimensional free particle problem.*

The complete set of commuting observables comprises energy E and linear momentum p . There is a two-fold degeneracy with each state being characterized by energy eigenvalue E and momentum eigenvalues $p = \pm \sqrt{2mE}$. Thus the general state can be written as $\psi = c_+ (E, +) + c_- (E, -)$

where $|E, \pm\rangle$ denotes the state with momentum eigenvalue $\sqrt{2mE}$ or $-\sqrt{2mE}$ respectively and $c_{\pm}$ are the coefficients of superposition.

(ii) *The 3-D hydrogen atom problem*

The complete set of commuting observable consists of E, L^2, L_z, S^2, S_z (the symbols have their usual meanings). Thus the general state can be represented as $|\psi\rangle = \sum C_{n,l,m_l,s,m_s} |n, l, m_l, S, m_s\rangle$ where for

each n, l can have values from 0 to $n-1$, m_l from $-l$ to l , $m_s = +\frac{1}{2}$ or $-\frac{1}{2}$ respectively.

According to Panli's principle not all of these quantum numbers can have identical values.

Going back to the problem of ascertaining if any matrix A is diagonalisable we write,

$$A = \left(\frac{A + A^\dagger}{2} + \frac{A - A^\dagger}{2} \right)$$

where $A^\dagger$ is the Hermitian adjoint of A . For A to be diagonalisable, there must exist a single similarity transformation that simultaneously diagonalises $A + A^\dagger$ and $A - A^\dagger$.

According to the discussion which we have made here, the two operators must commute.

Hence $(A + A^\dagger, A - A^\dagger) = 0$, which gives $(A, A^\dagger) = 0$. Such a matrix A for which A commutes with $A^\dagger$ is called a normal matrix. So, we can state that the necessary and sufficient condition of diagonalisability of a matrix is that it be a normal matrix. Many commonly used matrices (unitary, Hermitian, etc.) are normal matrices. It is true in general that a normal matrix possesses a complete set of orthogonal vectors.

(iii) Triangular matrix, *Jordan canonical form*

It is evident that not every matrix can be diagonalised by a similarity transformation. But every matrix can be transformed to the following forms that are simpler than the original matrix, albeit not as simple as the diagonal form.

- (a) *Triangular form* – It can be proved that all matrices can be converted to the upper or the lower triangular matrices by similarity transformations. Thus for the original matrix A and a suitable matrix M we can write $M^\dagger A M = T$, where T for the upper triangular matrix is given by $T_{ij} = 0$ for $i > j$ (Fig. 3.1(a)) and for the lower triangular matrix T is given by $T_{ij} = 0$ for $i < j$ (Fig. 3.1(b)). The diagonal elements of T are the eigenvalues of the matrix.

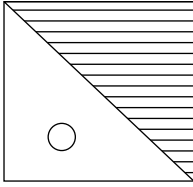


Fig. 3.1(a)

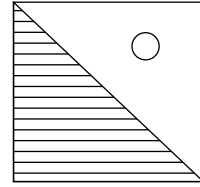


Fig. 3.1(b)

- (b) *Jordan Canonical Form* – An eigenvector u of the matrix A is of rank k if $(A - \lambda_k I)^{k-1} u \neq 0$ but $(A - \lambda_k I)^k u = 0$. Thus, we obtain a set of linearly independent generalised eigenvectors $u, (A - \lambda_k I) u, \dots, (A - \lambda_k I)^{k-1} u$.

[A diagonalisable matrix has eigenvectors of rank one]. The set of generalised eigenvectors of different ranks constitutes the Jordan basis J such that $J^{-1} A J = M$, where M is the Jordan canonical form. The matrix M can be shown to be of the form

$$M(A) = \begin{vmatrix} \square^{(1)} & 0 & \dots & 0 \\ 0 & \square^{(2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \square^{(k)} \end{vmatrix}, \text{ where } m_1, m_2, \dots, m_k \text{ are the multiplicities of the eigenvalues } \lambda_1, \dots, \lambda_k$$

respectively and

$\square^{(i)}$ is a $m_i \times m_i$ matrix given by

$$\square^{(i)} = \begin{vmatrix} \square^{(i,1)} & 0 & \dots & 0 \\ 0 & \square^{(i,2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \square^{(i,k)} \end{vmatrix} \text{ and}$$

$\square^{(i,j)}$ is the block corresponding to the i th eigenvector of rank j and is of the form

$$\square^{(i,j)} = \begin{vmatrix} \lambda_1 & \dots & 0 \\ 1 & \lambda_1 & \dots & 0 \\ 0 & 1 & \lambda_1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \lambda_1 \end{vmatrix} \quad (\text{lower triangular}) \text{ or of the form,}$$

$$\square^{(i,j)} = \begin{vmatrix} \lambda_1 & 1 & \dots & 0 \\ 0 & \lambda_1 & 1 & \dots & 0 \\ 0 & 0 & \lambda_1 & 1 & \dots & 0 \\ 0 & 0 & \dots & \lambda_1 & 1 \\ 0 & 0 & 0 & \dots & \lambda_1 \end{vmatrix} \quad \text{respectively. (upper triangular)}$$

[We do not give proofs of the statements given here. The proof can be found in standard text books on matrices or mathematical methods.]

3.8 CAYLEY-HAMILTON'S THEOREM

We will deduce a theorem that enables us to simplify expressions involving polynomials of matrices of a high order. The theorem can be stated as follows.

Every matrix satisfies its own characteristic equation, i.e., the equation satisfied by its eigenvalues.

Proof: - The characteristic polynomial of the matrix A given by $\phi(\lambda) = \det(A - \lambda I)$ is of the form $\phi(\lambda) = c_0 + c_1 \lambda + \dots + c_n \lambda^n$, $c_0, c_1, \dots, c_n$ being the coefficients in the expansion.

$$\text{Now, } (A - \lambda I)^{-1}_{ij} = \text{cofactor of } \frac{(A - \lambda I)_{ji}}{\phi(\lambda)} \quad [\phi(\lambda) \neq 0] \quad (3.9)$$

The cofactor of $(A - \lambda I)_{ji}$ is an $(n-1)$ th. order determinant which is a polynomial of degree $n-1$ if $i=j$ and of degree $(n-2)$ if $i \neq j$.

We can write Eq. (3.9) as,

$$(A - \lambda I)^{-1}_{ij} = \frac{1}{\phi(\lambda)} [t_{0ij} + t_{rij} \lambda + \dots + t_{(n-1)ij} \lambda^{n-1}], \quad (3.10)$$

where

$$t_{(n-1)ij} = 0 \text{ if } i \neq j.$$

Writing Eq. (3.10) as a matrix we obtain,

$$(A - \lambda I)^{-1} = \frac{1}{\phi(\lambda)} [T_0 + T_1 \lambda + \dots + T_{n-1} \lambda^{n-1}], \quad (3.11)$$

where

$$(T_r)_{ij} = t_{rij} \quad (r = 0, \dots, n-1)$$

Multiplying Eq. (3.11) by $\phi(\lambda) \cdot (A - \lambda I)$ we get,

$$\phi(\lambda) \cdot I = (A - \lambda I) [T_0 + T_1 \lambda + \dots + T_{n-1} \lambda^{n-1}].$$

Comparing the coefficients of 1, λ , λ^2 , ... we get,

$$c_0 I = A T_0, c_1 I = A T_1 - T_0, \dots, c_n I = -T_{n-1}.$$

Multiplying the second by A , the third by A^2 , ... and taking the sum we get,

$$c_0 + c_1 A + c_2 A^2 + \dots + c_n A^n = 0$$

Thus $\phi(A) = 0$ and hence the theorem is proved.

Example Suppose a (2×2) matrix A has two eigenvalues 1 and 2. $\phi(\lambda) = \lambda^2 - 3\lambda + 2 = 0$

$$\text{Let } f(A) = A^6 + 7A^5 + 6A^4 + 4A^3 + 2A^2 + A$$

$$= A^4 (A^2 - 3A + 2) + 10A^3 (A^2 - 3A + 2) + 34A^2 (A^2 - 3A + 2) + 86A (A^2 - 3A + 2) + 192 (A^2 - 3A + 2) + 405A - 384$$

$$= 405A - 384 \text{ [Since } A^2 - 3A + 2 = 0 \text{ by the previous theorem) and is easily determined.}$$

We discuss below a few matrices of great use in physical problems.

- (a) **Orthogonal matrix** – Such a matrix A satisfies the equality $A^T A = A A^T = I$, where A^T denotes the transpose of A . Suppose we operate the matrix A on a vector x in a finite dimensional vector space to change it to x' . Then $(x', x') = (x, A^T A x)$. For this to equal (x, x) , $A^T A = I$. The matrix A is thus orthogonal. As an example, the matrix A may represent a rotation in the vector space, where (a) it may rotate the vector itself keeping the basis vectors fixed or rotate the basis vectors keeping the vector fixed thereby changing the representation of the vector. The transformation (a) is called active and (b) passive respectively. In either case, the magnitude of the vector which is $(x, x)^{1/2}$ or the inner product of two vectors (x, y) remains unaltered, the basis being chosen to be Cartesian.

A rotation about any coordinate axis, say, the z -axis in the anticlockwise sense in the three-dimensional

vector space is represented by $\begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ which is an orthogonal matrix. Any general

rotation in 3-dimension from $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ to $\begin{pmatrix} x^1 \\ y^1 \\ z^1 \end{pmatrix}$ can be executed by three successive rotations about three

coordinate axes and hence can be represented by the product of three orthogonal transformation. This general rotation is, thus, represented by an equivalent orthogonal matrix which is the product of three orthogonal matrices parametrized by three angles of rotation called the Euler angles. The choice of the Euler angles is not unique and depends on the order of rotations about three respective coordinate axes.

This general transformation is important in describing the transformation from the space-fixed set of axes to the body-fixed set in the rigid body dynamics in classical as well as quantum mechanics.

For orthogonal transformation $\det(A^T A) = \det(A A^T) = 1$

or, $(\det A^T)(\det A) = |\det A|^2 = 1$, which gives, $\det A = \pm 1$.

The transformations with $\det A = +1$ are called proper orthogonal transformations where the handedness of the system of axes is preserved (for a pure rotation or a transformation by two successive reflections), while those with $\det A = -1$ do not preserve the handedness of the axes (for an odd number of reflections).

- (b) *Unitary matrix* – The unitary matrix U satisfies the criterion $U^\dagger U = U U^\dagger = I$, where $U^\dagger$ is the Hermitian adjoint of U . Such a matrix can be said to be the counterpart of the orthogonal matrix in the complex vector space. If a vector x changes to x' under the operation of the matrix U , so that $x' = Ux$, we have $(x', x') = (x, U^\dagger U x)$. Clearly $(x', x') = (x, x)$ if the matrix U be unitary, i.e., $U^\dagger U = U U^\dagger = I$. Thus like an orthogonal matrix, the magnitude of a vector or the inner product of two vectors remains unaltered also here.

If the matrix U has an eigenvalue λ corresponding to the eigenvector X we have, $UX = \lambda X$, (3.12)

Taking the Hermitian adjoint of Eq. (3.12) and multiplying the two, we get, $X^\dagger U^\dagger UX = X^\dagger \lambda^* \lambda X$.

Since $U^\dagger U = I$ we get, $|\lambda|^2 = 1$ and hence the eigenvalues of a unitary matrix are unimodular. Since the determinant of a unitary matrix is the product of the eigenvalues, the unitary matrix has the modulus of the determinant also equal to unity. These properties are also shared by the orthogonal matrices. The unitary matrices play important roles in quantum mechanics in the transformation from one representation to another. For example, in the transition from the coordinate representation, where the basis vectors are the coordinate eigenstates to the momentum representation where the basis vectors are the

momentum eigenstates, the elements of the transformation matrix are $\exp\left(\frac{i \bar{p} \cdot \bar{r}}{\hbar}\right)$ and vary continuously as functions of $\bar{r}$ or $\bar{p}$. We can write here,

$$\psi(\bar{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int \phi(\bar{p}) \exp\left(\frac{i \bar{p} \cdot \bar{r}}{\hbar}\right) d\bar{p}$$

The norm of the state vector, which is a measure of the probability of the system in the coordinate or the momentum space remains unchanged, i.e., $\int |\psi(\bar{r})|^2 d\bar{r} = \int |\phi(\bar{p})|^2 d\bar{p}$, an equality that has to be satisfied from physical considerations.

- (c) *Hermitian Operator* – An operator H is said to be hermitian if $(\psi, H\phi) = (H\psi, \phi)$, where ψ and ϕ are two vectors in a vector space on which H acts. In a finite dimensional vector space the above condition becomes $\psi^\dagger H\phi = (H\psi)^\dagger \phi$. Thus for a hermitian matrix H , $H^\dagger = H$. In a function space, briefly introduced in Section 3.10, the condition can be written as,

$$\int_a^b \psi^* H\phi dx = \int_a^b (H\psi)^* \phi dx, \text{ where the functions are considered to be functions of a single variable } x \text{ in}$$

the interval (a, b) . The definitions can be extended to functions of any number of variables.

The condition of hermiticity of an operator is intricately related to the boundary conditions we impose on the functions. This will be illustrated below by examples.

Examples Suppose we want to test if the operator $-i\hbar \frac{\partial}{\partial x}$ that is the x -component of the momentum operator in the Schrodinger representation is hermitian. Putting $H = -i\hbar \frac{\partial}{\partial x}$ we get,

$$\begin{aligned} \int_v \psi^* \left(-i\hbar \frac{\partial}{\partial x} \right) \phi d\bar{r} &= \int -i\hbar \left[\frac{\partial}{\partial x} (\psi^* \phi) - \frac{\partial \psi^*}{\partial x} \phi \right] d\bar{r} \\ &= -i\hbar \int_v [\psi^* \phi]_{x_1}^{x_2} dy dz + i\hbar \int \frac{\partial \psi^*}{\partial x} \phi d\bar{r} \end{aligned}$$

Clearly $\int_v \psi^* \left(-i\hbar \frac{\partial}{\partial x} \right) \phi d\bar{r} = \int_v \left(-i\hbar \frac{\partial \psi}{\partial x} \right)^* \phi d\bar{r}$, i.e., the operator $-i\hbar \frac{\partial}{\partial x}$ is hermitian in the space of functions ψ, ϕ if the first term on the *r. h. s.* vanishes. This can happen for example, if (ψ, ϕ) vanishes at the end points x_1, x_2 of the interval of x or attains equal values at the boundary points.

(The first term can vanish also by imposing other types of boundary conditions). It can be easily verified that by adjusting appropriate boundary conditions the operators $i\hbar \frac{\partial}{\partial t}$ or ∇^2 can be made hermitian.

Let ψ and ϕ be two eigenstates of the operator H belonging to eigenvalues λ and μ respectively.

$$H\phi = \mu\phi, H\psi = \lambda\psi.$$

Due to hermiticity $(\psi, H\phi) = (H\psi, \phi)$

or, $(\psi, \mu\phi) = (\lambda\psi, \phi)$

or, $(\lambda^* - \mu)(\psi, \phi) = 0$

If μ and λ are unequal, $(\psi, \phi) = 0$.

Hence ψ and ϕ are orthogonal to each other. Thus the eigenvectors of a hermitian operator belonging to different eigenvalues are orthogonal to each other.

Putting $\psi = \phi$, $\lambda = \mu$ and hence $(\lambda - \lambda^*)(\psi, \psi) = 0$.

Since $(\psi, \psi) \neq 0$ [$(\psi, \psi) = |\psi_1|^2 + |\psi_2|^2 + \dots + |\psi_n|^2$ for a n -dimensional vector space and $= \int \psi^* \psi dx dy dz$ in a space of functions] $\lambda = \lambda^*$. Hence, all the eigenvalues of a hermitian operator are real.

Note: It is to be noted that the non-hermitian matrices may also have all the eigenvalues real. As a simple example we consider the non-hermitian matrix $\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$, which has got the eigenvalues determined by the

equation $\det \begin{vmatrix} \lambda - 1 & -2 \\ 1 & \lambda - 4 \end{vmatrix} = 0$, or, $\lambda^2 - 5\lambda + 6 = 0$, which gives $\lambda = 2, 3$, which are real. But if the eigenvalues

are real and the eigenvectors are orthogonal the matrix must be Hermitian. (You can check that the eigenvectors are not orthogonal in this case.)

Let U be the diagonalising matrix for a Hermitian matrix H so that we can write $U^{-1}HU = D$, D being the diagonalised matrix with the diagonal elements all real. Taking the Hermitian adjoint we get,

$$(U^{-1}HU)^\dagger = U^\dagger H^\dagger (U^{-1})^\dagger = D [D = D^\dagger \text{ here}]$$

or, $U^{-1}HU = U^\dagger H (U^{-1})^\dagger$, since $H = H^\dagger$.

Or, $U^{-1} = U^\dagger$, which is the necessary and sufficient condition for the matrix U to be unitary. So, a hermitian matrix can be diagonalised by a similarity transformation with a unitary matrix. A unitary matrix also can be seen to have another unitary matrix as the diagonalising matrix.

Quadratic Form

An expression of the form $x_i^\dagger H_{ij} x_j$ is called the hermitian quadratic form. In matrix form this reads $x^\dagger Hx$. If we now subject x to a unitary transformation so that $x = Ux'$ we get,

$x^\dagger Hx = (x')^\dagger U^\dagger H U x'$. If the unitary matrix U be adjusted to be the same as the diagonalising matrix for the hermitian matrix H we get, $U^{-1} H U = U^\dagger H U = D$, D being the diagonalising matrix with real diagonal elements.

or, $x^\dagger H x = (x')^\dagger D x' = \sum_1 \lambda_j |x'_j|^2$, (λ_j 's are the eigenvalues), which involves the squares of $|x'_j|$ only.

If in $x^\dagger H x$ we replace H by a real symmetric matrix S , say, the diagonalising matrix will be orthogonal in this case. So that we can write $x^\dagger S x = \sum_j \lambda_j x_j^2$, where the x 's are all real.

The reduction of a quadratic form to a sum of squares finds application (i) in the theory of small oscillations where by a similarity transformation with a single orthogonal transformation (i.e., a rotation of coordinate axes) the kinetic energy and the potential energy, both being quadratic forms in the generalized velocity and the generalized coordinates respectively, can be transformed to the sum of the respective squares, (ii) in the expression for the kinetic energy of a rigid body that can be reduced to the form given by $\frac{1}{2} \sum_i I_i \omega_i^2$, the I_i 's, and the ω_i 's being the moments of inertia and the angular velocities about the principal axes.

[It is noteworthy that any quadratic form $\sum_{ij} A_{ij} \cdot x_i \cdot x_j$, can be written as $\sum_{ij} \left[\frac{1}{2} (A_{ij} + A_{ji}) + \frac{1}{2} (A_{ij} - A_{ji}) \right] x_i x_j$. $A_{ij} - A_{ji}$ being anti-symmetric with respect to i and j , the expression $\sum_{ij} A_{ij} \cdot x_i \cdot x_j$, can be reduced to $\sum_{ij} A_{ij}^s x_i x_j$, where A_{ij}^s , given by $\frac{1}{2} (A_{ij} + A_{ji})$ is symmetric with respect to i and j , because the other part involving the anti-symmetric part $\frac{1}{2} (A_{ij} - A_{ji})$ vanishes on taking the sum over i and j . Hence any quadratic form $\sum_{i,j} A_{ij} x_i x_j$, can, irrespective of the nature of A_{ij} , be transformed to a sum of squares, since a real symmetric matrix is diagonalisable.

3.9 RANK, LINEAR ALGEBRAIC EQUATIONS AND MATRICES

The rank of a matrix is the dimension of its highest order minor with non-vanishing determinant. For a $(m \times n)$ matrix the rank r is clearly less than or equal to the smaller of m and n . The knowledge of the rank of a matrix enables us to ascertain if a set of linear algebraic equations homogeneous or inhomogeneous has a consistent set of solutions. Below we state the conditions without giving any proof which can be found in any standard book on matrices and determinants.

- (a) *Homogeneous linear equations.* Let us consider a set of homogeneous equations involving n unknowns of the form,

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n &= 0 \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n &= 0 \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots & \\ a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n &= 0 \end{aligned} \quad (3.12)$$

The set possesses a trivial solution $x_1 = \dots = x_n = 0$

The non-trivial solution exists if the rank r of the matrix formed out of the coefficients of the unknowns in the equation be less than the number of unknowns n . The general solution can then be written as a linear combination of these $(n - r)$ solutions which may be said to constitute the basis of the solution space of the homogeneous set. It is therefore non-unique.

- (b) *Inhomogeneous linear equations.* Let us consider a set of in homogeneous linear equations of the form

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

The set of equations has a consistent solution if the rank r of the matrix of the coefficients $\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{22} & \dots & a_{2n} \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$

and that of the augmented matrix $\begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} & b_n \end{pmatrix}$ be the same.

If the rank r be the same as the number of unknowns the set possesses a unique solutions. The general solution of the set of linear equations is a superposition of the solutions of the inhomogeneous and the homogeneous equation. (If the rank r = number of unknowns, the homogenous equation has a trivial solution with all the variables equal to zero).

3.10 INFINITE DIMENSIONAL VECTOR SPACE

We conclude this chapter with a brief introduction to the vector space of an infinite number of dimensions, the function space, for example. In many areas we come across an infinite dimensional function space, i.e., the space covered by an infinite number of functions $\psi_n(x)$, where the $\psi_n(x)$'s represent the eigenfunctions, in general, infinite in number, that arise as a result of the boundary conditions imposed on any problem. Here for each n the function $\psi_n(x)$ can be regarded as the vector having a continuously infinite number of components because x can take up a continuously infinite number of values in any interval, while n can also have an infinite number of values. The definitions of the inner product, orthogonality, completeness, linear operators follow in a manner analogous to the case of finite dimensional vector space outlined briefly as follows.

- (a) *Inner Product* – The inner product of two functions $\psi_n(x)$ and $\psi_m(x)$ is defined as (ψ_n, ψ_m)

$$= \int_a^b \psi_n^*(x) \psi_m(x) w(x) dx \text{ where } x \text{ is defined in the interval } (a, b) \text{ and } w(x) \text{ represents the weight factor}$$

which may be taken to be unity by a suitable redefinition of the functions.

- (b) *Orthogonality* – Two functions $\psi_n(x)$ and $\psi_m(x)$ are orthogonal if $(\psi_n, \psi_m) = 0$ in the interval (a, b) of x .

- (c) *Completeness* – The functions $\psi_n(x)$ are said to be complete if the infinite sum $\sum_n \psi_n(x)$ converges to

function $\psi(x)$ which belongs to the space of functions $\psi_n(x)$. The problem of convergence and hence completeness is more involved in the infinite dimensional space than in a finite dimensional vector space where the completeness is ensured.

- (d) *Linear Operators* – The idea of linear operator carries over analogously to the infinite dimensional function space. Thus if ψ and ϕ be two functions connected by a linear operator K we can write $\psi = K\phi$, i.e., in a more explicit form as

$$\psi(x) = \int_a^b K(x, x') \phi(x') \cdot d'x', \text{ where } (a, b) \text{ denotes the interval of definition of } x.$$

In the next section we give a very brief introduction to the application of the vector space formalism to quantum mechanics

3.11 APPLICATION

Dirac utilised the linear vector space formalism in his formulation of quantum mechanics. He introduced the state vector $|\psi\rangle$ which he called the ket and which represents the state of a system in a coordinate-independent way. When a particular basis is chosen we get the corresponding representation, $\psi(\bar{r})$, for example in the coordinate representation defined to be the wavefunction, which is in general complex. Corresponding to every ket $|\psi\rangle$ there exists a vector $\langle\psi|$ called the bra, which spans the dual space such that $\langle\psi|\phi\rangle$ is a scalar.

Thus if $|\psi\rangle$ be represented by $\begin{vmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{vmatrix}$, the corresponding bra vector $\langle\psi|$ is given by $|\psi_1^* \psi_2^* \dots|$,

Thus $\langle\psi|\phi\rangle$ is given by $\psi_1^* \phi_1 + \psi_2^* \phi_2 + \dots$.

The operators that act on the state vectors in a complex Hilbert space are linear, the linearity of the operators allowing one to apply the principle of superposition that plays a vital role in the formulation of quantum mechanics. Moreover, the operators representing dynamical variables in quantum mechanics are all hermitian because the variables are all real and so are eigenvalues of a hermitian operator. If $A|\psi\rangle$ be repre-

sented by $\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix}$ then $\langle\psi|A$ is given by $(\psi_1^* \psi_2^* \dots \psi_n^*) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$.

SOLVED PROBLEMS

Problem 4 Prove that $\det(e^A) = \exp(\text{Tr } A)$

Solution: The eigenvalue equation is $\det(\lambda I - A) = 0$. If A be a $(n \times n)$ matrix, this equation is a n th. degree algebraic equation in λ , the determinant of the matrix A multiplied by $(-1)^n$ being the term independent of λ . Thus from the theory of equation $\det(A) = \lambda_1 \dots \lambda_n$. Since e^A has eigenvalues equal to $e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n}$, we get by following the same logic that $\det(e^A) = e^{\lambda_1} \cdot e^{\lambda_2} \dots e^{\lambda_n} = \exp(\lambda_1 + \dots + \lambda_n) = \exp(\text{Tr } A)$.

Problem 5

A matrix A is given by $\begin{pmatrix} 2 & 4 \\ 5 & 6 \end{pmatrix}$ in the standard basis $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Find A with respect to the basis $\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ -7 \end{pmatrix}$.

Solution: Let $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be the matrix that transforms $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to $\begin{pmatrix} 3 \\ -7 \end{pmatrix}$ respectively

$$\text{or, } \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ giving } a = 1, c = 2$$

$$\text{Similarly } \begin{pmatrix} 3 \\ -7 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ giving } b = 3, d = -7.$$

so that the transformation matrix is given by $\begin{pmatrix} 1 & 3 \\ -2 & -7 \end{pmatrix} = S$ (say).

The matrix A becomes transformed to $S^{-1}AS$, and is given by $\begin{pmatrix} -63 & -235 \\ 19 & 71 \end{pmatrix}$.

Problem 6

$$(A) \text{ Diagonalise } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad (B) \begin{vmatrix} \alpha & i\beta \\ -i\beta & \alpha \end{vmatrix}$$

Solution: The eigenvalue equation is $\det \begin{vmatrix} \lambda - 1 & 0 & 0 \\ 0 & \lambda & -1 \\ 0 & -1 & \lambda \end{vmatrix} = 0$

$$\text{or, } (\lambda - 1)(\lambda^2 - 1) = 0 \text{ giving } \lambda = 1, 1, -1.$$

For $\lambda = -1$ the eigenvalue equation is,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = - \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\text{or, } x = -x \text{ giving } x = 0 \text{ and } y + z = 0$$

$$\text{The normalised eigenvector} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

For $\lambda = 1$, the eigenvalue equation is,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ giving } y = z \text{ and } x \text{ arbitrary.}$$

We take one eigenvector = $\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

This is clearly orthogonal to the eigenvalue for $\lambda = -1$
For the second linearly independent eigenvector.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \text{ we choose it to be orthogonal to } \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

The condition is, $y' + z' = 0$.

Again since $y' = z'$ we get, $y' = z' = 0$.

Thus the normalized eigenvector is $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

The diagonalising matrix is thus

$$S = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \text{ so that } S^{-1}AS = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Note: Clearly, the choice of the eigenvectors for $\lambda = 1$ is not unique since there can be an infinite number of solutions satisfying the eigenvalue condition. But this is not unexpected because the degenerate eigenvectors corresponding to $\lambda = 1$ can be any two perpendicular eigenvectors (if they are chosen to be orthogonal to each other) in the invariant degenerate subspace which is a plane perpendicular to the direction of the eigenvector for $\lambda = -1$.

(B) The eigenvalue equation is $\begin{vmatrix} \lambda - \alpha & -i\beta \\ i\beta & \lambda - \alpha \end{vmatrix} = 0$, or, $(\lambda - \alpha)^2 = \beta^2$ giving $\lambda = \alpha \pm \beta$. For $\lambda = \alpha + \beta$, the eigenvector is given by $\begin{vmatrix} \alpha & i\beta \\ -i\beta & \alpha \end{vmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (\alpha + \beta) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ yielding $x_1 = ix_2$ or the eigenvector as $x_1 \begin{pmatrix} 1 \\ -i \end{pmatrix}$. The normalisation gives, $|x_1| = \frac{1}{\sqrt{2}}$ so that the normalised eigenvector is $\frac{1}{\sqrt{2}} e^{i\phi_1} \begin{pmatrix} 1 \\ -i \end{pmatrix}$. Similarly for $\lambda = \alpha - \beta$, the normalised eigenvector is $\frac{1}{\sqrt{2}} e^{i\phi_2} \begin{pmatrix} 1 \\ i \end{pmatrix}$, (ϕ_1, ϕ_2 are arbitrary).

Problem 7

Solve $x_1 + x_2 - x_3 - x_4 = 3, x_1 - x_2 + x_3 - x_4 = 1$.

Solution: Putting $x_3 = x_4 = 0$, one particular solution of the inhomogeneous equation is obtained as, $x_1 = 2, x_2 = 1, x_3 = x_4 = 0$. The rank of the coefficient matrix of the homogeneous equation is 2 and hence the

homogeneous equation has two independent solutions chosen as $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ respectively. The general solution of the inhomogeneous equation is given by

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \alpha \text{ and } \beta \text{ being two arbitrary constants. This provides us with the general}$$

solution because any other solution obtained by choosing a different set of values of x_1, x_2, x_3, x_4 for the solution of the inhomogeneous equation can be seen to be contained in the above solution for a new set of values of α and β . (the difference of any two apparently different solutions is a solution of the homogeneous equation).

Problem 8

A 3×3 matrix satisfies the equation $M^2 - 3M + 2I = 0$. (I is the identity matrix).

Find out the determinant of the matrix if its trace is 6.

Solution: By Caley Hamilton's theorem, the eigenvalues satisfy the equation $\lambda^2 - 3\lambda + 2 = 0$ so that the two eigenvalues are 1, 2. Since the trace is 6, the third eigenvalue is 3. Hence the determinant which is the product of the eigenvalues is 6.

Problem 9

The matrix A is given by

$$A = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{vmatrix}$$

Find out the eigenvalues of $A^3 + 4A^2 + 3A + 2I$

Solution: The eigenvalues of the matrix A are determined by solving the equation $\det$

$$\begin{vmatrix} \lambda - 1 & -2 & -3 \\ 0 & \lambda - 4 & -7 \\ 0 & 0 & \lambda - 5 \end{vmatrix} = 0$$

or $(\lambda - 1)(\lambda - 4)(\lambda - 5) = 0$ giving $\lambda = 1, 4, 5$,

The matrices A^3, A^2 will also be of the same form as A (i.e., upper triangular) with the eigenvalues $(1^2, 4^2, 5^2)$ and $(1^3, 4^3, 5^3)$ respectively. The eigenvalues of $A^3 + 4A^2 + 3A + 2I$ are thus $(1^3 + 4 \cdot 1^2 + 3 \cdot 1 + 2, 4^3 + 4 \cdot 4^2 + 3 \cdot 4 + 2, 5^3 + 4 \cdot 5^2 + 3 \cdot 5 + 2)$, i.e., $(10, 142, 242)$.

Problem 10

If a unitary matrix U be written as $U = A + iB$, A and B being Hermitian and having non-degenerate eigenvalues, find out the conditions satisfied by A and B .

Solution: Since U is unitary, $(U, U^\dagger) = 0$.

or, $(A + iB, A - iB) = 0$, A and B being Hermitian

or, $(A, B) = 0$ and hence A and B commute.

Since the matrices A and B have non-degenerate eigenvalues; they have simultaneous eigenstates. (Refer to section 3.7).

Since $UU^\dagger = U^\dagger U = I$ we obtain, $(A + iB)(A - iB) = (A - iB)(A + iB) = I$ yielding, $A^2 + B^2 = I$.

Problem 11

Find out the lengths and the orientations of the axes of the ellipse

$$15x^2 + 10y^2 + 10\sqrt{2}xy = 1.$$

Solution: The equation of the ellipse referred to the coordinate axes as the axes is $\lambda_1 x'^2 + \lambda_2 y'^2 = 1$,

where λ_1, λ_2 are the eigenvalues of the matrix $\begin{pmatrix} 15 & 5\sqrt{2} \\ 5\sqrt{2} & 10 \end{pmatrix}$, i.e., 5 and 20 respectively.

Thus we get $5x'^2 + 20y'^2 = 1$, the lengths of the major and the minor axes being given by $\frac{1}{\sqrt{5}}$ and $\frac{1}{2\sqrt{5}}$ respectively.

The eigenvectors of the matrix are given by

$$\begin{pmatrix} 15 & 5\sqrt{2} \\ 5\sqrt{2} & 10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 5 \begin{pmatrix} x \\ y \end{pmatrix} \text{ giving } \frac{y}{x} = \sqrt{2}$$

and
$$\begin{pmatrix} 15 & 5\sqrt{2} \\ 5\sqrt{2} & 10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 20 \begin{pmatrix} x \\ y \end{pmatrix} \text{ giving } \frac{y}{x} = \frac{1}{\sqrt{2}}$$

The major and the minor axes are thus inclined at angles 45° and 135° respectively to the x -axes.

SUMMARY

- A vector space comprises a set of element which satisfy certain properties enlisted in the text. Apart from the ordinary vectors represented by directed line segments there are other entities, the set of even integers, complex numbers, solutions of homogeneous differential equations etc. for example, that also are members of the vector space.
- The inner product of any two members a and b of the vector space, denoted by (a, b) is a scalar satisfying certain properties discussed in the text.
- A set of vectors is said to be linearly independent if the relation $C_1 \bar{A}_1 + C_2 \bar{A}_2 + \dots + C_n \bar{A}_n = 0$ admits of the only solution $C_1 = C_2 = \dots = C_n = 0$. An identical definition follows also for the set of functions. The condition for linear independence (or dependence) of the set of vectors or of functions is discussed in the text.
- The linearly independent vectors form the basis of the vector space. The maximum number of linearly independent vectors is called the dimension of the space. In a n -dimensional vector space the n linearly independent basis vectors form a complete set so that any vector belonging to the vector space can be expanded in terms of them.
- A set of mutually orthogonal vectors (or functions) can be constructed out of a set of linearly independent vectors (or functions) by Gram-Schmidt orthogonalisation method. These complete sets of orthogonal vectors (or functions) are used in many problems of physics.
- Matrices are arrays of numbers or functions acting on the elements of a vector space. They are linear operators and may or may not change the dimension of the vector space according as the array is rectangular or square.
- The matrix elements determining its representation depend on the basis vectors of the underlying vector space. If we change the basis vector (by rotation of the axes say) the matrix representation changes. For example, if A be the original matrix and S the rotation operator. A change to A' after rotation given by $A' = S^{-1} A S$. This is called similarity transformation and keeps the trace and the determinant of the matrix unaltered.
- By an appropriate choice of basis vectors a matrix can be chosen to have different representations a few of which we will discuss below.

(a) *Reducible and Irreducible representation*

If a similarity transformation converts the matrix to a block-diagonal form where the matrix is divided into blocks arranged along the diagonal, it is said to be in the completely reducible form. The blocks act on non-overlapping subspaces of the original vector space and leave them invariant. If no similarity transformation can reduce the matrix to such a form so that the matrix operates on the entire vector space it is said to be irreducible. An example from quantum mechanics illustrating these features has been given in the text.

- (b) *Diagonal representation* – The eigenvalue equation of a matrix A is given by $Ax = \lambda x$, λ being the eigenvalue and x the eigenvector. The diagonal representation corresponds to a similarity transformation to a diagonal form $A'_{ij} = (S^{-1} A S)_{ij} = \lambda_i \delta_{ij}$, where S is the diagonalising matrix formed by the eigenvectors. The new basis vectors are the eigenvectors of the matrix A . The condition of diagonalisability is that the number of independent eigenvectors is equal to that of

the eigenvalues. It can be shown that for distinct eigenvalues diagonalisation is always possible. But for degenerate case (some of the eigenvalues are equal) a matrix can be diagonalised if it is a normal matrix, where the commutator $(A, A^\dagger) = 0$.

- (c) *Canonical Forms* – Even when diagonalisation is not possible a matrix can, by similarity transformation, be reduced to the triangular (upper or lower) or the Jordan canonical form, as described in the text.
- It is often possible to reduce a polynomial containing high powers of the matrix to a much simpler form by using Cayley-Hamilton's theorem that states that a matrix equation satisfies the same equation as that satisfied by the eigenvalues. This has been explained in the text with an example.
- A few important matrices are discussed briefly below.
 - (a) *Orthogonal Matrix* – Such a matrix satisfies the equality $AA^T = A^T A = I$. The determinant is ± 1 . Rotation ($\det. = 1$) and reflection matrices ($\det. = -1$) belong to this category.
 - (b) *Hermitian Matrix* – For a hermitian matrix A , $A^\dagger = A$. This has all the eigenvalues real. The dynamical variables in quantum mechanics are represented by hermitian matrices.
 - (c) *Unitary Matrix* – A unitary matrix U satisfies the equality $U^\dagger U = U U^\dagger = I$. The eigenvalues as well as the determinant of this matrix are unimodular. The unitary matrices play important roles in the change of representation in quantum mechanics.

All the matrices above are normal matrices and, hence, can be diagonalised by similarity transformation with unitary matrices.

- The study of matrices enlightens us about the nature of solution of a set of linear algebraic equations, that has been briefly discussed in the text. The chapter ends with a brief introduction to the idea of infinite dimensional vector space (function space for example).

MULTIPLE CHOICE QUESTIONS

1. Which of the statements is (are) true?
The three vectors $\vec{i} + 2\vec{j}$, $2\vec{i} - 3\vec{j}$, $4\vec{i} + 5\vec{j}$ are
 - (a) linearly independent
 - (b) linearly dependent
 - (c) linearly independent if any two are taken together
 - (d) capable of forming a basis
2. A set of $(n - 1)$ vector in the n -dimensional vector space is
 - (a) linearly dependent
 - (b) linearly independent
 - (c) linearly dependent or independent depending on the part of the vector space from which the set has been taken
 - (d) linearly independent or dependent depending on n
3. You are given two vector spaces, one being finite dimensional and the other infinite dimensional.
 - (a) Both are complete
 - (b) Both are incomplete
 - (c) The first one is complete while the second one may be incomplete
 - (d) The first one is incomplete while the second one is complete
4. Of the two operations, namely, differentiation and complex conjugation
 - (a) both are linear

- (b) only the first one is linear
(c) only the second one is linear
(d) both are nonlinear
5. The boundedness of an operator
(a) depends on the members of the space on which the operator operates
(b) depends on whether the space is finite or infinite dimensional
(c) depends on the basis chosen
(d) depends only on the operator
(Hint: An operator A is bounded if the magnitude of $A|x\rangle$ is finite for all $|x\rangle$).
6. The eigenvalues of a matrix are real. The matrix
(a) is Hermitian
(b) not necessarily Hermitian
(c) hermitian or unitary
(d) may be of different types depending on the dimension of the space
7. The trace and the determinant of a 3×3 matrix are both 6, while one of the eigenvalues is 2. The others are
(a) 1, 6 (b) 1, 3 (c) 2, 2 (d) None of these
8. The eigenvalues of a matrix are i , $2i$ and $3i$ respectively. The matrix is
(a) unitary (b) anti-Hermitian (c) antiunitary (d) orthogonal
9. Two matrices A and B are related by $B = s^{-1}AS$. Which of the statements is not correct?
(a) They have the same determinant
(b) They have the same trace
(c) They have the same eigenvectors
(d) They have the same eigenvalues
10. The eigenvalues of a matrix are $1, \frac{1}{\sqrt{2}}(1+i)$ and $\frac{1}{\sqrt{2}}(1-i)$ respectively. The matrix is
(a) Hermitian (b) unitary (c) orthogonal (d) anti-Hermitian
11. The four vectors given by $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$
(a) are independent
(b) are orthonormal
(c) are linearly dependent
(d) cannot form a basis in four dimensions
12. The matrix $\begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}$
(a) is diagonalisable with eigenvalues a

- (b) is non-diagonalisable
- (c) has two independent eigenvectors
- (d) has rank equal to 1

13. The matrix $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$

- (a) has eigenvalues $-1, -1, 2$ and is diagonalisable
 - (b) has eigenvalues $-1, -1, 2$ and is non-diagonalisable
 - (c) has eigenvalues $1, 1, -2$ and is diagonalisable
 - (d) has eigenvalues $1, 1, -2$ and is non-diagonalisable
14. Which of the statements is/are false?

If a matrix A has the form $\begin{pmatrix} T & T & T \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{pmatrix}$, the rank is

- (a) 3 always
 - (b) 2 if $\alpha = \beta = \gamma$
 - (c) 1 if $\alpha = \beta = \gamma$
 - (d) may be 3 if $\alpha \neq \beta \neq \gamma$
15. A system of homogeneous linear algebraic equations involving 4 variables x_1, x_2, x_3, x_4 has the rank of the matrix formed by the coefficients equal to 1. The number of independent solutions is
- (a) 2
 - (b) 3
 - (c) 1
 - (d) 4
16. The relation $\text{Trace}(AB) = \text{Trace}(BA)$ is valid always if
- (a) A and B are finite dimensional
 - (b) A and B are infinite dimensional
 - (c) one is finite dimensional and the other infinite dimensional
 - (d) A and B be any pair of arbitrary matrices
- [Hint:** Consider the commutation relation $(x, p) = i\hbar$, where x, p are infinite dimensional matrices].

Answers to MCQs

1. (b)	2. (b)	3. (c)	4. (b)	5. (a)	6. (b)	7. (b)
8. (b)	9. (c)	10. (b)	11. (a)	12. (b)	13. (a)	14. (a),(b)
15. (b)	16. (a)					

EXERCISE

1. If a matrix A is represented by $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with respect to the basis vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ find out A when the basis vectors are chosen to be $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ respectively.

2. Diagonalise the matrices $\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

For the first one, transform the vector $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ in the standard basis to the new representation in which the matrix is diagonal. The standard basis is formed by $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

3. Prove that finite rotations about two different axes (take for example, rotations about the x and y -axes through angles θ and ϕ respectively) do not commute.

[**Hint:** Find out $[R_x(\theta) R_y(\phi) - R_y(\phi) R_x(\theta)]$, where $R_x(\theta)$ and $R_y(\phi)$ represent the respective rotation operators].

4. Prove that for an antihermitian matrix A (iA is hermitian) the eigenvalues are purely imaginary and the eigenvectors corresponding to different eigenvalues are orthogonal.

5. Find out the projection operator in the 3D-vector space on the set $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ respectively.

[**Hint:** Projection operator corresponding to an axis yields the resolved part of a vector along the axis when it acts on a vector].

6. Prove that $\text{tr} [|a\rangle \langle b|] = \langle b|a\rangle$
7. (a) Find out the condition that (i) the product of two unitary matrices is unitary, (ii) the product of two hermitian matrices is hermitian.
- (b) Prove that the eigenvectors of a unitary matrix corresponding to different eigenvalues are orthogonal.
8. Prove that the diagonalising matrix of a unitary matrix is unitary.

9. For the matrix $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ -1 & 1 & 3 \end{pmatrix}$ find out

(a) $\sum_{i=1}^3 \lambda_i$,

(b) $\sum_{i=1}^3 \lambda_i^2$,

(c) $\sum_{i=1}^3 \lambda_i^3$,

(d) $\lambda_1 \lambda_2 \lambda_3$ without explicitly determining the λ_i 's.

10. (a) Show that $(1, t, t^2)$ can be used as a basis for a linear vector space spanned by the real polynomials of order 2. Find out the matrix representation of the operator $\sum_{n=0}^2 \frac{1}{n!} \frac{d^n}{dt^n}$ with respect to the above basis.

- (b) If $A x_n(t) = \frac{dx_n(t)}{dt}$, $x_n(t)$ denoting a polynomials of degree less than or equal to 3 find out the matrix representation of A .
11. (a) For what values of α , the following matrices are invertible. (i) $\begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & \alpha \\ 1 & \alpha \end{pmatrix}$
- (b) Prove that the trace of a rotation matrix A is related to the angle of rotation θ by the relation $2 \cos \theta = 1 + \text{Tr}(A)$. Fine out also the axis of rotation.
 [Hint: By a similarity transformation, the axis of rotation can be made to coincide with any coordinate axis. The answer to the first part is then obvious. For the second part start from the fact that the axis of rotation remains unchanged on rotation].
- (c) If A is an idempotent matrix (i.e., $A^2 = A$) and is not the unit matrix show that A is singular.
- (d) If A is a unitary matrix and $B = AP$, P being non-singular show that PB^{-1} is unitary.
12. (a) $A(3 \times 3)$ matrix M satisfies the equation $M^2 - 6M + 9 = 0$. If the trace of the matrix be 10 find out the determinant.
 [Hint: Apply Cayley-Hamilton's theorem].
- (b) Find out the matrix which rotates by 45° about the x -axis and expands the y and z coordinates by a factor of $\sqrt{3}$.
13. (a) Prove that an odd-dimensional matrix with real elements has at least one real eigenvalue.
- (b) What can you say about the nature of A if $A^2 = A$?
14. For what values of λ will the following equations give a nontrivial solution?

$$\lambda x_1 - 2x_2 + x_3 = 0$$

$$x_1 - 3x_2 + x_3 = 0$$

$$x_1 - \lambda x_2 + 2x_3 = 0$$

Is the solution unique? Explain giving the number of independent solutions.

Hints and Answers to Exercise Questions

1. $\frac{1}{2} \begin{pmatrix} a+b+c+d & a-b+c-d \\ c+d-a-b & c-d-a+b \end{pmatrix}$

[Hint: Find out the transformation matrix S from $S \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $S \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Then apply similarity transformation].

2. (i) The matrix $\begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$ has eigenvalues $1, \pm 2$.

The diagonalising matrix $S = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}$.

(ii) The matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has eigenvalues ± 1 .

The diagonalising matrix $S = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$.

(iii) The matrix $\begin{vmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{vmatrix}$ has eigenvalues $0, \pm \sqrt{2}$.

The diagonalising matrix is $\begin{vmatrix} \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{i}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{2} & \frac{-1}{2} \end{vmatrix}$.

(iv) The matrix $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}$ has eigenvalues $1, 1, -1$,

The diagonalising matrix can be taken to be $\begin{vmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{vmatrix}$.

(It is non-unique because of the non-uniqueness of the eigenvectors).

For the answer to the last part, the transformation matrix is S for (i). Now operate S on $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ to get the new vector.

$$5. \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix},$$

(a) Always true, (b) The matrices must commute

9. (a) 6, (b) 14, (c) 36, (d) 6.

$$10. (a) \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{vmatrix}, (b) \begin{vmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & 0 \end{vmatrix}, \text{ where}$$

$$a_{11} = a_{12} = a_{13} = 0, a_{21} = \frac{b_1}{c_1}, a_{22} = a_{23} = 0$$

$$a_{33} = 0, a_{32} = \frac{2c_2}{b_1}, a_{31} = \left(b_2 - \frac{2c_2 a_1}{b_1} \right) / c.$$

$$a_{41} = b_3 - \left(\frac{2c_2 c_3 - 3 b_2 d_3}{b_1 c_2} \right) a_1 - \frac{3d_3}{c_2} a_2, a_{42} = \left(\frac{2c_2 c_3 - 3 b_2 d_3}{b_1 c_2} \right),$$

$$a_{43} = \frac{3d_3}{c_2}, \text{ the polynomials of order zero, one, two and three being } c, a_1 + b_1 t, a_2 + b_2 t + c_2 t^2 \text{ and } a_3$$

$$+ b_3 t + c_3 t^2 + d_3 t^3 \text{ respectively.}$$

11. (a) (i) Invertible for all α ,
(ii) Non-invertible for all α .

$$12. (a) 36, (b) \begin{vmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\frac{3}{2}} & \sqrt{\frac{3}{2}} \\ 0 & -\sqrt{\frac{3}{2}} & \sqrt{\frac{3}{2}} \end{vmatrix} \text{ (assuming anticlockwise rotation).}$$

13. (a) [**Hint:** Utilise the eigenvalue equation of an odd-dimensional matrix].
(b) Two eigenvalues are 0 and 1.

14. $\lambda^2 - 7\lambda + 5 = 0$; Rank of the matrix is 2 and hence there is one independent solution.

APPENDIX

(A) Direct Product of Two Matrices

The direct product of two matrices is formed by taking the product of each element of one matrix with all the elements of another matrix. Thus the direct product of two matrices A and B of dimensions m and n given by

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix} \text{ respectively is } AB \text{ given by}$$

$$AB = \begin{pmatrix} a_{11}B & \dots & \dots & a_{1m}B \\ a_{21}B & \dots & \dots & a_{2m}B \\ \dots & \dots & \dots & \dots \\ a_{m1}B & \dots & \dots & a_{mm}B \end{pmatrix} [B \text{ represents the entire matrix}] \text{ and of dimension } mn.$$

The direct product is used in the theory of group representation as will be explained in the chapter on Group Theory (Chapter 12).

(B) Examples of Vector Spaces

The following discussions are meant to illustrate certain vector spaces that often occur in mathematical physics

- (i) *Metric Space*: This is a linear space with the distance between any two elements defined by $d(x, y)$, also called the metric, that satisfies the following properties

(i) $d(x, y) \geq 0$, (ii) $d(x, y) = d(y, x)$ and $= 0$ if $x = y$.

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ (Triangle inequality)

A suitable example is $d(x, y) = |x - y|$

- (ii) *Euclidean Space*: A particular type of metric space with the distance function defined by the Euclidean metric is called the Euclidean space.

- (a) The distance between two functions $f(x)$ and $g(x)$ in an infinite dimensional Euclidean space is

$$\text{given by } d(f, g) = \left(\int_a^b |f(x) - g(x)|^2 dx \right)^{\frac{1}{2}}$$

- (b) The distance between two points in a finite dimension space is given by the Pythagorean formula.

$$\text{Such as, } d(x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_n) = \left[\sum_i (x_i - y_i)^2 \right]^{\frac{1}{2}}$$

- (iii) *Hilbert Space*: It is a linear vector space, (complex in general) which is complete and endowed with a well-defined inner product of any two elements. Completeness of the space means that every convergent infinite sequence of elements converges to an element which is a member of the space. The state vectors in the linear vector space of formulation of quantum mechanics belong to a complex Hilbert space.

Tensor Analysis

CHAPTER

4

4.1 INTRODUCTION

The discussion on tensor analysis is important because the different objects we encounter, such as scalars, vectors etc., belong to the world of tensors and all the mathematical equations appearing in the physical problems must relate tensors on both the sides. This is because they must obey condition of invariance under a general transformation of coordinates which tensor equations can be shown to satisfy.

The tensors are classified into different types given below according to the manner in which they transform under a change of basis $\bar{x}^i = \bar{x}^i(x^1, x^2, \dots, x^n)$.

1. **Contravariant:** The components $T^{mnp\dots}$ of a contravariant tensor transform under a change of basis as

$$\bar{T}^{mnp\dots} = \frac{\partial \bar{x}^m}{\partial x^\alpha} \frac{\partial \bar{x}^n}{\partial x^\beta} \dots T^{\alpha\beta\gamma\dots} \quad (4.1)$$

(the repeated indices $\alpha\beta\gamma\dots$ are assumed to be summed over). The number of indices is called the rank of the tensor. The coordinate differentials dx^i transform

as $\bar{dx}^i = \frac{\partial \bar{x}^i}{\partial x^\alpha} dx^\alpha$ and thus act as the components of a contravariant tensor of rank 1.

2. **Covariant:** The components of a covariant tensor $T_{mnp\dots}$ transform as

$$\bar{T}_{mnp\dots} = \frac{\partial x^\alpha}{\partial \bar{x}^m} \cdot \frac{\partial x^\beta}{\partial \bar{x}^n} \dots T_{\alpha\beta\gamma\dots} \quad (4.2)$$

The components of the gradient of a scalar ϕ transform under a change of basis

as $\frac{\partial \bar{\phi}}{\partial \bar{x}^m} = \frac{\partial \phi}{\partial x^\alpha} \cdot \frac{\partial x^\alpha}{\partial \bar{x}^m}$ (ϕ being a scalar, $\bar{\phi} = \phi$).

Thus the components transform as those of a covariant tensor of rank 1.

3. Mixed: The components $\bar{T}_{abc\dots}^{mnp\dots}$ of a mixed tensor transform under the change of basis as,

$$\bar{T}_{abc\dots}^{mnp\dots} = \frac{\partial \bar{x}^m}{\partial x^\alpha} \cdot \frac{\partial \bar{x}^n}{\partial x^\beta} \cdot \frac{\partial \bar{x}^p}{\partial x^\gamma} \dots \frac{\partial x^i}{\partial \bar{x}^a} \cdot \frac{\partial x^j}{\partial \bar{x}^b} \cdot \frac{\partial x^k}{\partial \bar{x}^c} T_{ijk\dots}^{\alpha\beta\gamma\dots}$$

The product of a contravariant tensor and a covariant tensor behaves as a mixed tensor. Thus a tensor of rank one behaves as a vector (contravariant or covariant) while one of rank zero as a scalar that remains unchanged under a change of basis (i.e., $\phi(x^\alpha) = \bar{\phi}(\bar{x}^\alpha)$).

If we take the direct product of two tensors, one of rank p and the other of rank q , we obtain a tensor of rank $p + q$. Thus if both A and B be contravariant tensors of ranks p and q , respectively the product AB is a contravariant tensor of rank $p + q$. If both be covariant the product is a covariant tensor of rank $p + q$ and if one be contravariant and the other covariant we obtain a mixed tensor of rank $p + q$. This can easily be demonstrated by taking two tensors, viz., A^{mn} , B^{pq} and finding out the transformation of the product.

$$\bar{A}^{mn} \bar{B}^{pq} = \frac{\partial \bar{x}^m}{\partial x^i} \cdot \frac{\partial \bar{x}^n}{\partial x^j} \cdot \frac{\partial \bar{x}^p}{\partial x^k} \cdot \frac{\partial \bar{x}^q}{\partial x^l} A^{ij} B^{kl} \quad [\text{sum over repeated indices is implied}].$$

The product tensor has rank $2 + 2$ or 4 and is contravariant in character.

Let us now take the product of the same two tensors, one being contravariant and the other being covariant. The product is a tensor of rank 4 .

$$\bar{A}^{mn} \bar{B}_{pq} = \frac{\partial \bar{x}^m}{\partial x^i} \cdot \frac{\partial \bar{x}^n}{\partial x^j} \cdot \frac{\partial x^k}{\partial \bar{x}^p} \cdot \frac{\partial x^l}{\partial \bar{x}^q} A^{ij} B_{kl}$$

Let us take $m = p$ and take the sum over the pair of these terms

$$\begin{aligned} \bar{A}^{mn} \bar{B}_{mq} &= \frac{\partial \bar{x}^m}{\partial x^i} \cdot \frac{\partial \bar{x}^n}{\partial x^j} \cdot \frac{\partial x^k}{\partial \bar{x}^m} \cdot \frac{\partial x^l}{\partial \bar{x}^q} A^{ij} B_{kl} \\ &= \frac{\partial x^k}{\partial x^i} \cdot \frac{\partial \bar{x}^n}{\partial x^j} \cdot \frac{\partial x^l}{\partial \bar{x}^q} A^{ij} B_{kl} \\ &= \frac{\partial \bar{x}^n}{\partial x^j} \cdot \frac{\partial x^l}{\partial \bar{x}^q} A^{ij} B_{il} \end{aligned} \quad (4.3)$$

Thus clearly the resultant product behaves like a tensor of rank $4-2$ or 2 . Thus making a pair of contravariant and covariant indices equal and taking the sum over the pair we get a tensor of rank reduced by 2 . By making each pair of contravariant and covariant indices same and summing over the pair, we can reduce the rank of the resultant tensor at each step by 2 . The reduction of the rank of tensor by this process is called contraction. Taking two tensors of rank 1 each, one contravariant and the other covariant, making the indices same and carrying out the sum over the equal index we get a scalar. The inner product of two vectors, giving a scalar, furnishes such an example. [The work done by a covariant force in producing an infinitesimal displacement is such an example].

4.2 SOME IMPORTANT PROPERTIES

There are a few important properties enlisted below.

- (i) If the components of a tensor vanish in one particular frame of reference they will vanish in all frames. This can be seen immediately from the transformation rule of the components of a tensor. This leads to

the result that any tensor equation of the form $A = B$ (or $A - B = 0$) is form-invariant under a change of axes.

(ii) The transformation rule gives,

$$\begin{aligned}\bar{A}^{mnp} &= \frac{\partial \bar{x}^m}{\partial x^i} \cdot \frac{\partial \bar{x}^n}{\partial x^j} \cdot \frac{\partial \bar{x}^p}{\partial x^k} \dots A^{ijk} \\ &= \frac{\partial \bar{x}^m}{\partial x^j} \cdot \frac{\partial \bar{x}^n}{\partial x^i} \cdot \frac{\partial \bar{x}^p}{\partial x^k} \dots A^{jik} \quad (\text{on interchanging } i \text{ and } j)\end{aligned}$$

If $A^{ijk\dots} = A^{jik\dots}$, we get,

$$\bar{A}^{mnp\dots} = \frac{\partial \bar{x}^m}{\partial x^j} \cdot \frac{\partial \bar{x}^n}{\partial x^i} \cdot \frac{\partial \bar{x}^p}{\partial x^k} \dots A^{ijk} = \bar{A}^{nmp} \quad (4.4)$$

Hence, if the components of the tensor $A^{ijk\dots}$ are symmetric with respect to the pair $i, j\dots$ (i.e., the first two elements) they will remain so even after transformation. This invariance of symmetry (or anti-symmetry), as can be easily seen from the transformation rule under transformation of axes, will hold good for any pair of indices irrespective of their position in the tensor.

(iii) The contraction with respect to pairs of components of two tensors one being symmetric and the other being anti-symmetric with respect to the same pair will lead to a null result.

Thus $A^{mnp\dots} B_{mn\alpha\beta\dots} = A^{nmp\dots} B_{mn\alpha\beta\dots}$ [interchanging in and n]

$$= -A^{mnp\dots} B_{mn\alpha\beta\dots}$$

[using the conditions of symmetry and anti-symmetry of A and B with respect to the pair respectively]

$$\text{Or, } A^{mnp\dots} B_{mn\alpha\beta\dots} = 0$$

This leads to the important consequences in problem of physics. Thus, the work done during any strain of an elastic body /vol. = $\frac{1}{2} T^{ik} \sigma_{ik}$, where T_{ik} , the stress tensor, is symmetric with respect to i, k by definition and σ_{ik} is the symmetric part of the strain tensor. A general α_{ik} can be written as,

$$\alpha_{ik} = \frac{1}{2}(\alpha_{ik} + \alpha_{ki}) + \frac{1}{2}(\alpha_{ik} - \alpha_{ki}).$$

The first part is symmetric with respect to i, k , and is called the strain tensor σ_{ik} here. Clearly, the contribution of the second part, that is antisymmetric with respect to i, k , to the work done, vanishes because of what has been proved just now.

(iv) Let us consider a tensor $C^{ijkl} = A^{ij} B^{kl}$ and let A^{ij} be a contravariant tensor of rank 2.

After transformation we obtain,

$$C^{ijkl'} = \frac{\partial x^{i'}}{\partial x^m} \cdot \frac{\partial x^{j'}}{\partial x^n} \cdot \frac{\partial x^{k'}}{\partial x^p} \cdot \frac{\partial x^{l'}}{\partial x^q} C^{mnpq}$$

$$\text{Or, } A^{ij'} B^{kl'} = \frac{\partial x^{i'}}{\partial x^m} \cdot \frac{\partial x^{j'}}{\partial x^n} \cdot \frac{\partial x^{k'}}{\partial x^p} \cdot \frac{\partial x^{l'}}{\partial x^q} A^{mn} B^{pq}$$

[Sum over the repeated indices is implied]

Since $A^{ij'} = \frac{\partial x^{i'}}{\partial x^m} \cdot \frac{\partial x^{j'}}{\partial x^n} A^{mn}$ we get, $B^{k'l'} = \frac{\partial x^{k'}}{\partial x^p} \frac{\partial x^{l'}}{\partial x^q} B^{pq}$ because the equation is valid for arbitrary A^{ij} .

Thus the object B^{pq} transforms like a contravariant tensor of rank 2. This example is an illustration of a general theorem known as the quotient rule.

If $C = AB \dots$, and C on the left hand side, as well as all other objects except one on the right hand side are tensors, the remaining object on the r. h. s. also behaves like a tensor of the appropriate rank and type. This is clearly valid for tensors A, B, C of any type and any rank.

Example

$$ds^2 = g_{ik} dx^i dx^k$$

ds^2 is the square of the infinitesimal distance between two points, and hence a scalar. On the right hand side, dx^i and dx^k behave like components of a contravariant tensor of rank 1.

By the quotient rule g_{ik} is a covariant tensor of rank 2. g_{ik} is called the metric tensor of the space. Since $dx^i dx^k$ gives the components of a contravariant tensor of rank 2, which is symmetric with respect to i and k , g_{ik} is also symmetric with respect to i and k .

Actually, we can write $ds^2 = d\bar{r} \cdot d\bar{r} = \frac{\partial \bar{r}}{\partial x^i} \cdot \frac{\partial \bar{r}}{\partial x^k} \cdot dx^i dx^k$

We can write $g_{ik} = \frac{\partial \bar{r}}{\partial x^i} \cdot \frac{\partial \bar{r}}{\partial x^k}$, $\frac{\partial \bar{r}}{\partial x^i}$ being the tangent vector along the direction of x^i , ($= \bar{e}_i$ say).

Let us define $\delta_j^i = \frac{\partial x^i}{\partial x^j} = 1$ if $i = j$ and $= 0$ if $i \neq j$.

Clearly, $\bar{\delta}_j^i = \frac{\partial \bar{x}^i}{\partial \bar{x}^j} = \frac{\partial \bar{x}^i}{\partial x^m} \cdot \frac{\partial x^m}{\partial \bar{x}^j} = \frac{\partial \bar{x}^i}{\partial x^m} \cdot \frac{\partial x^m}{\partial \bar{x}^j} \delta_n^m$

and $\bar{\delta}_j^i$, like $\delta_j^i = 1$ for $i = j$ and 0 for $i \neq j$

Thus the kronecker delta δ_j^i behaves like as invariant mixed tensor of rank 2.

4.3(A) NON-ORTHOGONAL CURVILINEAR COORDINATE SYSTEM

In the second chapter we dealt with the vector relations in orthogonal curvilinear coordinate systems. Here we address the general non-orthogonal curvilinear coordinate systems. In this system neither the coordinate surfaces intersect orthogonally nor the tangents drawn to the surfaces at the point of intersection are normals to the respective coordinate surfaces $q_1 = \text{constant}$, $q_2 = \text{constant}$, $q_3 = \text{constant}$ respectively.

Let $\bar{e}_i$ be the unit covariant basis vector along the q^i line defined by $\bar{e}_i = \frac{1}{h_i} \frac{\partial \bar{r}}{\partial q^i}$, h_i being the scale factor along this direction. Thus any infinitesimal displacement $d\bar{r}$ along the q^i line is given by $d\bar{r} = \bar{e}_i h_i dq^i$.

We can write a vector $\bar{A}$ as $\bar{A} = \sum_i \bar{e}_i A^i$, where the A^i 's are called the contravariant components of $\bar{A}$.

We define a triad of vectors $\bar{e}^1, \bar{e}^2, \bar{e}^3$, by the relations $\bar{e}^1 = \frac{\bar{e}_2 \times \bar{e}_3}{V}, \bar{e}^2 = \frac{\bar{e}_3 \times \bar{e}_1}{V}, \bar{e}^3 = \frac{\bar{e}_1 \times \bar{e}_2}{V}$ where

$V = \bar{e}_1 \cdot (\bar{e}_2 \times \bar{e}_3)$ is the volume of the parallelepiped formed by the triad. Clearly $\bar{e}^i \cdot \bar{e}_j = \delta_j^i$. The components A_i defined by $\bar{A} = \sum_i \bar{e}^i A_i$ are the covariant components of $\bar{A}$. The vectors $\bar{e}^i, \bar{e}^j, \bar{e}^k$ are the unit basis vectors in the space reciprocal to that formed by $\bar{e}_i, \bar{e}_j$ and $\bar{e}_k$ respectively, where $\bar{e}^i \cdot (\bar{e}^j \times \bar{e}^k) = V' = \frac{1}{V}$. The contravariant components A^i of $\bar{A}$ are the scalar products of $\bar{A}$ with the unit basis vectors in the reciprocal space, i.e., $A^i = \bar{A} \cdot \bar{e}^i$. Similarly for the covariant components $A_i = \bar{A} \cdot \bar{e}_i$. The contravariant and the covariant components of $\bar{A}$ are shown geometrically in Fig. 4.1 and discussed in the Section 4.3(B).

The square of the infinitesimal line element ds^2 given by $ds^2 = h_i h_j \bar{e}_i \cdot \bar{e}_j dq^i dq^j = g_{ij} dq^i dq^j$, where $g_{ij} = h_i h_j \bar{e}_i \cdot \bar{e}_j$ constitutes the covariant metric tensor components of the space. The contravariant components g^{ij} are defined by $g^{ij} = (\bar{e}^i \cdot \bar{e}^j) h^i h^j$ and related to the covariant components by the identity $g^{ij} g_{jk} = \delta_k^i$.

In a Cartesian coordinate system $g_{ij} = \delta_{ij}$, and hence $\bar{e}_i = \bar{e}^i$, $A_i = A^i$ for all i .

The infinitesimal surface element formed by q_2, q_3 is given by $d\bar{A} = d\bar{S}_2 \times d\bar{S}_3 = h_2 h_3 (\bar{e}_2 \times \bar{e}_3) dq_2 dq_3$. The magnitudes of the infinitesimal areas are given by $|d\bar{A}_1| = \sqrt{g_{22} g_{33} - g_{23}^2} dq_2 dq_3$, $|d\bar{A}_2| = \sqrt{g_{11} g_{33} - g_{13}^2} dq_1 dq_3$, $|d\bar{A}_3| = \sqrt{g_{11} g_{22} - g_{12}^2} dq_1 dq_2$. (Use the relation $(\bar{A} \cdot \bar{B})^2 = A^2 B^2 - (\bar{A} \times \bar{B})^2$).

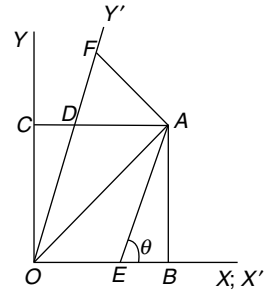


Fig. 4.1

The infinitesimal volume element is given by $dV = h_1 h_2 h_3 |\bar{e}_1 \cdot (\bar{e}_2 \times \bar{e}_3)| dq^1 dq^2 dq^3 = \sqrt{g} dq^1 dq^2 dq^3$, where g is the determinant of the g_{ij} matrix, given by $h_1^2 h_2^2 h_3^2 |\bar{e}_1 \cdot (\bar{e}_2 \times \bar{e}_3)|^2$.

4.3(B) CONTRAVARIANT AND COVARIANT COMPONENTS

We define g^{ij} by the relation $g^{ij} g_{jk} = \delta_k^i$. Since g_{jk} behaves as a covariant tensor of rank 2, by the quotient rule g^{ij} behaves as a contravariant tensor of rank 2. Clearly, the relations $A^i = g^{ij} A_j$ and $A_j = g_{jk} A^k$ are mutually consistent. A^i and A_i can be called the contravariant and covariant components of the vector A_i while g^{ij} and g_{ij} play the roles of the raising and lowering operators respectively. Thus, a vector A can be written as, $A = e_i A^i = \bar{e}^i A_i$ (e^i or $\bar{e}_i$ need not have unit magnitude).

Let us try to understand geometrically the contravariant and the covariant components of a vector (Fig. 4.1).

Let $O\bar{A}$ be the vector $\bar{A}$, OX and OY be the old axes and OX' , OY' the new axes.

OB and OC are the components parallel to OX and OY , respectively, while OE and OD are components parallel to OX' and OY' respectively of $\bar{A}$.

Let us name $A^X = OB$, $A^{X'} = OE$, $A^Y = OC$, $A^{Y'} = OD$

$$A^{X'} = OE = OB - BE = A^X - AB \cot \theta = A^X - A^Y \cot \theta$$

$$A^{Y'} = OD = OC \operatorname{cosec} \theta = A^Y \operatorname{cosec} \theta$$

Now, clearly $x' = x - y \cot \theta$, $y' = y \operatorname{cosec} \theta$

$$\text{or,} \quad A^{x'} = A^x - A^y \cot \theta = \frac{\partial x'}{\partial x} A^x + \frac{\partial x'}{\partial y} A^y$$

$$A^{y'} = A^y \operatorname{cosec} \theta = \frac{\partial y'}{\partial x} A^x + \frac{\partial y'}{\partial y} A^y$$

Thus the components parallel to the axes transform like the contravariant components of $\bar{A}$.
Let us draw perpendicular AF on OY' , the rotated axis.

$$\text{Let} \quad A_x = A'_x = OB, A_y = OC, A'_y = OF$$

$$\text{From} \quad x = x' + y' \cos \theta, y = y' \sin \theta, \frac{\partial x}{\partial x'} = 1, \frac{\partial x}{\partial y'} = \cos \theta$$

$$\text{and} \quad \frac{\partial y}{\partial x'} = 0, \frac{\partial y}{\partial y'} = \sin \theta$$

$$\text{Thus,} \quad A_{x'} = OB = \frac{\partial x}{\partial x'} A_x + \frac{\partial y}{\partial x'} A_y \left[\because \frac{\partial y}{\partial x'} = 0 \right] \quad \dots 4.5(A)$$

$$A_{y'} = OF = OD + DF = A_y \cos \theta + A_x \cos \theta - OD \cos^2 \theta$$

or, $A_{y'} = A_x \cos \theta + A_y \sin \theta$, since from geometry $OD = OC \operatorname{cosec} \theta = A_y \operatorname{cosec} \theta$.

Thus substituting $\frac{\partial x}{\partial y'} = \cos \theta, \frac{\partial y}{\partial y'} = \sin \theta$ we obtain,

$$A_{y'} = \frac{\partial x}{\partial y'} A_x + \frac{\partial y}{\partial y'} A_y \quad \dots 4.5(B)$$

Equations 4.5 (A) and 4.5 (B) show that OB and OF , the distances of the origins from the feet of the perpendiculars from A on the axes transform like the covariant components of a vector:

The vector $\bar{A}$ can also be written as $\bar{A} = \bar{e}_{oi} A_p(i), \bar{e}_{oi}$ being the unit vector $\left(\bar{e}_{oi} = \frac{e_i}{\sqrt{g_{ii}}} \right)$. $A_p(i)$ is called the physical component of the vector $\bar{A} = A^i \sqrt{g_{ii}} = A_i / \sqrt{g_{ij}}$

4.4 CARTESIAN TENSORS, PSEUDOTENSOR

The previous discussions referred to a general transformation of axes where the contravariant and the covariant tensors behave differently. Of particular significance is the case of transformation from a set of rectangular Cartesian axes to another such set; where the distinction between the contravariant and the covariant tensors disappears. This can be illustrated by taking the example of rotation about any axis (say-z-axis) in 3D space. We can write $x' = x \cos \theta + y \sin \theta, y' = -x \sin \theta + y \cos \theta, z' = z$ and $x = x' \cos \theta - y' \sin \theta, y = x' \sin \theta + y' \cos \theta, z = z'$. Comparing the coefficient of transformation in the case of a tensor of any rank, we can easily see that the contravariant and the covariant tensors behave identically. All the results obtained in the earlier discussions hold good with the exception that in the case of contraction we can contract between any pair of indices, whereas in the general transformation the contraction has to be between a pair of contravariant and covariant components. Tensors for which we confine our attention to transformation of Cartesian axes alone are called Cartesian tensors. The metric tensor $g_{ik} = g^{ik} = \delta_{ik}$ in this case.

Let us consider the Levi-civita symbol ϵ_{ijk} which in the orthonormal Cartesian basis is equal to 1 for i, j, k following a cyclic order of 1, 2, 3, changes sign for an odd number of permutations between the symbols and vanishes for any two of the indices being equal. If ϵ_{ijk} has to transform like a tensor of rank 3 for an orthogonal transformation from a set of Cartesian axes to another (including also inversion of axes when the determinant of the transformation matrix equals -1) we must have,

$\epsilon'_{ijk} = \alpha_{il} \alpha_{jm} \alpha_{kn} |\det \alpha| \epsilon_{lmn}$, the α 's being the coefficients of transformation [sum over the repeated indices is implied and $\det \alpha = \pm 1$ for proper and improper orthogonal transformations respectively]. ϵ'_{ijk} satisfies the same definition as ϵ_{ijk} .

The presence of $|\det \alpha|$ in the transformation makes ϵ_{ijk} behave like a pseudotensor or relative tensor of weight 1 in exception to the proper tensor where there is no such factor.

A tensor of rank one $C_i = \epsilon_{ijk} A^j B^k$, where A^j and B^k behave like ordinary tensors, transforms like a pseudovector. Thus the vector $\vec{A} \times \vec{B}$ transforms like a pseudovector.

Let us address the problem of general coordinate transformation. An object $T_{ij\dots}^{ab\dots}$ that transforms like $\bar{T}_{ij\dots}^{ab\dots} = \frac{\partial \bar{x}^a}{\partial x^m} \cdot \frac{\partial \bar{x}^b}{\partial x^n} \dots \frac{\partial x^p}{\partial \bar{x}^i} \cdot \frac{\partial x^q}{\partial \bar{x}^j} \dots T_{pq\dots}^{mn\dots} \left| \frac{\partial x}{\partial \bar{x}} \right|^w$, where $\left| \frac{\partial x}{\partial \bar{x}} \right|$ is the determinant of the transformation matrix, is said to behave like a relative tensor of weight w .

Examples

- (1) Let us take the covariant metric tensor g_{ij}

$$\bar{g}_{ij} = \frac{\partial x^m}{\partial \bar{x}^i} \cdot \frac{\partial x^n}{\partial \bar{x}^j} g_{mn}.$$

Taking the determinant of both the sides,

$$\bar{g} = \left| \frac{\partial x}{\partial \bar{x}} \right|^2 \cdot g \text{ leading to } \sqrt{\bar{g}} = \left| \frac{\partial x}{\partial \bar{x}} \right| \sqrt{g}.$$

Hence $\sqrt{g}$ behaves like a relative tensor of rank zero (i.e, a scalar) and weight 1.

- (2) The n -dimensional volume element is easily seen to transform like a relative tensor of weight -1 since

$$d\bar{x}^1 d\bar{x}^2 \dots d\bar{x}^n = \left| \frac{\partial \bar{x}}{\partial x} \right| dx^1 \dots dx^n$$

Specializing to the case of 3 dimensions, we can prove that on change of basis the triple product $\bar{e}_1 \cdot (\bar{e}_2 \times \bar{e}_3)$ behaves like a relative tensor of weight 1.

Thus $\sqrt{g} d^n x (= \sqrt{g} dx^1 \dots dx^n)$ behaves like a proper scalar under coordinate transformation.

- (3) $\bar{\delta}_j^i = \frac{\partial \bar{x}^i}{\partial \bar{x}^j} = \frac{\partial \bar{x}^i}{\partial x^m} \cdot \frac{\partial x^m}{\partial \bar{x}^j} \delta_n^m$

δ_j^i , therefore, behaves like a relative tensor of weight zero (or absolute tensor).

ϵ_{ikj} can be shown to behave like a relative tensor of weight -1 while ϵ^{ijk} behaves like one of weight $+1$ respectively.

4.5 COVARIANT DERIVATIVE

We have already seen that the derivative of a scalar behaves like a covariant tensor of rank 1 for general transformations. Let us now ascertain what happens to the derivatives of the components of a tensor, of arbitrary rank and type.

Now,

$$\frac{\partial \bar{A}^m}{\partial \bar{x}^i} = \frac{\partial}{\partial x^j} \left[\frac{\partial \bar{x}^m}{\partial x^l} A^l \right] \frac{\partial x^j}{\partial \bar{x}^i} = \frac{\partial \bar{x}^m}{\partial x^l} \cdot \frac{\partial x^j}{\partial \bar{x}^i} \cdot \frac{\partial A^l}{\partial x^j} + \frac{\partial^2 \bar{x}^m}{\partial x^j \partial x^l} \cdot \frac{\partial x^j}{\partial \bar{x}^i} A^l.$$

For general non-linear transformations of coordinate the second term on the right hand side does not vanish and hence the derivative of a tensor of rank 1 does not transform like a tensor. We investigate below the reasons behind such a non-tensorial behaviour and introduce the modifications necessary to make the derivatives of tensor behave like tensors for general coordinate transformations.

The derivatives of the components of a tensor imply the change of the components of the tensor with the change in position. The change of the components of the tensor comprises two parts. (i) change due to the different basis vectors at different points and (ii) change of the components of the tensor itself when referred to the same choice of basis vectors. The first change clearly occurs due to the space being intrinsically curved or due to the choice of non-Cartesian coordinates in a flat space. Thus, for a contravariant vector the change of the component of the vector between two points 1 and 2 [$\bar{x}$ and $\bar{x} + d\bar{x}$ respectively].

$$\begin{aligned} d\bar{A}^m &= \bar{A}_{(2)}^{m(2)} - \bar{A}_{(1)}^{m(1)} = \bar{A}^{m(2)}(\bar{x}_i + d\bar{x}_i) - \bar{A}^{m(1)}(\bar{x}_i) \\ &= \left\{ \bar{A}^{m(2)}(\bar{x}_i + d\bar{x}_i) - \bar{A}^{m(1)}(\bar{x}_i + d\bar{x}_i) \right\} + \left\{ \bar{A}^{m(1)}(\bar{x}_i + d\bar{x}_i) - \bar{A}^{m(1)}(\bar{x}_i) \right\} \end{aligned}$$

The term within the first curly bracket represents the contribution due to the second part while that within the second bracket gives the contribution due to the first term.

Thus

$$\begin{aligned} d\bar{A}^m &= D\bar{A}^m + \delta\bar{A}^m, \text{ where} \\ D\bar{A}^m &= \bar{A}^{m(2)}(\bar{x}_i + d\bar{x}_i) - \bar{A}^{m(1)}(\bar{x}_i + d\bar{x}_i) \text{ and} \\ \delta\bar{A}^m &= \bar{A}^{m(1)}(\bar{x}_i + d\bar{x}_i) - \bar{A}^{m(1)}(\bar{x}_i). \end{aligned} \quad (4.6)$$

$D\bar{A}^m$ is the difference between the components of the vector $\bar{A}^{m(2)}$ and $\bar{A}^{m(1)}$ evaluated at the point $\bar{x}_i + d\bar{x}_i$ with respect to the coordinate system located at that point and hence behaves as the component of a vector. In order to find out $D\bar{A}^m$ we have to displace the vector $\bar{A}$ at the point 1 parallel to itself to the point 2 and then take the difference between the components of the vector which is already located at 2 and the parallel-transported vector which was originally at 1. The term $\delta\bar{A}^m$ represents the difference between the component of the vector $\bar{A}^m$ at 1 taken with respect to the different basis vectors located at $\bar{x}$ and $\bar{x} + d\bar{x}$ respectively. $\delta\bar{A}^m$, the change in the component $\bar{A}^m$ due to the parallel transport from (1) to (2) depends on the vector itself and the measure of displacement $d\bar{x}$. It can be written as, $\delta\bar{A}^m = -\Gamma_{kl}^m A^k dx^l$, where Γ_{kl}^m is called the Riemann-Christoffel symbol evaluated at the point,

Or, $D\bar{A}^m = d\bar{A}^m - \delta\bar{A}^m$ from (4.6).

Since the change in the component of the displacement $d\bar{x}^i$ behaves as that of a contravariant vector

$$\frac{D\bar{A}^m}{dx^i} = \bar{A}_{;i}^m = \frac{d\bar{A}^m}{dx^i} - \frac{\delta\bar{A}^m}{dx^i} \text{ should behave as the component of a mixed tensor of rank 2.}$$

Thus, $\bar{A}^m_{;i}$, also called the covariant derivative can be expressed as, $\bar{A}^m_{;i} = \bar{A}^m_{,i} + \Gamma^m_{ki} A^k$

$$\text{Or,} \quad \bar{A}^m_{;i} = \bar{A}^m_{,i} + \Gamma^m_{ki} A^k \quad (4.7)$$

For a scalar ϕ since $\delta\phi = 0$, we can write, $\phi_{ji} = \phi_{ij}$ (covariant derivative = ordinary derivative).

Since $\bar{A}^m \bar{B}_m$ is a scalar,

$$\delta(\bar{A}^m \bar{B}_m) = 0 \text{ giving } \delta \bar{B}_m \cdot \bar{A}^m = -\delta \bar{A}^m \cdot \bar{B}_m$$

$$\text{Or,} \quad \bar{A}^m \cdot \delta \bar{B}_m = (\Gamma^m_{kl} \bar{A}^k d\bar{x}^l) \bar{B}_m$$

$$\text{Or,} \quad \delta \bar{B}_k = \Gamma^m_{kl} d\bar{x}^l \bar{B}_m$$

$$\text{Or,} \quad D\bar{B}_k = d\bar{B}_k - \Gamma^m_{kl} d\bar{x}^l \bar{B}_m$$

$$\text{Or,} \quad \bar{B}_{k;i} = \bar{B}_{k,i} - \Gamma^m_{ki} \bar{B}_m \dots \quad (4.8)$$

The covariant derivative of the product of two contravariant tensors of rank one is given by

$$(A^m B^n)_{;k} = A^m (B^n)_{;k} + (A^m)_{;k} B^n \dots \quad (4.9)$$

The result follows because by going over to a locally Cartesian system we can write $(A^m B^n)_{;k} = (A^m B^n)_{,k}$ giving (4.9) because the Christoffel indices vanish in that system, the change due to parallel transport being zero. The equation (4.9) being a tensor equation is therefore valid in every coordinate system and hence we can write Eq. (4.9) in all generality.

$$\begin{aligned} \text{Thus} \quad (A^m B^n)_{;k} &= A^m [(B^n)_{;k} + \Gamma^n_{jk} B^j] \\ &\quad + B^n [(A^m)_{;k} + \Gamma^m_{jk} A^j] \end{aligned}$$

Since any second rank contravariant tensor C^{mn} can be written as $A^m B^n$, the product of two such tensors of rank one each we can write,

$$(C^{mn})_{;k} = (C^{mn})_{,k} + \Gamma^m_{jk} C^{jn} + \Gamma^n_{jk} C^{mj} \dots \quad (4.10)$$

Similarly writing $C_{mn} = A_m B_n$ and $C^m_n = A^m B_n$,

$$\text{we can prove that } C_{mn;k} = C_{mn,k} - \Gamma^p_{mk} C_{pn} - \Gamma^p_{nk} C_{mp} \text{ and} \quad (4.11) \text{ (A)}$$

$$C^m_{n;k} = C^m_{n,k} + \Gamma^m_{in} C^i_k - \Gamma^p_{nk} C^m_p \dots \quad (4.11) \text{ (B)}$$

The analysis can similarly be extended to tensors of any type and any rank. The Christoffel symbols Γ^i_{jk} (also called 'affine connection coefficients') do not behave like tensors under coordinate transformations because if they were tensors, they would vanish in all the coordinate systems if they did so in any one of them. But here the coefficients vanish in Cartesian coordinate systems while remaining nonzero in other systems thereby implying the non-tensorial character.

4.6 PROPERTIES OF CHRISTOFFEL SYMBOLS

1. Symmetry Property

We can write from Eq. (4.8) that

$$B_{m;k} = B_{m,k} - \Gamma_{mk}^i B_i, \text{ and}$$

$$B_{k;m} = B_{k,m} - \Gamma_{km}^i B_i$$

$$\text{or, } (B_{m;k} - B_{k;m}) = (B_{m,k} - B_{k,m}) + (\Gamma_{km}^i - \Gamma_{mk}^i) B_i$$

Choosing $B_m = \phi_{,m}$ and $B_k = \phi_{,k}$, where ϕ is any scalar we get, $B_{m;k} - B_{k;m} = (\Gamma_{km}^i - \Gamma_{mk}^i) B_i$

Going over to a locally Cartesian coordinate system we obtain, $\Gamma_{km}^i = \Gamma_{mk}^i = 0$ for all m and k and hence $B_{m;k} - B_{k;m} = 0$ for any pair m and k .

Since this is a tensor equation the left hand side must vanish for any arbitrary coordinate system. Hence, we can write $\Gamma_{km}^i = \Gamma_{mk}^i$ for any basis chosen.

2. $\Gamma_{\beta\gamma}^\alpha$ in terms of the metric tensor

Taking $A_\alpha = g_{\alpha\beta} A^\beta$ we get, $DA_\alpha = D(g_{\alpha\beta} A^\beta)$

$$\text{Or, } DA_\alpha = g_{\alpha\beta} (DA^\beta) + (Dg_{\alpha\beta}) A^\beta \dots \quad (4.12)$$

DA_α and DA_β being components of a covariant tensor of rank 2 we can write $(DA_\alpha) = g_{\alpha\beta} (DA^\beta)$.

Using $DA_\alpha = g_{\alpha\beta} DA^\beta$ we get, $Dg_{\alpha\beta} = 0$

$$\begin{aligned} \text{From 4.11(A) } g_{\alpha\beta,\gamma} &= g_{i\beta} \Gamma_{\alpha\gamma}^i + g_{\alpha i} \Gamma_{\beta\gamma}^i \\ -g_{\beta\gamma,\alpha} &= -g_{i\gamma} \Gamma_{\beta\alpha}^i - g_{\beta i} \Gamma_{\gamma\alpha}^i \dots \end{aligned} \quad (4.13)$$

Taking half of the sum of the above two equations and using $\Gamma_{\alpha\beta}^i = \Gamma_{\beta\alpha}^i$ we get,

$$g_{i\alpha} \Gamma_{\beta\gamma}^i = \frac{1}{2} (g_{\alpha\beta,\gamma} + g_{\gamma\alpha,\beta} - g_{\beta\gamma,\alpha})$$

Contracting with $g^{\alpha j}$ we get,

$$\Gamma_{\beta\gamma}^j = \frac{1}{2} g^{\alpha j} (g_{\alpha\beta,\gamma} + g_{\gamma\alpha,\beta} - g_{\beta\gamma,\alpha})$$

The above equation can also be rewritten as,

$$\Gamma_{\beta\gamma}^\alpha = \frac{1}{2} g^{k\alpha} (g_{k\beta,\gamma} + g_{\gamma k,\beta} - g_{\beta\gamma,k}) \dots \quad (4.14)$$

3. Transformation property of $\Gamma_{\beta\gamma}^\alpha$

Using the transformation rule of the tensors T^i and T_{ij}^i we obtain, $\bar{T}^i = \frac{\partial \bar{x}^i}{\partial x^m} T^m$ and $\bar{T}_j = \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial x^j}{\partial \bar{x}^n} T_n^m$

$$\begin{aligned} \text{Using } \bar{T}^i &= \frac{\partial \bar{x}^i}{\partial x^m} T^m \text{ we get, } \bar{T}_{,j}^i = \frac{\partial}{\partial \bar{x}^j} \left(\frac{\partial \bar{x}^i}{\partial x^m} T^m \right) + \bar{\Gamma}_{kj}^j \bar{T}^k \\ &= \frac{\partial}{\partial x^l} \left(\frac{\partial \bar{x}^i}{\partial x^m} T^m \right) \frac{\partial x^l}{\partial \bar{x}^j} + \bar{\Gamma}_{kj}^i \bar{T}^k \end{aligned}$$

$$\begin{aligned}
&= \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial x^l}{\partial \bar{x}^j} \frac{\partial T^m}{\partial x^l} \\
&+ \left(\frac{\partial^2 \bar{x}^i}{\partial x^l \partial x^m} \frac{\partial x^l}{\partial \bar{x}^j} + \bar{\Gamma}_{kj}^i \frac{\partial \bar{x}^k}{\partial x^m} \right) T^m \dots
\end{aligned} \tag{4.15}$$

From the relation $\bar{T}_{,j}^i = \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial x^n}{\partial \bar{x}^j} T_{,n}^m$ we get,

$$\begin{aligned}
\bar{T}_{,j}^i &= \frac{\partial \bar{x}^i}{\partial x^m} \cdot \frac{\partial x^n}{\partial \bar{x}^j} \cdot \left(\frac{\partial T^m}{\partial x^n} + \Gamma_{qn}^m T^q \right) \\
&= \frac{\partial \bar{x}^i}{\partial x^q} \cdot \frac{\partial x^n}{\partial \bar{x}^j} \left(\frac{\partial T^q}{\partial x^n} + \Gamma_{qn}^q T^q \right) \dots
\end{aligned} \tag{4.16}$$

Comparing Eqs. (4.15) and (4.16) we get,

$$\frac{\partial \bar{x}^i}{\partial x^q} \cdot \frac{\partial x^n}{\partial \bar{x}^j} \Gamma_{mn}^q = \frac{\partial^2 \bar{x}^i}{\partial x^l \partial x^m} \cdot \frac{\partial x^l}{\partial \bar{x}^j} + \bar{\Gamma}_{qj}^i \cdot \frac{\partial \bar{x}^q}{\partial x^m} \dots$$

Multiplying both the sides by $\frac{\partial \bar{x}^j}{\partial x^n} \frac{\partial x^s}{\partial \bar{x}^i}$ we get,

$$\Gamma_{mn}^s = \frac{\partial^2 \bar{x}^i}{\partial x^l \partial x^m} \cdot \frac{\partial x^l}{\partial \bar{x}^j} \cdot \frac{\partial \bar{x}^j}{\partial x^n} \cdot \frac{\partial x^s}{\partial \bar{x}^i} + \bar{\Gamma}_{qj}^i \cdot \frac{\partial \bar{x}^q}{\partial x^m} \cdot \frac{\partial \bar{x}^j}{\partial x^n} \cdot \frac{\partial x^s}{\partial \bar{x}^i}$$

$$\text{or,} \quad \Gamma_{mn}^s = \bar{\Gamma}_{qj}^i \cdot \frac{\partial \bar{x}^q}{\partial x^m} \cdot \frac{\partial \bar{x}^j}{\partial x^n} \cdot \frac{\partial x^s}{\partial \bar{x}^i} + \frac{\partial^2 \bar{x}^i}{\partial x^n \partial x^m} \cdot \frac{\partial x^s}{\partial \bar{x}^i} \dots \tag{4.17}$$

We can invert the relation to get,

$$\bar{\Gamma}_{uv}^w = \Gamma_{mn}^s \cdot \frac{\partial x^m}{\partial \bar{x}^u} \cdot \frac{\partial x^n}{\partial \bar{x}^v} \cdot \frac{\partial \bar{x}^w}{\partial x^s} - \frac{\partial^2 \bar{x}^w}{\partial x^r \partial x^m} \cdot \frac{\partial x^m}{\partial \bar{x}^u} \cdot \frac{\partial x^r}{\partial \bar{x}^v} \dots \tag{4.18}$$

It can easily be seen from Eq. (4.14) that the coefficients $\Gamma_{\beta\gamma}^\alpha$ are non-zero even in a flat space with curvilinear coordinates. For example using 2D polar coordinates $\Gamma_{\theta\theta}^r = -\frac{1}{2}$ or 3 D spherical polar coordinates $\Gamma_{\theta\theta}^r = -\frac{1}{2} \sin \theta$.

The non-vanishing of the Christoffel symbols does not, therefore, ensure the non-flatness of the space under consideration. In order that a space be flat, it is necessary that any vector should remain unchanged if it is transported parallel to itself in a complete circuit around any closed curve; or equivalently if the change in a vector due to parallel transport between two fixed points is independent of the path connecting them. By calculating the change of a vector due to parallel transport around a closed curve (the calculation is straightforward and can be seen in any standard text book on Tensor analysis or General Relativity) we can show that the necessary and sufficient condition for flatness of a space is that the components of the curvature

tensor given by $R_{\beta\gamma\delta}^\alpha = \frac{\partial \Gamma_{\beta\delta}^\alpha}{\partial x^\gamma} - \frac{\partial \Gamma_{\beta\gamma}^\alpha}{\partial x^\delta} + \Gamma_{\epsilon\gamma}^\alpha \Gamma_{\beta\delta}^\epsilon - \Gamma_{\epsilon\delta}^\alpha \Gamma_{\beta\gamma}^\epsilon$ vanish at every point.

4.7 TENSOR CALCULUS IN A GENERAL COORDINATE SYSTEM

- (a) *Gradient* – The gradient of a scalar ϕ equals its covariant derivative because a scalar does not change due to parallel transport.

$$\text{Thus } \phi_{;m} = \phi_{,m} = \frac{\partial \phi}{\partial x^m} \dots \quad (4.19)$$

- (b) *Curl* – In defining the components of the curl of a contravariant or a covariant vector we replace the ordinary derivative by the covariant derivative so that the components of the curl transform like those of a tensor. Thus the covariant curl of a covariant vector A

$$= A_{m;n} - A_{n;m} = (A_{m,n} - \Gamma_{mn}^p A_p) - (A_{n,m} - \Gamma_{nm}^p A_p)$$

$$\text{Since } \Gamma_{mn}^p = \Gamma_{nm}^p \text{ we obtain, } A_{m;n} - A_{n;m} = A_{m,n} - A_{n,m} \dots \quad (4.20)$$

The covariant curl of a contravariant vector can be similarly calculated but the result will not be clearly as simple as in the previous cases.

- (c) *Divergence* – The covariant divergence of a contravariant vector

$$A^m_{;m} = \frac{\partial A^m}{\partial x^m} + \Gamma_{ml}^m A^l = A^m_{,m} + \Gamma_{ml}^m A^l.$$

$$\text{Now } \Gamma_{ml}^m = \frac{1}{2} g^{mp} [g_{pm,l} + g_{lp,m} - g_{ml,p}]$$

On interchanging m and p and using $g_{ml} = g_{lm}$, $g^{mp} = g^{pm}$ we get, $\Gamma_{ml}^m = \frac{1}{2} g^{mp} g_{pm,l}$

$$\begin{aligned} \text{Now, for an arbitrary matrix } M, \delta [\ln (\det M)] &= \ln \{ \det (M + \delta M) \} - \ln (\det M) \\ &= \ln \det [M^{-1} (M + \delta M)]. \end{aligned}$$

Using $\det (e^A) = e^{Tr A}$ and retaining terms containing the linear power of δM only for small δM we get,

$$\delta [\ln (\det M)] = Tr [M^{-1} \delta M]$$

$$\text{Or, } \frac{\partial}{\partial x^\lambda} \ln (\det M) = Tr \left(M^{-1} \frac{\partial}{\partial x^\lambda} M \right)$$

Or, taking M to be the metric tensor g_{pm} we get,

$$\Gamma_{ml}^m = \frac{1}{2} \frac{\partial}{\partial x^l} \ln g, \text{ where } g \text{ is the determinant of the matrix } g_{pm}.$$

$$\text{For } g < 0, \Gamma_{ml}^m = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^l} (\sqrt{-g}).$$

Thus, the covariant divergence $A^m_{;m}$, which behaves as a scalar, is given by $A^m_{;m} = A^m_{,m} + \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^l} (\sqrt{-g}) A^l$

$$\text{or, } A^m_{;m} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^l} (\sqrt{-g} A^l). \dots \quad (4.21)$$

The covariant divergence of a tensor $A^{mn}_{;m}$ is given by $A^{mn}_{;m} = A^{mn}_{,m} + \Gamma_{m\lambda}^m A^{\lambda n} + \Gamma_{m\lambda}^n A^{m\lambda}$

Putting $\Gamma_{m\lambda}^m = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\lambda} (\sqrt{-g})$ we get,

$$A_{;m}^{mn} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^m} (\sqrt{-g} A^{mn}) + \Gamma_{m\lambda}^n A^{m\lambda}$$

In the case of an antisymmetric A , where $A^{m\lambda} = -A^{\lambda m}$,

we get, $A_{;m}^{mn} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^m} (\sqrt{-g} A^{mn})$, the second term vanishing after contraction since

$$\Gamma_{m\lambda}^n = \Gamma_{\lambda m}^n.$$

For a symmetric tensor A_i^k we get, ($A_i^k = A_k^i$),

$$\begin{aligned} A_{i;k}^k &= A_{i,k}^k + \Gamma_{lk}^k A_i^l - \Gamma_{ik}^l A_l^k \\ &= \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^k} (\sqrt{-g} A_i^k) - \Gamma_{ki}^l A_l^k. \end{aligned}$$

Putting $\Gamma_{ki}^l = \frac{1}{2} g^{jl} (g_{jk,i} + g_{ij,k} - g_{ki,j})$ we get,

$$\begin{aligned} A_{i;k}^k &= \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^k} (\sqrt{-g} A_i^k) - \frac{1}{2} g^{jl} A_l^k (g_{jk,i} + g_{ij,k} - g_{ki,j}) \\ &= \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^k} (\sqrt{-g} A_i^k) - \frac{1}{2} A^{jk} g_{jk,i} \quad [\text{Since } A^{jk} = A^{kj}] \end{aligned}$$

For the Laplacian we can write, from Eq. (4.21)

$$\nabla^2 \phi = \text{div} \left(g^{nm} \frac{\partial \phi}{\partial x^m} \right) = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^n} \left(\sqrt{-g} \cdot g^{nm} \frac{\partial \phi}{\partial x^m} \right)$$

Or, $\nabla^2 \phi = A_{;n}^n$, where $A^n = g^{nm} \frac{\partial \phi}{\partial x^m}$

Integral Relations

We can extend the integral relations of vector calculus (namely, Stokes theorem, Gauss theorem) to the general curvilinear coordinates in the following way. (The proof is omitted here).

(a) The generalization of Stokes theorem to curvilinear coordinates for any covariant vector with

components A_α is, $\int_s A_\alpha \frac{dx^\alpha}{ds} ds = \iint \epsilon^{\alpha\beta\gamma} A_{\beta,\alpha} ds_\gamma$, where

$$\epsilon^{\alpha\beta\gamma} = + \frac{1}{\sqrt{|g|}} \text{ if } \alpha, \beta, \gamma \text{ be an even and odd permutation of } 1, 2, 3 \text{ respectively. } \dots \quad (4.22)$$

$[\epsilon_{\alpha\beta\gamma} = \pm \sqrt{|g|}]$ and $\epsilon^{\alpha\beta\gamma} = \pm \frac{1}{\sqrt{|g|}}$, the + and the – signs for α, β, γ forming an even and odd permutation of 1, 2, 3 respectively, are the Levi Civita symbols and can be shown to behave like pseudo tensors of rank 3]

(b) The equivalent of Gauss divergence theorem is,

$$\oint A^i \sqrt{|g|} ds_i = \int_v dv_i \sqrt{|g|} A^i_{;i} \dots \quad (4.23)$$

These relations are valid for tensors of any rank. For a second rank tensor we can write $\int T^{ij} \sqrt{|g|} ds_j = \int dv \sqrt{|g|} T^{ij}_{;j}$.

Applications

Tensor analysis finds manifold applications in various branches of physics. Below are given briefly two applications, one in the field of Fluid Mechanics and the other in General Relativity.

(a) *Fluid Mechanics* – The equation of motion of an incompressible fluid is,

$$\int_v \rho \left(\frac{\partial \bar{u}_i}{\partial t} + \bar{u} \cdot \nabla u_i \right) dv = \int \rho g_i dv + \int T_{ij} \cdot ds_j \text{ where}$$

$\int_s T_{ij} \cdot ds_j$ represents the i th component of the stress force acting on the boundary of the surface enclosing the volume V , the normal to the surface pointing along the direction $\bar{j}$. Using Gauss divergence theorem we get,

$$\rho \left(\frac{\partial V_i}{\partial t} + (\bar{V} \cdot \bar{\nabla}) V_i \right) = \rho g_i + \frac{\partial T_{ij}}{\partial x_j}$$

For a perfect fluid $T_{ij} = p \cdot \delta_{ij}$, p being the pressure inside the fluid. Hence we get,

$$\rho \left(\frac{\partial \bar{V}}{\partial t} + \bar{V} \cdot \nabla \bar{V} \right) = \rho \bar{g} - \bar{\nabla} p, \text{ also called the Euler's equation} \quad (4.24)$$

(b) *Path of a free particle* – In flat space (where the curvature tensor components vanish) a particle moves along a straight line which is the path of extremum length. Here, we want to determine the path of extremum length in a space that is curved in general.

The condition of extremum length is, $\delta \int_1^2 ds = \delta \int_1^2 (g_{\mu\nu} d\xi^\mu d\xi^\nu)^{\frac{1}{2}} = 0$

or, $\delta \int_1^2 ds = \delta \int_1^2 \left[g_{\mu\nu} \frac{d\xi^\mu}{d\lambda} \cdot \frac{d\xi^\nu}{d\lambda} \right]^{\frac{1}{2}} d\lambda = 0$ where λ is any arbitrary parameter along the path.

Using the variational principle we obtain Euler-Lagrange equation as given by $\frac{d}{d\lambda} \left[\frac{\partial L}{\partial \left(\frac{d\xi^\mu}{d\lambda} \right)} \right] = \frac{\partial L}{\partial \xi^\mu}$

$$= \left(\frac{\partial L}{\partial g_{\alpha\beta}} \right) g_{\alpha\beta,\mu}, \text{ where } L = \left[g_{\mu\nu} \frac{d\xi^\mu}{d\lambda} \cdot \frac{d\xi^\nu}{d\lambda} \right]^{\frac{1}{2}}$$

Thus we get,
$$\left(\frac{d\lambda}{ds} \right) \left[g_{\mu\nu,\gamma} \frac{d\xi^\nu}{d\lambda} \cdot \frac{d\xi^\gamma}{d\lambda} + g_{\mu\nu} \frac{d^2 \xi^\nu}{d\lambda^2} + \frac{g_{\mu\nu} \frac{d\xi^\nu}{d\lambda} \cdot \frac{d^2 \lambda}{ds^2}}{\left(\frac{d\lambda}{ds} \right)^2} - \frac{1}{2} g_{\alpha\beta,\mu} \frac{d\xi^\alpha}{d\lambda} \cdot \frac{d\xi^\beta}{d\lambda} \right] = 0$$

Multiplying by $g_{\mu\rho}$ and symmetrizing we get,

$$\frac{d^2 \xi^\rho}{d\lambda^2} + \frac{1}{2} g^{\mu\rho} (g_{\mu\alpha,\beta} + g_{\mu\beta,\alpha} - g_{\alpha\beta,\mu}) \frac{d\xi^\alpha}{d\lambda} \cdot \frac{d\xi^\beta}{d\lambda} + \frac{\frac{d\xi^\rho}{d\lambda} \cdot \frac{d^2 \lambda}{ds^2}}{\left(\frac{d\lambda}{ds} \right)^2} = 0$$

Putting $\lambda = s$ we get,

$$\frac{d^2 \xi^\rho}{ds^2} + \frac{1}{2} g^{\mu\rho} (g_{\mu\alpha,\beta} + g_{\mu\beta,\alpha} - g_{\alpha\beta,\mu}) \frac{d\xi^\alpha}{ds} \cdot \frac{d\xi^\beta}{ds} = 0$$

as $\frac{d^2 \xi^\rho}{ds^2} + \Gamma_{\beta}^{\rho} \frac{d\xi^\alpha}{ds} \cdot \frac{d\xi^\beta}{ds} = 0 \dots$ giving the equation of the geodesic. (4.25)

SOLVED PROBLEMS

Problem 1

Prove by using the property of Levi-Civita symbol that

- (a) $\bar{A} \cdot (\bar{B} \times \bar{C}) = B \cdot (\bar{C} \times \bar{A}) = \bar{C} \cdot (\bar{A} \times \bar{B})$
- (b) $\bar{A} \cdot (\bar{A} \times \bar{B}) = 0$

Solution:

- (a) $\bar{A} \cdot (\bar{B} \times \bar{C}) = A^i (\bar{B} \times \bar{C})_i = A^i \epsilon_{ijk} B^j C^k$
or, $\bar{A} \cdot (\bar{B} \times \bar{C}) = \epsilon_{ijk} A^i B^j C^k = \epsilon_{jki} B^j C^k A^i$
or, $\bar{A} \cdot (\bar{B} \times \bar{C}) = \bar{B} \cdot (\bar{C} \times \bar{A}) = \epsilon_{kij} C^k A^i B^j = \bar{C} \cdot (\bar{A} \times \bar{B})$
- (b) $\bar{A} \cdot (\bar{A} \times \bar{B}) = A^i \cdot (\bar{A} \times \bar{B})_i = \epsilon_{ijk} A^i A^j B^k$
 $= -\epsilon_{ijk} A^i A^j B^k$ (using $\epsilon_{ijk} = -\epsilon_{jik}$)

$$\text{or, } \bar{A} \cdot (\bar{A} \times \bar{B}) = \epsilon_{ijk} A^i A^j B^k = 0$$

Problem 2

Express the components of the metric tensor in the Cartesian and the spherical polar coordinates respectively. Find out the contravariant and the covariant components of a vector A in the spherical polar coordinates given that $T^x = x^2$, $T^y = y^2$ and $T^z = z^2$. Verify that the results obtained are mutually consistent.

Solution: The square of an infinitesimal line element in the Cartesian coordinates is given by $ds^2 = dx^2 + dy^2 + dz^2$ so that $g_{xx} = g_{yy} = g_{zz} = 1$. Since $g^{ij} g_{jk} = \delta^i_k$, the corresponding contravariant components are given by $g^{xx} = g^{yy} = g^{zz} = 1$. The covariant basis vectors in the spherical polar coordinates are given by $\bar{e} \ (x = r \sin \theta \cos \phi, y = r \sin \theta \sin \phi, z = r \cos \theta)$

$$\begin{aligned} \bar{e}_r &= \frac{\partial \bar{r}}{\partial r} = \bar{i} \sin \theta \cos \phi + \bar{j} \sin \theta \sin \phi + \bar{k} \cos \theta \\ \bar{e}_\theta &= \frac{\partial \bar{r}}{\partial \theta} = \bar{i} r \cos \theta \cos \phi + \bar{j} r \cos \theta \sin \phi - \bar{k} r \sin \theta \\ \bar{e}_\phi &= \frac{\partial \bar{r}}{\partial \phi} = -\bar{i} r \sin \theta \sin \phi + \bar{j} r \sin \theta \cos \phi \end{aligned}$$

The covariant components of the metric tensor in the (r, θ, ϕ) coordinates are,

$$[g_{ij}] = \bar{e}_i \cdot \bar{e}_j = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix}$$

The corresponding contravariant components are given by

$$g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/r^2 & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2 \theta} \end{pmatrix}$$

The contravariant components of the vector in the (r, θ, ϕ) coordinates are

$$\begin{aligned} T^r &= \frac{\partial r}{\partial x} T^x + \frac{\partial r}{\partial y} T^y + \frac{\partial r}{\partial z} T^z \\ T^\theta &= \frac{\partial \theta}{\partial x} T^x + \frac{\partial \theta}{\partial y} T^y + \frac{\partial \theta}{\partial z} T^z \\ T^\phi &= \frac{\partial \phi}{\partial x} T^x + \frac{\partial \phi}{\partial y} T^y + \frac{\partial \phi}{\partial z} T^z \end{aligned}$$

so that these are

$$T^r = x^2 \sin \theta \cos \phi + y^2 \sin \theta \sin \phi + z^2 \cos \theta = r^2 \sin^3 \theta \cos^3 \phi + r^2 \sin^3 \theta \sin^3 \phi + r^2 \cos^3 \theta$$

$$T^\theta = \frac{x^2}{r} \cos \theta \cos \phi + \frac{y^2}{r} \cos \theta \sin \phi - \frac{z^2 \sin \theta}{r} = r \sin^2 \theta \cos \theta \cos^3 \phi + r \sin^2 \theta \cos \theta \sin^3 \phi - r \cos^2 \theta \sin \theta$$

$$T^\phi = -\frac{x^2 \sin \phi}{r \sin \theta} + \frac{y^2 \cos \phi}{r \sin \theta} = -r \sin \theta (\sin \phi \cos^2 \phi - \sin^2 \phi \cos \phi)$$

The corresponding covariant components are,

$$T_x = x^2, T_y = y^2, T_z = z^2, \text{ Since } g_{xx} = g_{yy} = g_{zz} = 1$$

The covariant components in the (r, θ, ϕ) coordinates are,

$$T_r = \frac{\partial x}{\partial r} \cdot T_x + \frac{\partial y}{\partial r} \cdot T_y + \frac{\partial z}{\partial r} \cdot T_z = x^2 \sin \theta \cos \phi + y^2 \sin \theta \sin \phi + z^2 \cos \theta,$$

$$T_\theta = \frac{\partial x}{\partial \theta} T_x + \frac{\partial y}{\partial \theta} T_y + \frac{\partial z}{\partial \theta} T_z = x^2 r \cos \theta \cos \phi + y^2 r \cos \theta \sin \phi - z^2 r \sin \theta$$

$$T_\phi = \frac{\partial x}{\partial \phi} T_x + \frac{\partial y}{\partial \phi} T_y + \frac{\partial z}{\partial \phi} T_z = -x^2 r \sin \theta \sin \phi + y^2 r \sin \theta \cos \phi$$

Substituting the expressions for x, y, z in terms of r, θ, ϕ respectively we can verify that

$$T^r = g^{rr} T_r, T^\theta = g^{\theta\theta} T_\theta, T^\phi = g^{\phi\phi} T_\phi$$

Problem 3

(a) Prove that $|\bar{g}| = |g| \left| \frac{\partial x}{\partial \bar{x}} \right|^2$, where g and $\bar{g}$ are the determinants of the metric

tensor g in two systems of coordinates and $\left| \frac{\partial x}{\partial \bar{x}} \right|$ is the Jacobian.

(b) Prove that $\frac{\partial g^{ik}}{\partial x^j} = -g^{hk} \Gamma_{hj}^i - g^{hi} \Gamma_{hj}^k$

Solution:

(a) We have, $\bar{g}_{ij} = g_{\alpha\beta} \frac{\partial x^\alpha}{\partial \bar{x}^i} \cdot \frac{\partial x^\beta}{\partial \bar{x}^j}$.

Taking the determinant of both sides we get,

$$|g| = |\bar{g}| \left| \frac{\partial x}{\partial \bar{x}} \right|^2$$

(b) Since the covariant derivative $g_{ij}^{ik} = 0$ we get,

$$g_{ij}^{ik} = -g^{hk} \Gamma_{hj}^i - g^{hi} \Gamma_{hj}^k.$$

Problem 4

Prove that $A_{i,j} - A_{j,i}$ transforms like a tensor.

Solution: $A_{i,j} - A_{j,i} = (A_{i,j} - A_{j,i}) - (\Gamma_{ij}^m - \Gamma_{ji}^m) A_m$

Since $\Gamma_{ij}^m = \Gamma_{ji}^m$ we have, $A_{i,j} - A_{j,i} = A_{i,j} - A_{j,i}$

Since the covariant derivatives $A_{i,j}$ and $A_{j,i}$ transform like the components of a tensor, so does the *l.h.s.* $A_{i,j} - A_{j,i}$.

SUMMARY

- Scalars, vectors and other more-involved objects belong to the category of a general class of mathematical entities called tensors that are characterized by the way they change under transformation of coordinates.

- (a) *Contravariant* – Tensors where components transform under a change of coordinates from x^i to $\bar{x}^i$ given by $\bar{x}^i = \bar{x}^i(x^1, \dots, x^n)$ as,

$$\bar{T}^{ijk} = \frac{\partial \bar{x}^i}{\partial x^\alpha} \frac{\partial \bar{x}^j}{\partial x^\beta} \frac{\partial \bar{x}^k}{\partial x^\gamma} \cdot T^{\alpha\beta\gamma} \text{ are called contravariant tensors.}$$

- (b) *Covariant* – Covariant tensors transform as

$$\bar{T}_{ijk} \dots = \frac{\partial x^\alpha}{\partial \bar{x}^i} \frac{\partial x^\beta}{\partial \bar{x}^j} \frac{\partial x^\gamma}{\partial \bar{x}^k} \dots T_{\alpha\beta\gamma} \dots$$

- (c) *Mixed* – Mixed tensors transform as,

$$\bar{T}_{mn}^{ij} = \frac{\partial \bar{x}^i}{\partial x^\alpha} \cdot \frac{\partial \bar{x}^j}{\partial x^\beta} \dots \frac{\partial x^\alpha}{\partial \bar{x}^m} \cdot \frac{\partial x^\beta}{\partial \bar{x}^n} \dots T_{ab}^{\alpha\beta}.$$

The total number of indices in the tensor is called the rank of the tensor. Thus, a vector and a scalar behave as tensors of ranks 1 and 0 respectively. The coordinate differential dx^i behaves as a contravariant vector while the components of the gradient of a scalar behaves as those of a covariant vector.

- Tensors satisfy the properties similar to those enjoyed by the vectors in a vector space. The characteristic features that emerge from the transformation properties of tensors are as follows.
 - The sum and the difference of two tensors of the same type and the same rank give a tensor of the same type and rank.
 - The direct product of two tensors yields a tensor whose rank is the sum of the ranks of the individual tensors and the type of the product tensor depends on the types of the individual tensors. The type is (a) contravariant, if both be contravariant; (b) covariant, if both be covariant; and (c) mixed, if one be contravariant and the other covariant.
 - If we multiply two tensors, one contravariant of rank m and the other covariant of rank n , make p pairs of contravariant and covariant indices equal and then sum over these, the rank of the resultant mixed tensor becomes $m + n - 2p$. This reduction in rank is called contraction. Examples of contraction are in plenty in different disciplines of physics. For example, the K. E. of a rigid body given by $T = \frac{1}{2} I_{ij} \omega^i \omega^j$ is an example of a scalar obtained by contraction.

- (d) The symmetry properties of a tensor remain invariant under transformation of coordinates. If we multiply two tensors, one contravariant with symmetry (or anti symmetry) between two elements a and b say and the other covariant with antisymmetry (or symmetry) between two elements m and n say and carry out the process of contraction (i.e., make $m = a$, $n = b$ and carry out the sum over the equal pair) the resultant product vanishes.
- (e) If the components of a tensor vanish in one coordinate system they must do so in all the systems. This leads to the result that any tensor equation must remain form-invariant under transformation of coordinates.
- (f) The quotient rule may be stated very crudely as “if $AB = C$ denotes a tensor equation and A , C are tensors then B must be a tensor of type and rank determined by those of A and C ,” for example, if $A_i B^{ik} = C^k$ represents a tensor equation and A_i , C^k denote components of covariant and contravariant tensors of rank 1 respectively, B^{ik} must be the components of a second rank contravariant tensor.
- (g) The square of the infinitesimal line element between two points, a scalar, is given by $ds^2 = g_{ik} dx^i dx^k$, the g_{ik} 's constituting the elements of a second rank covariant symmetric tensor. The contravariant counterparts g^{ik} are related to g_{ik} by $g^{ik} g_{ki} = \delta^i_i$. The g_{ik} 's and g^{ik} 's (called the metric tensor of the space) act as the lowering and the raising operators as illustrated by $A_i = g_{ik} A^k$ and $A^i = g^{ik} A_k$ respectively. The components A_i and A^i , called the covariant and the contravariant components of the vector A are explained in the text.
- (h) A special case worth mentioning here, refers to Cartesian tensors where we are confined to transformations of the components of a tensor under transformation of Cartesian axes. The contravariant and the covariant tensors behave identically in this case.
- (i) The elements of tensor calculus are briefly described in the text. Under general nonlinear coordinate transformation, the derivatives of the components of a tensor do not in general behave like a tensor. The reason is the change of the directions of the basis vector in going from one point to another in spaces. When this change is accounted for in the transformation we obtain the expressions for the covariant derivatives of the components of a tensor which behave like tensors. The expressions for the derivatives of the different types of tensors involve the Christoffel symbols (also called affine connection coefficients) which do not themselves behave like tensors and are given in the text. The necessary and sufficient condition for a space being intrinsically curved is that the curvature tensor components constructed out of the Christoffel symbols and their derivatives vanish everywhere in space.
- (j) The expressions for differential operators of tensor calculus in general coordinate systems are discussed. Also, the integral relations that are equivalents of Stokes and the Gauss theorems of vector calculus are discussed for tensor fields in general coordinate systems without derivations of the formulae that are given in standard texts.

MULTIPLE CHOICE QUESTIONS

1. Let there be two vector $\bar{\nabla}\phi$ and the velocity vector $\bar{v}$. Which of the following statements is correct?
 - (a) $\bar{\nabla}\phi$ behaves like a contravariant and $\bar{v}$ like a covariant vector.
 - (b) $\bar{\nabla}\phi$ behaves like a covariant and $\bar{v}$ like a contravariant vector.
 - (c) Both behave like contravariant vectors.
 - (d) Both behave like covariant vectors.

2. If A^{ij} and B_{kl} denote the contravariant and the covariant tensors of rank 2 respectively, then $A^{ij} A^{kl}$ and $A^{ij} B_{kl}$ transform
 - (a) like the contravariant and the mixed tensors of rank 4 respectively,
 - (b) like the contravariant and the covariant and the covariant tensors of rank 4 respectively,
 - (c) both like the contravariant tensors of rank 4
 - (d) both like the covariant tensors of rank 4.
3. $\sum_{ij} [(A^i B^j + A^j B^i) C^k] [(D_i E_j - D_j E_i) F_l]$ gives
 - (a) a second rank mixed tensor.
 - (b) a second rank contravariant tensors
 - (c) a second rank covariant tensor
 - (d) zero.
4. The spatial derivative of a contravariant vector behaves
 - (a) like a contravariant tensor of rank 2
 - (b) like a mixed tensor of rank 2 if the transformation be linear
 - (c) always like a non-tensorial object
 - (d) like a covariant tensor of rank 2.
5. The number of independent components of a symmetric and an antisymmetric tensor of rank 4 is
 - (a) 6 in both the cases
 - (b) 10 and 6 respectively
 - (c) 6 and 10 respectively
 - (d) none of these.
6. There are three objects $\bar{A} \cdot \bar{B}$, $(\bar{A} \times \bar{B})$, $\bar{A} \cdot (\bar{B} \times \bar{C})$. Which of the following statements is correct?
 - (a) The first two are proper tensors of rank 0 while the last is a proper tensor of rank 1.
 - (b) The first one is a proper tensor of rank 0 while the last two are pseudotensors of ranks 1 and 0 respectively.
 - (c) The first two are proper tensors of rank 0 while the last one is a pseudotensor of rank 1.
 - (d) The three are proper tensors of ranks 0, 0 and 1 respectively.
7. $\sum_i A^{ijl} B_{ik}$ behaves like
 - (a) a mixed tensor of rank 5.
 - (b) a contravariant tensor of rank 3.
 - (c) a mixed tensor of rank 3
 - (d) a covariant tensor of rank 5.
8. The non-vanishing components of a metric tensor g_{ik} are given by

$$ds^2 = g_{ik} dx^i dx^k, \text{ where } ds^2 = c^2 dt^2 - \frac{1}{(1 - kr^2)} [dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)]$$

The components g^{ik} are given by

$$(a) \quad g^{11} = -(1 - kr^2), \quad g^{22} = \frac{-(1 - kr^2)}{r^2}, \quad g^{33} = \frac{-(1 - kr^2)}{r^2 \sin^2 \theta}, \quad g^{44} = \frac{1}{c^2}$$

$$(b) \quad g^{11} = -\frac{1}{(1-kr^2)}, g^{22} = -\frac{r^2}{(1-kr^2)}, g^{33} = -\frac{(1-kr^2)}{r^2 \sin^2 \theta}, g^{44} = \frac{1}{c^2}$$

$$(c) \quad g^{11} = -\frac{1}{(1-kr^2)}, g^{22} = -\frac{(1-kr^2)}{r^2}, g^{33} = -\frac{(1-kr^2)}{r^2 \sin^2 \theta}, g^{44} = c^2$$

$$(d) \quad g^{11} = -(1-kr^2), g^{22} = -\frac{(1-kr^2)}{r^2}, g^{33} = -\frac{r^2 \sin^2 \theta}{(1-kr^2)}, g^{44} = \frac{1}{C^2}$$

9. The components of a tensor of rank 2 are $T_{11} = 1$, $T_{22} = 3$, $T_{12} = 2$, $T_{21} = 0$. The coordinate axes are rotated through 60° about the z -axis. The new components T'_{11} and T'_{22} are given by

$$(a) \quad T'_{11} = 5/2, T'_{22} = 3/2$$

$$(b) \quad T'_{11} = \frac{5+\sqrt{3}}{2}, T'_{22} = \frac{3-\sqrt{3}}{2}$$

$$(c) \quad T'_{11} = \frac{5-\sqrt{3}}{2}, T'_{22} = \frac{3+\sqrt{3}}{2}$$

$$(d) \quad T'_{11} = -\left(\frac{5+\sqrt{3}}{2}\right), T'_{22} = \left(\frac{3-\sqrt{3}}{2}\right)$$

10. The metric tensor g^{ik} and the Christoffel symbol Γ_{jk}^i are such that

$$(a) \quad g^{ik} = g^{ki}, \Gamma_{jk}^i = -\Gamma_{kj}^i \text{ and both are tensors}$$

$$(b) \quad g^{ik} = g^{ki}, \Gamma_{jk}^i = \Gamma_{jk}^i \text{ and both are tensors.}$$

$$(c) \quad g^{ik} = g^{ki}, \Gamma_{jk}^i = \Gamma_{jk}^i \text{ and } \Gamma_{jk}^i \text{ does not behave like a tensor while } g_{ik} \text{ does}$$

$$(d) \quad g^{ik} = g^{ki}, \Gamma_{jk}^i = \Gamma_{jk}^i \text{ and neither of them behaves like a tensor}$$

Answers to MCQs

1. (b)	2. (a)	3. (d)	4. (b)	5. (b)	6. (b)	7. (c)
8. (a)	9. (b)	10. (c)				

EXERCISE

- (a) The Cartesian components of a tensor are $T_{11} = T_{22} = T_{33} = 1$ and $T_{12} = 2$; the other T_{ij} 's being zero. Determine the components after
 - the 30° rotation about the y -axis,
 - the transformation $x' = -x$, $y' = -y$, $z' = -z$.
- (b) Let $x' = \frac{1}{\sqrt{2}}(x+y)$, $y' = z$, $z' = \frac{1}{\sqrt{2}}(x-y)$, find out T_{22} , T_{12} if $T'_{11} = T'_{22} = T'_{33} = 1$, the other T'_{ijk} being zero.
- Break up the components of a second rank tensor T_{ij} into symmetric and antisymmetric parts. Prove that they transform independently under homogeneous linear transformations.

3. Prove that $T_{ij,k} + T_{jk,i} + T_{ki,j}$ transforms like a tensor.
4. Find out the components of the metric tensor and the Christoffel symbols in (i) Cartesian (ii) Cylindrical and (iii) Spherical polar coordinates.
5. (a) Prove that the Minkowski space time is flat.

[Hints: Prove that the components of the curvature tensor given in section 4.5(3) vanish].

- (b) Prove that $(g_{\alpha\beta} T^\alpha T^\beta)_{ij}$

$$= g_{\alpha\beta} T_{ij}^\alpha T^\beta + g_{\alpha\beta} T^\alpha T_{ij}^\beta.$$

6. Show that the geodesic on a spherical surface is a great circle.

[Hints: Write $ds^2 = R^2 (d\theta^2 + \sin^2\theta d\phi^2)$, R being the radius. Expressing ds^2 as

$$R^2 \left[\left(\frac{d\theta}{d\lambda} \right)^2 + \sin^2\theta \left(\frac{d\phi}{d\lambda} \right)^2 \right] (d\lambda)^2,$$

λ being an arbitrary parameter along the world line find out the extremum of $\left(\frac{d\theta}{d\lambda} \right)^2 + \sin^2\theta \left(\frac{d\phi}{d\lambda} \right)^2$ i.e., solve the equations,

$$\frac{d}{d\lambda} \left[\frac{\partial f}{\partial \left(\frac{d\theta}{d\lambda} \right)} \right] = \frac{\partial f}{\partial \theta} \text{ and } \frac{d}{d\lambda} \left[\frac{\partial f}{\partial \left(\frac{d\phi}{d\lambda} \right)} \right] = \frac{\partial f}{\partial \phi}, \text{ where}$$

$$f = R^2 \left(\frac{d\theta}{d\lambda} \right)^2 + R^2 \sin^2\theta \left(\frac{d\phi}{d\lambda} \right)^2.$$

7. Discuss a few examples in physics where you can illustrate the contraction of tensors. (search for such examples in the fields of rigid body dynamics, continuum mechanics, relativity etc).

Hints and Answers to Exercise Questions

$$1. (a) (i) T'_{11} = \frac{3}{4}T_{11} + \frac{\sqrt{3}}{4}T_{12}, T'_{12} = \frac{1}{4}(3T_{22} - \sqrt{3}T_{11}),$$

$$T'_{21} = \frac{1}{4}(\sqrt{3}T_{22} - T_{12} - \sqrt{3}T_{11}), T'_{22} = \frac{1}{4}[(T_{11} + 3T_{22} - \sqrt{3}(T_{12} + T_{21}))],$$

the other components being zero

$$(ii) T'_{11} = -T_{11}, T'_{22} = -T_{22}, T'_{33} = -T_{33} \text{ the other components being zero.}$$

$$(b) T_{12} = 0, T_{22} = 1.$$

Infinite Series

CHAPTER

5

5.1 INTRODUCTION

The study of infinite series is highly relevant from the point of view of physical applications, because (i) frequently we come across expressions involving an infinite series of numbers or functions and to make some meaningful conclusion we are required to know about the convergence or divergence of the series, (ii) the solution of a differential equation in the form of an infinite series has to be necessarily investigated to ascertain the convergence of such a series, (iii) the question of convergence is important for carrying out a term-by-term integration or differentiation of an infinite series. (iv) the convergence or divergence of integrals also deserves attention in many situations. In this chapter we will discuss the different types of convergence arising in the infinite series of numbers or functions, omitting the proofs in many cases for the sake of brevity of space and citing the references when needed. The topic (iv) is briefly discussed in the appendix.

Let us first consider a sequence of numbers $t_0, t_1, \dots, t_n, \dots$ and consider the partial sums $s_0 (= t_0), s_1 (= t_0 + t_1), \dots, s_n = \sum_{m=0}^n t_m$. The infinite sequence (s_n) is said to converge to a unique sum s as $n \rightarrow \infty$ if given a number ϵ such that for $|s - s_n| < \epsilon$

we can find out a $N(\epsilon)$ for which $n > N(\epsilon)$. The remainder after n terms thus should tend to zero as $n \rightarrow \infty$. A necessary condition for this to happen is that $t_n \rightarrow 0$ as $n \rightarrow \infty$. But this is not clearly sufficient because examples can be given of the infinite series for which $t_n \rightarrow 0$ but which do not converge. The series $\sum_{n=1}^{\infty} \frac{1}{n}$ offers such an example. The series $\sum_n \frac{1}{n^2}$ or $\sum_n \frac{(-1)^{n-1}}{n}$ can be proved to be convergent on the other hand.

Before going into a detailed analysis regarding the different types of convergence of a series of numbers or functions let us discuss Cauchy's test that gives a necessary

and sufficient condition for the convergence of a series. The necessary and sufficient condition for convergence according to Cauchy is that if for two integer numbers n and $m > N(\epsilon)$ where N is any integer depending on ϵ , a preassigned small positive number, $|s_n - s_m| < \epsilon$ and $|s_n - s_m| \rightarrow 0$ as $N \rightarrow \infty$, s_n and s_m being the n th and the m th partial sums respectively, the series is convergent.

Let us take, for example, the series $\sum_n \frac{1}{n}$.

The m th partial sum $s_m = 1 + \frac{1}{2} + \dots + \frac{1}{m}$ can be written as, $s_m = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots$

or, $s_m > 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots$ to m terms or $> \frac{(m+1)}{2}$

Clearly, $|s_m - s_p|$ can be made to increase as m and p increase and cannot be made less than any small number ϵ . The series is, therefore, divergent. It is to be noted that according to the definition of convergence, the behaviour of the first few terms of an infinite series does not matter in deciding the convergence.

5.2 CONVERGENCE OF A SERIES OF NUMBERS

There are two different types of convergence relevant in the discussion of the convergence of a number-series.

1. *Absolute convergence*: If $\sum_{n=1}^{\infty} |t_n|$ is convergent the series $\sum_n t_n$ is said to be absolutely convergent.
2. *Conditional convergence*: If $\sum_{n=1}^{\infty} t_n$ converges while the series $\sum_{n=1}^{\infty} |t_n|$ does not converge, the series

is said to be conditionally convergent. We will present first the test for conditional convergence because for absolute convergence we have a lot of tests to choose from, which we will discuss briefly later.

Leibnitz put forward a test for ascertaining the conditional convergence of an alternating series (i.e., a series with alternately positive and negative terms).

The condition can be expressed as: If the terms of an alternating series progressively decrease in magnitude, i.e., if $|t_n| > |t_{n+1}|$ and $\lim_{n \rightarrow \infty} |t_n| = 0$, the series is convergent.

The test is of importance because in many cases a series with all terms (not of the same sign) can be transformed to an alternating series by a rearrangement of terms.

Proof: (i) Let us consider the even partial sum.

$$\begin{aligned} \text{Let } s &= t_1 - t_2 + t_3 - t_4 + \dots \\ &= t_1 - (t_2 - t_3) - (t_4 - t_5) + \dots \end{aligned}$$

If $t_2 > t_3$, $t_4 > t_5$ etc., $s < t_1$.

Again, the series can be rewritten as $(t_1 - t_2) + (t_3 - t_4) + \dots$

Here, the sum $s > 0$, progressively increases with the increase of the number of terms but is bounded above by t_1 . Thus the series is convergent. Similarly we can prove that the odd partial sum is also convergent.

Applying Leibnitz criterion to the case of the series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$, we can see that this series is convergent.

For a conditionally convergent series, the sum, even if it exists, is not unique in general.

For example, if we write the sum $\sum_n \frac{(-1)^{n-1}}{n}$ by rearrangement of its terms as,

$$\begin{aligned} s &= \left(1 - \frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{6} - \frac{1}{8}\right) + \dots \\ &= \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{6} - \frac{1}{8}\right) + \dots \end{aligned}$$

The terms on the right hand side of the equality sign clearly converge to a sum $\frac{1}{2}s$. This is not all.

A conditionally convergent series may, by a suitable rearrangement of its terms, be made to approach any desired value and even to diverge. But for an absolutely convergent series the sum is unique irrespective of the ordering of the terms.

5.2.1 Tests of Absolute Convergence

- (1) **Comparison Test** – If $\sum_n t_n$ be a series of positive (or negative) terms and $\sum_n u_n$ be another such series which is known to be convergent, then the series $\sum_n t_n$ converges if $|t_n| \leq |u_n|$ for all n . Nothing can be said about the convergence of $\sum_n t_n$ if $|t_n| > |u_n|$.

Again if $\sum_n u_n$ be a divergent series and $|t_n| > |u_n|$ for all n , $\sum_n t_n$ is also divergent. Again nothing can be said regarding the convergence of $\sum_n t_n$ if $|t_n| < |u_n|$ in this case.

Example We consider the geometric series $\sum_{n=\infty}^{\infty} r^n$ and the series $\sum_n 1$. Clearly for $|r| \geq 1$, $|r^n| \geq |t_n|$ (> 1) for all n . Since the series $\sum_n 1$ is known to be divergent, the series $\sum_n r^n$ diverges for $|r| \geq 1$, $\left|\frac{\sin nx}{n^2}\right| \leq \frac{1}{n^2}$ for all n . Since $\sum_A \frac{1}{n^2}$ is convergent, $\sum_n \frac{\sin nx}{n^2}$ also converges. The test below shows that $\sum_n r^n$ converges for $|r| < 1$.

- (2) **Ratio Test** – For a series $\sum_n u_n$ of positive or negative terms the series is convergent or divergent according as $\lim_{n \rightarrow \infty} \left|\frac{u_{n+1}}{u_n}\right| < \text{or} > 1$ respectively. If $\lim_{n \rightarrow \infty} \left|\frac{u_{n+1}}{u_n}\right| = 1$, the test fails to predict anything.

The logic behind such a statement can be understood by writing the remainder after N terms as:

$R_N = u_N + u_{N+1} + \dots = u_N(1 + m + m^2 + \dots)$, where we put $u_{N+1} = mu_N$, $u_{N+2} = mu_{N+1}$, ... or, $R_N = u_N(1 + m + m^2 + \dots)$ ($m = \text{any constant for } n > N$)

The geometric series $1 + m + m^2 + \dots$ converges if $m < 1$. (This can be independently proved by writing the expression for the sum of a geometric series $a + ar + \dots + ar^{n-1} = \frac{a(1-r^n)}{1-r}$ because as $n \rightarrow \infty$ the series converges for $r < 1$ to $\frac{a}{1-r}$).

Thus, R_N converges if $m < 1$ and clearly tends to zero as $N \rightarrow \infty$ because $u_N \rightarrow 0$. If $m > 1$, the series clearly diverges.

Example

(1) Let $u_n = \frac{1}{n!}$

$$\left| \frac{u_{n+1}}{u_n} \right| \sim \frac{1}{n+1} \text{ and } \rightarrow 0 \text{ as } n \rightarrow \infty$$

If we take $\left| \frac{u_{n+1}}{u_n} \right| < \epsilon$ for $n > N$, we can take $N \geq \frac{1}{\epsilon} - 1$.

(2) For $u_n = \frac{1}{n^p}$ (p being any number independent of n) we get, $\left| \frac{u_{n+1}}{u_n} \right| \sim \left(\frac{n}{n+1} \right)^p$ and $\rightarrow 1$ as $n \rightarrow \infty$.

The ratio test therefore fails to predict anything regarding the convergence of $\sum_n u_n$. Below, we discuss a test called the integral test that comes to our rescue in such cases.

Integral test — Let the series be represented by $\sum_n t_n$. Cauchy's integral test states that if we consider t_n to be replaced by $t(n)$, the function of a continuous variable n and $\int_c^\infty t(n) dn$ exists, the series $\sum_n t_n$ converges.

The lower limit C is arbitrary except that the integral must not diverge at the lower limit.

Proof:

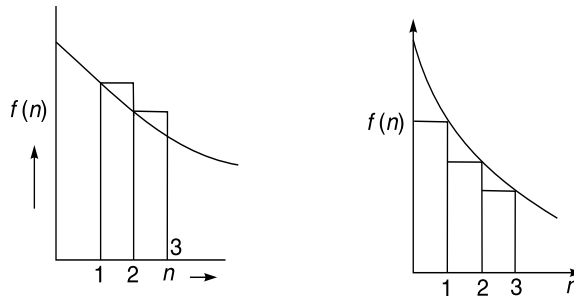


Fig. 5.1

We see from Fig. 5.1 that,

$$\sum_{n=1}^{\infty} t_n > \int_1^{\infty} f(n) dn \text{ and also } \sum_{n=1}^{\infty} t_n - t_1 < \int_1^{\infty} f(n) dn$$

Combining the two we obtain that

$$\int_1^{\infty} f(n) dn \leq \sum_{n=1}^{\infty} t_n \leq \int_1^{\infty} f(n) dn + t_1$$

Hence, if the integral $\int_1^{\infty} f(n) dn$ exists, $\sum_{n=1}^{\infty} t_n$ clearly converges.

There are other tests regarding convergence or divergence of a series of numbers, but since the tests described here suffice in most of the cases to ascertain the convergence or the divergence of a series, we do not discuss these here.

5.3 CONVERGENCE OF A SERIES OF FUNCTIONS

The problem of convergence of a series of functions is more varied compared to that of a series of numbers. The problem can be divided into two classes.

1. Point convergence
2. Convergence in the mean

It is also to be remembered that this discussion applies, in general, to the function of a complex variable.

1. Point Convergence

Here, we consider the problem of convergence of the series for every point in the interval of definition of the variable concerned (or more than one variable in the case of a function of several variables). Apart from the usual problems of convergence discussed in Section (5.2) the discussion includes the problem of uniform convergence that has no analogue in the case of a series of numbers. The ratio test which often fails to predict anything regarding convergence of numbers applies almost universally in the case of a series of functions.

Examples

- (i) The series $\sum_n \frac{z^n}{n}$ converges absolutely, as can be seen by the application of the ratio test for $|z| < 1$.

For z , a complex number, the inequality $|z| < 1$ implies the inside of a circle of radius 1, centered at the origin. The circle is called the circle of convergence with radius of convergence unity. Considering z to

be a function of x and y , the region implies $|x| < 1$, and for each x in that region, $|y| < \sqrt{1-x^2}$. The series

$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{z^{n-1}}{n}$, converges absolutely for $|z| < 1$ and is easily seen to be conditionally convergent for $z = 1$.

- (ii) The series $\sum_n \frac{\sin nx}{n^2}$ is seen to be absolutely convergent for all finite x by comparison with the series

$\sum_n \frac{1}{n^2}$ which is known to be absolutely convergent and each term of which is greater than or equal to that of the first series.

Let us now consider the following two functional sequences

- (i) $f_n(x) = \frac{\sin nx}{n^2}$ for all x and (ii) $f_n(x) = x^n$ for $0 \leq x \leq 1$.

The first one converges absolutely to zero for all x while the second one converges to zero for $0 < x < 1$ but becomes 1 for $x = 1$.

Since $|\sin nx| \leq 1$, we can choose $|f_n(x)| < \epsilon$ for any $n > N(\epsilon)$, where ϵ is any preassigned positive number however small it may be and $N(\epsilon) = \frac{1}{\sqrt{\epsilon}}$ for all x .

Such a type of sequence where $f_n(x) \rightarrow f(x)$ for $n \rightarrow \infty$, so that $|f_n(x) - f(x)| < \epsilon$ for any $n > N(\epsilon)$, $N(\epsilon)$ being independent of x for any chosen ϵ is called a uniformly convergent sequence. (An identical definition for a uniformly convergent series also follows similarly). $f(x)$ is equal to zero for the sequence

$$f_n(x) = \frac{\sin nx}{n^2} \quad (f(x) = \lim_{n \rightarrow \infty} f_n(x)).$$

- (ii) The second sequence x^n differs from the one discussed in (i) in the sense that here there is no finite $N(\epsilon)$ independent of x for $0 < x \leq 1$ so that $|x^n| < \epsilon$ because n has to be greater than $\frac{\ln \epsilon}{\ln x}$. The sequence is therefore non-uniformly convergent, N being dependent on x .

However, if we exclude the point $x = 1$ for any η so that $0 < x \leq 1 - \eta$, we can choose $N = \frac{\ln \epsilon}{\ln(1 - \eta)}$ which is clearly independent of x for a given η . The sequence becomes then uniformly convergent.

The study of a uniformly convergent series (or sequence) is significant because it enlightens us about the algebraic properties of the infinite series (or sequence) of function stated below. (The proofs, omitted here, can be found in standard texts).

- (1) The sum of an infinite series of continuous and uniformly convergent functions is continuous at every point where each of the functions is continuous and uniformly convergent, and for every such point $x = x_0$ we have,

$$\lim_{x \rightarrow x_0} \sum_n f_n(x) = \sum_n \lim_{x \rightarrow x_0} f_n(x) = \sum_n f_n(x_0)$$

This equality is valid also for a function of many variables.

- (2) The sum of an infinite series of uniformly convergent functions is integrable in a domain of definition of the arguments where each of the functions is integrable and the magnitude of the integral equals that obtained by integrating term-by-term the functions and taking the sum.

Thus, $\int_a^b \sum_n f_n(x) dx = \sum_n \int_a^b f_n(x) dx$, where $f_n(x)$ for all n is integrable within the interval. The integrated

series has a domain of uniform convergence equal to that for every series.

Example Let $f(x) = \frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots$

Since the series $\sum_n (-1)^{n-1} x^{2n}$ is uniformly convergent in the interval $0 < x < 1$, we have from (2) above,

$$\int_0^1 \frac{1}{1+x^2} dx = \left[x - \frac{x^3}{3} + \frac{x^5}{5} - \dots \right]_0^1 = 1 - \frac{1}{3} + \frac{1}{5} - \dots$$

The series $\sum_n (-1)^{n-1} \frac{x^{2n-1}}{2n-1}$ is uniformly convergent in the same interval and converges to $f(x) = \tan^{-1} x$

$$\text{Thus } \int_0^1 \frac{dx}{1+x^2} = \tan^{-1} x \Big|_0^1 = \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \dots$$

- (3) However, in the case of differentiation of an infinite series of uniformly convergent functions the condition is more stringent. Thus if $\sum_n f_n(x)$ and $\sum_n \frac{df_n(x)}{dx}$ are uniformly convergent for all x in an interval (a, b) , then for $f(x) = \sum_n f_n(x)$,

$$\frac{d}{dx} f(x) = \sum_n \frac{df_n(x)}{dx}.$$

The uniformly convergent series expansion of a function in a certain interval of values of the argument enabling us to carry out a termwise integration of the series facilitates the evaluation of the integral of the function (when the integral cannot be determined in a closed form) to a certain extent of accuracy.

2. Tests of Uniform Convergence

We will enlist below certain tests of ascertaining if a series is uniformly convergent without giving proofs,

- (a) *Weirstrass test* – If we can find out a series of numbers $\sum_n s_n$ which is convergent and if $|u_n(x)| \leq |s_n|$ for all x in a certain interval, the series $\sum_n u_n(x)$ is uniformly convergent in the same interval.

Thus $\sum_n \frac{\sin nx}{n^2}$ ($-\infty < x < \infty$) and $\sum_n \frac{(-1)^{n-1} x^n}{n}$ ($0 \leq x \leq 1$) are both uniformly convergent because the first one is dominated by the series $\sum_n \frac{1}{n^2}$ while the second one by $\sum_n \frac{(-1)^{n-1}}{n}$, both of them being convergent in the specified interval.

- (b) *Abel's test* – If we write $u_n(x) = a_n f_n(x)$, where $\sum_n a_n = \text{finite}$ and $f_n(x)$ is a bounded monotonically non-increasing sequence so that $|f_n(x)| \leq p$, some fixed number for all x in the interval, $\sum_n u_n(x)$ converges uniformly for x lying in the interval. Following the test we can thus say that

(1) the power series $\sum_n \frac{x^n}{n!}$ converges uniformly to e^x for all $|x| < 1$ and

(2) the series $\sum_n \frac{\sin nx}{n^s}$ is uniformly convergent for all x and $s > 1$.

- (c) *Cauchy's test* – Cauchy's test of uniform convergence is very similar to that discussed at the outset in connection with the test for convergence of a series of numbers with the exception that the partial sums become functions of the variable x also and the condition has to be similarly modified. The Abel's test can be shown to be a consequence of Cauchy's test.

5.4 POWER SERIES

An infinite series of the form $\sum_n a_n (x - x_0)^n$, the a_n 's being constant coefficients is termed a power series about the point $x = x_0$. The power series expansion occurs frequently in many problems and also the discussions in the earlier sections are exemplified by referring to this expansion.

Every power series is convergent in general, (see **Appendix** at the end of the chapter) within a finite or an infinite region specified by $|x - x_0| \leq R$, where R depends on the nature of the coefficients a_n and is called the radius of convergence of the series. From the ratio test we can easily see that R for the series $\sum_n a_n (x - x_0)^n$

is given by $\lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right|$. Thus R for the exponential series $\sum_n \frac{x^n}{n!}$ is infinite while for the binomial series

$\sum_n x^n$ is less than unity. The power series expansion is valid for a function of many variables and specifically for a function of a complex variable. The region of convergence is called the region of analyticity of the function. If we take $f(z) = \frac{1}{z-1}$ and expand it about $z = 0$, we obtain,

$$f(z) = -1 (1 - z)^{-1} = -(1 + z + z^2 + \dots).$$

The series expansion is valid for $|z| < 1$. The series diverges at $|z| = 1$, where the function is no longer analytic and is said to have a singularity. There may be other types of singularities of a function which we will discuss later in the chapter on complex variables. If we expand the same function $\frac{1}{z-1}$ about $z = 2$, we obtain, $f(z) = \frac{1}{(z+h)-1}$, where we have put $z = 2 + h$, and get, $f(z) = (1+h)^{-1}$

or, $f(z) = 1 - (z-2) + (z-2)^2 - \dots$ and the expansion is valid for $|z-2| < 1$ or $1 < z < 3$. Clearly, the expansion about any point $z = z_0$ is valid within a circle of convergence $|z - z_0| < 1$, i.e., upto the point of singularity $|z| = 1$. By varying the point $z = z_0$ we would obtain different regions of analyticity of the function. By choosing the point $z = z_0$ in different ways we can extend the region of analyticity to the entire z -plane except

$z = 1$. If we want to expand a function $f(z)$ about $z = z_0$ where $z_0 \rightarrow \infty$, we put $z = \frac{1}{z'}$, and expand $\phi(z') = f\left(\frac{1}{z'}\right)$ about $z' = 0$. If we have more than one points of singularity of $f(z)$, the region of analyticity will be determined by the distance of z_0 from the nearest singularity. Thus if we expand $f(z) = \frac{1}{(z-1)(z-2)}$ about

$z = z_0 = 3$, say, the nearest singularity is at $z = 2$ and hence the region of analyticity will be $|z-3| < 1$, i.e., it will lie within a circle centered at $z = 3$ and radius unity. The determination of the region of validity of the power series expansion is important because it can be shown that within this region the series is uniformly convergent and term by term integration or differentiation (when the derivative series is also uniformly convergent in the same region) of the series is thereby valid (See "Appendix" at the end of the chapter).

It can be shown from elementary calculus that a single-valued function $f(x)$ possessing all order derivatives at a point $x = a$ can be expanded in the form of a series about the point in a certain region of convergence as,

$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!} f''(a) + \dots + R_{n+1}$, where R_{n+1} is the remainder after $n+1$ terms given as,

$$R_{n+1} = \frac{(x-a)^n}{n!} f^n \{a + \theta(x-a)\}, 0 < \theta < 1.$$

If $\lim_{n \rightarrow \infty} R_{n+1} = 0$, we have the expansion given by

$f(x) = \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!} f^n(a)$, known as the Taylor's series (A brief derivation of the Taylor's series expansion is given at the end). Thus, $f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$, where $c_n = \frac{f^n(a)}{n!}$, gives a power series expansion of $f(x)$ about $x = a$.

Since different powers of $(x-a)$ are linearly independent, it is obvious that the Taylor series is the unique power series representation of $f(x)$ about a point. $x = a$, [if $f(x) = \sum_n a_n (x-x_0)^n = \sum_n b_n (x-x_0)^n$,

$$\sum_n (a_n - b_n) (x-x_0)^n = 0, \text{ yielding } a_n = b_n \text{ for all } n].$$

Symbolically, $f(x) = \left\{ \exp \left[(x-a) \frac{d}{dx} \right] f(x) \right\}_{x=a}$ for $|x-a| < \text{some } R$, R being the radius of convergence of the series.

For a function of many variables we can similarly write,

$$f(x) = \left\{ \exp \left[(x-a) \frac{\partial}{\partial x} + (y-b) \frac{\partial}{\partial y} + \dots \right] f(x, y, \dots) \right\}_{x=a, y=b}$$

for $|x-a| \leq R_1, |y-b| \leq R_2, \dots$ ($R_1, R_2, \dots$ denote the corresponding radii of convergence)

If we put $a, b, c, \dots = 0$ we get the Maclaurin's series expansion of the function, $f(x, y, \dots)$ about the origin. The Maclaurin's series expansion of $\sin x, e^x, (1+x)^{-1}$ give respectively,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \quad (0 < x < \infty)$$

$$e^x = 1 + x + \frac{x^2}{2!} + \dots \quad (0 < x < \infty)$$

$$(1+x)^{-1} = 1 - x + x^2 - \dots \quad (-1 < x < 1).$$

Note: For a multivalued function we have to (i) where it is singlevalued and (ii) avoid the branch point which is a point of singularity of the multivalued function. Thus for $z^{1/2}$ we avoid $z = 0$ and choose any branch, say, $0 < \theta < 2\pi$, where $z = re^{i\theta}$ choose a particular branch for Taylor's series expansion.

5.5 CONVERGENCE IN THE MEAN

In contrast to the pointwise convergence of an infinite series or sequence of functions (absolute, conditional or uniform) discussed earlier, we will be concerned here with the overall behaviour of the series or sequences over an interval of definition of the variable. If a series of functions $f_n(x)$ converges in the mean to a function

$f(x)$ we can write, $\int_a^b \left| f(x) - \sum_{n=1}^m f_n(x) \right|^2 dx \rightarrow 0$ as $m \rightarrow \infty$, where the interval of x is chosen to be $a \leq x \leq b$.

An identical definition follows in the case of convergence of a sequence of functions.

The series of functions convergent in the mean to a function defined over an interval does not necessarily converge to the function at every point of the interval. The function may even be discontinuous at several points of the interval as is exemplified in the Fourier series expansion of a piecewise continuous function in an interval to be discussed in a subsequent chapter on integral transforms. Similarly, a sequence of functions converging at every x may not converge in the mean. For example, the sequence of functions

$$f_n(x) = \sqrt{2nx} e^{-\frac{1}{2}nx^2} \rightarrow 0 \text{ for every } x \text{ as } n \rightarrow \infty \text{ in the closed interval } (0, 1), \text{ while the integral}$$

$$\int_0^1 [f_n(x) - f(x)]^2 dx = \int_0^1 2nx e^{-nx^2} dx = (1 - e^{-n}) \rightarrow 1 \text{ as } n \rightarrow \infty. \text{ Thus the sequence does not converge in the}$$

mean to $f(x)$. However, a uniformly convergent series or sequence can be shown to converge in the mean.

The idea of convergence in the mean plays an important role in the expansion of a function in an interval by a set of functions. (The Fourier series is an example in this regard). A set of functions $\phi_1(x), \phi_2(x) \dots$ is said to be

complete in the expansion of a function $f(x)$ in an interval $a \leq x \leq b$, if $\int_a^b \left| f(x) - \sum_{n=1}^m c_n \phi_n(x) \right|^2 dx \rightarrow 0$ if $m \rightarrow \infty$, where the c_n 's are certain constants depending on the forms of $f(x)$ and $\phi_n(x)$ respectively. If the $\phi_n(x)$'s are mutually orthogonal and normalized, i.e., $\int_a^b \phi_n^* \phi_m dx = \delta_{nm}$, then the c_n 's are given by $c_n = \int_a^b \phi_n^*(x) f(x) dx$.

5.6 SUMMABILITY OF A DIVERGENT SERIES

Although the sum of a divergent series does not exist in the conventional sense, there have been attempts made by C'esarò, Borel etc. to deduce an expression for the sum of such a series by consistently modifying the definition of the sum. We will briefly state here C'esarò's summability condition without giving the proofs that can be seen in standard references.

C'esarò's condition (C1) states that if $s_n = \sum_{m=1}^n t_m$ does not converge but $\frac{1}{n} (s_1 + \dots + s_n)$ converges to a sum s , then $\sum_n t_n \rightarrow s$. Therefore, not all types of divergent series can be summed by C'esarò's summability criterion.

Example We take the oscillatory series $s_n = 1 - 1 + 1 - 1 + \dots$. Clearly $s_n = \frac{1}{2} [1 - (-1)^n]$ so that the sum of an even number of the S_n 's converges to $\frac{1}{2}n$ while that of an odd number of the S_n 's converges to $\frac{1}{2}(n+1)$. Hence for n either odd or even, the sum $\frac{1}{n} \sum_{m=1}^n s_m$ converges to $\frac{1}{2}$.

Note: There are higher order C'esarò's summability conditions (Cr) but they are not discussed here.

5.7 COMBINATION OF TWO INFINITE SERIES

The rules regarding the composition of more than one infinite series which occurs frequently in many applications are the following:

- (i) The sum or difference of two infinite series is an infinite series with the region of convergence common to the regions of convergence of the two series and the result converging to the sum (or difference) of each individual series. The proof is straightforward and follows from the definition of convergence of an infinite series.
- (ii) Regarding the product or division of one infinite series by the other we need be more cautious. The product of two conditionally convergent series may yield a divergent series. But if we have two absolutely (or uniformly) convergent series the result of multiplying or dividing one series by the other yields a convergent series converging to a value equal to the product or ratio of two individual sums with the region of convergence where each series converges (In the case of the ratio, the denominator series must also not vanish at the point of interest).

The statements made above may be illustrated well by referring to a power series.

- (i) If we have two series $\sum_n a_n x^n$ and $\sum_n b_n x^n$ converging to s_1 and s_2 respectively for $|x| \leq R_1$ and $|x| \leq R_2$ respectively, then $\sum_n (a_n \pm b_n) x^n$ is an infinite series $\sum_n C_{n\pm} x^n$, convergent for $|x| \leq$ the smaller of R_1 and R_2 ($C_{n\pm} = a_n \pm b_n$) and converging to $s_1 \pm s_2$ respectively.
- (ii) For the product of two series similarly,

$$\sum_n a_n x^n \rightarrow s_1 \text{ and } \sum_n b_n x^n \rightarrow s_2 \text{ yields}$$

$$\left(\sum_n a_n x^n \right) \left(\sum_n b_n x^n \right) = \sum_n C_n x^n \text{ converging to } s_1 s_2, \text{ where}$$

$$C_0 = a_0 b_0, C_1 = a_0 b_1 + a_1 b_0, C_2 = a_0 b_2 + a_2 b_0 + a_1 b_1, \dots$$

The term by term equality of the coefficients is possible because the different powers of x are linearly independent. Similarly, $\sum_n a_n x^n / \sum_n b_n x^n = \sum_n d_n x^n$, with the ratio series converging to s_1/s_2 at every point where the denominator does not vanish in the region of convergence.

The coefficients are given by $d_0 = a_0/b_0$, $b_0 d_1 + b_1 d_0 = a_1$ and so on, as obtained by the term by term equality of the coefficients.

The discussions made above are valid for a function of many variables and specifically for a function of a complex variable.

5.8 ASYMPTOTIC SERIES

Occasionally we come across a series which is not convergent in the usual sense of the term, i.e., the remainder after n terms R_n does not tend to zero for a fixed value of the variable as $n \rightarrow \infty$ but $|R_n|$ is less than the $(n+1)$ th term and approaches zero for any fixed n as the variable x goes to infinity. Such a series is of practical importance because we can truncate the series after a certain number of terms depending on x and the extent of accuracy we aim to obtain. This type of series defines an asymptotic series.

Let us consider the integral $I = \int_x^\infty \bar{e}^{y^2} dy$.

Putting $y^2 = z$ we can transform I to the form.

$$I = \int_{x^2}^\infty \bar{e}^z \frac{dz}{2\sqrt{z}} = \left[-\frac{\bar{e}^{-z}}{2\sqrt{z}} \right]_{x^2}^\infty - \int_{x^2}^\infty \frac{\bar{e}^{-z}}{2^2 z^{3/2}} dz$$

or

$$I = \bar{e}^{-x^2} \left(\frac{1}{2x} \right) + \left[\frac{\bar{e}^{-z}}{2^2 z^{3/2}} \right]_{x^2}^\infty + \int_{x^2}^\infty \frac{1.3}{2^3} \frac{\bar{e}^{-s}}{5^{5/2}} ds$$

$$= \bar{e}^{-x^2} \left[\frac{1}{2x} - \frac{1}{2^2 x^3} + \frac{1.3}{2^3 x^5} \right] \dots + \int_{x^2}^\infty \frac{1.3.5.7}{2^5} \frac{\bar{e}^{-s}}{5^{7/2}} + ds$$

The last integral $< \frac{1.3.5.7}{2^5 x^9} \bar{e}^{-x^2}$ and hence $<$ the next term in the expansion.

Moreover, for $x \rightarrow \infty$, the remainder $\rightarrow 0$ if we truncate the series after a certain number of terms. Thus,

this infinite series representation of the integral exemplifies an asymptotic series for which $\lim_{n \rightarrow \infty} \left| \frac{u_{n+1}}{u_n} \right|$ diverges for any fixed x .

The Stirling's formula for $x!$ for large x offers another such example. It can be shown that

$$\ln x! = \left(x + \frac{1}{2} \right) \ln x - x + \frac{1}{2} \ln (2x) + \text{series in odd powers of } \frac{1}{x}.$$

Hence, for large x , $\ln x! \approx x \ln x - x$, an approximation frequently utilized in statistical mechanics.

5.9 APPLICATION TO PHYSICAL PROBLEMS

The infinite series occurs in various branches of physics. One application, simple but of profound importance, is in the determination of the frequency spectrum of black body radiation in thermodynamic equilibrium at a temperature T . Adopting planck's quantum hypothesis that a quantum of black body radiation can exist in discrete energy states of magnitudes $\hbar\omega$, $2\hbar\omega$, ... along with the idea that the probability of having an energy E at a temperature T in equilibrium is proportional to $\exp(-E/\hbar T)$ we obtain that the average energy of an

oscillator that acts as the quantum of radiation is
$$\frac{\hbar\omega e^{-\hbar\omega/kT} + 2\hbar\omega e^{-2\hbar\omega/kT} + \dots}{1 + e^{-\hbar\omega/kT} + e^{-2\hbar\omega/kT} + \dots}.$$

Each of the two series in the numerator and the denominator being absolutely convergent the ratio is well-

defined and converges to
$$\frac{\hbar\omega e^{-\hbar\omega/kT} (1 + 2e^{-\hbar\omega/kT} + 3e^{-2\hbar\omega/kT} + \dots)}{(1 + e^{-\hbar\omega/kT} + e^{-2\hbar\omega/kT} + \dots)}$$

or, $\frac{\hbar\omega e^{-\hbar\omega/kT}}{(1 - e^{-\hbar\omega/kT})}$ or, $\frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}.$

Multiplying the expression for the average energy of an oscillator at temperature T as given above by the number density of oscillators in the frequency interval w to $w + dw$ we obtain the energy density of oscillators in the specified interval.

The same expression for the average energy per quantum at a temperature T also appears in connection with the energy associated with the lattice vibration of a solid assuming it to be quantised and is applied to the calculation of the specific heat of solids. The calculation of cohesive energy in an ionic crystal also involves the calculation of the sum of an infinite series. Thus the total potential energy of an ionic crystal can be approximately written, as $\phi = \sum_{i+j} \phi_{ij}$ by taking nearest neighbour interactions into account. This series is

often found to be very slowly convergent. For details, the reader may consult the standard text books on solid state physics.

The study of nonlinear dynamics and chaos, a topic of current interest and having diverse applications requires a knowledge of dealing with the problems of infinite series. In most of the problems an exact solution is impossible and a solution is sought in the form of an infinite series expansion in powers of the nonlinearity parameter. A very simple example below illustrates the point. Though we deal with the problem of differential equations in a great detail in chapter 7 we discuss the problem and its solution very briefly here.

Let us have an equation $\frac{d^2x}{dt^2} + w^2x = \epsilon x^3$. with $(x)_0 = a$, $(\dot{x})_0 = 0$. If we assume that $x = x_0 + \epsilon x_1 + \dots$ we obtain by equating powers of ϵ ,

$$\frac{d^2x_0}{dt^2} - w^2x_0 = 0 \text{ giving } x_0 = a \cos wt.$$

$$\frac{d^2x_1}{dt^2} + w^2x_1 = x_0^3 = \frac{a^3}{4} [\cos 3wt + 3 \cos wt], \text{ the solution of which yields a term proportional to } t \cos wt$$

[it will be shown in detail in the chapter 7]. Thus clearly this term diverges as $t \rightarrow \infty$ leading to a divergence of the perturbation series when we calculate $x_2, x_3, \dots$ and substitute in x . To avoid this unwanted divergence

which is physically absurd we also rescale the independent variable by using $\tau = \frac{2\pi t}{T}$, where $T = 1 + \epsilon \alpha_1 + \epsilon^2 \alpha_2 + \dots$.

Putting in the equation and equating the coefficient of ϵ we get, $\frac{d^2x}{d\tau^2} + x_1$

$$= -2\alpha_1 a \cos \tau + \frac{a^3}{w^2} \cos^3 \tau = -2\alpha_1 a \cos \tau + \frac{a^3}{4w^2} [\cos 3\tau + 3 \cos \tau]$$

Clearly for removal of the terms containing $\cos \tau$ that leads to divergence we must have $\alpha_1 = \frac{3a^2}{8w^2}$.

Similarly adjusting the values of $\alpha_2, \alpha_3, \dots$ we can eliminate the divergence producing terms in $x_2, x_3, \dots$.

Now $x_1 = -\frac{a^2}{32w^2} \cos 3\tau$. Thus putting $x_1, x_2, \dots$ we obtain a perturbation series that yields convergence and

can be summed approximately to obtain a result valid to a certain extent of accuracy.

SOLVED PROBLEMS

Problem 1

Test the convergence of the following series

- (a) $\sum \frac{1}{5^n}$, (b) $\sum 3^n (\cos x)^n$, (c) $\sum \frac{\ln n}{n}$,
 (d) $\sum \frac{n}{2^n}$, (e) $\sum \frac{1}{n^2 + 3n + 4}$, (f) $\sum \frac{n}{n^2 + 1}$,
 (g) $\sum 2^n \sin \frac{x}{3^n}$.

Solution:

(a) Using the ratio test, $\left| \frac{u_{n+1}}{u_n} \right|_{n \rightarrow \infty} = \frac{1}{5}$.

Hence the series converges

(b) From ratio test, $\left| \frac{u_{n+1}}{u_n} \right| = |3 \cos x|$.

Hence the series converges for $-\frac{1}{3} < \cos_k < \frac{1}{3}$

(c) Using the integral test we see that $\int_i^{\infty} \frac{\ln n}{n} dn = \frac{1}{2} (\ln n)^2 |_c^{\infty}$ and hence does not exist. So, the series diverges.

(d) From ratio test, $\left| \frac{u_{n+1}}{u_n} \right|_{n \rightarrow \infty} = \frac{1}{2}$. Hence the series converges

(e) Here u_n for every n is less than $\frac{1}{n^2}$. Hence $\sum \frac{1}{n^2 + 3n + 4} < \sum \frac{1}{n^2}$. Since $\sum \frac{1}{n^2}$ converges, from comparison test the series converges. (Also, the convergence can be proved by utilizing directly the integral test).

(f) Using the integral test we get, $I = \int_c^{\infty} \frac{n}{n^2 + 1} dn = \frac{1}{2} \ln(n^2 + 1) \Big|_0^{\infty}$. Thus, I does not exist and hence the series diverges

(g) $\left| \frac{u_{n+1}}{u_n} \right|_{n \rightarrow \infty} = 2 \left| \frac{\sin \frac{x}{3^{n+1}}}{\sin \frac{x}{3^n}} \right|_{n \rightarrow \infty} = 2 \left| \frac{x}{3^{n+1}} \cdot \frac{3^n}{x} \right|_{n \rightarrow \infty} = \frac{2}{3} < 1$. Hence, from the ratio test the series

converges for $-\infty < x < \infty$.

Problem 2 Test the uniform convergence of

- (a) $\sum_n (-1)^{n-1} \frac{\sin nx}{n}$
- (b) the sequence $\frac{\sin nx}{n}$.

Solution:

- (a) Using Abel's test, since $\sum_n a_n = \sum_n \frac{(-1)^{n-1}}{n}$ is convergent and $\sin nx$ is a non-monotonically increasing function we obtain the result that $\sum_n (-1)^{n-1} \frac{\sin nx}{n}$ converges uniformly for all x .
- (b) Here $|f_n(x) - f(x)|$, where $f_n(x) = \frac{\sin nx}{n}$ and $f(x) = \lim_{n \rightarrow \infty} f_n(x) = 0$, is equal to $\left| \frac{1}{n} \sin nx \right| \leq \frac{1}{n} < \epsilon$ for $n > N(\epsilon) = \frac{1}{\epsilon}$ for $-\infty < x < \infty$. Hence the sequence $f_n(x)$ converges uniformly to zero for $-\infty < x < \infty$.

Problem 3 Expand stating the region of validity of the expansion.

- (a) $f(x) = \frac{1}{(x-1)(x-4)}$ about $x = 3$,
- (b) $\frac{x}{x-3}$ about $x \rightarrow \infty$,
- (c) $\frac{1}{x^2}$ in powers of $(x+1)$,
- (d) e^x in powers of $x-a$.

Solution:

- (a) let $x = 3 + h$ so that $f(x) = \frac{1}{(2+h)(h-1)}$

$$\text{or, } f(h) = \left[\frac{1}{(h-1)} - \frac{1}{(h+2)} \right] \cdot \frac{1}{3} = \frac{1}{3} \left[\sum_n h^n - \frac{1}{2} \sum_n (-1)^n \left(\frac{h}{2} \right)^n \right]$$

The first series converges for $|h| < 1$ or $|x-3| < 1$ or $2 < x < 4$.

The second series converges for $|h| < 2$ or, $1 < x < 5$.

The resultant series therefore converges in the common region $2 < x < 4$.

The expansion about $x = 3$ is valid up to the point of nearest singularity, i.e., $x = 4$.

(b) We put $x = \frac{1}{y}$ so that $f(x) = \phi(y) = \frac{\frac{1}{y}}{\frac{1}{y} - 3} = \frac{1}{1 - 3y}$

The expansion about $x \rightarrow \infty$ thus boils down to an expansion about $y = 0$.

Thus $\phi(y) = \sum_n (-1)^n (3y)^n$ leads to $f(x) = \sum_n (-1)^n (3/x)^n$

The expansion is valid for $|3y| < 1$ or, $|x| > 3$, i.e., outside the point of singularity $x = 3$.

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

or $\frac{1}{x^2} = - \sum_n (x+1)^n$.

The expansion is valid for $|x+1| < 1$ and centered at $x = -1$. Clearly, the validity extends only up to the region of nearest singularity, i.e., $x = 0$

(d) $e^x = e^{(x-a+a)} = e^a \cdot e^{(x-a)}$.

The Taylor series expansion of e^{x-a} about $x = a$ gives, $e^x = e^a \sum_n \frac{(x-a)^n}{n!}$ and this is valid for $-\infty < x < \infty$.

Problem 4

Find out the Maclaurin's series expansion of the following:

(a) $e^{\sin x}$, (b) $\frac{x}{\sin x}$, (c) $\frac{\sin x}{1+x}$

Solution:

(a) $e^{\sin x} = (e^{\sin x})_0 + x \left(\frac{d}{dx} e^{\sin x} \right)_0 + \frac{x^2}{2!} \left(\frac{d^2}{dx^2} e^{\sin x} \right)_0 + \dots$

or, $e^{\sin x} = 1 + x [\cos e^{\sin x}]_0 + \frac{x^2}{2!} [\cos^2 x e^{\sin x} - \sin x e^{\sin x}]_0 + \dots$
 $= 1 + x + \frac{x^2}{2!} + \dots$

(b) $\frac{x}{\sin x} = \sum_n a_n x^n$ (say)

or, $x = \left(\sum_n a_n x^n \right) \sin x$

Using the series expansion of $\sin x$, we get

$$\text{or, } x = (a_0 + a_1x + a_2x^2 + \dots) \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} \dots \right)$$

Equating the coefficients of various power of x , we get

$$a_0 = 1, a_1 = 0, a_2 = \frac{a_0}{3!} = \frac{1}{3!}, a_3 = \frac{a_1}{3!} = 0$$

$$\text{or, } \frac{x}{\sin x} = 1 + \frac{x^2}{3!} + \dots$$

$$(c) \frac{\sin x}{1+x} = \sum_n a_n x^n \text{ (say)}$$

$$\text{or, } \sin x = (1+x) \sum_n a_n x^n = (1+x)(a_0 + a_1x + a_2x^2 + \dots)$$

Using the series expansion of $\sin x$ we get,

$$\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \right) = (1+x)(a_0 + a_1x + a_2x^2 + \dots)$$

Equating the coefficients of the various powers of x , we get

$$a_0 = 0, a_0 + a_1 = 1 \text{ giving } a_1 = 1, a_2 = -a_1 = 0, a_3 + a_2 = -1/3! \text{ giving } a_3 = -\frac{1}{3!}$$

$$\text{We get, } \frac{\sin x}{1+x} = x - \frac{x^3}{3!} + \dots$$

(it is instructive to carry out the sums (b) and (c) by directly applying Maclaurin's series expansion to the entire function)

Problem 5

Using the Maclaurin's series expansions of the functions find out the following

integrals by term by term integration indicating the interval of convergence.

$$(a) \int_0^x \frac{\sin x}{x} dx, \int_0^x e^x \ln(1+x) dx, \int_0^x \frac{\sin^{-1} x}{x} dx$$

Solution:

$$\begin{aligned} (a) \int_0^x \frac{\sin x}{x} dx &= \int_0^x \frac{1}{x} \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \right) dx \\ &= x - \frac{x^3}{3 \cdot 3!} + \frac{x^5}{5 \cdot 5!} - \dots = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} \cdot x^{2n-1}}{(2n-1)(2n-1)!} \end{aligned}$$

Using the ratio test we see that the series converges for $-\infty < x \leq \infty$

$$\begin{aligned}
 \text{(b)} \quad \int_0^x e^x \ln(1+x) dx &= \int_0^x \left(1+x+\frac{x^2}{2!}+\dots\right) \left(x-\frac{x^2}{2}+\frac{x^3}{3}-\dots\right) dx \\
 &= \int_0^x \left(x+\frac{x^2}{2!}+\frac{x^3}{3}+\dots\right) dx = \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{2.3!} + \dots
 \end{aligned}$$

The series converges for $-1 \leq x \leq 1$.

$$\begin{aligned}
 \text{(c)} \quad \int_0^x \frac{\sin^{-1} x}{x} dx &= \int_0^x \frac{1}{x} \left[x \left(\frac{1}{\sqrt{1-x^2}} \right)_{x=0} + \frac{x^2}{2!} \left[\frac{x}{(1-x^2)^{3/2}} \right]_0 - \frac{x^3}{3!} \left(\frac{(1-x^2)^{3/2} + 3x^2(1-x^2)}{(1-x^2)^3} \right) \int_0^1 + \dots \right] \\
 &= \int_0^x \frac{1}{x} \left[x + \frac{x^3}{3!} + \dots \right] dx = x + \frac{x^3}{3.3!} + \dots
 \end{aligned}$$

The expansion is valid for $-\infty \leq x \leq \infty$.

Problem 6

Solve the equation of motion of pendulum for any finite amplitude.

Solution: The equation of motion of a pendulum bob is, $I \frac{d^2\theta}{dt^2} = -mgl \sin \theta$, the symbols having their usual meanings. Multiplying by $2 \frac{d\theta}{dt}$ and integrating we get

$$\left(\frac{d\theta}{dt} \right)^2 = \frac{2mgl}{I} \cos \theta + \text{constant}.$$

With the boundary condition $\frac{d\theta}{dt} = 0$ at $\theta = \theta_0$ we obtain

$$\left(\frac{d\theta}{dt} \right)^2 = \frac{2mgl}{I} (\cos \theta - \cos \theta_0)$$

or, $\int_0^{T/4} dt = \sqrt{\frac{I}{2mgl}} \int_0^{\theta_0} \frac{d\theta}{\sqrt{\cos \theta - \cos \theta_0}}$, θ_0 and T being the amplitude and the time period of oscillation respectively.

Changing the variable from θ to ϕ given by $\sin \phi = \frac{\sin \theta/2}{\sin \theta_0/2}$ we obtain

$$T/4 = \frac{1}{2} \sqrt{\frac{I}{mgl}} \int_0^{\phi_0} \frac{2d\phi}{\left[1 - \left(\frac{\sin^2 \theta_0}{2} \right) \sin^2 \phi \right]^{1/2}}$$

The Maclaurin's series expansion of $(1 - \sin^2 \theta_0/2 \sin^2 \phi)^{-1/2}$ being uniformly convergent in the region of integration we can integrate by expanding the series term by term. Thus we get

$$T/4 = \frac{1}{2} \sqrt{\frac{I}{mgl}} \int_0^{\phi_0} 2 \left[1 + \frac{1}{2} \sin^2 \frac{\theta_0}{2} \sin^2 \phi + \dots \right] d\phi$$

Putting $\phi_0 = \pi/2$, we obtain

$$\begin{aligned} T &= 2 \sqrt{\frac{I}{mgl}} \left[\pi + \frac{\theta_0^2 \pi}{16} \right], \text{ using } \int_0^{\pi/2} \sin^2 \phi d\phi = \pi/4 \\ &= 2\pi \sqrt{\frac{I}{mgl}} \left(1 + \frac{\theta_0^2}{16} \right), \text{ keeping terms to } O(\theta_0^2) \end{aligned}$$

Putting $I = ml^2$ for a point bob, we get

$$T = 2\pi \sqrt{l/g} \left(1 + \frac{\theta_0^2}{16} \right),$$

To find out the motion of the point bob we replace $T/4$ by t and θ_0 by θ to get

$$\sqrt{\frac{g}{l}} t = \phi \left[1 + \frac{\theta^2}{16} \right] - \frac{\theta^2}{32} \sin 2\phi, \text{ keeping terms upto order } \phi^2.$$

Putting $\phi = \sin^{-1} \left(\frac{\sin \theta/2}{\sin \theta_0/2} \right) = \sin^{-1} \left(\frac{\theta}{\theta_0} \right)$ for small θ and θ_0 .

we get, $t = \sqrt{\frac{l}{g}} \left(1 + \frac{\theta_0^2}{16} \right) \sin^{-1} \left(\frac{\theta}{\theta_0} \right) - \frac{\theta_0}{16} \left(1 - \frac{\theta^2}{\theta_0^2} \right)^{1/2} \sqrt{\frac{l}{g}}$. This example illustrates the use of an infinite series expansion in physical problems.

SUMMARY

If $\lim_{m \rightarrow \infty} \sum_{n=1}^m f_n \left[\text{or } \sum_{n=1}^m f_n(x) \right] = s$, s being finite, the series is convergent. Otherwise it is divergent. There are

two types of convergence of an infinite series.

- *Point convergence*—Here we are concerned with the convergence of a series or a sequence of functions at every point within an interval of definition of the variable (or the variables). This is further subdivided into.

- (i) *Absolute convergence* – If the series $\sum_n |f_n|$, where the f_n 's may define a sequence of numbers or function converges the series $\sum_n f_n$ is said to be absolutely convergent. The sum is independent of the rearrangement of the terms.

For example the series $\sum_n \frac{1}{n^s}$ is seen to be absolutely convergent for $s > 1$.

- (ii) *Conditional convergence* – If the series $\sum_n f_n$ converges but $\sum_n |f_n|$ does not, the series is conditionally convergent. According to Leibnitz criterion, if for an alternating series $|f_{n+1}| < |f_n|$ and $\lim_{x \rightarrow \infty} f_n = 0$ the series is conditionally convergent. The series $\sum_n \frac{(-1)^{n-1}}{n}$ is conditionally convergent while $\sum_n \frac{1}{n}$ can be shown to diverge. Regarding absolute convergence there are several tests among which three, namely, the Cauchy's test, the ratio test and the comparison test have been discussed in the text.
- (iii) *Uniform convergence* – The problem of uniform convergence is relevant only for a series of functions (or a sequence of functions). Thus a series of functions $\sum_n s_n(x)$ converges uniformly to $s(x)$ if for a given preassigned number ϵ , $\left| \sum_{n=1}^m s_n(x) - s(x) \right| < \epsilon$ for $m > N$, where N depends

only on ϵ . A similar definition applies to a sequence of uniformly convergent functions. The study of uniform convergence is essential in many cases because a uniformly convergent series can be integrated term by term and the integrated series has the same region of uniform convergence as the original series.

Thus $\int_a^b \sum_n f_n(x) dx = \sum_n \int_a^b f_n(x) dx$, where $a \leq x \leq b$ denotes the interval of convergence. An infinite uniformly convergent series can be differentiated term by term and provided the differentiated series is uniformly convergent in the same interval of values of the independent variable,

$$\frac{d}{dx} \sum_n f_n(x) = \sum_n \frac{df_n(x)}{dx}$$

Thus even if a uniformly convergent series cannot be integrated in closed form, the magnitude of the integral can be obtained to an extent of accuracy by properly truncating the integrated series obtained by termwise integration of the original series.

We have stated in the text three different tests of uniform convergence, namely, (i) Weirstrass test, (ii) Abel's test and (iii) Canchy's test. It can be shown by applying these tests that the series

$\sum_n \frac{\sin nx}{n^s}$, ($s > 1$) and $\sum_n \frac{(-1)^{n-1} x^n}{n}$, ($0 \leq x \leq 1$) are two examples of a uniformly convergent series.

The power series expansion of a function about a point in ascending or descending powers of the variable (when expanded about the point at infinity) is important in this context, since it is easily shown to be uniformly convergent in the interval of convergence extended from the given point to the nearest singularity of the function. Thus, a power series can be integrated or differentiated term by term (provided the differentiated series is uniformly convergent) yielding a new series that is convergent in the same interval. The Taylor series expansion of a function about a point $x = a$ involving different order derivatives of the function at $x = a$ is given symbolically

$$\text{by } \exp\left[(x-a)\frac{d}{dx}\right]f(x)\bigg|_{x=a} \quad \text{and by } \exp\left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y} + \dots\right]f(x,y,\dots)\bigg|_{a,b,c} \quad \text{for a}$$

function of many variables. Within a definite interval of convergence it provides us with the unique power series expansion about the point and when $a = b = c, \dots = 0$, we obtain Maclaurin's series. A number of power series about a given point can be combined by the process of addition, subtraction, multiplication or division provided the denominator does not vanish in the last case, giving rise to the new series which converge in the region of convergence common to all the individual series.

- *Convergence in the mean* – This involves a situation where instead of point convergence of a series at every point of an interval we deal with the problem of overall convergence of a series within a given interval. A series $\sum_n f_n(x)$ is said to converge in the mean to a function $f(x)$ in an interval (a, b) if

$$\lim_{m \rightarrow \infty} \int_a^b \left| \sum_{n=1}^m f_n(x) - f(x) \right|^2 dx = 0. \text{ When this equality holds, the set of functions } f_n(x) \text{ is said to form a}$$

complete set in the given interval. The series (or sequence) of functions converging in the mean to a given function may not converge at every point to the function in the given interval. This is illustrated in the text by an example.

- *Divergent series* – Instead of rejecting outright a divergent series a few mathematicians, C'easaro, Borel etc. have given alternative definitions of ascertaining the sum of a series which is divergent in the usual sense. We have discussed C'easaro's summability condition in the text (C1) where a series $\sum_n s_n$ is said to converge to a unique sum s , if the average of the partial sums $\frac{1}{n} (s_1 + \dots + s_n)$ is convergent as $n \rightarrow \infty$. The problem of the oscillatory series $1 - 1 + 1 - 1 + \dots$, which is divergent, has been discussed in the text.
- *Asymptotic series* – A series which is divergent in the sense that $R_n(n, x)$, the remainder after n terms does not go to zero as $n \rightarrow \infty$ for a fixed x but for a fixed n $\lim_{x \rightarrow \infty} R_n(n, x) = 0$ is termed an asymptotic series. Thus for a large x the series can be truncated after a certain number of terms depending on the accuracy to which we want the result to be evaluated. This is explained in the text with reference to the approximate evaluation of an integral.

MULTIPLE CHOICE QUESTIONS

1. Consider the three series $\sum \frac{1}{n^{1/3}}$, $\sum \left(\frac{n+1}{n^p}\right)$, $\sum \frac{(-1)^n}{n^2 + 4n + 3}$. Which of the following statements is correct?

- (a) The three series are absolutely convergent
(b) The first series is divergent while the last two are absolutely convergent.
(c) The first one is divergent and the third one convergent while the second one converges absolutely for $p > 2$.
(d) The three series are divergent.
2. $\sum_n \frac{2 \cos n \theta}{n(n+2)}$ is
(a) divergent for all θ
(b) absolutely convergent for all θ
(c) convergent for $0 < \theta < \pi/2$
(d) convergent conditionally for all θ .
3. $\sum_n (-1)^n u_n$ is
(a) convergent if $|u_n| < |u_{n+1}|$ and $\lim_{n \rightarrow \infty} |u_n| = 0$
(b) Convergent if $|u_n| < |u_{n+1}|$
(c) Convergent if $\lim_{n \rightarrow \infty} |u_n| = 0$
(d) Convergent if $|u_{n+1}| < |u_n|$ and $\lim_{n \rightarrow \infty} |u_n| = 0$
4. Consider the four series $\sum \frac{1}{n \ln n}$, $\sum n e^{-n^2}$, $\sum \frac{n!}{n^2}$, $\sum \frac{(-1)^n}{n^{3/2}}$. Which of the following statements is correct?
(a) All the four series are convergent
(b) The first series diverges while the remaining three converge
(c) The first and the third series diverge while the other two converge
(d) All the four series diverge
5. The series $\sum_n (3 \cos x)^n$
(a) Converges for all x , (b) diverges for all x ,
(c) Converges for $|\cos x| < \frac{1}{3}$, (d) Converges for $|\cos x| > \frac{1}{3}$
6. You are given two series $\sum_n x^n n!$ and $\sum_n \frac{4.7 \dots (3n+1)}{n!} x^n$
Which of the statements is correct?
(a) Both the series diverge for all x
(b) The first series diverges for all x while the second one converges for $|x| < \frac{1}{3}$
(c) The first series converges only for $x = 0$ while the second one converges for $|x| < \frac{1}{3}$
(d) The first series converges only for $x = 0$ while the second one diverges for $|x| < \frac{1}{3}$

7. The series $\sum_n \frac{x^n}{(n+1)^n}$ and $\sum_n \frac{(-1)^{n+1} \cdot x^n}{1+x^n}$ are given
- The first one converges for all x while the second one does so for $|x| < 1$ only
 - Both the series converge for all x
 - Both the series converge for $|x| < 1$ only
 - The first one converges for all x while the second one diverges for $|x| < 1$
8. You are given an absolutely convergent and a conditionally convergent series. Which of the following statements is correct?
- The sum of the first series is independent of the arrangement of terms while that of the second one is not:
 - Both the sums are sensitive to the arrangement of the terms in the series
 - The first sum depends on the arrangement of terms while the second one is independent of it
 - Neither of the sums changes with the rearrangement of terms
9. The expansion of $\cos^2 x$ about $x = \pi/4$ is
- $\frac{1}{2} + (\pi/4 - x) + \frac{2}{3}(x - \pi/4)^3 + \dots$
 - $1 - x^2 + \frac{x^4}{3} + \dots$
 - $\frac{1}{2} + \frac{(\pi/4 - x)^2}{2} - \frac{(\pi/4 - x)^3}{3!} + \dots$
 - $1 + (\pi/4 - x) + \frac{(\pi/4 - x)^2}{2!} + \dots$
10. The expansion of $\ln(1 + \sin x)$ about $x = 0$ is,
- $1 - x + \frac{x^2}{2!} - \dots$
 - $x - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$
 - $x - \frac{x^2}{2!} + \frac{x^3}{3!} - \dots$
 - $x + \frac{x^2}{2!} - \frac{x^3}{3!} + \dots$
11. The expansion of $\frac{1}{x(1-x)}$ about $x = 0$ is,
- $\frac{1}{x} + 1 + x + \frac{x^2}{2!} + \dots$

- (b) $1+x+x^2+\dots$
- (c) $\frac{1}{x}+1+x+x^2+\dots$
- (d) $\frac{1}{x}-1-x-x^2-\dots$
12. The expansion of $\ln\left(\frac{1-x^2}{1+x^2}\right)$ about $x=0$ is,
- (a) $x^2+\frac{x^6}{3}+\frac{x^{10}}{5}+\dots$
- (b) $-2\left(x^2-\frac{x^6}{3}+\frac{x^{10}}{5}-\dots\right)$
- (c) $x^2-\frac{x^6}{3}+\frac{x^{10}}{5}-\dots$
- (d) $-2\left(x^2+\frac{x^6}{3}+\frac{x^{10}}{5}+\dots\right)$
13. The magnitude of $\int_0^x \sin x^2 dx$ is, correct to two decimals
- (a) 0.33 (b) 0.31 (c) 0.34 (d) 0.35
14. $\sum \frac{\cos nx}{n^4}$ is
- (a) a uniformly convergent series for $|x|<1$
- (b) a uniformly convergent series for $|x|>1$
- (c) nowhere uniformly convergent
- (d) uniformly convergent for all x
15. Consider the two integrals $\int_0^\infty \cos x^2 dx$ and $\int_0^\infty \frac{(\ln x) \sin x}{x} dx$
- (a) The first one is divergent while the second one is convergent
- (b) The first one is convergent while the second one diverges
- (c) Both the series are convergent
- (d) Both the series are divergent

Answers to MCQs

1. (c)	2. (b)	3. (d)	4. (c)	5. (c)	6. (c)	7. (a)
8. (a)	9. (a)	10. (c)	11. (c)	12. (d)	13. (b)	14. (d)
15. (b)						

EXERCISE

1. Test the convergence of the following series:

- (a) $\sum \frac{(-1)^{n-1}}{\sqrt{n}}$, (b) $\sum \left(1 - \frac{1}{n^2}\right)$,
 (c) $\sum \frac{e^n}{\sqrt{n!}}$, (d) $\sum \frac{n!}{(2n)!}$,
 (e) $\sum \frac{1}{n(n+1)}$ (f) $\sum_n (-1)^n 2^n (\sin x)^n$,
 (g) $\sum_n \frac{3^n (5x+1)^n}{n^2}$, (h) $\sum_n \frac{(n!)^2}{(3n)!} x^n$,
 (i) $\sum \frac{(n!)^3}{(3n)!} x^n$ (j) $\sum (a_n \cos n\theta + b_n \sin n\theta)$,
 (k) $\sum \frac{(-1)^{n-1} \sin nx}{n^3}$ (l) $\sum_n n \left(\frac{1}{2}\right)^{n-1}$

2. Expand the following stating the region of convergence.

- (a) $\cot x$ in powers of x , about $x = 0$ (b) $\frac{1}{z(z-2)}$ about $z = 3$,
 (c) $\tan x$ in powers of $x - d$, (d) $\frac{1}{(1+x)} \ln(1+x)$ about $x = 0$.
 (e) $\frac{e^{-x}}{\sin x}$ about $x = 0$ (f) $\frac{1}{(x-1)(x-2)}$ about the point at infinity

3. Find out the integrals by the term-by-term integration.

- (a) $\int_a^x \frac{1}{x^2} \ln(1+x) dx$, (b) $\int_0^{\pi/2} \frac{\cos x}{1+x^2} dx$ (c) $\int_0^{1/2} e^{\sqrt{x}} dx$ to two decimals,

4. Find out the C'easaro sum (C1) of $\sin \theta + \sin 3\theta + \dots +$

5. A particle is projected from the earth's surface (assumed to be spherical) vertically upwards with a speed u . Find out the time t to reach a given height z in powers of z . (Do not assume a constant gravitational field but take variation of g with height into account)

6. $\sum_{n=0}^{\infty} x^{2n} = \frac{1}{1-x^2}$ and $\sum_{n=-1}^{-\infty} x^{2n} = \frac{1}{x^2-1}$

Adding we get, $\sum_{n=-\infty}^{\infty} x^{2n} = 0$. But this is absurd. Explain the paradox.

Hints and Answers to Exercise Questions

1. (a) conditionally convergent, (b) divergent
 (c) absolutely convergent (d) absolutely convergent,
 (e) absolutely convergent, (f) absolutely convergent for $|\sin x| < 1/2$,
 (g) absolutely convergent for $|5x + 1| < 3$, (h) absolutely convergent for $x = 0$ only,
 (i) absolutely convergent for $x = 0$ only. (j) uniformly convergent if $\sum_n |a_n|$ and $\sum_n |b_n|$ converge,
 (k) absolutely convergent (l) absolutely convergent.
2. (a) $\frac{1}{x} - \frac{x}{3} + \frac{x^3}{40} - \dots$,
 (b) $\frac{1}{3} \left[1 - \frac{4}{3}(z-3) + \frac{13}{9}(z-3)^2 - \dots \right]$.
 (c) $\tan d + (x-d)\sec^2 d + (x-d)^2 \sec^2 d \tan d + \dots$.
 (d) $x - \frac{3x^2}{2} + \frac{11x^3}{6} - \dots$
 (e) $\frac{1}{x} \left(1 - x + x^2 - \frac{x^3}{3!} + \dots \right)$,
 (f) $\frac{1}{x^2} \left(1 + \frac{3}{x} + \frac{7}{x^2} + \dots \right)$ [put $y = 1/x$ and expand about $y = 0$]
3. (a) $\ln \frac{x}{a} - \frac{1}{2}(x-a) + \frac{(x^2 - a^2)}{6} + \dots$
4. (a) $\frac{1}{4} \cos \theta$ [Hint:- Use the definition of C'esarò sum given in the text]
5. (a) The exact time is for $u^2 < \frac{2GM}{R}$ given by, ($u < \text{escape velocity}$)

$$T = \frac{A}{B^{3/2}} \left[\sin^{-1} \sqrt{\frac{Bh}{A}} - \sqrt{\frac{Bh}{A}} \sqrt{1 - \frac{Bh}{A}} \right], \text{ where } A = 2GH, B = u^2 - \frac{2GM}{R}$$

Assuming that $\left| \frac{Bh}{A} \right| \ll 1$ make a power series expansion in powers of $\frac{Bh}{A}$.

For the opposite limit, i.e., $u > \sqrt{\frac{2GM}{R}}$,

$$T = \frac{A}{B^{3/2}} \left[\sqrt{\frac{Bz}{A}} \sqrt{1 + \frac{Bz}{A}} - \ln \left(\sqrt{\frac{Bz}{A}} + \sqrt{1 + \frac{Bz}{A}} \right) \right],$$

For $\left| \frac{Bz}{A} \right| \ll 1$ perform a similar expansion,

[**Hint:** Start from the equation $\frac{d^2 z}{dt^2} = -\frac{GM}{z^2}$ Multiply by $2 \frac{dz}{dt}$ and then integrate].

(b) The addition is not valid because the regions of convergence of the two series are different.

APPENDIX

(A) Brief derivation of the Taylor series expansion

A brief derivation of the Taylor series expansion of a function that is single-valued and possesses all order derivatives with respect to the variable in a certain interval can be given by following the idea of successive approximation.

We have,
$$f(x_0 + h) = f(x_0) + \int_{x_0}^{x_0+h} f'(x) dx \quad (\text{A4.1})$$

In the first approximation $f(x_0 + h) = f(x_0) + hf'(x_0)$ where we put $x = x_0$ in $f'(x)$. Similarly, $f'(x_0 + h) = f'(x_0) + hf''(x_0)$,

Substituting for $f'(x)$ in (A4.1),

$$f(x) = f(x_0) + \int_{x_0}^x \{f'(x_0) + (x - x_0)f''(x_0)\} dx$$

Or, $f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2!} f''(x_0)$ in the next approximation.

Continuing this process we obtain,

$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2!} f''(x_0) + \dots + \frac{(x - x_0)^n}{n!} f^n(x_0) + R_{n+1}$, where R_{n+1} is the remainder after n terms and given by,

$R_{n+1} = \frac{(x - x_0)^{n+1}}{(n+1)!} f^{n+1}(x_0 + \theta h)$, where $0 < \theta < 1$. Clearly for the series to converge R_{n+1} should go to zero as $n \rightarrow \infty$.

Example

(a) $\sin x = x - \frac{x^3}{3!} + \dots + \frac{(-1)^{n+1} x^{n+1}}{(n+1)!} + R_{n+1}$ [$x_0 = 0$ here]

$$|R_{n+1}| = \left| \frac{x^{n+1}}{(n+1)!} (-1)^{n+1} (x_0 + \theta h)^{n+1} \right| = \left| \frac{x^{n+1}}{(n+1)!} (-1)^{n+1} (\theta h)^{n+1} \right|$$

For all finite x , $\lim_{n \rightarrow \infty} \frac{x^n}{n!} = 0$ and hence $|R_{n+1}| \rightarrow 0$ as $n \rightarrow \infty$.

(b) $\frac{1}{1+x}$ about $x = 0 = 1 - x + x^2 - \dots + (-1)^{n+1} x^{n+1} + R_{n+1}$,

where $|R_{n+1}| = |x^n (\theta h)^n| \rightarrow 0$ as $n \rightarrow \infty$ for $|x| < 1$.

Thus the above represents the Taylor series expansion of $\frac{1}{1+x}$ about $x = 0$ (i.e., the Maclaurin series) for $|x| < 1$.

(B) Some Important Points Regarding a Power Series

- (a) It has been mentioned in the text that a power series is in general, convergent with a finite or an infinite radius of convergence. An exception is however the series $\sum_n n! x^n$. The ratio

$$\lim_{n \rightarrow \infty} \left| \frac{t_{n+1}}{t_n} \right| = |(n+1) \cdot x|_{n \rightarrow \infty}. \text{ Thus clearly the series converges only for } x = 0.$$

- (b) We state here some important characteristics of a power series that are relevant from the point of view of application of the series to the mathematical problems. Proofs will be omitted though hints will be provided so that the students do not have any difficulty in understanding.

- (i) A power series $\sum_{n=0}^{\infty} a_n x^n$ converging for $x = x_0$, converges absolutely for all x satisfying

$|x| < x_0$. [This results from the fact that for a series $\sum_n a_n x^n$ to be convergent, $a_n x^n \rightarrow 0$ as $n \rightarrow \infty$ such that for a const A , $|a_n x^n| \leq A$ for all n].

- (ii) A power series $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly in every closed subinterval within the radius of convergence. [This comes from applying Weirstrass M-test of uniform convergence].

- (iii) A power series, convergent at the boundary points of an interval, converges uniformly in the closed interval $(0, l)$ where the point $x = l$ is the end point.

- (iv) The power series expansion may be used to evaluate limits of functions.

$$\text{For example, } \lim_{x \rightarrow a} \frac{f(x)}{\phi(x)} = \lim_{x \rightarrow a} \frac{f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!} f''(a) + \dots}{\phi(a) + (x-a)\phi'(a) + \frac{(x-a)^2}{2!} \phi''(a) + \dots}$$

$$\text{If } f(a) = \phi(a) = 0, \lim_{x \rightarrow a} \frac{f(x)}{\phi(x)} = \lim_{x \rightarrow a} \frac{f'(a)}{\phi'(a)}$$

where n is the minimum integer for which $\phi^n(a) \neq 0$. When $\lim_{x \rightarrow a} f(x)$ and $\lim_{x \rightarrow a} \phi(x)$ diverge, the limit is obtained by writing $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{\bar{g}(x)}{\bar{f}(x)}$ where $\bar{g} = \frac{1}{g}$, $\bar{f} = \frac{1}{f}$.

For example, $\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{\cos x}{1} = 1$,

$$\lim_{x \rightarrow 0} \frac{\sin 3x}{\sin hx} = \lim_{x \rightarrow 0} \frac{3 \cos 3x}{\cos hx} = 3$$

$$\lim_{x \rightarrow \infty} \frac{a_1 x^3 + b_1 x^2 + c_1 x + d_1}{a_2 x^3 + b_2 x^2 + c_2 x + d_2} = \lim_{x' \rightarrow 0} \frac{a_1 x'^{-3} + b_1 x'^{-2} + c_1 x'^{-1} + d_1}{a_2 x'^{-3} + b_2 x'^{-2} + c_1 x'^{-1} + d_2},$$

where $x' = \frac{1}{x}$. Thus the limit = $\frac{a_1}{a_2}$.

(C) Convergence of Integrals

In an analogy to the problems of convergence of a series, we also encounter problems related to convergence or divergence of integrals. Such problems arise in the evaluation of improper integrals, i.e., the definite integrals

that cannot be determined by Riemann's definition of integrals, namely, $\int_a^b f(x)dx = \lim_{h \rightarrow 0} \sum_{r=0}^{n-1} hf(a+rh)$, where

$b-a = rh$. We briefly discuss below such problems omitting the proofs and giving appropriate examples.

The integral $\int_a^b f(x)dx$ can become improper in two ways, (1) when a or b both tend to infinity, (2) $f(x)$ diverges at a or b or both or at any point c , where $a < c < b$. In the case (i) the integral is written as $\int_a^\infty f(x)dx = \lim_{b \rightarrow \infty} \int_a^b f(x)dx$ or (ii) $\int_{-\infty}^a f(x)dx = \lim_{b \rightarrow -\infty} \int_b^a f(x)dx$ or (iii) $\int_{-\infty}^\infty f(x)dx = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b f(x)dx$ provided the limits exist in all the cases.

In the case (2) the improper integral is evaluated by writing $\int_a^b f(x)dx = \lim_{\epsilon \rightarrow 0} \int_{a+\epsilon}^b f(x)dx$ when $f(x) \rightarrow \infty$

as $x \rightarrow a$ or, $\int_a^b f(x)dx = \lim_{\epsilon \rightarrow 0} \int_a^{b-\epsilon} f(x)dx$ when $f(x) \rightarrow \infty$ as $x \rightarrow b$ or, $\int_a^b f(x)dx = \lim_{\substack{\epsilon \rightarrow 0 \\ \delta \rightarrow 0}} \int_{a+\epsilon}^{b-\delta} f(x)dx$ when

$f(x) \rightarrow \infty$ as $x \rightarrow a$ and $x \rightarrow b$ or, $\int_a^b f(x)dx = \lim_{\substack{\epsilon \rightarrow 0 \\ \delta \rightarrow 0}} \left[\int_a^{t-\epsilon} f(x)dx + \int_{c+\delta}^b f(x)dx \right]$ when $f(x) \rightarrow \infty$ as $x \rightarrow c$.

When the limiting values of the integrals exist, we say that the improper integrals exist. Below we shall state briefly the conditions of absolute or conditional convergence in these cases giving examples. The defini-

tions of absolute and conditional convergence are similar to those in the case of a series, i.e., (i) if $\int_a^b |f(x)| dx$ converges, the integral is *absolutely convergent*, while (ii) if $\int_a^b f(x) dx$ converges but $\int_a^b |f(x)| dx$ diverges the integral is *conditionally convergent*.

(D) Conditions of Absolute Convergence

- (a) *Comparison test* – If $|f(x)| \leq |g(x)|$ for x in the interval, then $\int_a^b f(x) dx$ is absolutely convergent if $\int_a^b g(x) dx$ be so.

Example: (i) $\int_0^\pi \frac{dx}{x^{1/2}} = \lim_{\epsilon \rightarrow \epsilon} \int_\epsilon^\pi \frac{dx}{x^{1/2}} = 2\sqrt{\pi}$. Since $\left| \frac{\sin x}{x^{1/2}} \right|$ or $\left| \frac{\cos x}{x^{1/2}} \right| \leq \frac{1}{x^{1/2}}$ and $\int_0^\pi \frac{dx}{x^{1/2}}$ converges, $\int_0^\pi \frac{\sin x}{x^{1/2}} dx$ and $\int_0^\pi \frac{\cos x}{x^{1/2}} dx$ are absolutely convergent.

By the same logic $\int_1^\infty \frac{\sin x}{x^2} dx$ or $\int_1^\infty \frac{\cos x}{x^2} dx$ is absolutely convergent $\left(\text{since } \int_1^\infty \frac{dx}{x^2} \text{ converges} \right)$.

- (b) Since $\int_a^\infty \frac{dx}{x^\alpha} = \lim_{b \rightarrow \infty} \int_a^b \frac{dx}{x^\alpha} = \frac{a^{1-\alpha}}{\alpha-1}$ for $\alpha > 1$ and $\rightarrow \infty$ for $\alpha \leq 1$ and $\alpha > 0$, the comparison test shows that if $|f(x)| < \frac{K}{x^\alpha}$ for large x and $k > 0$, $\alpha > 1$, the integral $\int_a^\infty f(x) dx$ converges absolutely.

The comparison test also leads to the conclusion that if $\int_a^b |g(x)| dx$ diverges and $|f(x)| \geq |g(x)|$ for

x very large, then $\int_a^b |f(x)| dx$ is divergent. Thus, if $|f(x)| > \frac{K}{x^\alpha}$ for $k > 0$ and $\alpha \leq 1$ the integral

$\int_a^\infty |f(x)| dx$ diverges.

(E) Condition of Conditional Convergence

Abel's test – If $\int_a^b f(x) dx$ is bounded for all $b \leq \infty$ and $g(x)$ continuously differentiable on the interval $a \leq x < \infty$ and a monotone decreasing function, which tends to zero for $x \rightarrow \infty$, the integral $\int_a^\infty f(x)g(x) dx$ is convergent.

For example, (i) $\int_0^{\infty} \sin(x^2) dx = \int_0^{\infty} \frac{\sin y}{2\sqrt{y}} dy$ or (ii) $\int_0^{\infty} \frac{(\ln x) \cos x}{x} dx$, converges by Abel's test $\left[g(y) = \frac{1}{2\sqrt{y}} \text{ in the first case and } = \frac{\ln y}{y} \text{ in the second} \right]$.

(F) Improper Integrals

We cite in the following few examples regarding the evaluation of certain improper integrals.

(a) The integral $\int_{\alpha}^{\beta} \frac{dx}{(x-a)^r}$ ($\alpha < \beta$) converges for $r < 1$ and diverges for $r \geq 1$. [This can be proved by calculating $\int_{\alpha}^{\beta} \frac{dx}{(x-\alpha)^r}$ and taking the limit $a \rightarrow \alpha$]. Thus $\int_0^{\infty} \frac{dx}{x^{1/2}}$ converges while $\int_0^{\infty} \frac{dx}{x^2}$ diverges.

(b) The integral $\int_{\alpha}^{\beta} \frac{dx}{x-r}$, where $\alpha < r < \beta$, is an improper integral because $\frac{1}{x-r}$ diverges at $x=r$. We write the integral as, $I = \int_{\alpha}^{\beta} \frac{dx}{x-r} = \int_{\alpha}^{r-\epsilon} \frac{dx}{x-r} + \int_{r+\delta}^{\beta} \frac{dx}{x-r}$ when ϵ and δ independently tend to zero.

Or, $I = \ln \frac{\beta-r}{r-\alpha} + \lim_{\substack{\epsilon \rightarrow 0 \\ \delta \rightarrow 0}} \ln \frac{\epsilon}{\delta}$ and diverges in general.

Only if we take $\epsilon = \delta$, then $I = \ln \frac{\beta-\gamma}{\gamma-\alpha} = \text{finite}$. Then, the integral I will be called the principal value

of the integral $\int_{\alpha}^{\beta} \frac{dx}{x-r}$. Thus $\int_1^3 \frac{dx}{x-2}$ has the principal value $\ln \frac{3-2}{2-1}$ or 0. Similarly, the principle value

of the integral $\int_{-\alpha}^{\alpha} \frac{dx}{x} = 0$.

(c) Convergence of Gamma Function: $\Gamma(x)$ defined by $\Gamma(x) = \int_0^{\infty} e^{-x} x^{p-1} dx$. For $p-1 < 0$, the integral is singular at $x=0$. We write,

$$\Gamma(x) = \int_0^1 e^{-x} x^{p-1} dx + \int_1^{\infty} e^{-x} x^{p-1} dx$$

By comparison test, since $|e^{-x} x^{p-1}| < |x^{p-1}|$ for $0 \leq x \leq 1$, and $\int_0^1 x^{p-1} dx$ converges for $1-p < 1$, the first

integral converges for $p > 0$. The integral $\int_1^{\infty} e^{-x} x^{p-1} dx$ converges for all p . Hence, we obtain that $\Gamma(x)$

converges for all $p > 0$. For real p replaced by a complex p the condition is that $\text{Re } p > 0$.

Complex Variables

CHAPTER

6

6.1 INTRODUCTION

In the last chapters we were concerned with the algebraic properties of a function of a single variable or of many variables. A function of a complex variable $z = x + iy$ is necessarily a function of two variables but all functions of two variables need not be functions of a single complex variable. In this chapter we will discuss the condition for which a function of two variables reduces to a function $f(z)$ of a single complex variable z and the essential features of $f(z)$.

6.1.1 Limit and Continuity

The limit of a function $f(z)$ as $z \rightarrow z_0$ is $l (= u_0 + iv_0)$, i.e., $\lim_{z \rightarrow z_0} f(z) = l$ if $\lim_{x \rightarrow x_0} u = u_0$ and $\lim_{y \rightarrow y_0} v = v_0$, where $f(z) = u + iv$. The formal definition of the limit leads to the

statement that for any preassigned positive quantity ϵ , however small, there exists a number δ such that for $|f(z) - l| < \epsilon$, $|z - z_0| < \delta$, where δ depends on ϵ . Clearly for the limit to exist l must be unique irrespective of the path by which we approach the point $z = z_0$. Thus $\sin z$, e^z , z^2 , to speak of a few functions, have well-defined limits

at every z in the complex plane. But for the function $e^{-\frac{1}{z^2}}$, if we approach the point $z = 0$ along the real axis, $\lim_{z \rightarrow 0} e^{-\frac{1}{z^2}} = 0$, while $\lim_{z \rightarrow 0} e^{-\frac{1}{z^2}}$ diverges if the origin is approached along the imaginary axis. Hence the limit does not exist at $z = 0$. (Such is the case for $\exp\left(\pm \frac{1}{z}\right)$, $\sin\left(\frac{1}{z}\right)$ etc. also). The function thus does not exist at $z = 0$

in the limiting sense. (The function $e^{-\frac{1}{z^2}}$ or $e^{-\frac{1}{z}}$ approaches a new limit every time the path is changed).

Continuity For a function $f(z)$ to be continuous at $z = z_0$, $\lim_{z \rightarrow z_0} f(z) = f(z_0)$. Thus $\sin z$, e^z , any polynomial

are continuous everywhere. If we define a function $f(z) = \frac{\sin z}{z}$ for $z \neq 0$ and $= 1$ for $z = 0$, $f(z)$ is continuous at $z = 0$

But if we simply state that $f(z) = \frac{\sin z}{z}$, it is undefined at $z = 0$ and hence $f(z)$, becomes discontinuous at $z = 0$.

The formulae pertaining to the functions of a complex variable are similar to those of a real variable, namely,

- (i) The sum, difference and the product of continuous functions of a complex variable at a point are also continuous functions at the same point in the complex plane.
- (ii) The ratio of two continuous functions at a point is also continuous at the same point provided the denominator does not vanish there.

Differentiation The derivative of a complex variable $f(z)$ at $z = z_0$ is defined in analogy with that of a real variable as $\left. \frac{df}{dz} \right|_{z_0} = \lim_{z \rightarrow z_0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}$, where the limit must exist and be unique independent of the way in which the point $z = z_0$ is approached. Writing $f(z) = u + iv$ we obtain,

$$\left. \frac{df}{dz} \right|_{z_0} = \frac{\Delta u + i\Delta v}{\Delta z}, \text{ where } \Delta u = u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0),$$

$$\Delta v = v(x_0 + \Delta x, y_0 + \Delta y) - v(x_0, y_0) \text{ and } \Delta z = \Delta x + i\Delta y.$$

$$\text{or, } \left. \frac{df}{dz} \right|_{z_0} = \frac{\Delta u + i\Delta v}{\Delta x + i\Delta y}$$

If the partial derivatives are continuous in a neighbourhood of the point z_0 , then, $\Delta u = \frac{\partial u}{\partial x} \Delta x + \frac{\partial u}{\partial y} \Delta y + \xi(x, y)$ and $\Delta v = \frac{\partial v}{\partial x} \Delta x + \frac{\partial v}{\partial y} \Delta y + \eta(x, y)$, where $\xi(x, y)$ and $\eta(x, y) \rightarrow 0$ as $\Delta x, \Delta y \rightarrow 0$ in such a way that $\frac{\xi(x, y)}{\Delta z}$ and $\frac{\eta(x, y)}{\Delta z}$ go to zero as $\Delta z \rightarrow 0$.

$$\text{or, } \left. \frac{df}{dz} \right|_{z_0} = \lim_{z \rightarrow z_0} \left[\frac{\left(\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) \Delta x + \left(\frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y} \right) \Delta y}{\Delta x + i\Delta y} + \frac{\xi(x, y) + i\eta(x, y)}{\Delta z} \right]$$

Now, let us approach the point $z = z_0$ along the real axis so that $\Delta y = 0$ and $\Delta x \rightarrow 0$. Since the second term on the r. h. s. $\rightarrow 0$ we get, $\left. \frac{df}{dz} \right|_{z_0} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}$.

Similarly approaching the point $z = z_0$ along the imaginary axis so that $\Delta x = 0$ and $\Delta y \rightarrow 0$, we get, $\left. \frac{df}{dz} \right|_{z_0} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y}$.

Since $\frac{df}{dz}$ must have a unique value at $z = z_0$ independent of the manner of approach to $z = z_0$ we get,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ and } \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \quad (6.1)$$

Students can check this equality by approaching long a straight line $y = mx$ for various these are values of m . called Cauchy-Riemann conditions. Thus if the partial derivatives $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}$ are continuous at $z = z_0$ and satisfy Cauchy-Riemann conditions at $z = z_0$, the function is differentiable at $z = z_0$. If Cauchy-Riemann conditions are satisfied and the partial derivatives continuous at every point in a region surrounding $z = z_0$, the function is said to be analytic everywhere in the region.

In polar coordinates, $\frac{\partial u}{\partial x} = \cos \theta \frac{\partial u}{\partial r} - \frac{\sin \theta}{r} \frac{\partial u}{\partial \theta}$, $\left[r = \sqrt{x^2 + y^2}, \theta = \tan^{-1} \frac{y}{x} \right]$, $\frac{\partial u}{\partial y} = \sin \theta \frac{\partial u}{\partial r} + \frac{\cos \theta}{r} \frac{\partial u}{\partial \theta}$,

$$\frac{\partial v}{\partial x} = \cos \theta \frac{\partial v}{\partial r} - \frac{\sin \theta}{r} \frac{\partial v}{\partial \theta}, \frac{\partial v}{\partial y} = \sin \theta \frac{\partial v}{\partial r} + \frac{\cos \theta}{r} \frac{\partial v}{\partial \theta}.$$

Cauchy-Riemann conditions give, on equating coefficients of $\cos \theta$ and $\sin \theta$ respectively, $\frac{\partial u}{\partial r} = \frac{1}{r} \frac{\partial v}{\partial \theta}$,

$$\frac{\partial v}{\partial r} = -\frac{1}{r} \frac{\partial u}{\partial \theta}. \quad (6.2)$$

From this polar form of Cauchy-Riemann conditions we see that for $f = \ln z = u + iv$, $u = \ln r$, $v = \theta$, they are satisfied everywhere except at $r = 0$. The function $\ln z$ is thus analytic everywhere except at $r = 0$. This function will be discussed in section 6.4.

Examples

- (1) $\bar{z} (= x - iy)$ is nowhere an analytic function of z , as can be seen from Cauchy-Riemann conditions.
- (2) Any differentiable function $f(|z|^2)$ of the form $\sum_n c_n |z|^{2u}$, can be seen from Cauchy-Riemann conditions to be analytic only at $z = 0$.
- (3) For the function $f(z) = \frac{\bar{z}^2}{z}$ the Cauchy-Riemann conditions are satisfied only at $z = 0$ but the partial derivatives are not continuous here, since they depend on the manner of approach to $z = 0$ and, hence, the function ceases to be analytic at $z = 0$. (This is left as an exercise for the students).
- (4) From Cauchy-Riemann conditions, we get

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0. \text{ Thus the real and the imaginary parts of the analytic function of a complex}$$

variable satisfy the two dimensional Laplace's equation and are called harmonic functions. But the converse is not necessarily true. For example, $f(z) = ax + iby$, a and b being constants represents a function where the real and the imaginary parts do not satisfy Cauchy-Riemann conditions but obey Laplace's equation for $a \neq b$.

Application: The Cauchy-Riemann conditions show that if we have a two-dimensional vector field with components u and v , then $\bar{\nabla}u \cdot \bar{\nabla}v = 0$. Thus the $u = \text{constant}$ and $v = \text{constant}$ lines are orthogonal to each other. In the two-dimensional problems in electrostatics in a charge-free space the potential ϕ satisfies Laplace's equation. Thus the $u = \text{const.}$ and $v = \text{const.}$ lines (or vice-versa) represent the equipotentials and the electric field lines respectively. Similar is the case with the problems in fluid dynamics for the irrotational flow of an incompressible fluid, where in two dimensions, the $u = \text{const.}$ and the $v = \text{const.}$ lines can be made to coincide with the lines of the velocity field and those of equal potential respectively. These features leave the door open for applying the technique of complex analysis to the problems of physics which is briefly described below.

By utilizing a complex function $w = f(z)$ we transform the original potential problem in electrostatics (or fluid mechanics) to a simpler problem in the $u - v$ plane, where $w = u + iv$. Solving the problem in the $u - v$ plane and transforming back to the z -plane we can solve the original problem. Such a mapping $w = f(z)$ is called conformal mapping because the angle between two curves at the point of intersection remains invariant after mapping to the w -plane. The validity of the method is, however, restricted only to those problems that have sufficient symmetry enabling us to reduce them to two dimensions. Moreover, its efficacy also depends on the appropriate choice of w . (See 'Appendix')

6.2 INTEGRAL THEOREMS

We will describe below a few theorems that will be useful in the integration of the function of a complex variable along the prescribed contours. [The detailed proofs will be sometimes omitted].

1. Cauchy's integral theorem

Let $f(z)$ be the function of a complex variable analytic within and on a closed contour bounding a simply connected region inside which by definition any two points can be connected by a curve that lies entirely within the region.

Then $\oint_C f(z) dz$, where C symbolizes the contour

$$= \oint_C (u + iv)(dx + idy), (f(z) = u + iv)$$

$$= \oint_C [(udx - vdy) + i(udy + vdx)].$$

By applying Green's theorem in a plane,

$$\oint_C f(z) dz = \int_C \left[\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) + i \left(\frac{\partial v}{\partial y} - \frac{\partial u}{\partial x} \right) \right] dx dy$$

= 0 from Cauchy-Riemann conditions.

Hence we obtain Cauchy's theorem that if $f(z)$ be analytic within and on a closed curve enclosing a simply connected region, $\oint_C f(z) dz = 0 \dots$ (6.3)

The converse of Cauchy's integral theorem (called Morera's theorem) is also valid, i.e., if $\oint_C f(z) dz = 0$ for any closed curve in a simply connected region, $f(z)$ is analytic within and on C . (The derivation can be seen in any standard text book on complex variable or mathematical methods and is not given here).

Let us consider a multiply connected region bounded by the closed curves C_1 and C_2 (Fig. 6.1). Clearly, Cauchy's integral theorem cannot be straightforwardly applied here. We connect the two closed curves by the cuts AB and CD so that the region within $ABCD$ as shown by the arrows now becomes simply connected and the function $f(z)$ becomes analytic in the region. We can now apply Cauchy's theorem to get,

$$\oint_{C_1} f(z) dz + \int_A^B f(z) dz - \oint_{C_2} f(z) dz + \int_C^D f(z) dz = 0$$

The minus sign comes before $\int_{C_2} f(z) dz$ because the contour C_2 is traversed in a direction opposite to that in C_1 . Since the pairs of points, A, D and B, C are infinitesimally close to each other,

$$\int_A^B f(z) dz + \int_C^D f(z) dz = 0 \text{ and thus we obtain } \oint_{C_1} f(z) dz = \oint_{C_2} f(z) dz.$$

where both the contours C_1 and C_2 are covered in the same direction in the integral. This is an interesting result because it tells us that independent of the shape and the size of the two contours C_1 and C_2 , $\oint_{C_1} f(z) dz = \oint_{C_2} f(z) dz$ if the function remains analytic in between C_1 and C_2 . This we can choose any convenient contour for evaluating the integral.

If the multiply connected region encloses several closed contours $C_2, C_3, \dots, C_n$, we can proceed as above, construct an effectively simply connected region by cuts and obtain,

$$\int_{C_1} f(z) dz = \sum_{i=2}^n \int_{C_i} f(z) dz \dots \quad (6.4)$$

6.3 CAUCHY'S INTEGRAL FORMULA

The formula states: If a function $f(z)$ be analytic everywhere within and on a simple closed contour then $f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0} dz$, where the contour is described in an anticlockwise sense so that the region enclosed lies to the left of the contour. This is a result characteristic of complex analysis (having

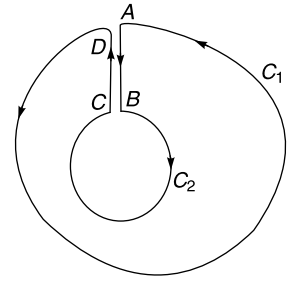


Fig. 6.1

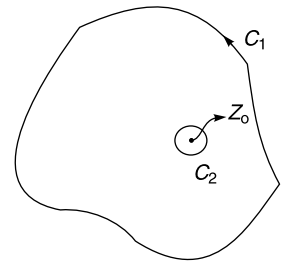


Fig. 6.2

no analogue in real analysis) where the function at a point within a region is known in terms of the values of the function at points on the contour bounding the region.

Surrounding the point z_0 by a small circle C_2 , we can say that the function $\frac{f(z)}{z - z_0}$ is analytic in the multiply connected region bounded by C_1 and C_2 (Fig. 6.2).

Following the result in the previous section we can say that

$$\oint_{C_1} \frac{f(z)}{z - z_0} dz = \oint_{C_2} \frac{f(z)}{z - z_0} dz = \oint_{C_2} \frac{f(z_0)}{z - z_0} dz + \oint_{C_2} \frac{f(z) - f(z_0)}{z - z_0} dz.$$

Now,

$$\oint_{C_2} \frac{f(z_0)}{z - z_0} dz = f(z_0) \oint_{C_2} \frac{dz}{z - z_0} = f(z_0) \oint_{C_2} \frac{ir e^{i\theta}}{re^{i\theta}} d\theta = 2\pi i f(z_0),$$

where on C_2 we put $z - z_0 = re^{i\theta}$.

Now, if $|f(z) - f(z_0)| < \epsilon$. for $|z - z_0| < \eta$ we can easily show that $\left| \oint \frac{f(z) - f(z_0)}{z - z_0} dz \right| < 2\pi\epsilon$ and $\rightarrow 0$ as $\epsilon \rightarrow 0$.

Thus we can write, $f(z_0) = \frac{1}{2\pi i} \oint_{C_1} \frac{f(z)}{z - z_0} dz. \dots$ (6.5)

Starting from $f(z_0)$ we can calculate higher order derivatives at $z = z_0$.

$$\text{Thus } f'(z_0) = \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}$$

Or,

$$f'(z_0) = \lim_{\Delta z \rightarrow 0} \frac{1}{2\pi i} \left[\oint_{C_1} \left\{ \frac{f(z_0)}{z - z_0 - \Delta z} - \frac{f(z)}{z - z_0} \right\} dz \right] \cdot \frac{1}{\Delta z}$$

Or,

$$f'(z_0) = \frac{1}{2\pi i} \oint_{C_1} \frac{f(z)}{(z - z_0)^2} dz$$

Proceeding similarly we can write $f^n(z_0) = \frac{n!}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{n+1}} dz$

Thus, if $f(z)$ be analytic at $z = z_0$, all order derivatives exist at $z = z_0$ and are analytic there.

We have from Cauchy's formula, $f(z) = \frac{1}{2\pi i} \oint_C \frac{f(z')}{z' - z} dz'$.

Taking a point ξ such that $\left| \frac{z - \xi}{z' - \xi} \right| < 1$ we get,

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(z')}{(z' - \xi) - (z - \xi)} dz' = \frac{1}{2\pi i} \oint_C \frac{f(z')}{(z' - \xi)} \sum_n \left(\frac{z - \xi}{z' - \xi} \right)^n dz'$$

Thus $f(z) = \sum_{n=0}^{m-1} (z - \xi)^n \cdot \left[\frac{1}{2\pi i} \oint \frac{f(z')}{(z' - \xi)^{n+1}} dz' \right] + R_m(z)$, where

$R_m(z)$, the remainder after m terms, is given by the remaining terms in the series.

$$\text{Thus } R_m(z) = \frac{1}{2\pi i} \oint_C \frac{f(z')}{(z' - \xi)^{m+1}} (z - \xi)^m \frac{(z' - \xi)}{(z' - z)} dz'$$

$$\text{Or, } R_m(z) = \frac{(z - \xi)^m}{2\pi i} \oint \frac{f(z')}{(z' - \xi)^m (z' - z)} dz'.$$

If $f(z')$ is bounded, $|R_m(z)|$ can be shown to tend to zero as $m \rightarrow \infty$. (This is because $\left| \frac{z - \xi}{z' - \xi} \right| < 1$ and $f(z)$ is bounded)

Thus, the infinite series $\sum_{n=0}^{m-1} \frac{(z - \xi)^n}{n!} f^n(\xi)$ converges to $f(z)$

$$\text{as } m \rightarrow \infty \left[\therefore f^n(\xi) = \frac{n!}{2\pi i} \oint \frac{f(z')}{(z' - \xi)^{n+1}} dz' \right].$$

$$\text{We can thus write, } f(z) = \sum_{n=0}^{\infty} \frac{(z - \xi)^n}{n!} f^n(\xi) \dots \quad (6.6)$$

This is the Taylor series expansion of $f(z)$ about $z = \xi$ with a radius of convergence R such that the series converges for $|z - \xi| < R$. A Taylor series is uniformly convergent within the region $|z - \xi| < R$.

Some common Taylor series expansions about $z = 0$ are:

$$e^z = 1 + z + \frac{z^2}{2!} + \dots (R \rightarrow \infty) \dots \quad 6.7 (a)$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots (R \rightarrow \infty) \dots \quad 6.7 (b)$$

$$\frac{1}{1 - z} = 1 + z + z^2 + \dots (R = 1) \dots \quad 6.7 (c)$$

Thus an analytic function can always be expanded in a unique power series, called the Taylor series, about any point with a non-zero radius of convergence. The converse is also seen to be valid.

Analytic continuation: The function $\frac{1}{1 - z}$ has a power series representation given by Eq. 6.7(c) that converges for $|z| < 1$. The infinite series thus represents an analytic function for $|z| < 1$. But since $\frac{1}{1 - z}$ exists for $|z| > 1$ also, it should be possible to have infinite series representations that would be valid for $|z| > 1$ also.

Indeed if we expanded $\frac{1}{1 - z}$ about any other point $z = \xi$, we would obtain power series representations that

would converge for $|z - \xi| < 1$ and hence would represent an analytic function in the region. Clearly this region extends beyond the first one given by $|z| < 1$ having a region of overlap with the first, where both the power series representations would converge to the same magnitude. By varying the point ξ we can obtain different regions of analyticity of the different power series representations of the function till we exhaust the entire z -plane. It is to be noted that all these circles of convergence have a common point of intersection

at $z = 1$, where $f(z) \left(= \frac{1}{1-z} \right)$ diverges. This process of extending the region of analyticity of the function is

called analytic continuation and widely used in complex analysis. It can be proved that all possible analytic continuations of a function can be obtained by the power series method. Thus the analytic continuation can be represented by defining a function $F(z)$ which is given by $f_1(z)$ for z in domain D_1 (say), $f_2(z)$ for z in D_2 and so on. Thus in the example discussed here $F(z) = \frac{1}{1-z}$ and $f_1(z), f_2(z) \dots$ represent different power series

representations of $\frac{1}{1-z}$ valid in the different regions $D_1, D_2, \dots$

The Taylor series expansion of a function about a point $z = z_0$ requires that the function be analytic in a certain region extending upto the nearest singularity of the function. But suppose that the region under consideration contains certain isolated singularities of the function at certain points. Now the problem is the expansion of a function in such cases. We will discuss in the next section the expansion of a function which is analytic in the annular region between two circles and also on the boundaries of the two circles.

Note: There exist methods of analytic continuations other than those of power series expansions. For example, $\Gamma(z)$ defined by the integral $\int_0^\infty e^{-t} t^{z-1} dt$ exists only for $\text{Re } z > 0$, while if we use the recurrence

relation $\Gamma(z+1) = z \Gamma(z)$, then the domain of analyticity is extended to all z , except for $z = 0, -1, -2, \dots$. The integral representations of certain special functions also serve to extend the domain of analyticity. Moreover the analytic continuation by power series expansions along different paths will lead to a unique result if the paths do not enclose any singularity.

6.4 LAURENT'S EXPANSION, RESIDUE

Let $f(z)$ be analytic in the annular region bounded by C_1 and C_2 and also on the boundaries of C_1 and C_2 (Fig. 6.3). Considering Γ to be an infinitesimally small circle surrounding the point z we can write,

$$\oint_{\Gamma} \frac{f(z')}{z' - z} dz' = 2\pi i f(z), \text{ where we put } z' = z + re^{i\theta}.$$

Applying Cauchy's integral theorem to the function $\frac{f(z')}{z' - z}$, which is analytic in the simply connected region bounded by C_1, C_2 and Γ we

$$\text{obtain, } \oint_{\Gamma} \frac{f(z')}{z' - z} dz' = 2\pi i f(z) = \oint_{C_1} \frac{f(z')}{z' - z} dz' - \oint_{C_2} \frac{f(z')}{z' - z} dz' \quad (6.8)$$

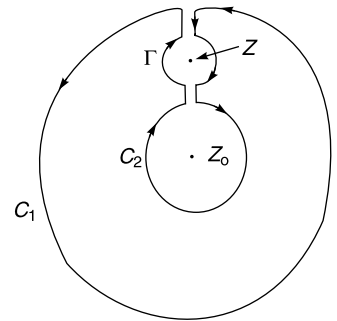


Fig. 6.3

For the first integral, $\oint_{C_1} \frac{f(z')}{z' - z} dz'$ is expressed as,

$$\oint_{C_1} \frac{f(z')}{z' - z} dz' = - \oint_{C_1} \frac{f(z')}{z - z'} dz' = - \oint_{C_1} \frac{f(z')}{(z - z_0) - (z' - z_0)} dz'$$

or,

$$\oint_{C_1} \frac{f(z')}{z' - z} dz' = \oint_{C_1} \frac{f(z')}{(z' - z_0) \left(1 - \frac{z - z_0}{z' - z_0} \right)} dz' = \sum_n (z - z_0)^n \oint_{C_1} \frac{f(z')}{(z' - z_0)^{n+1}} dz'$$

$$\begin{aligned} \oint_{C_2} \frac{f(z')}{z' - z} dz' &= - \oint_{C_2} \frac{f(z')}{(z - z')} dz' = - \oint_{C_2} \frac{f(z')}{[(z - z_0) - (z' - z_0)]} dz' \\ &= - \oint_{C_2} \frac{f(z')}{(z - z_0) \left[1 - \frac{z' - z_0}{z - z_0} \right]} dz' \\ &= - \oint_{C_2} \sum_n \frac{1}{(z - z_0)^{n+1}} (z' - z_0)^n f(z') dz' \end{aligned}$$

Putting in Eq. (6.8) we obtain

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \left[\sum_n (z - z_0)^n \oint_{C_1} \frac{f(z')}{(z' - z_0)^{n+1}} dz' + \sum_{n=0}^{\infty} \frac{1}{(z - z_0)^{n+1}} \oint_{C_2} (z' - z_0)^n f(z') dz' \right] \\ &= \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=0}^{\infty} \frac{b_n}{(z - z_0)^{n+1}}, \text{ where} \end{aligned} \quad (6.9)$$

$$a_n = \frac{1}{2\pi i} \oint_{C_1} \frac{f(z')}{(z' - z_0)^{n+1}} dz', \quad b_n = \frac{1}{2\pi i} \oint_{C_2} (z' - z_0)^n f(z') dz'$$

The series representation (6.9) is called the Laurent's series expansion of a function $f(z)$, the first term being the Taylor series expansion of an analytic function at $z = z_0$ and the second part being non-analytic at $z = z_0$.

The series (6.9) is rewritten as, $f(z) = f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n}$,

where $b_n = \frac{1}{2\pi i} \oint (z' - z_0)^{n-1} f(z') dz'$, the integral being in the positive direction along any contour about $z = z_0$.

The second term $\sum_n \frac{b_n}{(z - z_0)^n}$ is clearly non-analytic at $z = z_0$, where it diverges. The point $z = z_0$ is

a point of isolated singularity since there are no other singularities in its neighbourhood. These points of isolated singularity are called poles. If the second part contains terms up to $n = m$, the function $f(z)$ is said to have a pole of order m , all other terms in the infinite series being zero. In this case $\lim_{z \rightarrow z_0} (z - z_0)^m f(z)$ is finite

and equal to b_m . If the series comprises an infinite number of terms so that there exists no finite m for which $\lim_{z \rightarrow z_0} (z - z_0)^m f(z)$ is finite, the function $f(z)$ is said to possess an essential singularity at $z = z_0$.

b_1 , the coefficient of $\frac{1}{z - z_0}$, is said to be the residue of $f(z)$ at $z = z_0$. b_1 is also equal to $\frac{1}{2\pi i} \oint f(z') dz'$.

[by putting $n = 1$ in b_n]. A pole of order 1 is called a simple pole.

For a simple pole of $f(z)$ at $z = z_0$, the residue $b_1 = \lim_{z \rightarrow z_0} (z - z_0) f(z)$. Similarly the residue at a pole of order m at $z = z_0$ is given by $\frac{1}{(m-1)!} \cdot \frac{d^{m-1}}{dz^{m-1}} [(z - z_0)^m f(z)] \Big|_{z=z_0}$, as can be seen by direct multiplication of $f(z)$ by $(z - z_0)^m$,

Alternatively from Cauchy's integral formula, $(z - z_0)^m f(z) = \phi(z)$ say $= \frac{1}{2\pi i} \oint \frac{\phi(z')}{(z' - z)} dz'$

or,
$$(z - z_0)^m f(z) = \frac{1}{2\pi i} \oint \frac{(z' - z_0)^m f(z')}{(z' - z)} \cdot dz'$$

Clearly, $\frac{1}{(m-1)!} \left[\frac{d^{m-1}}{dz^{m-1}} \Big|_{z=z_0} [(z - z_0)^m f(z)] \right]$ is equal to $\frac{1}{2\pi i} \oint f(z') dz'$. The residue at an essential

singularity is found out by direct Laurent expansion and finding out the coefficient of $\frac{1}{z - z_0}$.

Another type of singularity occurs when the function $f(z)$ becomes multivalued, i.e., corresponding to a single point in the z -plane $f(z)$ attains different values. Thus, $f(z)$ becomes non-analytic and hence does not admit of a Taylor series expansion about the point. Let us consider the function $f(z) = z^{1/2}$. Thus if starting from a point in the z -plane we come back to it by traversing once around the origin the function $f(z)$ changes sign. It returns to its original value after two traversals around the origin. This problem of multivaluedness was taken care of by Riemann by the method described briefly below.

Putting $z = r \exp [i (\theta + 2\pi m)]$, m being an integer or zero, $z^{1/2} = r^{1/2} \exp \left[\frac{i}{2} (\theta + 2\pi m) \right]$, where $0 \leq \theta \leq 2\pi$ (or $-\pi \leq \theta \leq \pi$).

One particular value of m corresponds to a particular branch of the multivalued function. The complex z -plane was assumed to be divided into two Riemann sheets in this case, separated from each other by a cut, where each sheet represents a single branch of the multivalued function. The cut can be any line passing through the origin, which is called the branch point of the function. The point at infinity is another branch point of the function as can be seen by putting $z = \frac{1}{z'}$ and taking $z' \rightarrow 0$. Thus the cut is any line extending from the origin to the point at infinity. The bottom edge of the first sheet is connected to the

top edge of the second sheet while the bottom edge of the second one is joined to the top edge of the first sheet. These are so joined because the function $f(z) = z^{1/2}$ returns to itself after two complete traversals around the origin. To maintain single-valuedness so that we can utilize the theorems of complex variable related to analytic functions, we do not allow any contour to close round the origin by crossing the cut. Thus, during the first traversal round the origin the contour covers the points confined to the $m = 0$ branch

and the function takes up the values represented by $f_1(z) = r^{1/2} e^{i\theta/2}$ ($0 < \theta < 2\pi$ or $-\pi < \theta < \pi$) while

during the second traversal the function takes up the values represented by $f_2(z) = r^{1/2} e^{i(\theta+2\pi)/2}$ ($m = 1$).

Thus, the function remains single valued over the entire Riemann surface comprising the two sheets and the difficulty of multivaluedness is bypassed. Clearly, we have here,

$$\lim_{\epsilon \rightarrow 0} f_1(r, 2\pi - \epsilon) = \lim_{\epsilon \rightarrow 0} f_2(r, 0 + \epsilon) \text{ and}$$

$$\lim_{\epsilon \rightarrow 0} f_2(r, 2\pi - \epsilon) = \lim_{\epsilon \rightarrow 0} f_1(r, 0 + \epsilon).$$

The function $z^{\frac{1}{n}}$ has thus n Riemann sheets joined consecutively in the previous way with the bottom edge of the n th. one connected to the top edge of the first one. The function $\ln z$ requires thus an infinite number of sheets, because $\ln z = \ln r + i(\theta + 2m\pi)$, $m = 0, 1, \dots$ and never returns to its value.

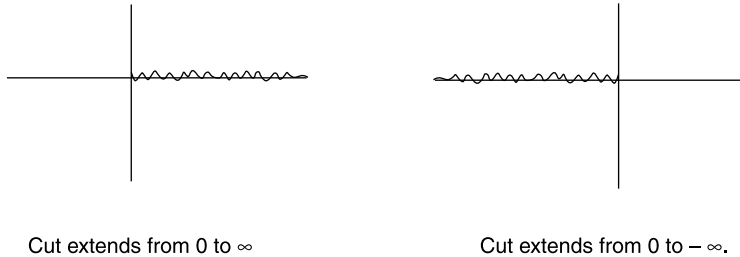


Fig. 6.4

6.5 THEOREM OF RESIDUES

If a function $f(z)$ is analytic everywhere within and on a contour except at a finite number of poles, $\oint_C f(z) dz = 2\pi i$ [Sum of the residues at the poles], where the contour is described in an anticlockwise sense.

Proof: Let a function $f(z)$ be analytic everywhere within and on C except at a finite number of poles $z_1, z_2, \dots$. Surrounding the poles by closed contour $C_1, C_2, \dots$ and inserting cuts at appropriate points we convert the region within C to a simply connected region inside and on which $f(z)$ is analytic (Fig. 6.5). Applying

the integral theorem we get, $\oint_C f(z) dz = \oint_{C_1} f(z) dz + \oint_{C_2} f(z) dz + \dots$

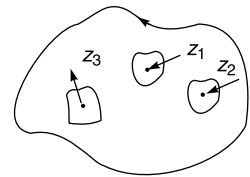


Fig. 6.5

From the definition of residues, $\oint_{C_1} f(z) dz = 2\pi i \sum \text{Residues} \dots\dots$ (6.10)

[All the integrals are described in the same sense.]

A few simple examples of complex integrations will be given below. A simple integration involving a multivalued function and an application of complex analysis to the study of dielectrics are provided in the appendix.

SOLVED PROBLEMS

Problem 1

Find the value of z for which (a) $e^z = a + ib$, (b) $z^4 = 5 + 6i$

Solution:

(a) $e^z = a + ib = r \exp [i(\theta + 2m\pi)]$, m being an integer or zero, where $r = (a^2 + b^2)^{1/2}$, $\theta = \tan^{-1} \frac{b}{a}$.

We, therefore, get, $z = \ln r + i(\theta + 2m\pi)$

(b) $z^4 = 5 + 6i = \sqrt{61} \exp \left[i \left(\tan^{-1} \frac{6}{5} + 2m\pi \right) \right]$, m being an integer or zero

or, $z = (61)^{1/8} \exp \left[\frac{i}{4} \left(\tan^{-1} \frac{6}{5} + 2m\pi \right) \right]$.

Problem 2

Test if the following functions are analytic

(a) $\sinh z$,

(b) $|z|^2$,

(c) $z^{1/2}$.

Solution:

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

$$\frac{\partial u}{\partial x} = \cosh x \cos y = \frac{\partial v}{\partial y}, \frac{\partial u}{\partial y} = -\sinh x \sin y = -\frac{\partial v}{\partial x}.$$

Thus Cauchy-Riemann conditions are satisfied everywhere leading to the analyticity of $\sinh z$ in the entire z -plane.

(b) $f(z) = |z|^2 = x^2 + y^2$.

Cauchy-Riemann conditions we see that $|z|^2$ is analytic only at $z = 0$.

$$\left[\frac{\partial u}{\partial x} = 2x = 0, \frac{\partial v}{\partial y} = 2y = 0, \frac{\partial v}{\partial y} = \frac{\partial v}{\partial x} = 0 \right]$$

(c) $f(z) = z^{1/2}$.

$f(z)$ has a branch point at $z = 0$ only and everywhere else it is analytic.

Problem 3

Prove that if Cauchy-Reimann conditions are satisfied, then any $f(x, y)$ is a function of z -alone.

Solution:

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

$$\frac{\partial f}{\partial z^*} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial z^*} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial z^*} = \frac{1}{2} \frac{\partial f}{\partial x} - \frac{1}{2i} \frac{\partial f}{\partial y}$$

Putting $f = u + iv$ we get, $\frac{\partial f}{\partial z^*} = \frac{1}{2} \frac{\partial}{\partial x} (u + iv) - \frac{1}{2i} \frac{\partial}{\partial y} (u + iv)$

Applying Cauchy-Riemann conditions, $\frac{\partial f}{\partial z^*} = 0$ and hence $f = f(z)$ only. $\left[\because \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \right]$

It can also be shown that if we expand $f(x, y) = u(x, y) + iv(x, y)$ in a Taylor series about $x = \xi, y = \eta$ and apply Cauchy-Reimann conditions we can express $f(x, y)$ as a Taylor series about $z = z_0$ of some function of z . [$z_0 = \xi + i\eta$]

Problem 4

Test the singularities of (a) $\frac{1}{z^2 + i}$, (b) $\frac{1 - \cos z}{z^2}$, (c) $\frac{\sin \frac{1}{z}}{z^2 + 9}$

Solution:

(a) The function has a pole of order 1 at each of the roots of the equation $z^2 + i = 0$ or, $z =$

$$\pm \exp\left[\frac{i}{2} \times \left\{(2n+1)\pi + \frac{\pi}{2}\right\}\right] \text{ for all integer } n, \text{ including } n = 0.$$

$$(b) \frac{1 - \cos z}{z^2} = \frac{2 \sin^2 \frac{1}{2} z}{z^2} = 2 \cdot \frac{\sin^2 \frac{1}{2} z}{\left(\frac{z}{2}\right)^2}.$$

The function $\frac{\sin^2 \frac{z}{2}}{\left(\frac{z}{2}\right)^2}$ has a removable singularity at $z = 0$ because $\lim_{z \rightarrow 0} \frac{\sin^2 \frac{z}{2}}{\left(\frac{z}{2}\right)^2} = 1$

(c) $\sin \frac{1}{z}$ has an essential singularity at $z = 0$ because it has an infinite series expansion in $\frac{1}{z}$. Moreover

$\lim_{z \rightarrow 0} \sin \frac{1}{z}$ is not unique, as can be proved by approaching $z = 0$ along the real or the imaginary

axis. (Actually at the point of essential singularity the function can be made to approach any magnitude by following a suitable path of approach). Also $\frac{1}{z^2+9}$ has poles at $z = \pm 3i$ of order unity each.

Problem 5

Find the residues at the singularities of

(a) $\frac{e^{iz}}{z^2 - a^2},$

(b) $\frac{e^{iz}}{(z^2 + a^2)^2}$

(c) $ze^{\frac{1}{z}}$

Solution:

- (a) The function has two poles at $z = \pm a$ of order one each.

$$\text{Residue at } z = a = \lim_{z \rightarrow a} (z - a) \frac{e^{iz}}{z^2 - a^2} = \frac{e^{ia}}{2a}$$

$$\text{Residue at } z = -a = \lim_{z \rightarrow -a} (z + a) \frac{e^{iz}}{z^2 - a^2} = -\frac{e^{-ia}}{2a}$$

- (b) The function has got two poles at $z = \pm ia$, each of order 2.

Applying the formula of calculating the residue we get,

$$\begin{aligned} \text{Residue at } z = ia \text{ is} &= \frac{d}{dz} \left[(z - ia)^2 \frac{e^{iz}}{(z^2 + a^2)^2} \right]_{z=ia} \\ &= \frac{d}{dz} \left[\frac{e^{iz}}{(z + ia)^2} \right]_{z=ia} = \left[\frac{(z + ia)^2 i e^{iz} - 2e^{iz} (z + ia)}{(z + ia)^4} \right]_{z=ia} \\ &= \left[\frac{-4a^2 i e^{-a} - 2e^{-a} (2ia)}{(2ia)^4} \right] = \frac{-4ia e^{-a} (1 + a)}{16a^4} = \frac{-i e^{-a} (1 + a)}{4a^3}. \end{aligned}$$

$$\begin{aligned} \text{Residue at } z = -ia \text{ is} &= \frac{d}{dz} \left[(z + ia)^2 \frac{e^{iz}}{(z^2 + a^2)^2} \right]_{z=-ia} \\ &= \frac{d}{dz} \left[\frac{e^{iz}}{(z - ia)^2} \right]_{z=-ia} = \left[\frac{(z - ia)^2 i e^{iz} - 2e^{iz} (z - ia)}{(z - ia)^4} \right]_{z=-ia} \\ &= \left[\frac{-4a^2 \cdot i e^a + 2e^a (2ia)}{16a^4} \right] = \frac{4ia e^a (1 - a)}{16a^4} = \frac{i e^a (1 - a)}{4a^3} \end{aligned}$$

(c) The residue of $ze^{\frac{1}{z}}$ at $z = 0$ is the coefficient of $\frac{1}{z}$ in the expansion of function and $= \frac{1}{2}$.

Problem 6

Find the Laurent or Taylor series expansions of

$$(a) f(z) = \frac{1}{(1+z)^{\frac{1}{3}}} \text{ about } z = 0,$$

$$(b) f(z) = \frac{1}{(z^2 - 5z + 6)} \text{ for } 2 < z < 3$$

Solution:

$$(a) f(z) = (1+z)^{-\frac{1}{3}} = 1 + z \left[\frac{d}{dz} (1+z)^{-\frac{1}{3}} \right]_{z=0} + \frac{z^2}{2!} \left[\frac{d^2}{dz^2} (1+z)^{-\frac{1}{3}} \right]_{z=0} + \dots$$

$$\text{or, } f(z) = 1 - \frac{z}{3} + \frac{1}{9} z^2 + \dots \text{ [The expansion is valid for } |z| < 1]$$

$$(b) f(z) = \frac{1}{z^2 - 5z + 6} = \frac{1}{(z-2)(z-3)}$$

$$\frac{1}{(z-3)} = \frac{1}{(3l-3)}, \text{ where } z = 3l$$

$$= -\frac{1}{3} \left[1 + \frac{z}{3} + \frac{z^2}{9} + \dots \right] \text{ [valid for } |z| < 3]$$

$$\frac{1}{(z-2)} = \frac{1}{\left(\frac{2}{m} - 2\right)}, \text{ where } z = \frac{2}{m}$$

$$= \frac{1}{2} \frac{m}{(1-m)} = \frac{1}{z} \left[1 + \frac{2}{z} + \frac{4}{z^2} + \dots \right] \text{ [valid for } |z| > 2]$$

$$\text{or } \frac{1}{(z-2)(z-3)} = \frac{1}{3z} \left[1 + \frac{z}{3} + \frac{z^2}{9} + \dots \right] \left[1 + \frac{2}{z} + \frac{4}{z^2} + \dots \right]$$

$$= -\frac{1}{3z} \left[\sum_n \left(\frac{z}{3}\right)^n \cdot \sum_m \left(\frac{2}{z}\right)^m \right] = -\sum_{\min} \frac{2^m}{3^{n+1}} z^{n-m-1}.$$

Since the two expansions are valid for $|z| < 3$ and $|z| > 2$ respectively, the product is valid in the common region $2 < |z| < 3$.

Problem 7Find the branch points of $(z^2 - 1)^{1/2}$, $(z + i)^{1/4}$.**Solution:**

- (a) For
- $f(z) = \{(z-1)(z+1)\}^{1/2}$
- , the branch points are at
- $z = \pm 1$
- .

The point $z \rightarrow \infty$ is a regular point as can be seen by putting $z = \frac{1}{x}$ in $f(z) = (z^2 - 1)^{1/2}$ which behaves as $\frac{1}{x}$ for $x \rightarrow 0$. The branch cut may be chosen to be a line joining the points $z = \pm 1$.

- (b) For
- $f(z) = (z + i)^{1/4}$
- , the branch points are at
- $z = -i$
- and at
- $z \rightarrow \infty$
- .

The existence of a branch point at $z \rightarrow \infty$ can be demonstrated by putting $z = \frac{1}{z'}$ and taking the limit $z' \rightarrow 0$ in $f\left(\frac{1}{z'}\right)$. The branch cut is any line joining $z = -i$ and $z \rightarrow \infty$.

Problem 8

- (a) Find the complex function if
- $u = e^x \cos y$
- (
- u
- = real part of
- $f(z)$
-)

- (b) Calculate

(i) $\oint \frac{dz}{z^2(z+5)}$ around the contour $|z| = 2$,

(ii) $\oint \frac{dz}{z(z-i)}$ around the circle with centre at $z = \frac{i}{2}$ and radius 1.

(iii) $\int_{-\infty}^{\infty} \frac{\sin mx}{x} dx \quad (m > 0)$

Solution:

- (a) Using Cauchy-Riemann condition we get,

$$\frac{\partial u}{\partial x} = e^x \cos y = \frac{\partial v}{\partial y} \text{ so that } v = e^x \sin y + \phi(x), \text{ where } \phi(x) \text{ is any function of } x. \text{ From } \frac{\partial u}{\partial y} = -e^x \sin y$$

$$= -\frac{\partial v}{\partial x} \text{ we get } \frac{d\phi}{dx} = 0 \text{ so that } \phi = \text{const.}$$

or, $f(z) = u + iv = e^x (\cos y + i \sin y) = e^{x+iy} = e^z$. The const. is ignorable.

- (b) (i) There are two poles, one at
- $z = 0$
- and the other at
- $z = -5$
- . The pole at
- $z = -5$
- lying outside the

$$\text{contour } |z| = 2, \text{ we obtain that } \oint \frac{dz}{z^2(z+5)} = 2\pi i \left[\text{Residue of } \frac{1}{z^2(z+5)} \text{ at } z = 0 \right] = -\frac{2\pi i}{25}.$$

[Apply the formula of calculating the residue at the pole of order 2 or simply find out the coefficient

$$\text{of } \frac{1}{z} \text{ by binomial expansion of } \frac{1}{(z+5)} = \frac{1}{5} \left(1 + \frac{z}{5} \right)^{-1} \text{ and multiplying by } \frac{1}{z^2} \Big]$$

(ii) These are two poles at $z = 0$ and $z = i$ of order 1 each.

$$\oint \frac{dz}{z(z-i)} = 2\pi i [\text{Sum of residues of } \frac{1}{z(z-i)} \text{ at } z=0 \text{ and } z=i \text{ respectively}] = 2\pi i \left[-\frac{1}{i} + \frac{1}{i} \right] = 0.$$

(iii) We consider $f(z) = \frac{e^{imz}}{z}$ we consider the closed contour

(Fig. 6.6) consisting of the semicircle C_s of infinite radius, the part of the real axis comprising the straight lines $-R$ to $-\epsilon$, ϵ to R and the infinitesimally small semicircle C_R of radius ϵ . (Finally we will take $R \rightarrow \infty$ and $\epsilon \rightarrow 0$).

By Jordan's lemma, which we will not prove here,

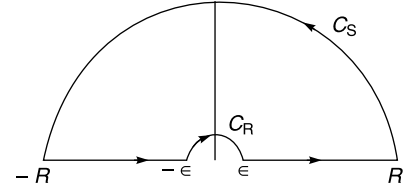


Fig. 6.6

$\int_{C_s} \frac{e^{imz}}{z} dz \rightarrow 0$ as $R \rightarrow \infty$. Hence we can write,

$$\int_{-R}^{-\epsilon} \frac{e^{imx}}{x} dx + \int_{\epsilon}^R \frac{e^{imx}}{x} dx + \int_{C_R} \frac{e^{imz}}{z} dz = 2\pi i. \quad [\text{Sum of the residues of } \frac{e^{imz}}{z} \text{ at the poles within the closed contour}]$$

Since there is no pole inside the closed contour, we obtain $\lim_{\substack{R \rightarrow \infty \\ \epsilon \rightarrow 0}} \left[\int_{-R}^{-\epsilon} \frac{e^{imx}}{x} dx + \int_{\epsilon}^R \frac{e^{imx}}{x} dx \right]$

$$= - \int_{C_R} \frac{e^{imz}}{z} dz = i\pi \quad \left[\text{Since } e^{imz} \rightarrow 1 \text{ as } \epsilon \rightarrow 0 \text{ and } \frac{dz}{z} = id\theta, z \text{ being } = re^{i\theta} \right]$$

or, $P \int_{-\infty}^{\infty} \frac{e^{imx}}{x} dx = i\pi$, where $P \int_{-\infty}^{\infty} \frac{e^{imx}}{x} dx$ indicates the principal value of the integral and

$$= \lim_{\substack{R \rightarrow \infty \\ \epsilon \rightarrow 0}} \left[\int_{-R}^{-\epsilon} \frac{e^{imx}}{x} dx + \int_{\epsilon}^R \frac{e^{imx}}{x} dx \right]. \quad \text{Equating the real and the imaginary parts we obtain}$$

$$P \int_{-\infty}^{\infty} \frac{\sin mx}{x} dx = \pi \text{ and } P \int_{-\infty}^{\infty} \frac{\cos mx}{x} dx = 0$$

SUMMARY

This chapter deals with the analytic properties of the functions of a complex variable and the methods of simple complex integration.

- The definitions of the limit and the continuity of the function of a complex variable are similar to those of a function of two variables, i.e., $\lim_{z \rightarrow z_0} f(z)$ must be unique irrespective of how the point $z \rightarrow z_0$ in the complex plane.
- For a function $f(z) = u + iv$ to be differentiable at $z \rightarrow z_0$ the partial derivatives $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$

should be continuous at $z = z_0$ and Cauchy-Reimann conditions $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$ and $\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$ must be satisfied there. If the partial derivatives are continuous in a region around $z = z_0$ and satisfy Cauchy-Riemann conditions there, the function $f(z)$ is analytic in that region.

- If $f(z)$ is analytic everywhere within and on a closed contour C , $\oint_C f(z) dz = 0$. Moreover, in this case,

$$f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0} dz, \text{ where } z_0 \text{ is any point inside the region bounded by the contour and the contour}$$

is described in the anticlockwise sense so that the region enclosed lies to the left of the contour.

- If $f(z)$ be analytic, everywhere within and on a contour C , then all order derivatives of the function exist at any point in the region and the n th order derivative at $z = z_0$ is given by $f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{n+1}} dz$.

The function $f(z)$ can be expanded in a Taylor series about $z = z_0$ with a finite radius of convergence in this case.

If $f(z)$ be analytic within and on the boundaries of the annular region bounded by C_1 and C_2 , the point z

$$= z_0 \text{ being at the centre of } C_1, f(z) = \sum_n a_n (z - z_0)^n + \sum_n \frac{b_n}{(z - z_0)^n}, \text{ where } a_n = \frac{1}{2\pi i} \oint \frac{f(z')}{(z' - z_0)^{n+1}} dz'$$

$$\text{and } b_n = \frac{1}{2\pi i} \oint (z' - z_0)^{n-1} f(z') dz'.$$

This is called the Laurent's series expansion of $f(z)$ about $z = z_0$.

If the series $\sum_n \frac{b_n}{(z - z_0)^n}$ contains terms upto $n = m$, the function $f(z)$ is said to have a pole of order

m at $z = z_0$. If the series does not terminate, $f(z)$ is said to have an essential singularity at $z = z_0$.

- The function $f(z)$ expanded about a point $z = z_0$ has a region of analyticity extending from $z = z_0$ to the nearest singularity of $f(z)$. Sometimes we can by choosing a different z_0 obtain an expansion of $f(z)$ valid in a region extending the previous one and having a region of overlap with it where the two expansions converge to the same value. This is called the principle of analytic continuation and applied almost everywhere in the analysis of complex variable.
- Apart from isolated singularities the function $f(z)$ may lose analyticity by multivaluedness, the functions $\sqrt{z}$, $\ln z$, being a couple of examples. The function is then made single valued over the complex plane (also called the Riemann Surface) by an ingenious construction of Riemann described briefly in the text.
- Cauchy's theorem of residues states that $\oint_C f(z) dz = 2\pi i$ [Sum of the residues of $f(z)$ at the poles inside

the contour C], where the contour C is described so that the region enclosed lies to the left of C . This theorem is widely used in the complex integrations and is exemplified in the text by a few illustrations. However, the integrations involving branch points are omitted in the main text and a simple example is given in the appendix.

The appendix also contains a simple application to the study of dielectric response to the electromagnetic waves.

MULTIPLE CHOICE QUESTIONS

1. The regions bounded by the loci of points satisfying the equations $|z - 5 - 3i| = 2$ and $|z + 1 - 2i| = 3$ respectively
 - (a) have only one common point;
 - (b) are such that the second one is completely enclosed by the first,
 - (c) have no common point.
 - (d) intersect over a finite region.
2. The equation $e^z = 2 + 3i$ has
 - (a) only one solution
 - (b) an infinite number of solution
 - (c) two distinct solutions
 - (d) no valid solution.
3. If the real part of $f(z)$ be $x^2 - y^2$, the corresponding imaginary part is
 - (a) $2xy$
 - (b) $x^2 + y^2$
 - (c) $-2xy$
 - (d) cannot be determined.
4. If $f(z)$ be analytic, $f^*(z^*)$ is
 - (a) analytic
 - (b) not analytic
 - (c) analytic only on the real axis
 - (d) none of these.
5. If the imaginary part of $f(z)$ be $e^y \cos x$, $f(z)$ is
 - (a) $ie^{(y+ix)}$
 - (b) $i e^{y-ix}$
 - (c) e^{x+iy}
 - (d) e^{x-iy}
6. If $|z_1 + z_2| = |z_1 - z_2|$, the difference of the phases of z_1 and z_2 is

(a) $\pi/4$	(b) $\pi/2$	(c) $3\pi/4$	(d) $\pi/6$
7. If the two functions be $f_1(z) = \sin z$ and $f_2(z) = \sin 1/z$, then
 - (a) f_1 and f_2 are analytic functions in the entire z -plane.
 - (b) Neither of them is everywhere analytic,
 - (c) f_1 is analytic everywhere but f_2 is not analytic at $z = 0$
 - (d) f_2 is analytic everywhere but f_1 is nowhere analytic
8. The function $\tan z$
 - (a) has simple poles at $z = (2n + 1)\pi/2$, ($n = 0, 1, \dots$) and no singularity at $z = \infty$.
 - (b) has simple poles at $z = (2n + 1)\pi/2$ and an essential singularity at $z = \infty$.
 - (c) has simple poles at $z = (2n + 1)\pi/2$ and also at $z = \infty$.
 - (d) has poles nowhere in the complex plane.

9. The function $e^{1/z}$ has
 (a) a simple pole at $z = 0$ and at $z = \infty$,
 (b) an essential singularity at $z = 0$ only
 (c) an essential singularity at $z = \infty$ also,
 (d) none of these.
10. Consider two functions $(z - a)^{1/3}$ and $\sqrt{(z - \alpha)(z - \beta)}$. Which of the statements is correct?
 (a) The first one has a branch point at $z = a$ while the second one has branch points at $z = \alpha$ and $z = \beta$
 (b) The first one has a pole at $z = a$ while the second one has poles at $z = \alpha$ and $z = \beta$
 (c) The first one has a branch point at $z = a$ while the second one has poles at $z = \alpha$ and $z = \beta$
 (d) The first one has a pole at $z = a$ while the second one has branch points at $z = \alpha$ and $z = \beta$
11. The residue of $1/z$ at $z = 0$ is,
 (a) 1 (b) $1/6$ (c) $1/3$ (d) $1/2$
12. The residue of $\frac{1}{z^2(z-2)}$ at $z = 0$ is
 (a) $\frac{1}{4}$ (b) $-\frac{1}{4}$ (c) $\frac{1}{2}$ (d) 1
13. The residue of $\frac{z^2 + 3z}{(z-1)^2(z-2)}$ at $z = 1$ is
 (a) -9 (b) 9 (c) -4 (d) 0
14. The residue of $z^2 e^{1/z}$ at $z = \infty$ is
 (a) $1/6$ (b) $-1/6$ (c) $1/3$ (d) 0
15. The magnitude of $\oint (z-1) e^{1/z} dz$, where c is the contour $|z| = 2$ is
 (a) πi (b) $-\pi i$ (c) $2\pi i$ (d) 0
16. The magnitude of $\int_c \frac{dz}{(z+2)^2(z+3)}$, where c is the circle $|z| = 1$ is
 (a) $2\pi i$ (b) $-2\pi i$ (c) 0 (d) πi
17. Which of the following functions are nowhere analytic?
 (a) z^{-2} (b) $Im z$ (c) e^z (d) $\log z$, (e) $|z|^4$
18. Which of the Laurent expansions of the following functions about $z = 0$ has the greatest radius of convergence?
 $\frac{1}{z^2(z-1)}$, $\frac{1}{(z-2)(z-3)}$, $\frac{e^z}{z}$, $\frac{e^{z-1}}{z(z-2)}$
19. The magnitude of $\int_0 \frac{dz}{z^2(z-1)(z-2)}$, where c is the circle $|z| = 3$ described clockwise is
 (a) $\frac{3\pi i}{2}$ (b) 0 (c) πi (d) $-\frac{3\pi i}{2}$
20. $\int_0^\infty \frac{\cos mx}{x^2 + 1} dx$ is equal to

(a) $\frac{\pi e^{-m}}{2}$

(b) $\frac{\pi e^m}{2}$

(c) 0

(d) πe^m

Answers to MCQs

1. (c)	2. (b)	3. (a)	4. (a)	5. (b)	6. (b)	7. (c)
8. (b)	9. (b)	10. (a)	11. (b)	12. (b)	13. (a)	14. (b)
15. (b)	16. (c)	17. (b)	18. (c)	19. (b)	20. (a)	

EXERCISE

- Construct the function $u + iv$ given that (i) $v = xy$, (ii) $u = \ln(x^2 + y^2)$.
- Test if the following functions are differentiable.

(i) $\sqrt{z-a}$,

(ii) $|z|^4$,

(iii) $\frac{\bar{z}^2}{z}$ ($\bar{z} = x - iy$),

(iv) $\sqrt{|xy|}$,

(v) $Re(z^2)$

- Find out the nature of the singularities of the following functions

(a) $\frac{\sin \frac{1}{z}}{z^2 + 9}$,

(b) $\frac{\sin z}{z^4}$,

(c) $\left(\frac{z+3}{z-2}\right)^{\frac{1}{4}}$,

(d) $(z + zi)^{1/4}$,

(e) $\ln(z + zi)$,

(f) $\frac{\cot z}{z^2 + z^3}$,

(g) $\sqrt[3]{z^2 - 1}$

- Find out the Laurent expansion of

(a) $\frac{1}{z^2}$ about $z = -1$,

(b) $\frac{1}{(z-2)(z-4)}$ for $2 < |z| < 4$,

(c) $\sin\left(\frac{1}{z-c}\right)$ about $z = a$,

(d) $\frac{1+z^2}{1-z}$ about $z \rightarrow \infty$,

(e) $\frac{1}{z(z-1)}$ about $z = 0$.

- Is it meaningful to calculate $\int_a^b |z|^2 dz$, where a and b are any two points in the z -plane? Explain.

- Calculate the residues of the following function at all the singularities

(a) $\cot z$,

(b) $\frac{1}{(z-2)^2(z+2)}$,

(c) $\operatorname{cosec}^2 z$,

$$(d) \frac{1}{z^2 \tan z}, \quad (e) \frac{\sin \frac{1}{z}}{z^2 + a^2}, \quad (f) \frac{1}{z(e^2 - 1)}$$

7. Integrate by the method of contour integration

$$(a) \int_0^\infty \left(\frac{\cos bx - \cos ax}{x^2} \right) dx, \quad a > b > 0, \quad (b) \oint \frac{dz}{z^2(z+3)} \text{ [along the circle } |z| = 2]$$

$$(c) \int_0^{2\pi} \frac{d\theta}{a + b \cos \theta} \quad (a > b > 0), \quad (d) \int \frac{\sin z}{z(z^2 + 16)} \text{ (along the circle } |z| = 2)$$

$$(e) \oint \frac{dz}{z(z-2)^3} \text{ along the circle } |z-2| = 1.$$

8. Find out the real and the imaginary parts of

$$(a) \sinh^{-1}(1+i), \quad (b) \ln(1-i), \quad (c) (2+5i)^{3-6i}.$$

9. Prove that if Cauchy-Riemann conditions are satisfied, $f(x, y)$ becomes a function of $z = x + iy$ by expanding $f(x, y)$ about any point $z = x_0, y = y_0$ and applying Cauchy-Riemann conditions.

Note: The residue of a function $f(z)$ at $z = \infty$ is the same as the residue of $-\frac{1}{z^2} f\left(\frac{1}{z}\right)$ at $z = 0$ (as can

be seen from the definition of the residue in terms of the integral). Alternatively it can also be defined

as $\frac{1}{2\pi i} \int_C f(z) dz$, where the contour C is now described along the clockwise direction and encloses all

the poles in the finite plane. A function may be analytic at $z = \infty$ but can have a residue there. Thus $f(z)$

$= \frac{1}{z}$ is analytic at $z = \infty$, but the residue of $f(z)$ at $z = \infty = \text{Residue of } \left(-\frac{1}{z}\right)$ at $z = 0$ and equals -1 .

Hints and Answers to Exercise Questions

- $\frac{1}{2} (x^2 - y^2) + lxy,$
 - $\ln(x^2 + y^2) + 2i \tan^{-1} y/x.$
- Everywhere except at $z = 0,$
 - only at $z = 0,$
 - everywhere except at $z = 0,$
 - nowhere,
 - only at $z = 0.$
- Essential singularity at $z = 0,$ simple poles at $z = \pm 3i,$
 - pole of order 3 at $z = 0,$
 - branch points at $z = -3, z = -2,$

- (d) branch point at $z = -2i$,
 (e) branch point at $z = -3i$,
 (f) pole of order 3 at $z = 0$, poles of order 1 at $z = -1, z = n\pi$ ($n = \pm 1, \pm 2, \dots$),
 (g) branch points at $z = \pm 1$
4. (a) $1 + 2(z+1) + 3(z+1)^2 + \dots$, (b) $-\frac{1}{4} \sum 2^{\{n-2(n+1)\}} z^{(n-m-1)}$
 (c) $\frac{1}{(z-a)} - \frac{1}{3!} \frac{1}{(z-a)^3} + \dots$ (d) $-z \left(1 + \frac{1}{z^2}\right) \left(1 + \frac{1}{z} + \frac{1}{z^2} + \dots\right):$
 (e) $-\frac{1}{z}(1+z+z^2+\dots).$
5. No. (**Hint:** check analyticity)
6. (a) 1 at $z = m\pi$, ($m = 0, \pm 1, \dots$) (b) $\frac{1}{16}$ at $z = -2$ and $-\frac{1}{16}$ at $z = 2$,
 (c) 0. (d) $-\frac{1}{3}$ at $z = 0$,
 (e) $-\frac{1}{2a} \sinh\left(\frac{1}{a}\right)$ at $z = \pm ia$, $\frac{1}{a^2}$ at $z = 0$ (f) $-\frac{1}{2}$ at $z = 0$, $\frac{1}{2\pi ni}$ at $z = 2\pi ni$ ($n \neq 0$),
7. (a) $-\frac{\pi}{2}(b-a)$, (b) $\frac{2\pi i}{16}$,
 (c) 0 if $a - \sqrt{a^2 - b^2} < 2b$, $\frac{2\pi}{b} \cdot \frac{1}{\sqrt{a^2 - b^2}}$ if $a - \sqrt{a^2 - b^2} < 2b < a + \sqrt{a^2 - b^2}$,
 and 0 if $a - \sqrt{a^2 - b^2} > 2b$
 (d) $\frac{2\pi i}{16}$, (e) If $\frac{\pi i}{4}$.
8. (a) Real part = x , Im. part = y , where $z = x + iy$ and

$$e^z = \left\{ \left[1 \pm 5^{\frac{1}{4}} \cos\left(\frac{1}{2} \tan^{-1} 2\right) \right] + i \left[1 \pm 5^{\frac{1}{4}} \sin\left(\frac{1}{2} \tan^{-1} 2\right) \right] \right\}.$$

 (b) $\frac{1}{2} \ln 2 - i \left(\frac{\pi}{4} \pm 2n\pi \right), n = 0, \pm 1, \dots$
 (c) Write $2 + 5i$ and $3 - 6i$ in the forms of $r_1 e^{i\theta_1}$ and $r_2 e^{i\theta_2}$ respectively. Then expand and separate the real and the imaginary parts respectively to get the result. The final answer looks cumbersome and is hence not given here.

APPENDIX

(A) Example of Conformal Mapping

We give below a very popular example (illustrated in many text books) of the application of conformal mapping to electrostatics. The problem is to find out the electrostatic potential in the inside of a cylinder of radius R whose two halves ($0 < \theta < \pi$ and $\pi < \theta < 2\pi$) are kept at potentials ϕ_1 and ϕ_2 respectively.

Thus $\phi(R, \theta) = \phi_1$ for $0 < \theta < \pi$ and $= \phi_2$ for $\pi < \theta < 2\pi$.

Let the mapping be effected by a transformation

$$w = i \frac{R - z}{R + z} = u + iv = \rho e^{i\psi}.$$

Putting $z = Re^{i\theta}$ we get, $w = i \left(\frac{1 - e^{i\theta}}{1 + e^{i\theta}} \right) = \tan \frac{\theta}{2}$, which shows that the circle $|z| = R$ gets mapped to the

real u -axis, the two semicircles being mapped to the positive and the negative u -axis respectively. Similarly from $w = u + iv$ it can be shown that the inside of the circle $|z| = R$ is transformed into the upper half w -plane.

The next task is to construct a function which satisfies Laplace's equation and also the boundary condition stated above.

Thus, we take a function $g(\rho, \psi)$ which is the imaginary part of an analytic function of w such that $g = \phi_1$ for $\psi = 0$ and $g = \phi_2$ for $\psi = \pi$. Clearly the choice $g(\rho, \psi) = \alpha + \beta\psi = \text{Im}(i\alpha + \beta \ln w)$ satisfies the boundary

condition $\alpha = \phi_1$ and $\alpha + \beta\pi = \phi_2$ giving $\beta = \frac{\phi_2 - \phi_1}{\pi}$. $g(\rho, \psi)$ clearly satisfies Laplace's equation and hence serves our purpose.

The required potential is $\phi(x, y) = \phi_1 + \left(\frac{\phi_2 - \phi_1}{\pi} \right) \psi(x, y)$.

Again from the expression of $w = i \left(\frac{1 - e^{i\theta}}{1 + e^{i\theta}} \right)$ and $x + iy = Re^{i\theta}$,

$$\text{we get, } u = \frac{2Ry}{(R+x)^2 + y^2}, \quad v = \frac{R^2 + x^2 + y^2}{(R+x)^2 + y^2} \cdot \tan \psi = \frac{R^2 - x^2 - y^2}{2Ry}.$$

$$\text{Thus, } \phi(x, y) = \phi_1 + \left(\frac{\phi_2 - \phi_1}{\pi} \right) \tan^{-1} \frac{R^2 - x^2 - y^2}{2Ry}.$$

It may appear that the choice of $g = \alpha + \beta\psi$ is not unique and hence may lead to ambiguity in the final solution. But the ambiguity disappears when we recall that the solution of Laplace's equation with a given choice of boundary condition (Dirichlet or Neumann) is unique.

The conformal mapping also finds application in problems of hydrodynamics where for a two dimensional flow the velocity function follow Laplace's equation. The applications can be seen in text books on continuum mechanics and are not given here.

(B) Kramers-Kronig Relation

An important application of the analysis of complex variables arises in connection with the study of response of a dielectric to an electromagnetic field.

Let the dielectric constant expressed as a function of frequency w be given by $\epsilon(w)$. If $\epsilon(w)$, analytically continued in the complex domain is an analytic function of w in the upper half plane we can write from Cauchy's theorem,

$$\epsilon(w) - 1 = \frac{1}{2\pi i} \oint_c \frac{[\epsilon(w') - 1]}{w' - w} dw' \dots\dots A(1)$$

The contour c consists of an infinite semicircle in the upper half plane and the entire real axis except for an infinitesimally small semicircular detour below the point $w' = w$, where the integrand has a pole. Using the fact that $\epsilon(w) - 1 \rightarrow 0$ as $w \rightarrow \infty$, we can make the contribution from the infinite semicircle go to zero so that we can write,

$\epsilon(w) - 1 = \frac{1}{2\pi i} p \int_{-\infty}^{\infty} \frac{[\epsilon(w') - 1]}{w' - w} dw' + \frac{1}{2} [\epsilon(w) - 1]$, where the second part comes from the small semicircular detour.

Equating the real and imaginary parts we get,

$$\text{Re } \epsilon(w) = 1 + \frac{1}{\pi} p \int_{-\infty}^{\infty} \frac{\text{Im } \epsilon(w')}{w' - w} dw' \text{ and}$$

$$\text{Im } \epsilon(w) = -\frac{1}{\pi} p \int_{-\infty}^{\infty} \frac{[\text{Re } \epsilon(w') - 1]}{w' - w} dw' \dots A(2)$$

The real and the imaginary parts of $\epsilon(w)$, as obtained from A(2), are said to be the Hilbert transforms of each other. Knowing empirically the behaviour of $\text{Im } \epsilon(w)$ we can predict about the dependence of $\text{Re } \epsilon(w)$ on w . The relations A(2), known as the dispersion relations are of importance in the disciplines of solid state physics, elementary particle physics etc.

(C) Integration Involving Multivalued Functions: An Example

Let us consider $I = \int_0^{\infty} \frac{x^{a-1}}{1+x} dx$ ($0 < a < 1$)

We replace the integrand by $f(z) = \frac{z^{a-1}}{1+z}$, which is clearly multivalued function having two branch points

at $z = 0$ and $z = \infty$. and has a simple pole at $z = -1$. The contour comprises an infinitely large circle c_1 , an infinitesimally small circle c_2 , the lines AB and CD respectively, (Fig. 6.7)

Along $c_1 \oint \frac{z^{a-1}}{1+z} dz \rightarrow 0$ since $|z^{a-1}| \rightarrow 0$ for $0 < a < 1$ and $\frac{dz}{1+z} \rightarrow ie^{i\theta} d\theta$ for $z = Re^{i\theta}$.

Similarly $\oint_{C_2} \frac{z^{a-1}}{1+z} dz \rightarrow \int_{2\pi}^0 r^{a-1} \cdot e^{i\theta(a-1)} \cdot ir e^{i\theta} d\theta$ and

hence goes to zero as $r \rightarrow 0$.

Along DC , $\int \frac{z^{a-1}}{1+z} dz = e^{2\pi ia} \int_{\infty}^0 \frac{x^{a-1}}{1+x} dx$

and along AB , $\int \frac{z^{a-1}}{1+z} dz = \int_0^{\infty} \frac{x^{a-1}}{1+x} dx$.

Summing up all the contributions we get,

$$\begin{aligned} I(1 - e^{2\pi ia}) &= 2\pi i \left(\text{Residue at } z = -1 \text{ of } \frac{z^{a-1}}{1+z} \right) \\ &= 2\pi i (-1)^{a-1}. \end{aligned}$$

Hence $I = \frac{2\pi i (-1)^{a-1}}{(1 - e^{2\pi ia})} = \frac{\pi}{\sin \pi a}$ after simplification.

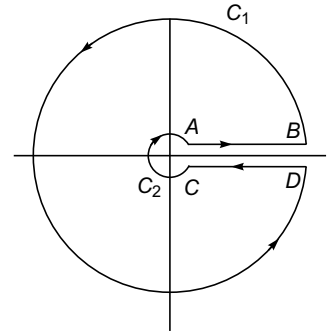


Fig. 6.7

Ordinary Differential Equations, Special Functions

CHAPTER

7

7.1 INTRODUCTION, FIRST ORDER ORDINARY DIFFERENTIAL EQUATIONS

Ordinary differential equations appear almost ubiquitously in problems of science, not to speak of physics alone. In this chapter we will dwell mainly on differential equations of second order after a very brief revision of the equations of first order.

A brief review of the first order differential equations.

A general first order differential equation in two variables x and y can be written as

$P(x, y)dx + Q(x, y)dy = 0$. This can be written as $\frac{dy}{dx} = -\frac{P(x, y)}{Q(x, y)}$, where $Q(x, y) \neq 0$. The easily solvable equations are given below.

(i) When $\frac{P(x, y)}{Q(x, y)}$ can be written as, $\frac{f(x)}{g(y)}$ we obtain, $f(x)dx + g(y)dy = 0$, on integrating which we get the solution as a family of surfaces of the form $F(x, y) = \text{constant}$, where $F(x, y) = \int f(x)dx + \int g(y)dy$.

(ii) When $\frac{P(x, y)}{Q(x, y)} = \frac{ax + by}{cx + dy}$, we get, (a, b, c, d are constants)

$\frac{dy}{dx} = -\left(\frac{ax + by}{cx + dy}\right)$. This is a homogeneous equation which we can

solve by putting $y = vx$ to obtain an equation of the form $x \frac{dv}{dx} = \phi(v)$ leading to a solution by the method of the separation of variables as in (i).

$\left(\text{Here, } \phi(v) = -\left(\frac{a + bv}{c + dv}\right) - v \right)$

(iii) When $\frac{P(x, y)}{Q(x, y)} = \frac{Ax + By + C}{Dx + Ey + F}$, (A, B, C, D, E, F are constants), we transform the equation to a

homogeneous form by putting $x = x + h, y = Y + k$ and adjusting the values of h, k so as to obey the equations $Ah + Bk + C = 0, Dh + Ek + F = 0$. We can then apply the method in (ii).

(iv) When $P(x, y)dx + Q(x, y)dy$ can be written as $dz(x, y)$ we can write by integrating, $z(x, y) = \int (Pdx + Qdy) = \text{constant}$. The first order differential $Pdx + Qdy$ is then said to be exact. Work done in an infinitesimal displacement $= \vec{F} \cdot d\vec{r}$. If this differential is exact the total work done in the displacement

from 1 to 2 can be written as $\int_1^2 \vec{F} \cdot d\vec{r}$ and is dependent only on the end points 1 and 2 irrespective of the path connecting 1 and 2.

If the differential $P(x, y)dx + Q(x, y)dy$ cannot be written as $dz(x, y)$ the differential is said to be inexact. In the case of two variables x and y we can find out a function $\lambda(x, y)$ so that $\lambda(Pdx + Qdy) = \text{exact}$. λ is called

the integrating factor. Since $\lambda(Pdx + Qdy) = dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy$.

For arbitrary dx and dy the condition is obviously, $\frac{\partial}{\partial y} (\lambda P) = \frac{\partial}{\partial x} (\lambda Q) \dots$ (7.1)

For an exact differential $Pdx + Qdy$ the condition $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ ($\lambda = 1$ here) is satisfied.

From Eq. (7.1) we get, $\lambda \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} \right) = Q \frac{\partial \lambda}{\partial x} - P \frac{\partial \lambda}{\partial y}$

(i) If $\frac{\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} \right)}{Q} = \text{function of } x \text{ alone}$, we obtain a solution by writing $\frac{1}{\lambda} \frac{d\lambda}{dx} = f(x)$ leading to

$$\lambda = \exp \left[\int f(x) dx \right].$$

Once we find out λ the solution of the equation $\lambda(Pdx + Qdy) = dF$ is obvious. [The solution is $dF = 0$ or $F(x, y) = \text{constant}$]

(ii) If $\frac{\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} \right)}{P} = \text{function of } y \text{ alone}$, we obtain a solution by writing $\frac{1}{\lambda} \frac{d\lambda}{dy} = g(y)$, so that we can find out λ by integration and proceed to obtain the solution of the differential equation as in (i).

For a general case when the conditions (i) and (ii) are not satisfied we can find out λ by solving the partial differential equation (7.1) which is always possible in two dimensions.

For a system of three variable x, y, z the corresponding equation satisfied by the integrating factor λ is

$\lambda(Pdx + Qdy + Rdz) = dF(x, y, z)$. We obtain, $\lambda P = \frac{\partial F}{\partial x}, \lambda Q = \frac{\partial F}{\partial y}, \lambda R = \frac{\partial F}{\partial z}$ for the differential to be exact.

Hence, $\frac{\partial}{\partial y} (\lambda P) - \frac{\partial}{\partial x} (\lambda Q) = \frac{\partial}{\partial z} (\lambda Q) - \frac{\partial}{\partial y} (\lambda R) = \frac{\partial}{\partial x} (\lambda R) - \frac{\partial}{\partial z} (\lambda P) = 0$

The required condition is, therefore,

$$P \left(\frac{\partial Q}{\partial z} - \frac{\partial R}{\partial y} \right) + Q \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z} \right) + R \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} \right) = 0 \dots \quad (7.2)$$

The required condition for the equation $Pdx + Qdy + Rdz = 0$ to have a unique solution $F(x, y, z) = \text{constant}$ leads to the criterion (the proof is omitted here) that there must be an infinite number of points in the neighborhood of any arbitrary point lying on the surface $F(x, y, z) = \text{constant}$ that are not accessible from the point by a path satisfying $dF = 0$. The general proof was given by Caratheodory and when applied to thermodynamics gives rise to an axiomatic formulation. It proves the existence of an integrating factor λ that is a function of temperature alone such that $\lambda dQ = \text{exact differential}$, where dQ is the infinitesimal heat transferred in a process and an inexact differential in general. ($\lambda = \frac{1}{T}$, T being the temp. in absolute scale)

7.2 LINEAR SECOND ORDER DIFFERENTIAL EQUATION

In this section we will be concerned with a few characteristics valid for the solutions of all differential equations in general. By the order of a differential equation we mean the order of the highest order derivative of the equation while the term linear implies that the derivatives and the dependent variable occur in the first power alone and no term involving their products appears in the equation. We will concentrate ourselves to the solution of the second order differential equations though the properties of the solutions are valid in general for higher order equations also. A second order differential equation is of the form $f(D)y = F(x)$,

where $D = \frac{d}{dx}$, $D^2 = \frac{d^2}{dx^2}$ and $f(D)$ is a linear quadratic expression in D . The general solution of this equation

is given by $y = y_h + y_p$, where y_h is the solution of the homogeneous equation $f(D)y_h = 0$ and y_p , a particular solution of the complete equation satisfying $f(D)y_p = F(x)$.

Let the homogeneous equation be of the form, $\frac{d^2 y}{dx^2} + P(x) \frac{dy}{dx} + Q(x)y = 0$

Let y_1 and y_2 be two functions of x satisfying the above equation so that we can write, $\frac{d^2 y_1}{dx^2} + P \frac{dy_1}{dx} + Q y_1 = 0$
and $\frac{d^2 y_2}{dx^2} + P \frac{dy_2}{dx} + Q y_2 = 0$.

Multiplying the first by y_2 , the second by y_1 and subtracting one from the other we get,

$$\frac{d}{dx} \left(y_2 \frac{dy_1}{dx} - y_1 \frac{dy_2}{dx} \right) + P \left(y_2 \frac{dy_1}{dx} - y_1 \frac{dy_2}{dx} \right) = 0,$$

or
$$y_1 \frac{dy_2}{dx} - y_2 \frac{dy_1}{dx} = W(y_1, y_2) = W_o \exp \left[- \int_{x_1}^x P dx \right], \dots \quad (7.3)$$

where W_o is the Wronskian of y_1 and y_2 evaluated at $x = x_1$.

Since $\exp\left[-\int_{x_1}^x P dx\right]$ is non-zero, $W(y_1, y_2)$ is not zero if $W_o \neq 0$. Hence for $W_o \neq 0$ the two functions y_1

and y_2 are linearly independent over the entire range of values of x where the equation is valid. Thus a second order differential equation possesses two linearly independent solutions. The equation (7.3) can be written as,

$$\frac{d}{dx}\left(\frac{y_2}{y_1}\right) = \left(y_1 \frac{dy_1}{dx} - y_2 \frac{dy_1}{dx}\right) \frac{1}{y_1^2} = \frac{W_o}{y_1^2} \exp\left[-\int_{x_1}^x P dx\right].$$

Or,
$$y_2 = y_1 \int \frac{W_o}{y_1^2} \exp\left(-\int_{x_1}^{x_2} P dx\right) dx \dots \quad (7.4)$$

Clearly, only two solutions are linearly independent because if we assume any third solution y_3 to satisfy the equation we can see from (7.4) that y_3 will be a linear combination of y_2 and y_1 . Any general solution of the homogeneous linear differential equation can then be written as $y = c_1 y_1 + c_2 y_2$ where c_1, c_2 are arbitrary

constants that can be determined from the conditions imposed on y . For example by choosing y and $\frac{dy}{dx}$ at

$x = x_o$ (x_o is any point in the interval) we can find out c_1 and c_2 . The general solution of the equation with the inhomogeneous term on the r. h. s. is thus written as $y = c_1 y_1 + c_2 y_2 + y_p(x)$, c_1 and c_2 being determined from the conditions imposed on the solution. The conditions can be any pair of independent conditions, such as,

specification of y and y' ($= \frac{dy}{dx}$) at any point x_o , or y_1, y_2 at $x = x_1$ and x_2 respectively for example. Thus for a

fixed pair of conditions imposed on the solution a second order differential equation has a unique solution in the entire range of the independent variable. This uniqueness holds good in the case of a differential equation of any finite order.

The solution of the inhomogeneous differential equation is, except for the case of constant coefficients on the l. h. s. and a few standard simple functions on the r. h. s. that we will deal with in a subsequent section, beset with algebraic difficulties. This needs some advanced techniques like, the Green's function or the integral transform techniques which are beyond the scope of our discussions. Below we will discuss a very useful method known as the method of the variation of parameters that enables us to determine the solution of the inhomogeneous equation provided we know the solution of the homogeneous part.

7.2.1 Variation of Parameters Method

Let the differential equation be $\frac{d^2 y}{dx^2} + P \frac{dy}{dx} + Qy = f(x)$ Suppose that y_1 and y_2 are two independent solutions of the homogeneous equation so that the general solution of the homogeneous equation is $y = c_1 y_1 + c_2 y_2$, c_1 and c_2 , being arbitrary constants. Considering c_1 and c_2 to be functions of x we assume that $y = c_1(x) y_1 + c_2(x) y_2$ is a solution of the full inhomogeneous equation.

We get,
$$\frac{dy}{dx} = c_1 y_1' + c_1' y_1 + c_2 y_2' + c_2' y_2.$$

Let us impose the condition $c_1' y_1 + c_2' y_2 = 0$ so that we obtain,

$$\frac{dy}{dx} = c_1 y_1' + c_2 y_2' \dots \quad (7.5)$$

This condition on the hitherto arbitrary functions $c_1(x)$ and $c_2(x)$ to simplify the method does not in any way lead to any loss of generality of the complete solution.

Putting in the original equation after differentiation of (7.5) we get,

$$c_1' y_1' + c_2' y_2' = f(x) \dots \quad (7.6)$$

Solving the Eq. (7.6) and $c_1' y_1 + c_2' y_2 = 0$ we get,

$c_1' = \frac{f(x) y_2}{(y_2 y_1' - y_2' y_1)}$ yielding $c_1 = - \int_{x_1}^x \frac{y_2 f(x)}{W(x)} dx$, where $W(x)$ is the Wronskian and x_1 is completely arbitrary.

Similarly, $c_2' = \frac{f(x) y_1}{y_1 y_2' - y_1' y_2}$ yielding $c_2 = \int_{x_2}^x \frac{y_1 f(x)}{W(x)} dx$, x_2 being arbitrary as before.

We have earlier seen that it is possible to obtain the second solution to the homogeneous equation provided the first solution is known. Again the method just described enables us to find out the complete solution of the inhomogeneous equation by knowing the solution of the homogeneous part alone. Hence we can know the full solution of the inhomogeneous equation if we can somehow find out one solution of the homogeneous equation.

7.3 LINEAR SECOND ORDER DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

Let us consider a linear second order differential equation with constant coefficients of the form

$$\frac{d^2 y}{dx^2} + a \frac{dy}{dx} + by = 0, \quad a, b \text{ being constants.}$$

The equation can be written in the form $(D^2 + aD + b)y = 0$, where $D = \frac{d}{dx}$. This can further be written as $(D - \alpha)(D - \beta)y = 0$, where $\alpha + \beta = -a$, $\alpha\beta = b$. One solution of this equation is the same as that of $(D - \beta)y = 0$. The equation $(D - \beta)y = 0$ has the solution given by $\frac{dy}{y} = \beta dx$, or, $y = \exp(\beta x)$. The equation can also be written as $(D - \beta)(D - \alpha)y = 0$. The solution of the equation $(D - \alpha)y = 0$ is clearly $y \sim e^{\alpha x}$. Thus $e^{\alpha x}$ and $e^{\beta x}$ are two independent solutions of the equation. The general solution of this equation is given by $y = c_1 e^{\alpha x} + c_2 e^{\beta x}$, where $\alpha, \beta = \frac{1}{2}(-a \pm \sqrt{a^2 - 4b})$. If $a^2 - 4b > 0$, we get two distinct real values of α and β . If $a^2 - 4b < 0$, we get two complex conjugate values of α, β and thus the general solution is of the form $y = e^{-ax/2} [c_1 \cos \gamma x + c_2 \sin \gamma x]$, where $\gamma = (4b - a^2)^{1/2}$. If $a^2 - 4b = 0$, the roots are identical and thus we get only one solution $y = e^{\alpha x}$ by this method. To find out the second solution we can use the method prescribed earlier or put $y = e^{\alpha x} f(x)$ in the equation to obtain,

$$\frac{dy}{dx} = e^{\alpha x} f'(x) + \alpha e^{\alpha x} f(x); \frac{d^2 y}{dx^2} = e^{\alpha x} f''(x) + 2\alpha e^{\alpha x} f'(x) + \alpha^2 e^{\alpha x} f(x).$$

Or,
$$f''(x) + (2\alpha + a)f'(x) + (\alpha^2 + a\alpha + b)f = 0$$

Since $\alpha = -a/2$ and $\alpha^2 + a\alpha + b = 0$ we get, $f''(x) = 0$ yielding $f(x) = Ax + B$, A and B being two constants. The second solution is thus $y = (Ax + B)e^{\alpha x}$, the two independent solutions being of the form $e^{\alpha x}$ and $xe^{\alpha x}$ respectively. We will now discuss briefly the method of obtaining solutions of a few inhomogeneous equations. The method of the variation of parameters of course provides a general method of finding the solution. Here we will instead show how to obtain solutions of some standard equations more simply.

- (i) Let us consider the equation $\frac{d^2 y}{dx^2} + b^2 y = \sin ax$ ($h \neq a$). If $b \neq a$, let $y = A \sin ax$ be a particular solution of the problem. Putting in the equation we get, $A(b^2 - a^2) = 1$ giving $A = \frac{1}{b^2 - a^2}$.

The complete solution is, $y = c_1 \sin bx + c_2 \cos bx + \frac{\sin ax}{(b^2 - a^2)}$, where the part containing c_1 and c_2 is the solution of the homogeneous equation, also called the complementary function.

Since $D^2(\sin ax)$ or $D^2(\cos ax) = -a^2 \sin ax$ or $-a^2 \cos ax$ respectively, the solution of an equation of the form $f(D^2)y = \sin ax$ or $\cos ax$ is also written as, $y = \frac{1}{f(-a^2)} \sin ax$ or $y = \frac{1}{f(-a^2)} \cos ax$ respectively.

If $b = a$ the above method falls through. To find out the general solution of the equation we substitute $y = f(x) \sin ax$ in the equation to get, $\left[\frac{dy}{dx} = f' \sin ax + a f \cos ax \text{ and } \frac{d^2 y}{dx^2} = f'' \sin ax + 2af' \cos ax - a^2 f \sin ax \right]$, $f'' \sin ax + 2af' \cos ax = \sin ax$

Putting $f' = z$ we get, $z' + 2a z \cot ax = 1$.

or, $z \exp \int 2a \cot ax \, dx = \int (\exp \int 2a \cot ax \, dx) dx + \text{const.}$

or, $z \exp [2 \ln \sin ax] = \int \exp (2 \ln \sin ax) dx + \text{const.}$

or, $z \sin^2 ax = \int \sin^2 ax \, dx + \text{const} = \frac{1}{2} \left[x - \frac{\sin 2ax}{2a} \right] + \text{const.}$

or, $z = \frac{1}{2} \left[x \operatorname{cosec}^2 ax - \frac{\cot ax}{a} \right] + c \operatorname{cosec}^2 ax$, c being an arbitrary constant.

or, $f = \frac{1}{2} \left[-\frac{x}{a} \cot ax - \int \frac{(-\cot ax)}{a} dx \right] - \frac{1}{2a^2} \ln (\sin ax) - \frac{c \cot ax}{a} + c_1$ ($c_1 = \text{arbitrary constant}$)

$$= -\frac{x}{2a} \cot ax + c_2 \cot ax + c_1, c_2 = -c/a$$

or $y = f(x) \sin ax = c_1 \sin ax + c_2 \cos ax - \frac{x}{2a} \cos ax$.

$$(ii) \quad \frac{d^2 y}{dx^2} - b^2 y = e^{ax} \quad (a \neq b)$$

To find out the particular solution we put $y = A e^{ax}$ so that we get, by putting in the equation

$$A = \frac{1}{a^2 - b^2}.$$

The complete solution is thus $c_1 e^{bx} + c_2 e^{-bx} + \frac{e^{ax}}{(a^2 - b^2)}$.

If $b = a$ the above method falls through. To find out the general solution we put $y = f(x) e^{ax}$ so that $y' = e^{ax} f' + a f e^{ax}$, $y'' = e^{ax} f'' + 2a e^{ax} f' + a^2 e^{ax} f$. We obtain by putting in the equation $f'' + 2a f' = 1$

Putting $f' = z$ we get, $z' + 2az = 1$, yielding $z e^{2ax} = \frac{e^{2ax}}{2a} + c$, c being an arbitrary constant.

or, $f' = \frac{1}{2a} + c e^{-2ax}$, giving $f = \frac{x}{2a} - \frac{c e^{-2ax}}{2a} + c_1$, c_1 being a constant.

or, $y = f e^{ax} = c_1 e^{ax} + c_2 e^{-ax} + \frac{1}{2a} x e^{ax}$, where $c_2 = -\frac{c}{2a}$.

$$(iii) \quad A \frac{d^2 y}{dx^2} + B \frac{dy}{dx} + Cy = \alpha x^2 + \beta x + \gamma$$

Let y_p be the particular integral given by $y_p = a x^2 + bx + c$, a, b, c being constants to be determined. Putting in the equation we get,

$$2Aa + B(2ax + b) + c(ax^2 + bx + c) = \alpha x^2 + \beta x + \gamma.$$

Equating the coefficients of x^2 , x and the x -independent term we get,

$$Ca = \alpha, Cb + 2aB = \beta, Cc + Bb + 2Aa = \gamma.$$

From these three equalities we can find out a, b, c and hence the general solution to the equation is given by $y = c_1 e^{\alpha_1 x} + c_2 e^{\alpha_2 x} + ax^2 + bx + c$, where α_1, α_2 are the roots of the equation $Am^2 + Bm + C = 0$ and C_1, C_2 the arbitrary constants.

We are not going to discuss these examples in a greater detail in this context because the variation of parameter method can in principle deal in a general way with these problems. Below will be, discussed briefly a few applications of the method described above to the solution of some physical problems.

Examples

(i) A CR circuit connected to an a.c. source (Fig. 7.1).

Let the e. m. f. be represented by $e = E_o \sin wt$.

The equation of e. m. f. is $Ri + Q/C = e = E_o \sin wt$

$$\text{or,} \quad \frac{dQ}{dt} + \frac{Q}{CR} = \frac{E_o}{R} \sin wt.$$

The integrating factor is seen to be $\exp \frac{t}{CR}$.

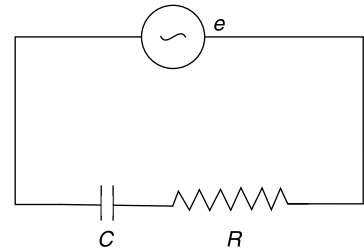


Fig. 7.1

[Using the method discussed in section 7.1]

The solution is, $Q \exp \left[\frac{t}{CR} \right] = \int \frac{E_o}{R} \sin wt e^{t/CR} dt + \text{const.}$

or,
$$Q \exp \left(\frac{t}{CR} \right) = \frac{E_o C \exp \left(\frac{t}{CR} \right)}{(1 + w^2 c^2 R^2)^{1/2}} \sin (wt - \alpha) + C_1$$

or,
$$Q = \frac{E_o C}{(1 + w^2 C^2 R^2)^{1/2}} \sin (wt - \alpha) + C_1 \exp \left(\frac{-t}{CR} \right) [C_1 = \text{constant}],$$

where $\alpha = \tan^{-1}(wCR)$.

The steady state charge developed is given by putting the second term on the right hand side to zero.

The steady state current flowing in the circuit is given by $I = \frac{dQ}{dt}$.

$$\text{Or, } I_{\text{steady}} = \frac{E_o C w}{(1 + w^2 C^2 R^2)^{1/2}} \cos (wt - \alpha).$$

Thus the current leads the e. m. f. in phase by an angle $\frac{\pi}{2} - \alpha$ or $\cot^{-1}(wCR)$

(ii) A series *LCR* circuit connected to an a. c. source (Fig. 7.2).

The equation of e. m. f. in the circuit is,

$$L \frac{dI}{dt} + RI + \frac{Q}{C} = e = E_o \sin wt.$$

Putting $I = \frac{dQ}{dt}$ we get,

$$\frac{d^2 Q}{dt^2} + \frac{R}{L} \frac{dQ}{dt} + \frac{Q}{LC} = \frac{E_o}{L} \sin wt \quad (7.7)$$

The steady state Q is given by $Q = Q_o \sin (wt - \alpha)$, where α has to be determined. Apart from this steady state Q , there will be an unsteady part of Q given by the complementary function, i.e., the solution of the homogeneous equation. The unsteady solution given by

$$Q = e^{\frac{-R}{2L}t} \left[A \exp \left(\sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} t \right) + B \exp \left(-\sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} t \right) \right], A \text{ and } B \text{ being constants, decays}$$

exponentially. Putting $Q = Q_o \sin (wt - \alpha)$ in (7.7) we get, (7.7)

$$\left(\frac{1}{LC} - w^2 \right) Q_o \sin (wt - \alpha) + \frac{wR}{L} Q_o \cos (wt - \alpha) = \frac{E_o}{L} \sin wt. \quad (7.8)$$

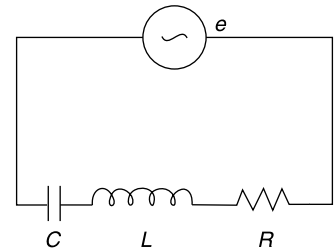


Fig. 7.2

Equating the coefficients of $\sin wt$ and $\cos wt$ respectively we get,

$$\left[\left(\frac{1}{LC} - w^2 \right) \cos \alpha + \frac{wR}{L} \sin \alpha \right] Q_o = \frac{E_o}{L} \text{ and } \dots \quad (7.9)$$

$$\left(\frac{1}{LC} - w^2 \right) \sin \alpha = \frac{wR}{L} \cos \alpha, \text{ giving } \tan \alpha = \frac{wR}{L \left(\frac{1}{LC} - w^2 \right)}$$

$$\text{Putting in (7.9) we get, } Q_o = \frac{E_o}{L \left[\left(\frac{1}{LC} - w^2 \right) \cos \alpha + \frac{wR}{L} \sin \alpha \right]} = \frac{E_o}{wz},$$

$$\text{where } z = \left[R^2 + \left(wL - \frac{1}{wc} \right)^2 \right]^{\frac{1}{2}}$$

$$\text{or, } I = \frac{dQ}{dt} = wQ_o \cos(wt - \alpha) = \frac{E_o}{z} \cos(wt - \alpha), \alpha = \tan^{-1} \frac{wR}{L \left(\frac{1}{LC} - w^2 \right)}$$

Thus the phase lead or lag of the current with respect to the e. m. f. is determined by α . When $w^2 = \frac{1}{LC}$, $\alpha = \frac{\pi}{2}$ and the current $I \left(= \frac{E_o}{z} \sin wt \right)$ is in the same phase as the e. m. f., determining the condition of resonance in the circuit. In this condition the p.d. across the capacitance and that across the inductance being equal in magnitude but opposite in phase cancel each other out leading to a purely resistive circuit, where the energy spent by the battery is used up in overcoming the resistance of the circuit.

- (c) The mechanical analogue of the series LCR circuit is the phenomenon of damped oscillatory motion subjected to an external periodic force. The equation of oscillation is given by,

$$m \frac{d^2 x}{dt^2} + k \frac{dx}{dt} + m w_o^2 x = F \sin wt, w_o \text{ being the natural frequency of oscillation. Dividing by } m \text{ we get}$$

$$\text{an equation of the form, } \frac{d^2 x}{dt^2} + 2b \frac{dx}{dt} + w_o^2 x = f \sin wt, \text{ where } b = \frac{k}{m} \text{ is the damping coefficient and } f = \frac{F}{m}.$$

Solving as before we get the solution of the steady part as,

$$x = \frac{f \sin(wt - \alpha)}{\left[(w_o^2 - w^2)^2 + 4b^2 w^2 \right]^{\frac{1}{2}}}, \alpha = \tan^{-1} \left(\frac{2wb}{w_o^2 - w^2} \right).$$

The velocity resonance (i.e., the peak value of the velocity $\frac{dx}{dt}$) occurs at $w = w_o$, when the velocity is in phase with the external periodic force. The role of resistance in the series LCR circuit is played here by

damping. (The amplitude resonance however occurs at a shifted value of frequency). At steady state the energy spent by the driving force is used to overcome the damping of oscillation.

7.4 SERIES SOLUTION METHOD

In this section we will consider a general method, the method of power series expansion, that is applicable to the second order linear differential equations with variable coefficients as well. Before going into the details of the method we will discuss a few general characteristics of the solution.

Let the differential equation be $\frac{d^2 y}{dx^2} + P(x) \frac{dy}{dx} + Q(x)y = 0$.

A point $x = x_o$ will be said to be an ordinary point of the equation if $P(x)$ and $Q(x)$ are analytic there, i.e. if $P(x)$ and $Q(x)$ can be expanded in Taylor series about $x = x_o$ having a finite region of convergence. If either $P(x)$ or $Q(x)$ (or both of them) ceases to be analytic at $x = x_o$ the point will be a singular point of the differential equation. We will restrict ourselves to isolated singularities only for our method to be applicable straightforwardly. If the singularity be such that $\lim_{x \rightarrow x_o} (x - x_o) P(x)$ and $\lim_{x \rightarrow x_o} (x - x_o)^2 Q(x)$ be both finite, the

singularity will be called a regular singular point of the equation. For a n th. order linear differential equation

of the form $\frac{d^n y}{dx^n} + p_1(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + p_n(x)y = 0$, the point $x = x_o$ is an ordinary point if $p_1 \dots p_n$ are analytic at $x = x_o$ and a regular singular point of the equation if $\lim_{x \rightarrow x_o} (x - x_o) p_1(x)$, $\lim_{x \rightarrow x_o} (x - x_o)^2 p_2(x) \dots$

$\lim_{x \rightarrow x_o} (x - x_o)^n p_n(x)$ are all finite. The definition applies when we consider the equation in a complex plane so that P and Q are functions of the complex variable z . If the above-mentioned conditions are not satisfied at the point, it is called an irregular singular point.

1. Solution about an Ordinary Point

When $P(x)$ and $Q(x)$ are analytic functions of x in a certain region and $x = x_o$ is any point in the region one can show that we can obtain a solution by assuming a power series about $x = x_o$ so that the solution will also be an analytic function in the same region. (The proof is straightforward but will not be given here for the sake of brevity). The region of convergence extends to the nearest singularity of $P(x)$ or $Q(x)$.

2. Solution about a Regular Singular Point

When $x = x_o$ is a regular singular point of the differential equation we can proceed to obtain a solution of the

form $y = \sum_n a_n (x - x_o)^{n+\rho}$, where the index ρ can be found out by solving a quadratic equation, called the

indicial equation that can be determined by putting y in the differential equations. This method, called the Frobenius method can be applied, as will be demonstrated in the text by an example, also to the case (a) above but in the case of an ordinary point the method leads to an unnecessary complication. The detailed nature of the solution can be ascertained from a theorem due to Fuchs and is stated below (the proof is omitted here and can be seen in any standard text).

- (i) If $\rho_1 \sim \rho_2$ (ρ_1, ρ_2 are two roots of the indicial equation) is neither zero nor an integer, the two solutions corresponding to two values of ρ can be found out by a straightforward series solution.

- (ii) If $\rho_1 = \rho_2$, one solution can be determined by the series solution. The other independent solution can be found out by the method described earlier of obtaining the second solution from the first (Section 7.2).
- (iii) If $\rho_1 \sim \rho_2 = \text{integer}$, at least one solution (generally corresponding to the larger root) can be determined by the method of series solution, while the other can be found out by the method described in the section 7.2.

In any one of the cases enlisted above, at least one of the two independent solutions becomes multivalued (or regarded as a function of complex variable develops a branch point type singularity).

3. Solution Near an Irregular Singular Point

In this case the straightforward application of the series solution often fails to yield any reasonable result. There is no well-defined prescription of tackling this situation. Sometimes the transformation of variables leads to an equation that is solvable by the usual methods. Here at least one of the solutions has an essential singularity at the point. We will deal with it further in the Section (7.10).

Now we will discuss certain standard differential equations and special functions arising out of their solution.

7.5 LEGENDRE DIFFERENTIAL EQUATION

The differential equation is, $(1 - x^2) \frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + \lambda y = 0$, λ being any real constant.

The equation has regular singular points in the finite domain at $x = \pm 1$. To check if the point at infinity is a singular point we put $x' = \frac{1}{x}$ so that $\frac{dy}{dx} = -\frac{1}{x^2} \frac{dy}{dx'} = -x'^2 \frac{dy}{dx'}$,

$$\frac{d^2 y}{dx^2} = \frac{d}{dx} \left(\frac{dy}{dx} \right) = \frac{d}{dx'} \left(-x'^2 \frac{dy}{dx'} \right) \cdot (-x'^2) = x'^4 \frac{d^2 y}{dx'^2} + 2x'^3 \frac{dy}{dx'},$$

Substituting in the equation we get,

$$\left(1 - \frac{1}{x'^2} \right) \left[x'^4 \frac{d^2 y}{dx'^2} + 2x'^3 \frac{dy}{dx'} \right] + 2x' \frac{dy}{dx'} + \lambda y = 0.$$

$$\text{or,} \quad (x'^4 - x'^2) \frac{d^2 y}{dx'^2} + 2x'^3 \frac{dy}{dx'} + \lambda y = 0$$

or, $\frac{d^2 y}{dx'^2} + \frac{2x'}{(x'^2 - 1)} \frac{dy}{dx'} + \frac{\lambda}{x'^2(x'^2 - 1)} y = 0$, so that $x' = 0$ is a regular singular point of this differential equation. Hence the point at infinity is a regular singular point of the original differential equation.

The point $x = 0$ is an ordinary point of the differential equation so that we can attempt a power series solution about $x = 0$ by writing $y = \sum_n a_n x^n$ in the equation. [If we want to have a power series solution about $x = \pm 1$, we have to put $y = \sum_n a_n (x \mp 1)^{n+\rho}$ in the equation. If the desired expansion is about the point at

infinity we have to write $y = \sum_n a_n \left(\frac{1}{x}\right)^{n+\rho}$ or $y = \sum_n a_n x^{\alpha-n}$, where $\alpha = -\rho$. We obtain after substitution,

$$(1-x^2) \sum_n n(n-1) a_n x^{n-2} - 2x \sum_n n a_n x^{n-1} + \lambda \sum_n a_n x^n = 0$$

or,

$$\sum_n n(n-1) a_n x^{n-2} + \sum_n [\lambda - n(n+1)] a_n x^n = 0$$

Equating the coefficients of x^n , we get

$$a_{n+2} = - \left[\frac{\lambda - n(n+1)}{(n+2)(n+1)} \right] a_n \dots\dots\dots (7.10)$$

Taking a_0 and a_1 as two arbitrary constants, we can express the coefficients $a_2, a_4, \dots$ in terms of a_0 and $a_3, a_5, \dots$ in terms of a_1 . [It is to be noted that we would obtain a similar recurrence relation if we substitute

$y = \sum_n a_n x^{n+\rho}$ in the equation. We would get, $\sum_n (n+\rho)(n+\rho-1) a_n x^{n+\rho-2} + \sum_n [\lambda - (n+\rho)(n+\rho+1)] a_n x^{n+\rho} = 0$.

Equating the coefficient of the lowest power of x , i.e., $x^{\rho-2}$ to zero, we get, $\rho(\rho-1) a_0 = 0$. Taking $a_0 \neq 0$, we get $\rho = 0, 1$.

Equating the coefficient of the next higher power of x , i.e., $x^{\rho-1}$ to zero we get, $(\rho+1)\rho a_1 = 0$. Thus for $\rho = 0$ we can take both a_0 and a_1 to be non zero. For $\rho = 1$ only $a_0 \neq 0$. Putting $\rho = 0$, we obtain the same recurrence relation as Eq. (7.10). The recurrence relation for $\rho = 1$ is basically the same as that for $\rho = 0$ and containing a_1 . Note that this happens to be the case whenever we attempt a series solution of the form $y = \sum_n a_n (x - x_0)^{n+\rho}$, x_0 being an ordinary point of the equation.

So the $\rho = 0$ recurrence relations lead to the general solution, which is also the same as that obtained by starting with $y = \sum_n a_n x^n$.

From (7.10), $a_2 = -\frac{\lambda}{2!} a_0, a_4 = -\frac{(\lambda-6)}{12} a_2 = \frac{\lambda(\lambda-6)}{4!} a_0 \dots$

$$a_3 = -\frac{(\lambda-2)}{3!} a_1, a_5 = \frac{(\lambda-2)(\lambda-12)}{5!} a_1 \dots$$

Thus, $y = a_0 \left[1 - \frac{\lambda}{2!} x^2 + \frac{\lambda(\lambda-6)}{4!} x^4 - \dots \right] + a_1 x \left[1 - \frac{(\lambda-2)}{3!} x^2 + \frac{(\lambda-2)(\lambda-12)}{5!} x^4 - \dots \right]$

If we replace λ by $l(l+1)$ we can write,

$$y = a_0 \left[1 - \frac{l(l+1)}{2!} x^2 + \frac{l(l-2)(l+1)(l+3)}{4!} x^4 - \dots \right] + a_1 x \left[1 - \frac{(l+2)(l-1)}{3!} x^2 + \frac{(l+2)(l+4)(l-1)(l-3)}{5!} x^4 - \dots \right] \quad (7.11)$$

The ratio of the successive terms $\left| \frac{t_{n+1}}{t_n} \right|$ as $n \rightarrow \infty$ goes as x^2 and hence converges for $|x| < 1$. This is

expected because the singularities of the differential equation occur at $|x| = 1$. Near $|x| = 1$, the series behaves like the series of $\ln(1 - x^2)$ [i.e., the ratio of the successive terms of the two series behaves similarly near $|x| = 1$] and hence diverges logarithmically near $|x| = 1$. In order that the solution remains finite for $|x| = 1$ as well, the series has to be terminated. This can be done if l is an integer. Two cases arise here – (a) l = even positive or odd negative integer when the first series terminates and (b) l = odd positive or even negative integer when the second series terminates. In either case one series reduces to a polynomial of degree l denoted by $P_l(x)$. The second linearly independent solution is given by equation (7.4), where P is equal to

$$-\frac{2x}{1-x^2} \text{ and hence } y_2 = y_1 W_o \int \frac{\exp\left(\int \frac{2x}{1-x^2} dx\right)}{y_1^2} dx, \text{ } y_1 \text{ being equal to } P_l(x)$$

$$\text{or, } y_2 = P_l(x) \int \frac{1}{y_1^2 (1-x^2)} dx = P_l(x) \int \frac{dx}{P_l^2 (1-x^2)} dx.$$

$$y_2 \text{ being denoted as } Q_L(x) \text{ we get, } Q_L(x) = P_l(x) \int \frac{dx}{P_l^2 (1-x^2)}$$

$$\text{A few simple } Q_1 \text{'s are :- } Q_o(x) = \frac{1}{2} \ln \frac{1+x}{1-x}, Q_1 = \frac{1}{2} \ln \frac{1+x}{1-x} - \frac{1}{x}.$$

The second linearly independent solution is seen after integration to have a multivaluedness of the form $\ln \frac{1+z}{1-z}$ for all l , when regarded as a function of the complex variable z . To make it effectively single-valued we draw a cut along the line joining $z = \pm 1$ on the real axis and define $Q_l(z)$ real for real $z > 1$. ($z = \pm 1$ are the two branch points). $Q_l(z)$ is also called the Legendre function of the second kind.

We define the Legendre polynomials by specifying suitable values of a_o, a_1 , the two arbitrary constants. We choose the values as, $a_o = (-1)^{l/2} \frac{l!}{2^l \left[\left(\frac{l}{2} \right)! \right]^2}$ (including $l = 0$) for even l and $a_o = (-1)^{(l-1)/2} \frac{(l+1)!}{2^l \left(\frac{l-1}{2} \right)! \left(\frac{l+1}{2} \right)!}$ for odd l .

These choices give, $P_o(x) = 1, P_1(x) = x, P_2(x) = \frac{1}{2} (3x^2 - 1), \dots$
Thus $P_l(1) = 1$ for all l with a_o given as above.

7.6 GENERATING FUNCTION

The generating function of Legendre polynomials is given by $F(x, t) = (1 - 2tx + t^2)^{-1/2} = \sum_n t^n P_n(x)$. [The same definition also applies in the complex plane with x replaced by z]. This can be verified by expanding the $l. h. s.$ in powers of t and equating the coefficients of different powers of t with the expressions for P_n 's

as given above. This can also be directly proved from the differential equation itself by the technique of complex variable but is omitted here.

Differentiating the relation $(1 - 2tx + t^2)^{-1/2} = \sum_n t^n P_n(x)$ with respect to t we obtain,

$$(x - t) (1 - 2tx + t^2)^{-3/2} = \sum_n nt^{n-1} P_n(x).$$

Or,
$$(x - t) \sum_n t^n P_n(x) = (1 - 2tx + t^2) \sum_n nt^{n-1} P_n(x).$$

Or,
$$\sum_n x (1 + 2n) t^n P_n(x) - \sum_n (1 + n) t^{n+1} P_n(x) - \sum_n nt^{n-1} P_n(x) = 0.$$

Equating the coefficient of t^n we get,

$$x(1 + 2n) P_n = n P_{n-1}(x) + (n + 1) P_{n+1}(x) \dots \quad (7.12)$$

Differentiating the relation $(1 - 2tx + t^2)^{-1/2} = \sum_n t^n P_n(x)$ with respect to x we get,

$$t \sum_n t^n P_n' = (1 - 2tx + t^2) \sum_n t^n P_n'(x).$$

Or,
$$\sum_n t^{n+1} P_n'(x) = \sum_n t^n P_n'(x) - 2x \sum_n t^{n+1} P_n'(x) + \sum_n t^{n+2} P_n'(x)$$

Equating the coefficient of t^n we get,

$$P_{n-1}(x) + 2x P_{n-1}'(x) = P_n'(x) + P_{n-2}'(x).$$

Or
$$P_n(x) + 2x P_n'(x) = P_{n+1}'(x) + P_{n-1}'(x) \dots \quad (7.13)$$

Combining Eqs. (7.10) and (7.11) we obtain again a recurrence relation involving the derivatives alone and numerous other relations.

7.6.1 Orthogonality Relation

we have,
$$\frac{1}{(1 - 2tx + t^2)^{1/2}} = \sum_n t^n P_n'(x)$$

Taking the square of both the sides we get on integrating with respect to x ,
$$\int_{-1}^1 \frac{dx}{(1 - 2tx + t^2)} = \int_{-1}^1 \sum_{n,m} t^{n+m} P_n P_m dx$$

or,
$$-\frac{1}{2t} \ln(1 - 2tx + t^2) \Big|_{-1}^1 = \sum_{n,m} t^{n+m} \int_{-1}^1 P_n P_m dx.$$

or,
$$\frac{1}{t} \ln \left(\frac{1+t}{1-t} \right) = \sum_{n,m} t^{n+m} \int_{-1}^1 P_n P_m dx. \dots \quad (7.14)$$

Using the differential equation itself we can write,

$$(1 - x^2)P_n'' - 2x P_n' + n(n+1)P_n = 0$$

and

$$(1 - x^2)P_m'' - 2x P_m' + m(m+1)P_m = 0$$

Multiplying the first by P_m , the second by P_n and subtracting we get,

$$\frac{d}{dx} [(1 - x^2)(P_m P_n' - P_n P_m')] = (m - n)(m + n + 1)P_m P_n$$

Integrating with respect to x between -1 and 1 we get,

$$(m - n)(m + n + 1) \int_{-1}^1 P_m P_n dx = 0$$

$$\text{Thus for } m \neq n, \int_{-1}^1 P_m P_n dx = 0 \quad \dots\dots (7.15)$$

Utilizing (7.15) in (7.14) we get,

$$\frac{1}{t} \ln \left(\frac{1+t}{1-t} \right) = \sum_n t^{2n} \int_{-1}^1 P_n^2 dx.$$

$$\text{Or, } \int_{-1}^1 P_n^2 dx = \frac{2}{2n+1}, \text{ which we get by equating the coefficient of } t^{2n} \text{ from both the sides....} (7.16)$$

$$\text{We can thus write, } \int_{-1}^1 P_n P_m dx = \frac{2}{2n+1} \delta_{n,m} \text{ (combining (7.15) and (7.16))} (7.17)$$

The relation (7.17) constitutes the orthogonality as well as the normalization of the Legendre polynomials.

7.6.2 Completeness

The Legendre polynomials form a complete set of orthogonal polynomials in the interval $(-1, 1)$ so that a bounded function, which satisfies Dinchilet conditions stated as, (i) it is piecewise continuous in the interval, (ii) it has at most a finite number of finite discontinuities, (iii) it has a finite number of maxima and minima, can be expanded in a series of Legendre polynomials so that we can write,

$$f(x) = \sum_l a_l P_l(x).$$

Using the orthonormality condition (7.17) we can write,

$$a_l = \left(\frac{2l+1}{2} \right) \int_{-1}^1 f(x) P_l(x) dx \quad \dots (7.18)$$

Clearly a function $f(x)$ which is a polynomial of n th degree requires Legendre polynomials of order upto $l = n$.

For example, if $f(x) = a + bx + cx^2$,

$$a_o = \frac{1}{2} \int_{-1}^1 (a + bx + cx^2) dx = \left(a + \frac{c}{3} \right),$$

$$\begin{aligned}
a_1 &= \frac{3}{2} \int_{-1}^1 (a + bx + cx^2) x \, dx = \frac{3}{2} \left[\frac{2b}{3} \right] = b. \\
a_2 &= \frac{5}{2} \int_{-1}^1 (a + bx + cx^2) \frac{1}{2} (3x^2 - 1) \, dx \\
&= \frac{5}{4} \int_{-1}^1 [(3a - c)x^2 - bx + 3bx^3 + 3cx^4 - a] \, dx \\
&= \frac{5}{4} \left[\frac{2}{3} (3a - c) + \frac{6c}{5} - 2a \right] = \frac{5}{4} \left[\frac{8}{15} c \right] = \frac{2c}{3}.
\end{aligned}$$

Rodrigue's Formula

We can write, $P_l(x)$ using the series expansion representation (7.11) as

$$P_l(x) = \sum_{n=0}^{[l/2]} (-1)^n \frac{(2l-2n)!}{2^l n! (l-n)! (l-2n)!} x^{l-2n}, \text{ where } [l/2] = l/2 \text{ for } l \text{ even and } = \frac{l-1}{2} \text{ for } l \text{ odd.}$$

$$\text{We can binomially expand } (x^2 - 1)^l \text{ to write, } (x^2 - 1)^l = \sum_{m=0}^l (-1)^m \frac{l!}{m! (l-m)!} x^{2(l-m)}$$

$$\text{or, } \frac{d^l}{dx^l} (x^2 - 1)^l = \sum_{m=0}^{[l/2]} (-1)^m \frac{l!}{m! (l-m)!} \frac{(2l-2m)!}{(l-2m)!} x^{l-2m} = 2^l l! P_l(x)$$

$$\text{Thus } P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l, \text{ which is Rodrigue's formula.} \quad \dots \quad (7.19)$$

7.7 ASSOCIATED LEGENDRE POLYNOMIAL

The Legendre differential equation satisfies, the equation

$$(1-x^2)P_l'' - 2xP_l' + l(l+1)P_l = 0$$

Differentiating the equation m times we get,

$$(1-x^2)z'' - 2(m+1)xz' + [l(l+1) - m(m+1)]z = 0, \text{ where } z = \frac{d^m}{dx^m} P_l(x)$$

Writing $y = (1-x^2)^{m/2} z$ we see that y satisfies the equation

$$(1-x^2)y'' - 2xy' + \left[l(l+1) - \frac{m^2}{1-x^2} \right] y = 0.$$

$$y = (1-x^2)^{m/2} \frac{d^{|m|}}{dx^{|m|}} P_l(x), \text{ where } m \text{ is an integer and } -l \leq m \leq l,$$

is called the associated Legendre polynomial.

The recurrence relation involving P_l^m 's is given by, (it is left as an exercise for the students).

$$(l - m + 1) P_{l+1}^m(x) - (2l + 1) x P_l^m(x) + (l + m) P_{l-1}^m(x) = 0 \dots \quad (7.20)$$

The orthogonality relation of the associated Legendre polynomials is given by

$$\int_{-1}^1 P_n^m(x) P_{n'}^{m'}(x) dx = \frac{2}{(2n+1)} \frac{(n+m)!}{(n-m)!} \delta_{nn'} \delta_{mm'} \dots \quad (7.21)$$

For $m = 0$ we get back the orthogonality relation of the Legendre polynomials $P_n(x)$,

Multiplying $P_l^m(\cos \theta)$ by $e^{im\phi}$, ϕ being the azimuthal angle and normalizing the product we obtain the set of functions called the spherical harmonics $Y_{lm}(\theta, \phi)$ given by

$$Y_{lm}(\theta, \phi) = \left[\frac{(2l+1)}{4\pi} \frac{(l-|m|)!}{(l+|m|)!} \right]^{\frac{1}{2}} P_l^m(\cos \theta) e^{im\phi} \times \begin{cases} (-1)^m & \text{for } m \geq 0, \\ 1 & \text{for } m < 0. \end{cases}$$

The orthonormality relation satisfied by $Y_{lm}(\theta, \phi)$ is given by

$$\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} Y_{lm}^*(\theta, \phi) Y_{l'm'}(\theta, \phi) \sin \theta d\theta d\phi = \delta_{ll'} \delta_{mm'} \dots \quad (7.22)$$

The relation (7.22) is easily seen to follow from (7.21).

The orthogonality relations (7.21) or (7.22) will not be deduced here (though it is straightforward) and can be seen in standard references.

7.8 APPLICATIONS OF THE LEGENDRE POLYNOMIALS TO PHYSICAL PROBLEMS

7.8.1 Multipole Expansion in Electrostatics

The electrostatic potential $\phi(\vec{r})$ due to a charge distribution at P is, $\phi(\vec{r}') = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}'$

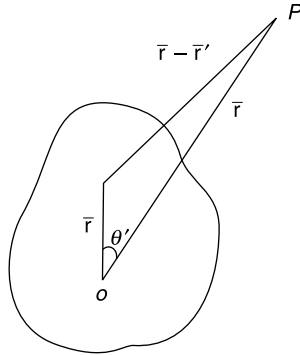


Fig. 7.3

Putting $|\bar{r} - \bar{r}'| = (r^2 + r'^2 - 2rr' \cos \theta)^{1/2}$ we obtain, (Fig. 7.3)

$$\begin{aligned}\phi(\bar{r}) &= \frac{1}{4\pi\epsilon_o} \int \frac{\rho(\bar{r}')}{(r^2 + r'^2 - 2rr' \cos \theta)^{1/2}} d\bar{r}' \\ &= \frac{1}{4\pi\epsilon_o} \left[\frac{1}{r} \int \rho(\bar{r}') d\bar{r}' + \frac{1}{\bar{r}^2} \int \rho(\bar{r}') r' \cos \theta dv' + \dots \right] \text{ for } r > r'\end{aligned}$$

Thus for $r > r'$ we can write, $\phi(\bar{r}) = \frac{1}{4\pi\epsilon_o} \sum_l \phi_l$, where

$\phi_l = \frac{1}{r^{l+1}} \int r'^l \rho(\bar{r}') P_l(\cos \theta) d\bar{r}'$, the integral representing the multipole moments of various orders.

If the origin be so chosen that $r < r'$, we obtain, $\phi(\bar{r}) = \frac{1}{4\pi\epsilon_o} \sum_l \phi_l$, where $\phi_l = \frac{1}{r^{l+1}} \int r'^l \rho(\bar{r}') P_l(\cos \theta') d\bar{r}'$

These two results are sometimes written by a single expression,

$$\phi_l = \frac{1}{r_{>}^{l+1}} \int r_{<}^l \rho(\bar{r}') P_l(\cos \theta') d\bar{r}' \quad (r_{<}, r_{>} \text{ means the smaller or the larger of } r \text{ and } r' \text{ respectively}) \dots\dots (7.23)$$

The spherical harmonics $Y_{lm}(\theta, \phi)$ occur in the multipole expansions of the radiation fields in electrodynamics. The calculations are involved and will not be given here. Another useful application arises in the quantum mechanical problem of a particle in a central potential and will be dealt with briefly in the chapter on partial differential equation.

7.9 HERMITE DIFFERENTIAL EQUATION

The Hermite differential equation is, $y'' - 2x y' + 2ny = 0$ ($n = \text{constant}$)

The equation has no singular points for the finite values of x . The equation has, however, an irregular singular point at infinity as can be proved by putting $x' = \frac{1}{x}$ and transforming the differential equation to an equation in x' .

The solution is obtained by writing $y = \sum_m a_m x^m$, in the equation. We obtain,

$$\sum_m m(m-1) a_m x^{m-2} - 2 \sum_m m a_m x^m + 2n \sum_m a_m x^m = 0$$

We get, by equating the coefficient of x^m , $a_{m+2} = \frac{2(m-n)}{(m+2)(m+1)} a_m$.

The solution is, $y = a_0 \left[1 - \frac{2n}{2!} x^2 + \frac{2^n n(n-2)}{4!} x^4 \dots \right] + a_1 \left[x - \frac{2(n-1)}{3!} x^3 + \frac{2^2 (n-1)(n-3)}{5!} x^5 \dots \right]$

a_0 and a_1 being two arbitrary constants. The regions of convergence of the two series are found out by

calculating $\lim_{m \rightarrow \infty} \left| \frac{t_{m+1}}{t_m} \right|$.

It is easily seen that the series converge for all finite x . This is consistent with the fact that there is no singular point for any finite value of x . If n is an integer, one of the two series reduces to a polynomial and is called the Hermite polynomial $H_n(x)$ of order n . The constants are chosen to be, $a_o = (-1)^{m/2} \frac{m!}{\left(\frac{m}{2}\right)!}$ for even m and $= (-1)^{(m-1)/2} \frac{(2m)!}{\left(\frac{m-1}{2}\right)!}$ for odd m .

$$\begin{aligned} \text{This gives, } H_m(x) &= (2x)^m - \frac{m(m-1)}{1!} (2x)^{m-2} + \frac{m(m-1)(m-2)(m-3)}{2!} (2x)^{m-4} - \dots \\ &= \sum_{r=0}^{[m/2]} (-1)^k \frac{m!}{(m-2r)! r!} (2x)^{m-2r} \dots \end{aligned} \quad (7.24)$$

Generating Function

$$\text{The generating function for Hermite polynomials is given by, } e^{2tx-t^2} = \sum_m \frac{t^m}{m!} H_m(x), \dots \quad (7.25)$$

This is verified by expanding the left hand side and equating the coefficients of different powers of t .

7.9.1 Recurrence Relations

Differentiating Eq. (7.25) with respect to t we get,

$$2(x-t) \sum_m \frac{t^m}{m!} H_m(x) = \sum_m \frac{t^{m-1}}{(m-1)!} H_m(x)$$

Equating the coefficient of t^m we obtain,

$$\frac{2x H_m}{m!} - \frac{2 H_{m-1}}{(m-1)!} = \frac{H_{m+1}}{m!}$$

$$\text{Or, } H_{m+1} - 2x H_m + 2m H_{m-1} = 0 \dots \quad (7.26)$$

Similarly differentiating Eq. (7.25) with respect to x we get,

$$2t \sum_m \frac{t^m}{m!} H_m(x) = \sum_m \frac{t^m}{m!} H'_m(x)$$

$$\text{Or, } H'_m(x) = 2m H_{m-1}(x) \dots \quad (7.27)$$

7.9.2 Orthogonality Relations and Completeness

$$\text{We have, } e^{2tx-t^2} = \sum_m \frac{t^m}{m!} H_m(x) \text{ and } e^{2sx-s^2} = \sum_n \frac{s^n}{n!} H_n(x)$$

$$\text{or,} \quad \int_{-\infty}^{\infty} e^{[2(t+s)x - (t^2 + s^2)]} e^{-x^2} dx = \sum_{m,n} \frac{t^m s^n}{m! n!} \int_{-\infty}^{\infty} e^{-x^2} H_m H_n dx$$

$$\text{or,} \quad \int_{-\infty}^{\infty} e^{-[x - (t+s)]^2} e^{2st} dx = \sum_{m,n} \frac{t^m s^n}{m! n!} \int_{-\infty}^{\infty} e^{-x^2} H_m H_n dx$$

$$\text{or,} \quad e^{2st} \sqrt{\pi} = \sum_{m,n} \frac{t^m s^n}{m! n!} \int_{-\infty}^{\infty} e^{-x^2} H_m H_n dx$$

$$\left[\text{Using } \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}, \text{ a relation which will be proved in the discussion on gamma function} \right]$$

Writing $e^{2st} = \sum \frac{(2st)^p}{p!}$ and comparing the two sides we obtain,

$$\int_{-\infty}^{\infty} e^{-x^2} H_n H_m dx = 2^n n! \sqrt{\pi} \delta_{m,n} \dots \quad (7.28)$$

The orthogonality relation (7.28) enables us to write, $f(x) = \sum_{n=0}^{\infty} a_n H_n(x)$, where $f(x)$ satisfies the Dirichlet conditions stated in sec. 7.6 (c), in the interval $(-\infty, \infty)$.

Or, $\int_{-\infty}^{\infty} e^{-x^2} H_m(x) f(x) dx = 2^m m! \sqrt{\pi} a_m$ yielding.

$$a_m = \frac{1}{2^m m! \sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2} H_m(x) f(x) dx \dots \quad (7.29)$$

In expanding $f(x)$ in terms of Hermite polynomials we note that for $f(x)$ a polynomial of order p we will have non-vanishing a_m 's for $m = 1, 2, \dots p$, because of the orthogonality relation.

[A few of the Hermite polynomials are $H_0(x) = 1, H_1(x) = 2x, H_2(x) = 4x^2 - 2 \dots$]

To speak of one of the applications to physical problems we can mention that the well-behaved bound state solution of a one-dimensional harmonic oscillator problem is expressed in terms of Hermite polynomials while that of a three-dimensional problem in Cartesian coordinate is given in terms of the product of three Hermite polynomials, each one corresponding to a particular dimension.

7.10 BESSEL DIFFERENTIAL EQUATION

This is an equation frequently occurring in diverse physical problems. The equation is,

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - m^2) y = 0. \dots \quad (7.30)$$

In this case the point $x = 0$ is the only singular point, regular in nature. The point at infinity is seen to be an irregular singular point as can be seen by putting $x' = \frac{1}{x}$ and exploring the singularity at $x' = 0$

Writing the solution in the form $y = \sum_n a_n x^{n+\rho}$ and substituting in the equation we obtain,

$$x^2 \sum_n (n+\rho)(n+\rho-1) a_n x^{n+\rho-2} + x \sum_n (n+\rho) a_n x^{n+\rho-1} + (x^2 - m^2) \sum_n a_n x^{n+\rho} = 0$$

or,
$$\sum_n [(n+\rho)^2 - m^2] a_n x^{n+\rho} + \sum_n a_n x^{n+\rho+2} = 0$$

Equating the coefficient of the lowest power of x to zero we get, $\rho^2 - m^2 = 0$ assuming that $a_0 \neq 0$.

Thus the two roots of the indicial equation are $\rho = \pm m$. Now if the difference of the roots be non-integral, a case arising for non-integral value of m except for $m = \pm \frac{1}{2}, \pm \frac{3}{2}, \dots$, which will be treated separately, the general solution is given by two independent series combined together. For integer m one solution will be obtained by the series expansion, and the other by the method described in sec 7.2. Equating the coefficient of the next higher power of x to zero we get, $a_1 [(\rho+1)^2 - m^2] = 0$.

Since $\rho = \pm m$, $a_1 = 0$ except when $m = \pm \frac{1}{2}$, the case that will be considered separately. Equating the coefficient of $x^{n+\rho}$ to zero we obtain,

$$a_{n+2} [(n+\rho+2)^2 - m^2] + a_n = 0 \text{ giving } a_{n+2} = - \frac{a_n}{(n+\rho+2)^2 - m^2}.$$

Thus for $\rho = m$, $a_{n+2} = - \frac{a_n}{(n+m+2)^2 - m^2}$, which leads to

$$a_2 = - \frac{a_0}{2^2(m+1)}, a_4 = \frac{a_0}{2^5(m+1)(m+2)} \text{ and so on.}$$

Hence $a_{2n} = (-1)^n \frac{a_0}{2^{2n}(m+1)(m+2)\dots(m+n)n!}$.

For $\rho = -m$, $a_{n+2} = - \frac{a_n}{(n-m+2)^2 - m^2}$ which gives,

$$a_{2n} = (-1)^n \cdot \frac{a_0}{2^{2n}(1-m)\dots(n-m)n!}$$

Defining a_0 as $2^m m!$ We define the Bessel function $J_m(x)$ as,

$$J_m(x) = \sum_{n=0}^{\infty} (-1)^n \cdot \frac{1}{n!(m+n)!} \left(\frac{x}{2}\right)^{m+2n} \text{ for integer } m \dots \quad (7.31)(a)$$

and for general m we replace $(m+n)!$ by $\Gamma(m+n+1)$, [$\Gamma(n)$ is called the gamma function of order n and $= (n-1)!$ for integer n . This will be treated in detail in a subsequent section of this chapter].

Clearly for negative m , $J_{-m}(x) = \sum_n (-1)^n \cdot \frac{1}{n! \Gamma(n-m+1)!} \left(\frac{x}{2}\right)^{2n-m} \dots \quad (7.31)(b)$

Replacing n by $n' = n - m$ we get,

$$J_{-m}(x) = \sum_{n'=-m}^{\infty} \frac{(-1)^{n'+m}}{(n'+m)! \Gamma(n'+1)} \left(\frac{x}{2}\right)^{2n'+m}, \text{ where } m \text{ has been assumed to be an integer,}$$

Since $\Gamma(n'+1) \rightarrow \infty$ for n' a negative integer, we obtain,

$$J_{-m}(x) = (-1)^m \sum_{n'=0}^{\infty} \frac{(-1)^{n'}}{n'! \Gamma(n'+m+1)} \left(\frac{x}{2}\right)^{2n'+m} = (-1)^m J_m(x), \dots$$

proving according to our expectations that for integer m the two solutions of the differential equation are not linearly independent. For $m = \frac{1}{2}$ we define $Y = \sqrt{x} y$ so that we get,

$$\frac{dy}{dx} = \frac{d}{dx} \left(\frac{Y}{\sqrt{x}} \right) = \frac{1}{\sqrt{x}} \frac{dY}{dx} - \frac{1}{2} x^{-3/2} Y,$$

$$\frac{d^2 y}{dx^2} = \frac{1}{\sqrt{x}} \frac{d^2 Y}{dx^2} - x^{-3/2} \frac{dY}{dx} + \frac{3}{4} x^{-5/2} Y \text{ so that we obtain, by putting in Eq. (7.30), and simplifying,}$$

$$\frac{d^2 Y}{dx^2} + Y = 0 \text{ yielding the solution } Y = A \sin x + B \cos x \text{ which leads to the solution}$$

$$y = (A \sin x + B \cos x) / \sqrt{x}, \text{ } A \text{ and } B \text{ being arbitrary constants. For } m = -\frac{1}{2} \text{ we obtain an identical solution.}$$

For $m = \pm \frac{3}{2}, \pm \frac{5}{2} \dots$ we obtain $J_m(x)$ by similar transformation of variables or by using the recurrence relations discussed later.

Second Solution for Integer m

$$\text{Writing the Bessel equation as } \frac{d^2 y}{dx^2} + \frac{1}{x} \frac{dy}{dx} + \left(1 - \frac{m^2}{x^2}\right) y = 0 \text{ we get, } p(x) = \frac{1}{x}$$

Thus, using the formula (Eq. 7.4) we get,

$$y_2 = c y_1 \int \frac{dx}{x y_1^2}, \text{ where } c \text{ is a constant equal to a } (W)_a, \text{ where the Wronskian}$$

$$W = (W)_a \exp \left[- \int_a^x \frac{dx}{x} \right] = a(W)_a \left(\frac{1}{x} \right). \quad [\text{a being any constant}]$$

Substituting the power series for y_1 given by 7.31(a) we can use the binomial expansion to write

$$y_2 = y_1 \left[\ln x + \sum_{n=1}^{\infty} c_n x^{2n} \right], \text{ the coefficients } c_n \text{ being determined from the expansion. Obviously it would be}$$

appropriate if we can write the second solution in a form that would be valid for $m = \text{integer or non-integer.}$

$$\text{We define the Neumann function as } N_m(x) = \frac{\cos m \pi J_m(x) - J_{-m}(x)}{(\sin m \pi)},$$

For non-integral m ($m = \frac{1}{2}, \frac{3}{2}, \dots$ are excluded), $N_m(x)$ is independent of $J_m(x)$. For integer m , $N_m(x)$ is an

indeterminate form $0/0$. Applying L' Hospital's rule we get, $N_m(x) = \lim_{m \rightarrow \text{integer}} \frac{\frac{\partial}{\partial m} [\cos m \pi J_m(x) - J_{-m}(x)]}{\frac{\partial}{\partial m} (\sin m \pi)}$

or,

$$N_m(x) = \lim_{m \rightarrow \text{integer}} \left[\frac{(-1)^m \frac{\partial J_m}{\partial m} - \frac{\partial J_{-m}}{\partial m}}{\pi (-1)^m} \right]$$

From the series expansions (7.31)(a) and (7.31)(b) we can see that for integer m , $N_m(x)$ is clearly independent of $J_m(x)$, because $\frac{\partial J_m(x)}{\partial m}$ and $\frac{\partial J_{-m}(x)}{\partial m}$ are independent of $J_m(x)$.

7.10.1 Generating Function and Recurrence Relation

The generating function of Bessel functions is given by $\exp \left[\frac{x}{2} \left(t - \frac{1}{t} \right) \right] = \sum_{n=-\infty}^{\infty} t^n J_n(x)$ as can be verified by expanding the exponential, equating the coefficients of t^n and comparing with the series expression of Bessel function.

Differentiating Eq. (7.32) with respect to t we get,

$$\frac{x}{2} \left(1 + \frac{1}{t^2} \right) \sum_n t^n J_n(x) = \sum_n n t^{n-1} J_n(x). \quad (7.32)$$

Equating the coefficient of t^n from the two sides we get,

$$J_{n+1}(x) + J_{n-1}(x) = \frac{2n}{x} J_n(x) \dots \quad (7.33)$$

Differentiating Eq. (7.32) with respect to x we get,

$$\frac{1}{2} \left(t - \frac{1}{t} \right) \sum_n t^n J_n(x) = \sum_n t^n J_n'(x).$$

Equating the coefficient of t^n from the two sides we get,

$$J_{n-1}(x) - J_{n+1}(x) = 2J_n'(x) \dots \quad (7.34)$$

Combining (7.33) and (7.34) we can write, $J_{n+1}(x) = \pm J_n'(x) + \frac{n}{x} J_n(x) \dots$ (7.35)

The recurrence relations (7.33) and (7.34) can also be deduced by using the series expression of Bessel's function.

$$\begin{aligned}
 \text{Thus, } J_{n+1}(x) + J_{n-1}(x) &= \sum_{r=0}^{\infty} \frac{(-1)^r}{r! \Gamma(n+r+2)} \left(\frac{x}{2}\right)^{n+2r+1} + \sum_{r=0}^{\infty} \frac{(-1)^r}{r! \Gamma(n+r)} \left(\frac{x}{2}\right)^{n+2r-1} \\
 \text{or, } J_{n+1}(x) + J_{n-1}(x) &= \left(\frac{x}{2}\right)^{n-1} \left[\frac{1}{\Gamma(n)} + \sum_{r=1}^{\infty} \left\{ \frac{(-1)^r}{r! \Gamma(n+r)} \left(\frac{x}{2}\right)^{2r} - \frac{(-1)^r \left(\frac{x}{2}\right)^{2r}}{(r-1)! \Gamma(n+r+1)} \right\} \right] \\
 &= \left(\frac{x}{2}\right)^{n-1} \left[\frac{1}{\Gamma(n)} + \sum_{r=1}^{\infty} (-1)^r \left(\frac{x}{2}\right)^{2r} \cdot \frac{n}{r! \Gamma(n+r+1)} \right] \\
 &= \frac{2n}{x} \sum_{r=1}^{\infty} \frac{(-1)^r}{r! \Gamma(n+r+1)} \left(\frac{x}{2}\right)^{n+2r} = \frac{2n}{x} J_n(x).
 \end{aligned}$$

(The derivation of the other one is left as an exercise). Putting $t = e^{i\theta}$ in the relation (7.32) we obtain,

$$\exp(ix \sin \theta) = \sum_n e^{in\theta} J_n(x).$$

$$\text{Thus, } J_n(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp[i(x \sin \theta - n\theta)] d\theta \quad (\text{for integer } n).$$

We can further write it as,

$$J_n(x) = \frac{1}{2\pi} \int_0^{\pi} [\exp\{i(x \sin \theta - n\theta)\} + (-1)^n \exp\{-i(x \sin \theta + n\theta)\}] d\theta$$

$$\text{For even } n, J_n(x) = \frac{1}{\pi} \int_0^{\pi} \cos(x \sin \theta) \cos n\theta d\theta \quad \dots \quad (7.36)$$

$$\text{and for odd } n, J_n(x) = \frac{1}{\pi} \int_0^{\pi} \sin(x \sin \theta) \sin n\theta d\theta$$

Thus (7.35) gives an integral representation of the Bessel functions for integer n . For general n finding out the integral representation is quite involved and omitted from our discussions.

7.10.2 Orthogonality of Bessel Functions and their Completeness

Let us take two Bessel functions $J_n(px)$ and $J_n(qx)$. They satisfy the equations, $\frac{d^2 f_1}{dx^2} + \frac{1}{x} \frac{df_1}{dx} + \left(p^2 - \frac{n^2}{x^2}\right) f_1 = 0$

and ...

(7.37(a), 7.37(b))

$$\frac{d^2 f_2}{dx^2} + \frac{1}{x} \frac{df_2}{dx} + \left(q^2 - \frac{n^2}{x^2}\right) f_2 = 0, \text{ where } f_1 \text{ and } f_2 \text{ are } J_n(px) \text{ and } J_n(qx) \text{ respectively.}$$

Multiplying 7.36 (a) by f_2 , 7.36 (b) by f_1 respectively and subtracting we get,

$(f_2 f_1'' - f_1 f_2'') + \frac{1}{x} (f_2 f_1' - f_1 f_2') + (p^2 - q^2) f_1 f_2 = 0$, where the prime denotes differentiation with respect to x .

Or,
$$\frac{d}{dx} [x(f_2 f_1' - f_1 f_2')] + (p^2 - q^2) f_1 f_2 x = 0.$$

Or,
$$(p^2 - q^2) \int_a^b x f_1 f_2 dx = [x(f_1 f_2' - f_2 f_1')]_a^b.$$

We choose the end points a and b of the integration such that the right hand side vanishes, [This can be done in different ways, for example, by taking a and b so that (i) $J_n(pa) = J_n(pb) = J_n(qa) = J_n(qb) = 0$ or, (ii) $J_n(pa) = J_n(qa) = 0, J_n'(pb) = J_n'(qb) = 0$ etc.]

Hence we obtain,
$$\int_a^b x J_n(px) J_n(qx) dx = 0 \text{ for } p \neq q. \dots \quad (7.38)$$

For $p = q$,
$$\int_a^b x J_n(px) J_n(qx) dx = \int_a^b x J_n^2(px) dx$$

Let $I = \int_a^b x J_n^2(px) dx = \frac{1}{p^2} \int_{pa}^{pb} z J_n^2(z) dz$, where $z = px$.

Thus
$$I = \frac{1}{p^2} \left[\frac{z^2}{2} J_n^2(z) - \int \{ n^2 J_n(z) - z J_n'(z) - z^2 J_n''(z) \} J_n'(z) dz \right]_{pa}^{pb}$$

[Integrating by parts and using Bessel's equation]

or,
$$I = \frac{1}{p^2} \left[\frac{z^2}{2} J_n^2(z) - \frac{1}{2} n^2 J_n^2(z) + \frac{z^2}{2} (J_n'(z))^2 \right]_{pa}^{pb}$$

If $J_n(z) = 0$ at $z = pb$ and pa respectively, we obtain,

$$I = \left[\frac{x^2}{2} [J_n'(px)]^2 \right]_{x=a}^{x=b}$$

Using the recurrence relation (7.35) we obtain,

$$I = \left[\frac{x^2}{2} [J_{n+1}^2(px)] \right]_{x=a}^{x=b} \text{ if } J_n(pa) = J_n(pb) = 0 \text{ by a suitable choice of } a \text{ and } b. \dots \quad (7.39)$$

Thus we can write by combining Eqs. (7.37) and (7.38), the orthogonality relation

$$\int_a^b x J_n(px) J_n(qx) dx = \left[\frac{x^2}{2} J_{n+1}^2(px) \right]_a^b \delta_{pq}, \text{ where } J_n(pa) = J_n(pb) = 0. \dots \quad (7.40)$$

Completeness

Since the Bessel functions form a set of orthogonal and hence linearly independent functions in the interval (a, b) suitably chosen, we expect them to form a complete set so that we can expand a function $f(x)$, that satisfies Dirichlet conditions in the interval in terms of them. Thus we can write, $f(x) = \sum_{p=1}^{\infty} a_p J_n(k_p x)$, where a_p stands for the expansion coefficient.

Since for a suitable choice of a, b the orthogonality relation (7.40) is valid we can write,

$$a_p = \frac{\int_a^b f(x) J_n(k_p x) x dx}{\left[\frac{x^2}{2} J_{n\pm 1}^2(k_p x) \right]_a^b}, \text{ where } J_n(k_p a) = J_n(k_p b) = 0$$

Since $J_n(0) = 0$ for all n we can write also,

$$a_p = \frac{\int_0^a f(x) J_n(k_p x) x dx}{\left[\frac{a^2}{2} J_{n\pm 1}^2(k_p a) \right]} \dots \quad (7.41)$$

The Neumann functions denoted by $N_m(x)$ satisfy the same recurrence relations as the Bessel functions $J_m(x)$ since they are formed out of a linear combination of the Bessel functions $J_m(x)$ and $J_{-m}(x)$. Another function suitable for use is the Hankel functions $H_c^{(1,2)}(x) = J_c(x) \pm i N_c(x)$. The other related functions that arise in the solution of some problems in spherical polar coordinates are spherical Bessel functions, spherical Neumann and spherical Hankel functions given respectively by,

$$j_l(kr) = \sqrt{\frac{2}{\pi kr}} J_{l+\frac{1}{2}}(kr), n_l(kr) = \sqrt{\frac{2}{\pi kr}} N_{l+\frac{1}{2}}(kr), h_l(kr) = \sqrt{\frac{2}{\pi kr}} H_{l+\frac{1}{2}}(kr).$$

There are numerous applications of the Bessel functions in physics. But a familiarity with these requires a knowledge of partial differential equations. We will deal with these in a subsequent chapter on partial differential equations.

7.11 EQUATION HAVING AN IRREGULAR SINGULAR POINT

We will give the following examples illustrating the efficacy of the series method, if at all, in the solution of a differential equation near an irregular singular point.

Example 1 $\frac{d^2 y}{dx^2} + \frac{1}{x^2} \frac{dy}{dx} + \frac{2y}{x^3} = 0$

Let us try a series solution $y = \sum_n a_n x^{n+\rho}$ where $x = 0$ is an irregular singular point. Substituting it in the equation we get,

$$\sum_n (n + \rho) (n + \rho - 1) a_n x^{n+\rho-2} + \sum_n (n + \rho) a_n x^{n+\rho-3} + 2 \sum_n a_n x^{n+\rho-3} = 0$$

or,

$$\sum_n (n + \rho) (n + \rho - 1) a_n x^{n+\rho-2} + \sum_n (n + \rho + 2) a_n x^{n+\rho-3} = 0$$

Equating the coefficient of $x^{\rho-3}$ to zero we get,
 $(\rho + 2) a_0 = 0$ giving $\rho = -2$ since $a_0 \neq 0$ by choice.

Equating the coefficient of $x^{n+\rho-2}$ to zero we get,

$$a_{n+1} (n + \rho + 3) = -a_n (n + \rho) (n + \rho - 1)$$

Putting $\rho = -2$ we get,

$$a_{n+1} = -a_n \frac{(n-2)(n-3)}{(n+1)}$$

Clearly all the a_m 's starting from a_3 would vanish. Thus we can write one solution as, $y = a_0 [1 - 6x + 6x^2] x^{-2}$. The other solution can be determined by the formula (7.4) and is seen to have a term of the form $e^{1/x}$ that implies the presence of an essential singularity at $x = 0$.

Example 2 $\frac{d^2 y}{dx^2} + \frac{y}{x^3} = 0$

Attempting a solution $y = \sum_n a_n x^{n+\rho}$ we get,

$$\sum_n (n + \rho) (n + \rho - 1) a_n x^{n+\rho-2} + \sum_n a_n x^{n+\rho-3} = 0$$

Equating the coefficient of $x^{\rho-3}$ we get, $a_0 = 0$.

From the recurrence relation obtained by equating the coefficient of $x^{n+\rho-2}$ we get, $a_1 = a_2 = \dots = a_n = 0$. Thus the only solution obtained by the series expansion method is $y = 0$ which is a trivial solution. Here the Frobenius method fails to give any non-trivial result.

7.12 BETA FUNCTION, GAMMA FUNCTION, DIRAC DELTA FUNCTION

This chapter dwells on some special functions other than those discussed previously. These special functions are of immense use in physical problems. The gamma function $\Gamma(z)$ is defined in several ways.

- (1) $\Gamma(z)$ is defined by the integral $\int_0^\infty e^{-t} t^{z-1} dt$, $\text{Re } z > 0$, for the integral to converge. (The proof of the convergence of the integral is given in the appendix of Chapter 5).

$$\text{Now, } \Gamma(z+1) = \int_0^\infty e^{-t} t^z dt = -t^z e^{-t} \Big|_0^\infty + \int_0^\infty z t^{z-1} e^{-t} dt = z \Gamma(z) \dots \quad (7.42)$$

Thus $\Gamma(z) = \Gamma(z+1)/z$ is defined for $\text{Re } z > -1$. By continuing the process we can define $\Gamma(z)$ in the entire z -plane by following this recurrence relation of $\Gamma(z)$. This is an example of analytic continuation discussed in the Chapter 6.

Now, $\Gamma(1) = \int_0^{\infty} e^{-t} dt = 1$ and by following the definition (Eq. 7.41) we can see that $\Gamma(z)$ has simple poles at $z = 0, -1, -2, \dots$. Thus it can be said that $\Gamma(z)$ is analytic in the entire z -plane except having simple poles at $z = -1, -2, \dots$.

For positive integral values of z , $\Gamma(z+1) = z \cdot (z-1) \dots 1$. $\Gamma(1) = z!$

(2) Euler gave the definition of $\Gamma(z)$ as,

$$\Gamma(z) = \lim_{n \rightarrow \infty} \frac{n!}{z(z+1) \dots (z+n)} n^z. \dots \quad (7.43)$$

(3) Weirstrass definition of $\Gamma(z)$ is.

$$\frac{1}{\Gamma(z)} = ze^{\gamma z} \prod_{p=1}^{\infty} \left(1 + \frac{z}{p}\right) e^{-z/p}, \text{ where the symbol } \prod_{p=1}^{\infty} \text{ stands for the infinite product and}$$

$$\gamma = \lim_{m \rightarrow \infty} \left(1 + \frac{1}{2} + \dots + \frac{1}{m} - \ln m\right) \approx 0.572. \text{ (Euler-Mascheroni constant)}$$

The three definitions of $\Gamma(z)$ can be shown to be equivalent to one another. Equivalence of (2) and (3).

$$\text{From Eq. (7.43), } \frac{1}{\Gamma(z)} = \lim_{n \rightarrow \infty} \frac{z(z+1) \dots (z+n)}{n!} n^{-z} = z \cdot \lim_{n \rightarrow \infty} \left(1 + \frac{z}{1}\right) \dots \left(1 + \frac{z}{n}\right) n^{-z}$$

$$\text{or, } \frac{1}{\Gamma(z)} = \left[\left\{ \prod_{r=1}^n \left(1 + \frac{z}{r}\right) e^{-z/r} \right\} z \exp \left[z \left(1 + \frac{1}{2} + \dots + \frac{1}{n} - \ln n\right) \right] \right]_{n \rightarrow \infty}$$

$$\text{or, } \frac{1}{\Gamma(z)} = ze^{\gamma z} \prod_{r=1}^{\infty} \left(1 + \frac{z}{r}\right) e^{-z/r}, \text{ where } \gamma = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{2} + \dots + \frac{1}{n} - \ln n\right).$$

Equivalence of (1) and (2)

Using the limit $\lim_{k \rightarrow \infty} \left(1 - \frac{t}{k}\right)^k = e^{-t}$ we obtain,

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt = \lim_{k \rightarrow \infty} \int_0^{\infty} \left(1 - \frac{t}{k}\right)^k t^{z-1} dt.$$

By repeated integration by parts of $\int \left(1 - \frac{t}{k}\right)^k \left(\frac{t}{k}\right)^{z-1} d\left(\frac{t}{k}\right)$ and taking $\lim_{k \rightarrow \infty}$ of the integral we can

deduce (this is left as an exercise for the students) that $\Gamma(z) = \lim_{n \rightarrow \infty} \frac{n!}{z(z+1) \dots (z+n)} n^z$.

$$(4) \quad \Gamma(x) \Gamma(y) = \int_0^{\infty} e^{-t} t^{x-1} dt \int_0^{\infty} e^{-r} r^{y-1} dr$$

Let $t + r = s$, so that we get,

$$\Gamma(x) \Gamma(y) = \int_0^{\infty} ds \int_0^s dt e^{-s} t^{x-1} (s-t)^{y-1}.$$

$$\text{Or, } \Gamma(x) \Gamma(y) = \int_0^1 dp \cdot p^{x-1} (1-p)^{y-1} \int_0^{\infty} e^{-s} s^{x+y-1} ds, \text{ where } t = ps.$$

Now, let us call $B(x, y) = \int_0^1 p^{x-1} (1-p)^{y-1} dp$ so that we can write,

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \dots \quad (7.44)$$

$B(x, y)$ called the Beta-function is clearly defined by this definition for all x, y in the complex plane, the integral converging however for $\text{Re } x, \text{Re } y > 0$.

$$\text{We frequently use a relation } \Gamma(z) \Gamma(1-z) = \frac{\pi}{\sin \pi z} \dots \quad (7.45)$$

which we will not prove here because this involves handling of complex integrations which is beyond our scope here.

$$\text{Using this we can prove, } \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

$$\text{Alternatively, } \Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} e^{-t} \cdot t^{-\frac{1}{2}} dt = 2 \int_0^{\infty} e^{-x^2} dx = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}, \text{ a relation frequently used.}$$

7.12.1 Dirac. Delta Function

This special function introduced by Dirac in his studies on quantum mechanics and symbolized by $\delta(x)$ finds wide applications in varied branches of physics. This function was assumed to have the properties

$\delta(x) = 0$ for $x \neq 0$ and $\int_a^b \delta(x) dx = 1$, where the interval (a, b) includes the point $x = 0$. Clearly this is not at all

a well-behaved function near $x = 0$ where it is highly discontinuous. (actually it has an infinite discontinuity at this point). The mathematicians call such a function a generalized function and attribute a meaning to it by considering it to be the limit of an infinite sequence of functions such that the above mentioned properties

are satisfied. The sequence of functions $F_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2}$, $n = 1, 2, \dots$ furnishes such an example in the

sense that for any well-behaved function $f(x)$, we obtain, $\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x) f(x) dx = f(0)$. By analogy we endow

the generalized function $\delta(x)$ with a meaning in the sense of an integral such that $\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0)$ or

more generally, $\int_{-\infty}^{\infty} \delta(x-a) f(x) dx = f(a)$. [The limits of the integration can be any two numbers a and b surrounding the point $x = 0$]. In three dimensions, similarly we define $\delta(\bar{r})$ by the following.

$$\int_{-\infty}^{\infty} f(\bar{r}) \delta(\bar{r} - \bar{a}) d\bar{r} = f(\bar{a}).$$

Some other properties

$$(1) \int_{-\infty}^{\infty} f(x) \frac{d\delta(x)}{dx} dx = [f(x)\delta(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{df}{dx} \delta(x) dx = -f'(0)$$

$$(2) \int_{-\infty}^{\infty} f(x) \delta[F(x)] dx, \text{ where } F(x) = (x - c_1)(x - c_2) \dots (x - c_n) \text{ can be determined as in the following.}$$

Clearly the contributions will come from the neighborhoods of the points $x = c_1, c_2, \dots, c_n$,

$$\text{Or, } \int_{-\infty}^{\infty} f(x) \delta[F(x)] dx = \int_{c_1 - \epsilon_1}^{c_1 + \epsilon_1} f(x) \delta[F(x)] dx + \int_{c_2 - \epsilon_2}^{c_2 + \epsilon_2} f(x) \delta[F(x)] dx + \dots + \int_{c_n - \epsilon_n}^{c_n + \epsilon_n} f(x) \delta[F(x)] dx$$

$$\text{The first integral } \int_{c_1 - \epsilon_1}^{c_1 + \epsilon_1} f(x) \delta[f(x)] dx = \int_{c_1 - \epsilon_1}^{c_1 + \epsilon_1} f(x) \delta[(x - c_1)(x - c_2) \dots (x - c_n)] dx$$

$$\text{This behaves as } \int_{c_1 - \epsilon_1}^{c_1 + \epsilon_1} f(x) \delta[(x - c_1)(c_1 - c_2) \dots (c_1 - c_n)] dx$$

$$= \int_{c_1 - \epsilon_1}^{c_1 + \epsilon_1} dx f(x) \delta[\alpha_1(x - c_1)], \text{ where } \alpha_1 = (c_1 - c_2) \dots (c_1 - c_n)$$

$$\text{Now } \int_{c_1 - \epsilon_1}^{c_1 + \epsilon_1} f(x) \delta[\alpha_1(x - c_1)] dx = \frac{1}{\alpha_1} \int_{-\epsilon_1}^{\epsilon_1} f(z/\alpha_1 + c_1) \delta(z) dz, \text{ by putting } z = (x - c_1)\alpha_1$$

$$\text{and letting } \epsilon_1 \rightarrow 0 = \frac{1}{\alpha_1} f(c_1) = \frac{1}{\alpha_1} \int_{-\infty}^{\infty} f(x) \delta(x - c_1) dx.$$

$$\text{Now, } \alpha_1 = \left(\frac{dF}{dx} \right)_{x=c_1} \text{ similarly we can carry out the other integrals}$$

to write finally,
$$\int_{-\infty}^{\infty} f(x) \delta[f(x)] dx = \sum_j \frac{1}{\left| \frac{dF}{dx} \right|_{x=c_j}} \int_{-\infty}^{\infty} f(x) \delta(x - c_j) dx$$

We replace $\frac{dF}{dx}$ in the final line by $\left| \frac{dF}{dx} \right|$ because we can easily show that when $\frac{dF}{dx} \Big|_{x=c_j}$ becomes negative for any j we have to replace $\frac{dF}{dx} \Big|_{c_j}$ by $\left| \frac{dF}{dx} \right|_{x=c_j}$. Thus we can write that for any well-behaved function $F(x)$ that can be expressed as $\prod_{j=1}^n (x - c_j)$, $\delta[F(x)]$ behaves as $\sum_j \frac{\delta(x - c_j)}{\left| \frac{dF}{dx} \right|_{c_j}}$.

For a three-dimensional delta function, we can write in

(1) Cartesian coordinates: $\int f(\bar{r}) \delta(\bar{r} - \bar{r}_0) d\bar{r} = f(\bar{r}_0)$.

Or,
$$\int f(\bar{r}) \delta(\bar{r} - \bar{r}_0) d\bar{r} = \int f(x, y, z) \delta(x - x_0) \delta(y - y_0) \delta(z - z_0) dx dy dz$$

Thus $\delta(\bar{r} - \bar{r}_0)$ behaves as $\delta(x - x_0) \delta(y - y_0) \delta(z - z_0)$

(2) Spherical polar coordinates.

$$\begin{aligned} \int f(\bar{r}) \delta(\bar{r} - \bar{r}_0) d\bar{r} &= \int f(r, \theta, \phi) \delta(r - r_0) \delta(\theta - \theta_0) \delta(\phi - \phi_0) dr d\theta d\phi \\ &= \int f(r, \theta, \phi) \delta(\bar{r} - \bar{r}_0) r^2 dr \sin \theta d\theta d\phi. \end{aligned}$$

Thus $\delta(\bar{r} - \bar{r}_0)$ behaves as $\frac{1}{(r^2 \sin \theta)} \delta(r - r_0) \delta(\theta - \theta_0) \delta(\phi - \phi_0)$.

or in terms of $\cos \theta$ as $\frac{1}{r^2} \delta(r - r_0) \delta(\cos \theta - \cos \theta_0) \delta(\phi - \phi_0)$.

Similarly we can have a n -dimensional representation of the delta function by writing the corresponding volume elements in the different coordinate systems.

Application

To mention a few examples we can write

- (i) The charge density of a localized point charge of strength e by $f(\bar{r}) = e \delta(\bar{r})$,
- (ii) Then current density of a uniformly moving charge of strength e moving with a constant velocity $\bar{u}$ by $\bar{J}_{(\bar{r}, t)} = e \bar{u} \delta(\bar{r} - \bar{u} t)$.
- (iii) The wave function of a particle localized at $\bar{r} = \bar{a}$ by $\psi(\bar{r}) = N \cdot \delta(\bar{r} - \bar{a})$, where N is the normalization constant.

Also the integrals involving delta functions appear frequently in calculations in various branches of physics, for example in quantum field theory. It is better to have some workable knowledge of handling these integrals. We will deal in the exercise with some simple integrals.

7.13 SYSTEM OF SIMPLE COUPLED DIFFERENTIAL EQUATIONS

A differential equation of high order with constant coefficients can be often written as a combination of equations with smaller order. Thus a second order differential equation of the form

$\frac{d^2x}{dt^2} + a\frac{dx}{dt} + bx = 0$ can be split up into two equations of the form. $\frac{dx}{dt} = \alpha x + \beta y$, $\frac{dy}{dt} = \gamma x + \delta y$ so that $a = -(\alpha + \delta)$, $b = -(\beta\gamma + \alpha\delta)$. This can be easily verified by differentiating the first one and using the second. Let us now see how to solve the couple of differential equations utilizing the simple matrix method. The two equations can be expressed as

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}, \text{ where } A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

The eigenvalues of the matrix A are λ_1, λ_2 given by

$$\text{dct} \begin{vmatrix} \lambda - \alpha & -\beta \\ -\gamma & \lambda - \delta \end{vmatrix} = 0 \text{ or, } \lambda^2 - \lambda(\alpha + \delta) + (\alpha\delta - \beta\gamma) = 0.$$

The two roots of the equation are λ_1 and λ_2 so that the independent eigenfunctions are given by $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$. The general solution is given by $x = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$. [The eigenfunctions form a complete set so that any solution can be written as a combination of two functions].

From the equation $\frac{dx}{dt} = \alpha x + \beta y$, y can be determined. The two constants c_1 and c_2 can be found out once we are given two initial conditions, x_o and $(\dot{x})_o$ for example. (or x_o and y_o)

An exceptional case occurs if $\lambda_1 = \lambda_2$ for in that case we have only one solution of the form $e^{\lambda_1 t}$. The other independent solution can be found out by substituting in the original differential equation, $x = f(t)e^{\lambda_1 t}$ and comes out to be of the form $x \sim te^{\lambda_1 t}$.

Thus the general solution is, $x = (c_1 + c_2 t) e^{\lambda_1 t}$ and

$$y = \left(\frac{dx}{dt} - \alpha x \right) \cdot \frac{1}{\beta} = [(c_1 \lambda_1 + c_2) - \alpha c_1] e^{\lambda_1 t} + (c_2 \lambda_1 - \alpha c_2) t e^{\lambda_1 t} \cdot \frac{1}{\beta}$$

$$= [\{c_1 (\lambda_1 - \alpha) + c_2\} e^{\lambda_1 t} + c_2 (\lambda_1 - \alpha) t e^{\lambda_1 t}] \cdot \frac{1}{\beta}.$$

The constants c_1, c_2 can be determined from the initial conditions.

If the coupled differential equations be inhomogeneous we can write them in the form, $\frac{dx}{dt} = \alpha x + \beta y + \gamma$ and $\frac{dy}{dt} = ax + by + c$, where $\alpha, \beta, \gamma, a, b, c$ are constants. Assuming the solution to be given by $x = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} + x_o$ and $y = d_1 e^{\lambda_1 t} + d_2 e^{\lambda_2 t} + y_o$ we substitute them in the equations. Equating the coefficients of $e^{\lambda_1 t}$, $e^{\lambda_2 t}$ and the term independent of t respectively from both the sides we obtain the values of λ_1, λ_2 and x_o, y_o . Finally c_1, c_2 (and d_1, d_2) can be found out as before by the initial conditions imposed on x, y .

The pair of coupled first order linear differential equations having more general coupling terms (containing derivatives on the $r. h. s.$ say) may be transformed to the above form by simple manipulation. The case of three first order coupled equations in three independent variables can be dealt with similarly.

Example

$$\frac{dx}{dt} = 3x + 2y, \quad \frac{dy}{dt} = x + 2y.$$

We can write $\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$

The eigenvalue equation is, $(\lambda - 3)(\lambda - 2) - 2 = 0$ giving $\lambda = 1, 4$.

Thus the general solution is, $x = c_1 e^t + c_2 e^{4t}$ and

$$y = [(c_1 e^t + 4 c_2 e^{4t}) - 3(c_1 e^t + c_2 e^{4t})]/2 = \frac{1}{2} (c_2 e^{4t} - 2 c_1 e^t).$$

The constants c_1 and c_2 can be found out from the initial conditions of the problem.

Note: The two equations can be combined to give,

$$\frac{d^2 x}{dt^2} - 5 \frac{dx}{dt} + 4x = 0.$$

This second order equation can also be split up into the following two first order equations: $\frac{dx}{dt} = y, \frac{dy}{dt} = 5y - 4x$ giving rise to the same solution as before. For three coupled first order differential equations with constant coefficients involving three independent variables the representative homogeneous equations are:

$$\frac{dx}{dt} = a_1 x + b_1 y + c_1 z.$$

$$\frac{dy}{dt} = a_2 x + b_2 y + c_2 z.$$

$$\frac{dz}{dt} = a_3 x + b_3 y + c_3 z.$$

The eigenvalues of the evolution matrix are found out by solving the eigenvalue equation, $\det |\lambda I - A| = 0$,

where the matrix A is given by $A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$. The eigenvectors corresponding to the three eigenvalues

of A are determined by the usual method and the general solution is obtained by the superposition of the individual eigenfunctions of the form $e^{\lambda_1 t}$, $e^{\lambda_2 t}$ and $e^{\lambda_3 t}$ satisfying the equations above, where λ_1 , λ_2 and λ_3 are the respective eigenvalues.

In the case of a set of coupled second order differential equations the situation is more complicated and depends on the type of coupling present.

In spite of the different available methods of integration and solution of differential equations neither all integrations can be exactly evaluated nor the equations are amenable to exact solution. In the appendix a few commonly used methods of numerically solving these problems are briefly introduced without derivations of the results used.

SOLVED PROBLEMS

Problem 1

First order differential equation

Solution:

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

Solution:

Dividing by y^2 we get, $\frac{d}{dx}\left(\frac{1}{y}\right) - \frac{1}{y \cdot x} = -1$.

or, $\frac{1}{y} \exp \int -\left(\frac{dx}{x}\right) = -1 \int \left[\exp \int \left(\frac{-dx}{x}\right) \right] dx + c$, c being any constant

or, $\frac{1}{y} \cdot \frac{1}{x} = - \int \frac{dx}{x} + c = c - \ln x$ leading to $y = \frac{1}{x(c - \ln x)}$

The equation is of the form $\frac{dy}{dx} + Py = Qy^n$, where P and Q are functions of x and called the Bernoulli equation.

(b) Solve. $\frac{dy}{dx} = \frac{a^2}{(x+y)^2}$. Putting $x+y=z$ we get,

Solution: $1 + \frac{dy}{dx} = \frac{dz}{dx}$ and hence $\frac{dz}{dx} = \frac{a^2}{z^2} + 1$

or, $\frac{z^2}{a^2 + z^2} dz = dx$ which leads on integration to

$$z - a \tan^{-1} z/a = x + \text{const.}$$

or, $y = a \tan^{-1} \left(\frac{x+y}{a} \right) + c$, c being the constant.

(c) Solve: $\frac{dy}{dx} = \frac{x^2(1+y^2)}{y^2(1+x^2)}$

Solution:

or, we obtain, $\frac{y^2}{1+y^2} dy = \frac{x^2}{1+x^2} dx$, which on integration yields,

$y - \tan^{-1} y = x - \tan^{-1} x + c$, c being the constant.

(d) Solve:- $y = px + \phi(p)$, $p = \frac{dy}{dx}$.

Solution:

Differentiating with respect to x we obtain, $x \frac{dp}{dx} + \phi'(p) \frac{dp}{dx} = 0$

$$\text{or, } \frac{dp}{dx} (x + \phi'(p)) = 0$$

Thus, $\frac{dp}{dx} = 0$ gives $p = c$, c being any constant.

Hence, $y = cx + \phi(c)$ is the general solution of the equation.

The other possibility results from $x + \phi'(p) = 0$, and substituted in $y = px + \phi(p)$ gives rise to another solution called the singular solution. This singular solution is seen to be the envelope of the family of curves given by the general solution. This type of equation is known as Clairut's equation.

Problem 2

Simple higher order equations and some simple functions. Solve the following

equations:

$$(a) \frac{d^2 y}{dx^2} + \left(\frac{dy}{dx} \right)^2 + 1 = 0$$

Solve: Putting $\frac{dy}{dx} = p$ we get, $\frac{dp}{dx} = -(1 + p^2)$

or, $\frac{dp}{1 + p^2} = -dx$, giving on integration, $\tan^{-1} p = c_1 - x$ ($c_1 = \text{constant}$)

or, $p = \frac{dy}{dx} = \tan(c_1 - x)$ leading to the solution

$$y = \int \tan(c_1 - x) dx + c_2 \quad (c_2 = \text{constant})$$

$$= \ln \{ \cos(c_1 - x) \} + c_2$$

$$(b) \text{ Solve: } \frac{d^3 y}{dx^3} - 2 \frac{d^2 y}{dx^2} - \left(\frac{dy}{dx} \right) + 2y = \sin x$$

The complementary function is given by $y = c_1 e^{m_1 x} + c_2 e^{m_2 x} + c_3 e^{m_3 x}$, where m_1, m_2 and m_3 are the roots of $m^3 - 2m^2 - m + 2 = 0$

or, $m_{1,2,3} = \pm 1, 2$.

The particular integral is found out by assuming a solution $y = A \cos x + B \sin x$, giving

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

Equating coefficients of $\cos x$ and $\sin x$ respectively we get,

$$4A - 2B = 0 \text{ and } 2A + 4B = 1$$

Solving these we get, $A = \frac{1}{10}$ and $B = \frac{1}{5}$

Hence the general solution is, $y = c_1 e^x + c_2 e^{-x} + c_3 e^{2x} + \frac{1}{10} \cos x + \frac{1}{5} \sin x$.

$$(c) x^2 y'' - 6y = x$$

Solution:

The homogeneous equation is, $x^2 y'' - 6y = 0$.

Let, $y = \sum_n a_n x^{n+p}$ be the solution of the homogeneous equation we obtain by substituting in the equation,

$$\sum_n [(n+p)(n+p-1) - 6] a_n x^{n+p} = 0.$$

Equating the coefficient of x^p we get, $a_0 [\rho(\rho-1) - 6] = 0$.

or, $\rho = 3, -2$ assuming that $a_0 \neq 0$.

Equating the coefficient of $x^{p+1}, x^{p+2}, \dots$ to zero we get, $a_1 = a_2 = \dots = 0$.

The two independent solutions of the homogeneous equation are x^3 and x^{-2} .

The Wronskian formed from these solution, $W = -5$

The general solution is, $y = c_1 y_1 + c_2 y_2$, where $y_1 = x^3, y_2 = x^{-2}$ and

$$c_1' = - \int_{x_1}^x \frac{x^{-2} x}{5} dx = - \frac{1}{5} \ln x + c_1$$

$$c_2' = \int \frac{1}{5} x^4 dx = \frac{x^5}{25} + c_2, \text{ where } c_1, c_2 \text{ are arbitrary constants.}$$

The general solution is, $y = c_1 y_1 + c_2 y_2 - \frac{x^3}{5} \ln x + \frac{x^3}{25}$

Problem 3

(a) Prove that $P'_{n+1} - P'_{n-1} = (2n+1) P_n$.

(b) Using the generating function and the orthogonality relation of Legendre polynomials find out (i) $P_l(1)$, (ii) $P_l(-1)$, (iii) $\int_{-1}^1 P_l dx$.

Solution:

(a) Differentiating (7.10) with respect to x we get,

$$(2n+1)(P_n + xP_n') = (n+1)P'_{n+1} + nP'_{n-1}$$

Multiplying (7.11) by $(2n+1)$ we obtain,

$$(2n+1)(P'_{n+1} + P'_{n-1}) = (2n+1)(P_n + 2xP_n')$$

Multiplying the first equation by 2 and adding to the second we obtain, $P'_{n+1} - P'_{n-1} = (2n+1)P_n$,

(b) (i) Putting $x = 1$ in the generating function we obtain,

$$\frac{1}{(1-t)} = \sum_l t^l P_l(1)$$

Equating coefficients of t^l we get, $P_l(1) = 1$

(ii) Putting $x = -1$ in the generating function we obtain,

$$\frac{1}{1+t} = \sum_l t^l P_l(-1)$$

Equating coefficients of t^l we obtain, $P_l(-1) = (-1)^l$

(iii) Since $\int_{-1}^1 P_l P_m dx = \frac{2}{(2l+1)} \delta_{lm}$ we obtain

$$\int_{-1}^1 P_l dx = 2 \delta_{l0}, \text{ since } P_0(x) = 1.$$

Problem 4

(a) Prove that $J_p(u+v) = \sum_r J_r(u) J_{p-r}(v)$

(b) Prove that $1 = 2 \sum_{n=1}^{\infty} J_n^2(x) + J_0^2(x)$.

Solution:

(a) We have, $\exp \left[u/2 \left(t - \frac{1}{t} \right) \right] = \sum_n t^n J_n(u)$

Thus, $\exp \left[\frac{(u+v)}{2} \left(t - \frac{1}{t} \right) \right] = \sum_n t^n J_n(u+v)$

or, $\sum_n t^n J_n(u) \cdot \sum_m t^m J_m(v) = \sum_s t^s J_s(u+v)$.

Equating the coefficient of t^p from both the sides we get,

$$J_p(u+v) = \sum_r J_r(u) J_{p-r}(v).$$

(b) Using the generating function we can write,

$$\exp \left[\frac{x}{2} \left(t - \frac{1}{t} \right) \right] = \sum_n J_n(x) t^n$$

$$\text{and } \exp \left[\frac{-x}{2} \left(t - \frac{1}{t} \right) \right] = \sum_m t^m J_m(-x).$$

or, multiplying the two we obtain, $\sum t^{n+m} J_n(x) J_m(-x) = 1$

Since $J_m(-x) = (-1)^m J_m(x)$ from the infinite series expression of $J_m(x)$, we obtain, $J_0^2(x) + 2J_1^2(x) + 2J_2^2(x) + \dots = 1$ [Since $J_{-m}(x) = (-1)^m J_m(x)$]

or, $1 = 2 \sum_{n=1}^{\infty} J_n^2(x) + J_0^2(x)$.

Problem 5

(a) Calculate $\int_{-\infty}^{\infty} x e^{-x^2} H_n H_m dx$

(b) Expand $(a + bx + cx^2)$ in a series of (i) Legendre and (ii) Hermite polynomials.

Solution:

$$(a) \quad I = \int_{-\infty}^{\infty} x e^{-x^2} H_n H_m dx = -\frac{1}{2} \int_{-\infty}^{\infty} \left[\frac{d}{dx} (e^{-x^2} H_n H_m) - e^{-x^2} (H'_n H_m + H_n H'_m) \right] dx$$

The first part vanishes on integration so that we obtain,

$$I = \frac{1}{2} \int_{-\infty}^{\infty} [e^{-x^2} (H'_n H_m + H_n H'_m)] dx$$

Using the recurrence relation (7.25) we obtain,

$$\begin{aligned} I &= \frac{1}{2} \int_{-\infty}^{\infty} [e^{-x^2} (H'_n H_m + H_n H'_m)] dx \\ &= n [2^m_m! \sqrt{\pi} \delta_{m, n-1} (m+1) + 2^n_n! \sqrt{\pi} \delta_{n, m-1} (n+1)] \\ &= n \sqrt{\pi} [2^m (m+1)! \delta_{m, n-1} + 2^n (n+1)! \delta_{n, m-1}]. \end{aligned}$$

[Using the relation (7.28)]

- (b) (i) Since the expression is quadratic in x we will obtain Legendre polynomials upto order 2 in the expansion as is clear from the orthogonality relation. Thus we can write, $a + bx + cx^2 = \sum_l c_l P_l(x)$, c_l 's being the expansion coefficients.

$$\text{Now, from (7.18), } c_l = \left(\frac{2l+1}{2} \right) \int_{-1}^1 (a + bx + cx^2) P_l(x) dx.$$

$$\text{Hence, } c_0 = \frac{1}{2} \cdot \left(a + \frac{c}{3} \right) \cdot 2 = \left(a + \frac{c}{3} \right) \quad c_1 = b, \quad c_2 = 5/2 \int_{-1}^1 (a + bx + cx^2) \frac{1}{2} (3x^2 - 1) dx$$

$$\begin{aligned} \text{or, } c_2 &= 5/4 \int_{-1}^1 (a + bx + cx^2) (3x^2 - 1) dx \\ &= 5/4 \left[3c \cdot \frac{2}{5} - \frac{2c}{3} \right] = \frac{2c}{3} \quad c; \text{ resulting in } a + bx + cx^2 = \left(a + \frac{c}{3} \right) P_0 + bP_1 + \frac{2}{3} CP_2 \end{aligned}$$

(ii) Here we will use the relation (7.29) to get,

$$c_l = \frac{1}{2^l l! \sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2} H_l(x) (a + bx + cx^2) dx$$

By the same reasoning as in (i) here will have non-vanishing c 's for $L = 0, 1, 2$. Using $H_0(x) = 1, H_1(x) = 2x, H_2(x) = 4x^2 - 2$ we get,

$$c_o = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2} H_o(x) \left[aH_o + \frac{1b}{2} H_1 + \frac{c}{4} (2 + H_2) \right] dx.$$

Using the orthogonality relation we can write,

$$c_o = \frac{1}{\sqrt{\pi}} \left[a + \frac{c}{2} \right] \sqrt{\pi} = \left(a + \frac{c}{2} \right), \quad c_1 = b, \quad c_2 = 2c.$$

Problem 6

Prove that $\frac{d^m}{dx^m} H_n(x) = \frac{2^m n!}{(n-m)!} H_{n-m}(x)$

Solution:

From the recurrence relation (7.25) $H'_n(x) = 2n H_{n-1}(x)$

We obtain by successive differentiation,

$$H''_n = 2_n H'_{n-1} = 2_n \cdot 2 (n-1) H_{n-2} = \dots$$

$$\text{Thus, } H_n^{(m)} = \frac{d^m}{dx^m} H_n(x) = 2_n \cdot [2 (n-1)] [2 (n-2)] \dots [2(n-m)] H_{n-m}(x) = \frac{2^m n!}{(n-m)!} H_{n-m}(x).$$

Note: Similar to the Legendre polynomials, we can deduce the following Rodrigues formula for the Hermite polynomials either by straightforward differentiation or by employing Cauchy's integral formula of the complex variable analysis.

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}).$$

Dirac delta function, Beta function and Gamma function.

(a) Show that $\delta'(x)$ behaves as $-\frac{\delta(x)}{x}$.

Solution:

$\int_{-\infty}^{\infty} x \delta'(x) f(x) dx$ for any well-behaved function $f(x)$ can be written as

$$\left[x f(x) \delta(x) \int_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{d}{dx} (x f(x)) \delta(x) dx \right]$$

after integration by parts. The second term = $-\int_{-\infty}^{\infty} f'(x) x \delta(x) + f(x) \delta(x) dx$

Thus $\int_{-\infty}^{\infty} x \delta'(x) f(x) dx = -f(0)$ since $\int_{-\infty}^{\infty} x f'(x) \delta(x) dx = 0$

Thus $\int_{-\infty}^{\infty} x \delta'(x) f(x) dx = -\int_{-\infty}^{\infty} f(x) \delta(x) dx$

Hence we can say that $x \delta'(x)$ behaves as $-\delta(x)$ or $\delta'(x)$ as $-\frac{\delta(x)}{x}$.

(b) $\int_{-4}^6 \sin(\alpha x) \delta(x^2 - 9) dx$

Using the formula $\int_{-\infty}^{\infty} f(x) \delta[F(x)] dx = \sum_{j=1}^n \frac{\delta(x - c_j)}{\left| \frac{dF}{dx} \right|_{x=c_j}} f(x)$, where

$F(x) = (x - c_1)(x - c_2) \dots (x - c_n)$ we obtain

$$\begin{aligned} &= \int_{-4}^6 \sin(\alpha x) \delta[(x - 3)(x + 3)] dx \\ &= \int_{-9}^6 \frac{f(x) \delta(x - 3)}{6} dx + \int_{-4}^6 \frac{f(x) \delta(x + 3)}{6} dx, \text{ where } f(x) = \sin(\alpha x) \\ &= \frac{1}{6} [\sin(3\alpha) - \sin(3\alpha)] = 0 \end{aligned}$$

[Note:] The result depends on the limits of the integral. If the integration extends from -2 to 6 , the result would be $\frac{1}{6} \sin 3\alpha$ since the second integral would vanish in that case].

(c) Calculate $\int_0^{\infty} e^{-ax^2} dx$ and hence find out $\int_0^{\infty} x^2 e^{-x^2} dx$ and $\int_0^{\infty} x^4 e^{-x^2} dx$

Solution:

$$I = \int_0^{\infty} e^{-ax^2} dx = \frac{1}{2} \int_0^{\infty} e^{-z} a^{-\frac{1}{2}} z^{-\frac{1}{2}} dz \text{ [Putting } ax^2 = z, dz = 2ax dx = 2a \cdot \sqrt{\frac{z}{a}} dx]$$

$$\text{or, } I = \frac{1}{2} \Gamma\left(\frac{1}{2}\right) \frac{1}{\sqrt{a}} = \frac{1}{2} \sqrt{\frac{\pi}{a}}$$

$$\frac{dI}{da} = - \int_0^{\infty} x^2 e^{-ax^2} dx = - \frac{1}{4} \sqrt{\pi} a^{-3/2}$$

$$\text{Putting } a = 1, \int_0^{\infty} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{4}$$

$$\frac{d^2 I}{da^2} = \int_0^{\infty} x^4 e^{-ax^2} dx = \frac{3}{8} \sqrt{\pi} a^{-5/2}$$

$$\text{or, } \int_0^{\infty} x^4 e^{-x^2} dx = \frac{3}{8} \sqrt{\pi}$$

$$(d) \text{ Calculate } \int_a^b (b-x)^{m-1} (x-a)^{n-1} dx, m, n > 0.$$

Solution:

$$I = \int_a^b (b-x)^{m-1} (x-a)^{n-1} dx = (b-a)^{m+n-2} \int_a^{\pi/2} \cos^{2(m-1)} \theta \sin^{2(n-1)} \theta (b-a) \sin 2\theta d\theta$$

[on putting $x = a \cos^2 \theta + b \sin^2 \theta$]

$$\text{or, } I = 2 (b-a)^{m+n-1} \int_a^{\pi/2} \cos^{2(m-1)} \theta \sin^{2(n-1)} \theta d\theta \dots$$

$$\text{Now } B(p, q) = \int_a^1 x^{p-1} (1-x)^{q-1} dx = \int_a^{\pi/2} \sin^{2(p-1)} \theta \cos^{2(q-1)} \theta \sin 2\theta d\theta$$

[where $x = \sin^2 \theta$]

$$\text{or, } B(p, q) = 2 \int_a^{\pi/2} \sin^{2p-1} \theta \cos^{2q-1} \theta d\theta$$

The required integral $I = 2(b-a)^{m+n-1} B(m, n)$.

$$(e) \text{ Calculate (i) } \Gamma\left(\frac{1}{2} - n\right) \Gamma\left(\frac{1}{2} + n\right), \text{ (ii) } \int_0^{\pi/2} \sin^5 \theta d\theta, \text{ (iii) } \int_0^{\pi/2} \cos^5 \theta d\theta, \int_0^1 x^2 \left[\ln\left(\frac{1}{x}\right) \right]^3 dx$$

Solution:

$$(i) \Gamma\left(\frac{1}{2} - n\right) \Gamma\left(\frac{1}{2} + n\right) = \frac{\pi}{\sin \pi \left(\frac{1}{2} - n\right)} \left[\text{Using } \Gamma(z) \Gamma(1-z) = \frac{\pi}{\sin \pi z} \right]$$

(ii) Putting $p = 3, q = \frac{1}{2}$ we get,

$$\int_0^{\pi/2} \sin^5 \theta d\theta = \frac{1}{2} B\left(3, \frac{1}{2}\right) = \frac{1}{2} \frac{\Gamma(3)\Gamma(1/2)}{\Gamma(7/2)} = \frac{\Gamma\left(\frac{1}{2}\right)}{\frac{5}{2} \cdot \frac{3}{2} \cdot \frac{1}{2} \Gamma\left(\frac{1}{2}\right)}$$

$$\text{or, } \int_0^{\pi/2} \sin^5 \theta d\theta = \frac{8}{15}$$

$$\text{(iii) } \int_0^{\pi/2} \cos^5 \theta d\theta = \frac{1}{2} B\left(\frac{1}{2}, \frac{3}{2}\right) = \frac{8}{15}.$$

$$\text{(iv) } I = \int_0^1 x^2 \left[\ln\left(\frac{1}{x}\right) \right]^3 dx$$

Putting $\ln\left(\frac{1}{x}\right) = z$ we get, $x = e^{-z}$

$$\text{or } I = \int_{\infty}^0 e^{-2z} z^3 (-e^{-z} dz) = - \int_{\infty}^0 z^3 e^{-3z} dz$$

$$\text{Substituting } 3z = t, I = \int_0^{\infty} \left(\frac{t}{3}\right)^3 e^{-t} dt/3 = \frac{1}{3^4} \Gamma(4) = \frac{2}{27}.$$

SUMMARY

In this chapter we have discussed the method of solving differential equations and the important characteristics of the solutions, concentrating ourselves mainly to the first and the second order linear equations.

- The solutions of some simple first order differential equations have been discussed. The method of finding out the integrating factors, that is always feasible in two variables, has been discussed. The condition of obtaining the integrating factors in the case of differentials comprising three variables, namely Caratheodory's principle, has been stated.
- Regarding the second order differential equations we have focused on the solutions of a few simple equations with constant coefficients. The solution consists of the solution of the homogeneous part (also called the complementary functions) that will contain two constants of integration along with that of the inhomogeneous part, also called the particular integral. The constants of integration can be determined from the boundary conditions imposed on the solution. The application to two physical problems dealing with resonances in electrical circuit and mechanical oscillations have been discussed. The method of the variation of parameters that enables us to know the solution of the complete differential equation in a general way, given the solution of the homogeneous part, has been discussed in the text.
- We have dealt with the method of the series expansion in connection with the solution of linear differential equations of the second order and exemplified this method with the examples of three differential equations, namely, Legendre, Hermite and Bessel. A second order differential equation of the form

$\frac{d^2 y}{dx^2} + P(x) \frac{dy}{dx} + Q(x) y = 0$ has $x = x_o$, an ordinary point if $P(x)$ and $Q(x)$ are analytic functions of x at $x = x_o$, $x = x_o$ is a regular singular point of the differential equation if $P(x)$ or $Q(x)$ or both has isolated singularities at $x = x_o$ and $\lim_{x \rightarrow x_o} (x - x_o) p(x)$ and $\lim_{x \rightarrow x_o} (x - x_o)^2 Q(x)$ are both finite. For a general n th, order differential equation $x = x_o$, is a regular singular point if $\lim_{x \rightarrow x_o} (x - x_o) P_1(x)$, $\lim_{x \rightarrow x_o} (x - x_o)^2 P_2(x)$... $\lim_{x \rightarrow x_o} (x - x_o)^n P_n(x)$ are all finite. The definition is equally valid in the complex plane with x replaced

by z . For the Legendre and the Hermite equation $x = 0$ is an ordinary point while it is a regular singular point for the Bessel equation.

For $x = x_o$, an ordinary point of the second order equation we can obtain a solution of the form $y = \sum_n a_n$

$(x - x_o)^n$, the a_n 's being constants by substituting in the equation to have a recurrence relation between the coefficients. Using this relation we will have two independent series solutions, are containing a_o and

the other a_1 . For $x = x_o$, a regular singular point of the equation we can attempt a solution of the form

$y = \sum_n a_n (x - x_o)^{n+\rho}$, the index ρ being determined from the indicial equation obtained by equating the

coefficient of the lowest power of x to zero. For two different values of ρ we will obtain two recurrence relations. If the difference of two roots ρ_1 and ρ_2 be non-integral, we will obtain two independent series solutions while if the difference be zero or integer we are sure to obtain one solution by the series method, the other being found out by a method given in the text. Of course, we can obtain two independent

series solutions in the first case, namely when $x = x_o$ is an ordinary point by putting $y = \sum_n a_n (x - x_o)^{n+\rho}$

$(\rho$ being zero and one). But this leads to unnecessary complications. When $x = x_o$ is neither an ordinary point nor a regular singular point, it is called an irregular singular point of the equations. In this case the series solution may not yield any non-trivial solution at all. This case has also been illustrated in the text by an example.

- For certain values of the parameters of the equation and the arbitrary constants occurring in the series we obtain the special functions arising out of a differential equation.

Thus we get Legendre polynomials, Hermite polynomials, Bessel functions from the solutions of the respective differential equations. One of the motivations for studying these equations is that these special functions have their orthonormalisation properties in certain intervals of values of the arguments and hence provide us with suitable basis functions in the expansion of functions satisfying Dirichlet conditions in the appropriate intervals. For example, the interval for Legendre polynomials is given by $-1 \leq x \leq 1$, for Hermite polynomials by $-\infty \leq x \leq \infty$, for Bessel functions by $0 \leq x \leq a$ where a is one of the zeroes of the Bessel function.

The orthonormality, generating functions and the recurrence relations for these different special functions have been discussed in the text.

- These are other types of special functions that play important roles in various applications but do not come out of differential equations. Their properties are given briefly in the following.

(a) Gamma functions—These symbolized by $\Gamma(z)$ have got various representations given in the text.

For example, the integral representation is $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt (Re\ z > 0)$. This can be integrated

by parts to yield, $\Gamma(z+1) = z \Gamma(z)$. Thus $\Gamma(n+1) = n!$ for integer n and $\Gamma(z) = \left(= \frac{\Gamma(z+1)}{z} \right)$ is defined in the entire z -plane except for poles at $z = 0, -1, \dots$.

- (b) Beta function $B(x, y)$ —This is related to the gamma function by $B(x, y) = \frac{\Gamma(x) \Gamma(y)}{\Gamma(x+y)}$ and thus is defined for x, y in the entire complex plane. $B(x, y)$ has an integral representation given by

$$B(x, y) = \int_0^1 p^{x-1} (1-p)^{y-1} dp \text{ converging for } \operatorname{Re} x > 0, \operatorname{Re} y > 0.$$

- (c) Dirac Delta function—This function introduced by Dirac in his studies on quantum mechanics has the property $\delta(x) = 0$ for $x \neq 0$ and $\delta(x) \rightarrow \infty$ for $x = 0$. This is an infinitely discontinuous function at $x = 0$. This can however be obtained as the limit of an infinite sequence of some well-behaved functions, and belongs to the class of generalized functions. This function is defined in the sense of an integral in the sense that for any well-behaved function $f(x)$, $\int_{-\infty}^{\infty} f(x) \delta(x-a) dx = f(a)$. [The

limits of the integral can be any two numbers surrounding the point $x = a$]. The properties of this function as well as its extension to higher dimensions have been discussed.

- We conclude this chapter with a brief introduction to the solution of the coupled linear first order differential equations.

MULTIPLE CHOICE QUESTIONS

- The integrating factor of $4x dy + 5(x+y) dx$ is
 (a) $x^{1/4}$ (b) $x^{-3/4}$
 (c) $x^{-1/4}$ (d) impossible to find out.
- The solution of $(x+1)dx - (y+1) dy = 0$ gives
 (a) a family of straight lines
 (b) a family of circles
 (c) a family of hyperbolas
 (d) a family of ellipses
- The solution of $x^2 dy + y(x+y) dx = 0$ is,
 (a) $2x - y = cx^3$ ($c = \text{constant}$)
 (b) $2x + y = cx^2 y$ (c) $2x + y = cx^3 y$ (d) $2x^2 y^2 = c(2x - y)$
- The order and the degree of the equation $y = x^5 \frac{d^3 y}{dx^3} + \left(\frac{dy}{dx} \right)^4$ are
 (a) 3, 4 (b) 4, 3 (c) 5, 4 (d) 5, 3
- The solution of the equation $(x+y^3)dy - ydx = 0$ is
 (a) $y^2 = x^2 + cy$ ($c = \text{constant}$)
 (b) $y^3 = x^3 + cy$
 (c) $y^3 = 2(cy + x)$
 (d) $y^3 - x = c$

6. The solution of the equation $\frac{dy}{dx} + y = e^x$ with the conditions $y'(x=0) = 1$ is,
 (a) $y = \sin x$ (b) $y = \cos hx$
 (c) $y = \sinh x$ (d) $y = \cos x$
7. The solution of the equation $y'' - 2y' + y = 0$ with the condition $(y)_0 = 1, y'_0 = 0$ is,
 (a) $y = e^x (2 + x)$
 (b) $y = e^{-x} (1 + x)$
 (c) $y = e^x (4 + x)$
 (d) $y = e^x (1 - x)$
8. The solution of the equations $\frac{dy}{dx} = ay + bz$, $\frac{dz}{dx} = cy + dz$ is,
 (a) $y = Ae^{\lambda x} + Be^{\mu x}, z = \frac{1}{b}[(a + \lambda)Ae^{\lambda x} + (\mu + a)Be^{\mu x}]$
 $[\lambda, \mu \text{ are the eigenvalues of the matrix } \begin{pmatrix} a & b \\ c & d \end{pmatrix}],$
 (b) $y = Ae^{\lambda x} + Be^{\mu x}, z = \frac{1}{b}[(\lambda - a)Ae^{\lambda x} + (\mu + a)Be^{-\mu x}]$
 (c) $y = Ae^{\lambda x} + Be^{-\mu x}, z = \frac{1}{b}[(\lambda - a)Ae^{\lambda x} + (\mu - a)Be^{-\mu x}]$
 (d) $y = Ae^{\lambda x} + Be^{\mu x}, z = \frac{1}{b}[(\lambda - a)Ae^{\lambda x} + (\mu - a)Be^{\mu x}]$
9. The solution of $y'' + y = 0$ with $(y)_0 = (y')_0 = 0$ has
 (a) a unique solution
 (b) an infinite number of solutions
 (c) no nontrivial solution
 (d) two independent solutions
10. The solution of $\frac{d^3 y}{dx^3} - y = x$ is,
 (a) $y = Ae^x + e^{-x/2} \left(B \cos \frac{\sqrt{3}}{2} x + C \sin \frac{\sqrt{3}}{2} x \right) - x$
 (b) $y = Ae^{-x} + Be^{-x/2} \cos \frac{\sqrt{3}}{2} x + Ce^{x/2} \sin \frac{\sqrt{3}}{2} x - x$
 (c) $y = Ae^{-x} + e^{-x/2} \left(B \cos \frac{\sqrt{3}}{2} x + C \sin \frac{\sqrt{3}}{2} x \right) + x$
 (d) $y = Ae^{-x} + e^{-x/2} \left(B \cos \frac{\sqrt{3}}{2} x - C \sin \frac{\sqrt{3}}{2} x \right) - x$

where A, B, C are constants.

11. If in the equation $y'' + p(x)y' + q(x)y = 0$, $xp(x) = 6 + x + x^2 + \dots$, and $q(x) = \frac{4}{x^2} + \frac{5}{x} + 7$ we attempt Frobenius method of solution, the roots of the indicial equation are
 (a) 1, 4 (b) 2, 3 (c) -1, -4 (d) -2, -3
12. The series solution of Bessel differential equation $x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - m^2)y = 0$ about $x = -1$ will be valid upto
 (a) $x = 1$ (b) $x = 0$ (c) $x \rightarrow \infty$ (d) $x = -2$
13. The order and the degree of the differential equation $\frac{d^3 y}{dx^3} = \left\{ y + \left(\frac{dy}{dx} \right)^3 \right\}^{1/4}$ is
 (a) 4, 3 (b) 3, 4 (c) 1, 4 (d) 4, 2
14. $x^2 + 2x + 3$ can be expressed as
 (a) $P_2 + \frac{2}{3}P_1 + 3P_0$ (b) $\frac{2}{3}P_2 + 2P_1 + \frac{10}{3}P_0$
 (c) $\frac{2}{3}P_2 + P_1 + 3P_0$ (d) $\frac{2}{3}P_2 + P_1 + 3P_0$
15. If $\int_{-1}^1 x P_n(x) dx = 2/3$ then n is
 (a) 1 (b) 2
 (c) 0 (d) -1
16. $\sin x$ is equal to
 (a) $2J_1 + 3J_2 + 4J_3 + \dots$
 (b) $J_1 - J_3 + J_5 \dots$
 (c) $J_1 - 2J_3 + 3J_5 \dots$
 (d) None of these
17. $J'_0(x)$ is equal to
 (a) $-J_{-1}(x)$ (b) $J_0(x) + J_2(x)$
 (c) $-J_1(x)$ (d) $J_0(x) - J_0(x)$
18. $H_n(-x)$ is equal to
 (a) $H_n(x)$ (b) $(-1)^n H_n(x)$
 (c) $-H_n(x)$ (d) None of these
19. $\int_{-1}^1 x P_n(x) \frac{dP_n(x)}{dx} dx$ is equal to
 (a) 0 (b) $\frac{2n}{2n+1}$
 (c) $\frac{1}{n+1}$ (d) $2(-1)^n$

20. $B\left(\frac{1}{2}, \frac{1}{2}\right)$ is equal to
- (a) π (b) $\sqrt{\pi}$ (c) $2\sqrt{\pi}$ (d) $\frac{1}{2}\sqrt{\pi}$
21. If $\int_0^{\pi/2} \cos^{1/2}\theta = I_1$ and $\int_0^{\pi/2} \sin^2\theta d\theta = I_2$ then
- (a) $I_1 = 2I_2$ (b) $I_1 = I_2$ (c) $I_2 = 2I_1$ (d) None of these
22. If $f(x)$ be even, $\int_{-\infty}^{\infty} f(x)\delta(x^2 - a^2)dx$ is equal to
- (a) $\frac{f(a)}{a}$ (b) $af(a)$ (c) $2f(a)$ (d) $\frac{2f(a)}{a}$
23. $\int_{-\pi/3}^{\pi/3} f(x)\delta'\left(x - \frac{2\pi}{3}\right)dx$ is equal to
- (a) $f\left(\frac{2\pi}{3}\right)$ (b) 0 (c) $f(\pi/3)$ (d) $f(-\pi/3)$
24. A particle moves along the x -direction under the action of a force $\lambda \delta(t - t_0)$, λ being a constant. The particle continues to move
- (a) with its velocity unchanged
 (b) with an acceleration λ/m
 (c) with a uniform velocity $u + \lambda/m$, u being the initial velocity
 (d) None of these
25. The general solution of $\frac{d^4 y}{dx^4} - 16y = 0$ is,
- (a) $y = c_1 e^{2x} + c_2 e^{-2x} + e^{2x} (c_3 \cos x + c_4 \sin x)$
 (b) $y = c_1 e^{2x} + c_2 e^{-2x} + c_3 \cos 2x + c_4 \sin 2x$
 (c) $y = c_1 e^{2x} + c_2 e^{-2x} + e^{-2x} (c_3 \cos x + c_4 \sin x)$
 (d) $y = (c_1 \cos x + c_2 \sin x) (c_3 e^{2x} + c_4 e^{-2x})$
26. If the generating function $G(x, \rho)$ of Hermite polynomials be defined by $G(x, \rho) = \sum_n \frac{P^n}{n!} H_n(x)$, then
- $$\int_{-\infty}^{\infty} x e^{-x^2} G(x, \rho_1) G(x, \rho_2) dx = \sqrt{\pi} A e^{2\rho_1 \rho_2}, \text{ where } A \text{ equals (a) } \rho_1 \rho_2 \text{ (b) } \frac{\rho_1}{\rho_0} \text{ (c) } \rho_1 - \rho_2 \text{ (d) } \rho_1 + \rho_2$$

Answers to MCQs

1. (a)	2. (c)	3. (b)	4. (a)	5. (c)	6. (a)	7. (d)
8. (d)	9. (c)	10. (a)	11. (c)	12. (b)	13. (b)	14. (b)
15. (a)	16. (b)	17. (a)	18. (b)	19. (b)	20. (a)	21. (b)
22. (a)	23. (b)	24. (c)	25. (b)	26. (d)		

EXERCISE

1. Solve the following differential equations

(a) $y' + y \sin x = \sin 2x$,

(b) $y'' = e^y$ (Hint: multiply by $2y'$).

(c) $y' = \frac{a^2}{(x+y)^2}$,

(d) $y'' + y = \sin x$,

(e) $y'' - y = e^x$,

(f) $\{(x+1)e^x - e^x\} dx = xe^y dy$.

(g) $\sec^2 y \frac{dy}{dx} + x \tan y = 2x$

[Hint: Solve the problems (d) and (e) by the method of the variation of parameters]

2. Solve the following equations by the method of expansion in series.

(a) $y'' + y = 0$,

(b) $y'' + 3y' + 2y = 0$,

(c) $y'' + xy = 0$

(d) $x^2 y'' - 2y = x$,

(e) $y'' + xy' + y = 0$.

3. (a) Given that one solution of the equation $y'' + 3y' + 2y = 0$ is e^{-x} find out the complete solution of the equation $y'' + 3y' + 2y = \sin x$ by the method of finding out the second solution and the variation of parameters.
 (b) One solution of the equation $x^2 y'' + xy' + x^2 y = 0$ is $J_0(x)$.

Find out the other independent solution.

4. The equation $x^2 y'' + (3x-1)y' + y = 0$ has an irregular singular point at $x = 0$. Show that the equation obtained by the method of expansion in series has a zero radius of convergence.
 5. Test the singularities of the following differential equations (include the point at infinity also).
 (a) $x(x-1)y'' + (\alpha x - \beta)y' + \gamma y = 0$
 (b) $x^2 y'' + (1-x)y' + \alpha y = 0$.
 (c) $x^2 y'' + 2y' + y = 0$
 (d) $e^x y'' + xy' + x^2 y = 0$.
 6. (a) Use the generating function of $P_n(x)$ to find out
 (i) $P_n(1)$,
 (ii) $P_n(-1)$,
 (iii) $P_{2n}(0)$,
 (iv) $P_{2n+1}(0)$.

- (b) Use the orthogonality relation of Legendre polynomials to find out $\int_{-1}^1 (ax^2 + bx + c) P_l(x) dx$.
- (c) Prove the recurrence relation (7.26) in the text.
7. (a) Expand the functions in the interval $(-1, 1)$ (i) $\theta(x)$ ($= 0$ for $x < 0$ and $= 1$ for $x > 0$), (ii) $|x|$ in terms of $P_l(x)$.
- (b) Prove that $|P_n(x)| \leq 1$ [Express $P_n(x)$ as a sum of cosines from the generating function].
- (c) Expand $a_0 + a_1 x + a_2 x^2$ in the interval $(-\infty, \infty)$ in terms of $H_n(x)$.
8. Prove using the recurrence relation of the Hermite polynomials that
- (a)
$$\int_{-\infty}^{\infty} x^2 e^{-x^2} H_n H_m dx = 2^{n-1} \pi^{1/2} (2n+1) n! \delta_{mn} + 2^n \pi^{1/2} (n+2) \delta_{n+2,m} + 2^{n-2} \pi^{1/2} n! \delta_{n-2,m}.$$
- (b)
$$\frac{d^m}{dx^m} H_n(x) = \frac{2^n n!}{(n-m)!} H_{n-m}(x)$$
9. Prove using the generating function of the Bessel functions, prove
- (a)
$$\cos x = J_0(x) + 2 \sum_{n=1}^{\infty} (-1)^n J_{2n}(x),$$
- (b)
$$\sin x = 2 \sum_{n=1}^{\infty} (-1)^{n+1} J_{2n+1}(x).$$
- (c)
$$1 = |J_0(x)|^2 + 2 \sum_{n=1}^{\infty} |J_n(x)|^2.$$
10. Calculate
- (a)
$$\int_0^{\infty} x^{-1/3} e^{-\beta x} dx,$$
- (b)
$$\int_0^1 \sqrt[3]{\ln x} dx,$$
- (c)
$$\int_0^3 \frac{x^2}{\sqrt{3-x}} dx,$$
- (d)
$$\int_0^{\infty} \frac{y^2}{(1+y)^{c_1}} dy,$$
- (e)
$$\int_0^1 x^2 (1-x^2)^{5/2} dx$$
11. Calculate
- (a)
$$\int_0^{\infty} e^{at} \delta(t+3) dt,$$
- (b)
$$\int_{-\pi}^{\pi} \sin 2x \delta'(x + \pi/3) dx,$$

(c) $\int_{-3}^6 \cos\left(\frac{\pi x}{8}\right) \delta(x^2 - 4) dx,$

(d) $\int_v \exp(i\vec{k} \cdot \vec{r}) \exp(-kr) \cos \theta \delta(\vec{r} - \vec{a}) d\vec{r}.$ (v is an infinite volume). (Hint) Use the r, θ, ϕ representation of δ -function.

12. A body projected upwards with a speed V_0 moves in a medium with resisting force proportional to the square of the speed V .
Find out the maximum height to which the body will rise.
13. A series LCR circuit is connected across an e. m. f. given by $V = V_0 e^{-\alpha t}$, α being a constant. Find out the current at any instant of term t .
14. Separate the following equations into first order equations and solve:

(a) $\frac{d^2 y}{dx^2} + 5 \frac{dy}{dx} + 6y = 0$, given that $(y)_o = 1, (y')_o = 0$ [Hint: $\frac{dy}{dx} = z$]

(b) $\frac{d^3 y}{dx^3} - 2 \frac{d^2 y}{dx^2} - \frac{dy}{dx} + 2 = 0$, given that
 $(y)_o = 1, (y')_o = 1, (y'')_o = 0.$ [Hint: $\frac{dy}{dx} = z, \frac{dz}{dx} = u$]

15. Two point charges q and $-q$ are placed at $z = \pm a$ respectively.
Find out the electric dipole potential at a point at a distance from the origin much greater than a in terms of the Legendre polynomial.
16. Find out the integrating factors where necessary and solve the following equations:
- (a) $6x dy + (5x + 3y) dx = 0$
- (b) $y(4x^2 y^2 + 1) dy + x(2y^4 + 1) dx = 0$
- (c) $dx + (e^{y \sin y} - x)(y \cos y + \sin y) dy = 0$
- (d) $x dy - (y + xe^{-y/x}) dx = 0.$
- (e) $(1 - x) dy + (y - (1 - x^2)) dx = 0$
17. Solve the following.

(a) Given that $y = x$ is one solution of the equation $y'' - \frac{2(1+x)}{x} y' + \frac{2(1+x)}{x^2} y = 0$, find out the other solution and hence solve the equation with x^3 on the right hand side.

(b) $y'' + 4y' + 3y = \cos(2x + 3)$, given that $(y)_o = 1, (y')_o = 1$

(c) $y'' + 4y' + 4y = 0$, given that $(y)_o = 1, (y')_o = 0.$

(d) $y'' - \frac{3y'}{x} + \frac{4y}{x} = 0$ by Frobenius method.

18. Show that a n th order linear differential equation has n independent solutions.

[Hint: If $f_1, f_2, \dots, f_{n+1} + 1$ be $(n + 1)$ solutions show that

$$\det \begin{vmatrix} f_1 & f_2 & \cdots & f_{n+1} \\ f_1' & f_2' & \cdots & f_{n+1}' \\ f_1^n & f_2^n & \cdots & f_{n+1}^n \end{vmatrix} = 0 \text{ thereby proving that } (n+1) \text{ solutions are linearly dependent.}$$

Hints and Answers to Exercise Questions

1. (a) $y = 4 \cos^2 x/2 + c e^{\cos x}$ ($c = \text{constant}$),

(b) $\sqrt{A+2e^y} - \sqrt{A} = c \exp \left(\frac{y}{2} + \frac{A^{\frac{1}{2}}}{2} x \right)$, ($A, c = \text{constants}$).

(c) $y = a \tan^{-1} \left(\frac{x+y}{a} \right) + \text{constant}$

(d) $y = c_1 \sin x + c_2 \cos x - \frac{1}{2} x \cos x$, (e) $y = c_1 e^x + c_2 e^{-x} + x e^x$

(f) $e^x(x+1) = e^y + c$ ($c = \text{constant}$), (g) $\tan y = 2 + c e^{-x/2}$ ($c = \text{constant}$)

2. (a) $y = A \left(x - \frac{x^3}{3!} + \cdots \right) + B \left(1 - \frac{x^2}{2!} + \cdots \right)$,

(b) $y = A \left(1 - x + \frac{x^2}{2!} - \cdots \right) + B(1 - 2x + 2x^2 - \cdots)$

(c) $y = a_0 \left(1 - \frac{x^2}{6} + \frac{x^6}{180} - \cdots \right) + a_1 \left(x - \frac{x^4}{12} + \cdots \right)$

(d) The solution of the homogeneous equation is, $y = Ax^2 + B/x$, A and B being constants: Now apply the technique of the variation of constants.

(e) $y = a_0 \left(1 - \frac{x^2}{2} + \cdots \right) + a_1 \left(x - \frac{x^3}{3} + \cdots \right)$

3. (a) $y = c_1 e^{-x} + c_2 e^{-2x} + \frac{1}{10} (\sin x - 3 \cos x)$

(b) $y_2 = J_0 \left[\ln x + \sum_n a_n x^n \right]$ [The coefficients are related to the coefficients of different powers of x in J_0].

5. (a) Regular singular at $x = 0, 1$ and $x \rightarrow \infty$.

(b) Regular singular at $x = 0$ and regular at $x \rightarrow \infty$.

(c) Irregular singular at $x = 0$ and regular singular at $x \rightarrow \infty$

(d) Regular in the finite domain but irregular singular at $x \rightarrow \infty$

6. (a) (i) 1, (ii) $(-1)^n$, (iii) $\frac{(2n)!}{2^{2n}(n!)^2} (-1)^n$, (iv) 0

(b) $2 \left(\frac{a}{3} + c \right) \delta_{l0} + \frac{2}{3} b \delta_{l1} + \frac{4a}{15} \delta_{l2}$

7. (a) (i) $\theta(x) = \frac{1}{2} P_0 + \frac{B}{4} P_1 + \dots$,

(ii) $|x| = \frac{1}{2} P_0 + \frac{5}{8} P_2 + \dots$

(iii) $a_0 + a_1 x + a_2 x^2 = \left(a_0 + \frac{a_1}{2} \right) H_0 + \frac{a_1}{2} H_1 + \frac{a_2}{4} H_2$

10. (a) $\frac{1}{\beta^{2/3}} \Gamma$, (b) $-\Gamma \left(\frac{4}{3} \right)$, (c) $\frac{48}{5}$,

(d) $\frac{1}{3}$, (e) $\frac{5}{256}$.

11. (a) 0, (b) 1 [**Hint:** $\int_a^b f(x) \delta'(x) dx = -f'(0)$ if $a < 0 < b$ and $= 0$ otherwise]

(c) $\frac{1}{2\sqrt{2}}$, (d) $\exp[k(ia \cos \theta_0 - a)] \cos \theta_0$, θ_0 being polar angle of $\vec{a}$.

12. $h_{\max} = R + \frac{1}{2K} \ln \left(1 + \frac{Kv_0^2}{g} \right)$, R being the radius of the earth.

13. $I = \left(\sqrt{\frac{R^2}{4C^2} - \frac{1}{LC}} - \frac{R}{2L} \right) \left[A \left\{ \exp \left(-\frac{R}{2L} + \sqrt{\frac{R^2}{4C^2} - \frac{1}{LC}} \right) \cdot t \right\} + B \left\{ \exp \left(-\frac{R}{2L} - \sqrt{\frac{R^2}{4C^2} - \frac{1}{LC}} \right) \cdot t \right\} \right]$
 $-aV_0 e^{-\alpha t} / \left(\frac{1}{C} - \frac{R}{L} \alpha + \alpha^2 \right).$

The arbitrary constants A and B can be determined once the initial values are prescribed. (for example, the initial current and charge).

14. (a) $y = 3e^{-2x} - 2e^{-3x}$, (b) $y = \frac{1}{2} \left(3e^x - \frac{1}{3}e^{-x} - \frac{2}{3}e^{2x} \right)$.
15. $\phi = 2q \sum_{l=odd} \frac{a^l p_l(\cos \theta)}{r^{l+1}}$, θ being the angle between $\vec{r}$ and the line joining two charges.
16. (a) $9x^{1/2}y + 5x^{3/2} = \text{constant}$.
 (b) $2x^2y^4 + y^2 + x^2 = \text{constant}$
 (c) $x \exp(\cos y - \sin y) + \int (y \cos y + \sin y) \exp[\cos y + (y-1) \sin y] dy = \text{const.}$
 (d) $e^{y/x} = \ln x + \text{constant}$, (e) $y - xy - x + x^3/3 = \text{constant}$.
17. (a) $y = \frac{x^4}{6} + \frac{1}{4}x^2(x+1) + \frac{1}{8}x$,
 (b) $y = c_1 e^{-x} + c_2 e^{-3x} + \frac{1}{65} [8 \sin(2x+3) - \cos(2x+3)]$, where
 $c_1 = 2 - \frac{1}{10}(\cos 3 + 2 \sin 3)$, $c_2 = \frac{1}{26}(2 \sin 3 + 3 \cos 3 - 26)$.
 (c) $y = e^{-2x}(2x+1)$
 (d) $y = a_0 \left(1 + \frac{4x}{3} + \frac{4}{3}x^2 + \dots \right) + b_0 x^4 \left(1 - \frac{4}{5}x + \frac{4}{15}x^2 - \dots \right)$, a_0 and b_0 being constants.

APPENDIX

(A) Numerical Methods Of Solving An Integral $\int_a^b f(x) dx$.

The subject of integration and the solution of differential equations being intimately connected, we give here a very brief outline of a few popular methods. Many important points of the vast world of numerical analysis being omitted, the discussion presented here falls short of the expectation of any serious reader.

1. Trapezoidal Method

The entire interval $b-a$ is divided into n equal parts and in any interval the function $f(x)$ is replaced by a linear function as that the magnitude of the integral in any such subinterval is replaced by the area of a trapezium. The

i th area $A_i = \frac{1}{2}h(f_i + f_{i-1})$, where $h = \frac{b-a}{n}$. Thus the total integral $= \sum_{i=0}^{n-1} A_i = \frac{h}{2}[f_0 + 2(f_1 + \dots + f_{n-1}) + f_n]$

where $f_i = f(x_0 + ih)$, $i = 0, 1, \dots$. The order of magnitude of the accuracy of the method is obtained by making a Taylor series expansion of the function and taking the difference from that evaluated by the above method. The error in estimation is seen to be proportional to h^2 and f'' , the average value of the second derivative over the interval.

2. Simpson's Rule

Simpson's formula replaces the linear approximation of the integrand in an interval by the quadratic approximation so that $f_{i+1} = f(x_i + h) = f_i + \alpha h + \beta h^2$, α and β being constants.

Now, $f_{i-1} = f(x_i - h) = f_i - \alpha h + \beta h^2$, so that $\beta h^2 = \frac{1}{2} (f_{i+1} + f_{i-1} - 2f_i)$ and $\int_{-h}^h f(x) dx = \int_{-h}^h (f_i + \alpha x + \beta x^2) dx$
 $= 2 (h f_i + \frac{1}{3} \beta h^3)$. Substituting for β we obtain, $\int_{-h}^h f(x) dx = \frac{h}{3} (4f_i + f_{i+1} + f_{i-1})$. For the entire interval

(a, b) the value of the integral according to Simpson's rule is equal to $\frac{1}{3} h [f_0 + f_n + 4(f_1 + f_3 + \dots) + 2(f_2 + f_4 + \dots)]$. The estimate of accuracy by following the method used in (1) leads to an error proportional to h^4 and the average value of $\frac{d^4 f}{dx^4}$ over the interval.

3. Gauss quadrature

Instead of dividing the entire interval into equal parts this method utilises the properties of the special functions in suitable intervals to express the integrals in terms of the values of the function at some appropriately selected points and suitable weight functions. Thus leaving aside the details of the calculation that can

be obtained in standard texts on Numerical analysis, the integral $\int_a^b f(x) dx$ can be written using the orthogonality properties of the Legendre polynomials as $\int_a^b f(x) dx = \int_{-1}^1 \left(\frac{b-a}{2} \right) g(x) dx$, where $g(x)$ is obtained from $f(x)$ as the limits are changed from (a, b) to $(-1, 1)$. The integral is also obtained as $\left(\frac{b-a}{2} \right) \sum_{i=1}^n f(x_i) w(x_i)$,

the x_i 's and the w_i 's being the zeroes of the Legendre polynomials and the weight factors respectively that are tabulated in the standard texts. For a three-point integration, the points are the zeroes of $P_3(x)$, i.e., 0 and ± 0.77460 while the weights are 8/9 and 5/9 respectively. The choices of the orthogonal polynomials vary according to the integrand and the limits of integration.

In the calculation of the infinite integrals (a) the trapezoidal or the Simpson's formula may be applied and the expansion continued till we arrive at the desired accuracy, (b) the Gaussian quadrature formula may be applied with a suitable weight formula.

Example Let us take a simple integral $\int_0^1 x^2 dx$ and apply the trapezoidal and the Simpson's formula respectively,

Trapezoidal formula:— Let us calculate $\int_0^1 x^2 dx$, where the values chosen are $h = 0.25, f_0 = 0, f_1 = \frac{1}{64}$,

$f_2 = \frac{1}{8}, f_3 = \frac{27}{64}, f_4 = 1$. Thus we obtain, $\int_0^1 x^2 dx = \frac{1}{8} \left[\frac{1}{32} + \frac{1}{4} + \frac{27}{32} + 1 \right] = 0.27$.

$$\text{Simpson's formula: } - \int_0^1 x^2 dx = \frac{0.25}{3} [0 + 1 + 4\{(0.25)^2 + (0.75)^2\} + 2(0.5)^2]$$

$$\text{or, } \int_0^1 x^2 dx = \frac{0.25}{3} \left[1 + 4 \left(\frac{1}{16} + \frac{9}{16} \right) + 1/2 \right] = \frac{1}{3}$$

The exact result is also $\frac{1}{3}$ thus agreeing exactly with Simpson's formula since $\frac{d^4 f}{dx^4} = 0$ for $f = x^2$ used here. (Follow the discussion given earlier).

Numerical Solution of Differential Equations

Since numerical solution of a differential equation involves integration, the solution of the equation is expected to bear resemblance with the results of numerical integration obtained earlier. In the following the results of Runge-Kutta method will be stated only without derivation.

The equation is $\frac{dy}{dx} = f(x, y)$ for example. If $x = x_0 + \delta x$, $(y)_{x_0 + \delta x} = y_0 + \frac{1}{6} [k_1 + k_4 + 2(k_2 + k_3)]$, where $y_0 = f(x_0, y_0)$, $k_1 = \delta x f(x_0, y_0)$, $k_2 = \delta x f\left(x_0 + \frac{1}{2}\delta x, y_0 + \frac{1}{2}k_1\right)$,

$k_3 = \delta x f\left(x_0 + \frac{1}{2}\delta x, y_0 + \frac{1}{2}k_2\right)$, $k_4 = \delta x f(x_0 + \delta x, y_0 + k_3)$. The result is now extended to calculate y at

$x_0 + 2\delta x$ in the same manner and continued onwards. The result is similar to Simpson's formula and seen to be accurate to $0(\delta x)^4$. Students may apply this method to solve some simple differential equations of the first order. In the case of second order differential equations the method may be used after breaking up the equation into two first order equations of the above form.

Let us take a couple of simple equations whose exact solutions are easily known so that a straightforward comparison is very simple.

1. $\frac{dy}{dx} = x^2$. Taking $x_0 = y_0 = 0$ and $\delta x = 0.1$ we obtain, that $k_1 = 0$, $k_2 = 0.00025$, $k_3 = 0.00025$, $k_4 = 0.001$,

$$\text{so that } (y)_{x_0 + \delta x} = \frac{1}{6}(0.001 + 2 \times 0.0005) = \frac{0.001}{3}. \text{ Exact calculation also gives } (y)_{x_0 + \delta x} = \frac{(\delta x)^3}{3} = \frac{0.001}{3}$$

2. $\frac{dy}{dx} = x + y$. Choosing $x_0 = y_0 = 0$, $\delta x = 0.1$ and following the above prescription we can see that $(y)_{x_0 + \delta x}$

$$= 0.0051. \text{ Exact calculation gives } (y)_{x_0 + \delta x} = (y)_{0.1} = e^{0.1} - (1 + 0.1) = 0.00516.$$

The numerical calculation thus yields results correct to $0(\delta x)^4$.

Sturm–Liouville Problem and Green’s Function

CHAPTER

8

8.1 BOUNDARY-VALUE PROBLEM AND EIGENFUNCTIONS

This chapter addresses the boundary value problems of homogeneous and inhomogeneous second order differential equations. Keeping the level of the students in view the discussion has been limited to ordinary differential equation only.

The boundary value problem implies the solution of a differential equation subjected to the relevant boundary conditions. Suppose, we solve the equation $\frac{d^2 y}{dx^2} + k^2 y = 0$, k being a constant with the condition that $(y)_{x=0} = (y)_{x=l} = 0$.

We obtain, $y = A \sin kx + B \cos kx$, A and B being constants. The condition $(y)_{x=0} = 0$ gives $B = 0$, while $(y)_{x=l} = 0$ gives, $k_n = \frac{n\pi}{l}$ ($n = \pm 1, \pm 2, \dots$). Thus for a particular n , the solution $y_n = A_n \sin \frac{n\pi x}{l}$ and the general solution is given

by $y = \sum_n A_n \sin \frac{n\pi x}{l}$, the k_n 's being called the eigenvalues. [Compare with the eigenvalues and the eigenstates in a vector space]. The eigenfunctions are linearly independent for different n and form a set of orthonormal functions in the interval

$(-L, L)$ so that $\int_{-L}^{+L} \sin \frac{n\pi x}{l} \sin \frac{m\pi x}{l} dx = L \delta_{n,m}$. Any function that satisfies Dirichlet

boundary conditions (stated in the earlier chapter and discussed elaborately in the subsequent chapter on Fourier series) in the same interval can be expanded in terms of the above trigonometric functions that satisfy the similar ortho-normalisation conditions in appropriate intervals and form suitable basis. The Legendre and the Hermite polynomials are a couple of such basis functions. This chapter will dwell briefly on the properties of these eigenfunctions and the nature of the differential equations that give rise to these functions as their solutions. These problems, called

the Sturm-Liouville problems can be said to be the equivalent of the vector space problems in the domain of functions.

The space of eigenfunctions is, in general, infinite dimensional whereas that of vectors may be finite or infinite dimensional. In the space of functions the range of the variation of the independent variable has to be specified. The operators acting in the space of functions are differential or integral operators while those in the space of the vectors are matrices.

To speak about the operators in the function space we recall the definition of the hermitian operators in the vector space, which is very similar to that in the function space. An operator F acting in the space of

functions is said to be hermitian if $\int_a^b \psi^* F \phi \, dx = \int_a^b (F \psi)^* \phi \, dx$, where a , and b denote the range of variation

of x and ψ, ϕ are the two functions on which F operates. (An identical definition holds good in more than one dimension). Actually the above equality is valid for a hermitian operator under certain boundary conditions on ψ, ϕ, ψ', ϕ' etc. This will be illustrated by some specific examples. If two operators F and G satisfy the condition (under certain boundary conditions on ψ, ϕ, ψ', ϕ').

$$\int_a^b \psi^* F \phi \, dx = \int_a^b (G \psi)^* \phi \, dx, \quad G \text{ will be said to be the hermitian adjoint of } F. \text{ It is usually denoted by } F^\dagger.$$

Examples

(a) Let us take $F = \frac{d}{dx}$.

$$\int_a^b \psi^* \frac{d\phi}{dx} \, dx = [\psi^* \phi]_a^b - \int_a^b \frac{d\psi^*}{dx} \phi \, dx$$

Clearly if the first term on the r. h. s. vanishes, a condition that can be satisfied under various possibilities

(for example, ψ and ϕ vanish at a, b or ψ, ϕ have equal magnitudes at a, b etc.) we obtain, $\int_a^b \psi^* \frac{d\phi}{dx} \, dx = - \int_a^b \frac{d\psi^*}{dx} \phi \, dx$, leading to the result that $-\frac{d}{dx}$ is the hermitian adjoint of $\frac{d}{dx}$.

$$\text{Let } F = \frac{d^2}{dx^2}$$

$$\begin{aligned} \text{(b) } \int_a^b \psi^* \frac{d^2 \phi}{dx^2} \, dx &= \left[\psi^* \frac{d\phi}{dx} \right]_a^b - \int_a^b \frac{d\psi^*}{dx} \frac{d\phi}{dx} \, dx \\ &= \left[\psi^* \frac{d\phi}{dx} \right]_a^b - \left[\phi \frac{d\psi^*}{dx} \right]_a^b + \int_a^b \frac{d^2 \psi^*}{dx^2} \phi \, dx \end{aligned}$$

Clearly, if the first two integrated parts vanish, the operator $\frac{d^2}{dx^2}$ is hermitian. Similarly, in three dimensions, ∇^2 is a hermitian operator. A general second-order differential equation can be expressed

as $p_o(x) y'' + p_1(x) y' + p_2(x) y = 0$ or $Ly = 0$. The operator L is thus $p_o \frac{d^2}{dx^2} + p_1 \frac{d}{dx} + p_2$. For the operator to be hermitian under certain conditions on the functions, it can be shown that $p_o'(x) = p_1(x)$. [This is left as an exercise for the students]. Then it follows that Ly is given by

$$Ly = \frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] + qy, \text{ where } p(x) = p_o(x) \text{ and } q(x) = p_2(x). \text{ The differential equation is then said to be}$$

in Sturm-Liouville form $Ly = 0$ (8.1)

For example, Legendre's differential equation $(1 - x^2)y'' - 2xy' + \lambda y = 0$ is already in Sturm-Liouville form with $p(x) = (1 - x^2)$, $q = \lambda$. Sometimes q is written as $q(x) - \lambda w(x)$, so that the Sturm-Liouville equation

reads, $\frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] + qy = \lambda wy$, w being called the weight factor and λ the eigenvalue of the eigenvalue equation which is the original equation with boundary conditions (often $w = 1$).

For Schrodinger equation, $q = \frac{2m}{\hbar^2} V(x)$, $\lambda = \frac{2m}{\hbar^2} E$, $w = 1$. Any general differential equation of the form $\left(p_o \frac{d^2}{dx^2} + p_1 \frac{d}{dx} + p_2 \right) y = 0$ can be brought to Sturm-Liouville form by multiplication with $\frac{1}{p_o(x)}$

$\exp \left[\int^x \frac{p_1(z)}{p_o(z)} dz \right]$. Then $\left(p_o \frac{d^2}{dx^2} + p_1 \frac{d}{dx} + p_2 \right) y$ transforms into $\frac{d}{dx} \left[\exp \left\{ \int^x \frac{p_1(z)}{p_o(z)} dz \right\} y' \right] + \frac{p_2}{p_o}$

$\exp \left[\int^x \frac{p_1(z)}{p_o(z)} dz \right] y$. This is clearly in Sturm-Liouville form with $p(x) = \exp \left\{ \int^x \frac{p_1(z)}{p_o(z)} dz \right\}$,

$q = \frac{p_2}{p_o} \exp \left[\int^x \frac{p_1(z)}{p_o(z)} dz \right]$. For Bessel equation the multiplying factor $= \frac{1}{x^2} \exp \left[\int^x \frac{z}{z^2} dz \right] = \frac{1}{x}$.

The equation is then of the form $x \frac{d^2 y}{dx^2} + \frac{dy}{dx} + \left(x - \frac{n^2}{x} \right) y = 0$

or, $\frac{d}{dx} \left(x \frac{dy}{dx} \right) + x y = \frac{n^2}{x} y$, Here $p(x) = x$, $q(x) = x$, $\lambda = n^2$, $w = \frac{1}{x}$.

The space of eigenfunctions satisfies properties similar in many respects to those of vectors in a vector space, as is exemplified below.

- (a) The eigenvalues of a Sturm-Liouville operator may be (i) entirely discrete (one dimensional infinite well in quantum mechanics), (ii) entirely continuous (scattering problem) or (iii) comprised of discrete and continuous sets. (One-dimensional finite well for example). In any one of the above three situations, the eigenfunctions form a complete set so that a well-behaved function can be expanded in an appropriate interval in terms of them. Of course, the question of completeness in an infinite dimensional function space is involved because of the problem of convergence and will not be taken up here. Thus, for the case (i) we can write, $\psi(x) = \sum_n c_n u_n(x)$. If the u_n 's are orthonormalised, we obtain, $c_n = \int u_n^*(x') \psi(x') dx'$

so that $\psi(x) = \int \left(\sum_n u_n(n) u_n^*(x') \right) \psi(x') dx'$, where the integration extends over the range of existence of the variable x' . Thus we obtain, $\sum_n u_n^*(x') u_n(x) = \delta(x - x')$ (8.2)

The eigenfunctions may be degenerate so that the summation above has to be taken over all the degenerate states corresponding to the eigenvalue.

A similar equality follows in the cases (ii) and (iii) respectively, with the difference that in the continuum case the summation sign in Eq. (8.2) has to be replaced by an integral. (Extension to the 3D case is obvious).

- (b) The eigenvalues of a hermitian operator are all real and the eigenfunctions of a hermitian operator corresponding to different eigenvalues are orthogonal. The proof is as follows and very similar to that in the case of vectors in a vector space.

For a hermitian operator F , $\int \psi^* F \phi dx = \int (F\psi)^* \phi dx$, ϕ and ψ being two functions belonging to the space. If $\phi = \psi$ and each corresponds to an eigenfunction of the operator F with an eigenvalue λ , we can write, $F\psi = \lambda\psi$, $F\phi = \lambda\phi$ and, hence, from the above $\int (\lambda - \lambda^*) \psi^* \psi dx = 0$ leading to $\lambda = \lambda^*$. If the respective eigenvalues corresponding to ψ and ϕ be λ and μ respectively, we get $\int (\lambda - \mu^*) \psi^* \phi dx = 0$ yielding the result that $\int \psi^* \phi dx = 0$, the integral extending over the domain of definition of x , thus proving orthogonality for $\lambda \neq \mu$.

- (c) The proof of Schwarz inequality $\left| \int_a^b f^* g dx \right|^2 \leq \left[\int_a^b f^* f dx \right] \left[\int_a^b g^* g dx \right]$ is exactly similar to Eq. (3.2)

and the method of Gram-Schmidt orthogonalisation has been discussed in Section (3.4) in the case of the space of functions.

- (d) The infinite dimensional space of functions is not necessarily complete, i.e., an infinite sum of functions may not coverage to a function of the same kind. For example, an infinite sum of trigonometric functions (Fourier series to be discussed in Chapter 9) which are continuous in any interval may coverage to a function that may have a discontinuity at some points in the interval. The condition of completeness of

a set of orthogonal functions $\phi_n(x)$ (i.e., $\int_a^b \phi_n^* \phi_m dx = \delta_{nm}$) in regard to the expansion of a function $\psi(x)$

is that, $\int_a^b |\psi(x)|^2 dx - \sum_{n=1}^m |a_n|^2 \geq 0$ where $a_n = \int_a^b \phi_n^* \psi dx$. For $m \rightarrow \infty$ $\int_a^b |\psi(x)|^2 dx = \sum_{n=1}^{\infty} |a_n|^2$.

8.2 INHOMOGENEOUS DIFFERENTIAL EQUATIONS, GREEN'S FUNCTION

The solution of differential equations of second order with an inhomogeneous term on the right hand side is of relevance in the various disciplines of physics, the problems of electrostatics, electrodynamics furnishing a few examples. The Sturm-Lionville equation with an inhomogeneous term on the right hand side is of the form,

$$Ly - \lambda y = f(x). \quad (8.3)$$

The boundary conditions are assumed to be homogeneous, i.e., of the form $(y)_{x=a} = (y)_{x=b} = 0$, for example. (A general homogeneous boundary condition is of the form $\alpha y + \beta y' = 0$ at $x = a$ and $x = b$, the two end points, α, β being constants).

The Green's function of the operator $(L - \lambda)$ subject to the same homogeneous boundary condition satisfies the equation

$$(L - \lambda) G(x, x') = \delta(x - x'). \quad (8.4)$$

Thus the Green's function represents the solution of the problem with a unit concentrated source at $x = x'$.

The solution of the inhomogeneous problem is thus $y = \int_a^b G(x, x') F(x') dx'$ (as can be seen by direct substitution also). The solution of the homogeneous equation can be added to this y to give the general solution provided the appropriate boundary conditions are satisfied.

Example

Let us solve the equation $\frac{d^2 y}{dx^2} + \lambda y = f(x)$, where the boundary conditions are $(y)_{x=0} = (y)_{x=L} = 0$. ($\lambda > 0$)

Solution: The Green's function $G(x, x')$ of this problem satisfies the equation $\frac{d^2 G}{dx^2} + \lambda G = \delta(x - x')$.

Thus for $x > x'$ as well as for $x < x'$, $G(x, x')$ obeys the equation $\frac{d^2 G}{dx^2} + \lambda G = 0$.

For $x < x'$ we can write the solution as

$$G_1 = a(x') \cos(\sqrt{\lambda} x) + b(x') \sin(\sqrt{\lambda} x). \quad (8.5)$$

For $x > x'$ the solution is similarly,

$$G_2 = C(x') \cos(\sqrt{\lambda} x) + d(x') \sin(\sqrt{\lambda} x), \quad (8.6)$$

Since $G(x, x')$ is continuous at $x = x'$ we obtain

$$a(x') \cos(\sqrt{\lambda} x') + b(x') \sin(\sqrt{\lambda} x') = C(x') \cos(\sqrt{\lambda} x') + d(x') \sin(\sqrt{\lambda} x'). \quad (8.7)$$

Again the equation $\frac{d^2 G}{dx^2} + G = \delta(x - x')$ on being integrated from $x = x' - \epsilon$ to $x' + \epsilon$ ($\epsilon \rightarrow 0$) gives,

$$\left. \frac{dG_2}{dx} \right|_{x=x'+\epsilon} - \left. \frac{dG_1}{dx} \right|_{x=x'-\epsilon} = 1.$$

$$\begin{aligned} \text{or,} \quad & \sqrt{\lambda} [b(x') \cos(\sqrt{\lambda} x') - a(x') \sin(\sqrt{\lambda} x')] \\ & - \sqrt{\lambda} [d(x') \cos(\sqrt{\lambda} x') - c(x') \sin(\sqrt{\lambda} x')] = -1, \end{aligned}$$

$$\text{or,} \quad [b(x') - d(x')] \cos(\sqrt{\lambda} x') - [a(x') - c(x')] \sin(\sqrt{\lambda} x') = -\frac{1}{\sqrt{\lambda}}. \quad (8.8)$$

Again, the homogeneous boundary condition give

$$a(x') = 0 \text{ and } c(x') \cos(\sqrt{\lambda} L) + d(x') \sin(\sqrt{\lambda} L) = 0 \dots \quad (8.9)$$

Thus
$$d(x') = -c(x') \cot(\sqrt{\lambda} l)$$

Equations (8.7), (8.8) and (8.9) enable us to find out the constants $a(x')$, $b(x')$, $c(x')$ and $d(x')$ respectively. The solution of the inhomogeneous equation is then,

$$y(x) = \int_0^x G_2(x, x') f(x') dx' + \int_x^l G_1(x, x') f(x') dx', \text{ where } G_1 \text{ and } G_2 \text{ are obtained by substituting the magni-}$$

tudes of b , c , d . A few points given below can be noticed in this regard.

- (1) The importance of the boundary condition on the evaluation of the Green's function can be easily understood. For example, if we alter the homogeneous boundary conditions, the relations (8.9) will change giving rise to a change in the Green's function for the same operator.
- (2) To a solution of the equation as determined above, we can always add a solution of the homogeneous equation to get the general solution. Only if the solution of the homogeneous equation can be made to vanish by a suitable choice of the boundary conditions the part involving the Green's function gives the full solution.

The Green's function can also be obtained by another method known as the eigenfunction expansion method, described as follows.

The equation determining the Green's function is,

$$(L - \lambda) G(x, x') = \delta(x - x'), \text{ with the homogeneous boundary conditions imposed on } G(x, x') \quad (8.10)$$

Let $u_n(x)$ be the eigenfunction of the operator L for the same boundary condition so that $L u_n(x) = \lambda_n u_n(x)$. Since the eigenfunctions form a complete set, we can write,

$G(x, x') = \sum_n a_n(x') u_n(x)$, the a_n 's being the expansion coefficients. Substitution in Eq. (8.10) gives,

$$\sum_n a_n (\lambda_n - \lambda) u_n(x) = \delta(x - x')$$

$$\text{or, } \int \sum_m a_m (\lambda_m - \lambda) u_m(x) u_n^*(x) dx = u_n^*(x')$$

Using the orthonormality of the eigenfunctions, we obtain

$$a_n = \frac{u_n^*(x')}{\lambda_n - \lambda} \text{ leading to } G(x, x') = \sum_n \frac{u_n^*(x') u_n(x)}{\lambda_n - \lambda} \dots\dots\dots (8.11)$$

SOLVED PROBLEMS

Problem 1

Show that the bound state eigenfunctions of a one-dimensional Schrodinger equation are non-degenerate.

Solution:

The one-dimensional time-independent Schrodinger equation is,
$$\frac{d^2 \psi}{dx^2} + \frac{2m}{\hbar^2} (E - V) \psi = 0.$$

Let us, for a particular value of E , have two different eigenfunctions ψ_1 and ψ_2 . We obtain,

$$\frac{d^2\psi_1}{dx^2} + \frac{2m}{\hbar^2} (E - V) \psi_1 = 0, \dots \quad (8.12) \text{ (A)}$$

and

$$\frac{d^2\psi_2}{dx^2} + \frac{2m}{\hbar^2} (E - V) \psi_2 = 0, \dots \quad (8.12) \text{ (B)}$$

Multiplying the first by ψ_2 , the second by ψ_1 and subtracting we get, $\frac{d}{dx} \left(\psi_2 \frac{d\psi_1}{dx} - \psi_1 \frac{d\psi_2}{dx} \right) = 0$

or,

$$\psi_2 \frac{d\psi_1}{dx} - \psi_1 \frac{d\psi_2}{dx} = \text{constant} \dots \quad (8.13)$$

Now, for a bound state problem, ψ_1 and ψ_2 go to zero as $x \rightarrow \pm \infty$

Utilizing the condition in (8.13) we obtain that constant = 0.

or,

$$\psi_2 \frac{d\psi_1}{dx} = \psi_1 \frac{d\psi_2}{dx} \text{ leading to } \frac{1}{\psi_1} \frac{d\psi_1}{dx} = \frac{1}{\psi_2} \frac{d\psi_2}{dx}.$$

Integrating we get, $\ln \frac{\psi_1}{\psi_2} = \text{const.}$

or, $\psi_1 = (\text{const.}) \psi_2$ implying that for a particular energy value there is one particular eigenfunction for the bound state problem.

Problem 2 Show that the Sturm-Liouville differential operator for the Legendre equation is hermitian. State the boundary condition.

Solution:

The Sturm-Liouville operator for the Legendre equation is

$\frac{d}{dx} \left[(1 - x^2) \frac{d}{dx} \right]$. Let ϕ and ψ be two functions of x in the interval $(-1, 1)$ on which the operator acts.

Then $\int_{-1}^1 \psi^* \frac{d}{dx} \left[(1 - x^2) \frac{d\phi}{dx} \right] dx = \left[\psi^* (1 - x^2) \frac{d\phi}{dx} \right]_{-1}^1 - \int_{-1}^1 \frac{d\psi^*}{dx} (1 - x^2) \frac{d\phi}{dx} dx$

If ψ, ϕ as well as their derivatives are finite at $x = \pm 1$ we get

$$\int_{-1}^1 \psi^* \frac{d}{dx} \left[(1 - x^2) \frac{d\phi}{dx} \right] dx = - \int_{-1}^1 \frac{d\psi^*}{dx} (1 - x^2) \frac{d\phi}{dx} dx \dots$$

or,

$$\int_{-1}^1 \psi^* \frac{d}{dx} \left[(1 - x^2) \frac{d\phi}{dx} \right] dx = - \left[(1 - x^2) \frac{d\psi^*}{dx} \phi \right]_{-1}^1 + \int_{-1}^1 \frac{d}{dx} \left[(1 - x^2) \frac{d\psi^*}{dx} \right] \phi dx$$

Using the same boundary conditions on ϕ, ψ and their derivatives as used in the previous relation (9.13) we get,

$$\int_{-1}^1 \psi^* \frac{d}{dx} \left[(1-x^2) \frac{d\phi}{dx} \right] dx = \int_{-1}^1 \frac{d}{dx} \left[(1-x^2) \frac{d\psi^*}{dx} \right] \phi dx.$$

Hence, the operator $\frac{d}{dx} \left[(1-x^2) \frac{d}{dx} \right]$ is hermitian in the space of functions which satisfy the boundary conditions stated above,

[Clearly if ϕ and ψ be the Legendre eigenfunctions these conditions are satisfied].

Problem 3

Find out the Green's function for the operator $\frac{d^2}{dx^2} + k^2$ subject to the condition that $y'(0) = y(l) = 0$. and solve the problem $\frac{d^2 y}{dx^2} + k^2 y = f(x)$, $f(x)$ being any arbitrary function of x .

Solution:

The Green's function of the problem satisfies the equation

$$\frac{d^2 G}{dx^2} + k^2 G = \delta(x-x').$$

Thus, for $x < x'$ we have $\frac{d^2 G}{dx^2} + k^2 G = 0$ giving as a solution $G(x, x') = A(x') \sin kx + B(x') \cos kx$.

For $x > x'$ similarly we have $G(x, x') = C(x') \sin kx + D(x') \cos kx$.

Putting the boundary conditions, we have

$$\begin{aligned} G_1(x, x') &= B(x') \cos kx \text{ for } 0 < x < x' \text{ and} \\ G_2(x, x') &= C(x') \sin kx + D(x') \cos kx \text{ for } x' < x < l. \end{aligned} \quad (8.14)$$

Since $G_2 = 0$ at $x = l$, we have $D = -C \tan k l$ leading to

$$G_2 = \frac{C(x')}{\cos kL} \sin k(x-L)$$

Imposing the conditions of continuity of the Green's function and the discontinuity of its derivative at

$x = x'$ we have, $B(x') \cos kx' = \frac{C}{\cos kL} \sin k(x' - L)$ and

$$\frac{kC}{\cos kL} \sin k(x' + L) + kB \sin kx' = 1 \dots \quad (8.15)$$

Using the two equations we get, $C = \frac{\cos kx'}{k}$ and $B = \frac{\sin k(x' - L)}{k \cos kL}$

Substituting the values of B and C in Eq. (8.14) we obtain G_1 and G_2 .

The solution of the inhomogeneous problem is,

$$y = \int_0^x G_2(x, x') f(x') dx' + \int_x^L G_1(x, x') f(x') dx',$$

SUMMARY

- The eigenfunctions of a self-adjoint (or hermitian) second order differential operator which are the solutions of a homogeneous Sturm-Liouville differential equation of the form $\frac{d}{dx} \left(p \frac{dy}{dx} \right) + qy = \lambda wy$, where p, q, w are in general functions of x , λ is the eigenvalue are briefly discussed in this chapter. Any linear second order differential operator can be brought to this form by multiplication with a suitable factor as described in the text. The space of eigenfunctions satisfies properties similar to those of vectors in an ordinary vector space (such as, completeness, orthogonality for different eigenvalues, degeneracy) with the exception that in the former the space is, in general, infinite dimensional and the operators are differential or integral operators instead of matrices in the ordinary vector space. The discussions are relevant in many problems of physics, particularly in quantum mechanics.
- The Green’s function method utilized for solving an inhomogeneous Sturm-Liouville equation of the form $\frac{d}{dx} \left[p \frac{dy}{dx} \right] + qy - \lambda wy = Ly - \lambda wy = f(x)$ with appropriate boundary conditions where L is the Sturm-Liouville operator $\frac{d}{dx} \left(p \frac{d}{dx} \right) + q$, is briefly described. The Green’s function satisfies the equation $(L - \lambda w) G = \delta(x - x')$, so that we can write the solution of the inhomogeneous problem as $y(x) = \int_a^b G(x, x') f(x') dx'$, where a and b depend on the domain of definition of x . The method of evaluating the Green’s function either by solving the differential equation or by expanding it in terms of the eigenfunctions of the homogeneous problem is discussed.

MULTIPLE CHOICE QUESTIONS

1. The energy eigenvalues of an otherwise free particle constrained to move inside a sphere of radius a are given by

(a) $j_n(ka) = 0$, where $k = \sqrt{\frac{2nE}{\hbar^2}}$, $n = 0, 1, \dots$,

(b) $E = \frac{n^2 \pi^2 \hbar^2}{8ma^2}$

(c) $E_n = \frac{\hbar^2 (2n+1)^2 \pi^2}{4ma^2}$

(d) $J_n(ka) = 0$

2. The condition of an operator to be Hermitian
 - (a) is independent of the boundary conditions imposed on the function on what it operates
 - (b) depends on the extent of the configuration space
 - (c) depends on the boundary conditions
 - (d) cannot be ascertained
3. The eigenfunctions of the one-dimensional Schrodinger equation

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V)\psi(x) = 0 \text{ with } V(x) = V(-x)$$

- (a) are of even or odd parity
 - (b) are of definite parity in the case of bound states only
 - (c) are always of even parity (d) are always of odd parity
4. The eigenfunctions of a three-dimensional Schrodinger equation $\nabla^2\psi + \frac{2m}{\hbar^2}(E - V)\psi = 0$ with $V = V(r)$
 - (a) are always of a definite parity
 - (b) are of a definite parity if $\psi \rightarrow 0$ as $r \rightarrow \infty$ (bound states)
 - (c) are never of a fixed parity
 - (d) are of a definite parity if initially so
5. The eigenfunctions of a differential equation $\frac{d^2y}{dx^2} + \alpha \frac{dy}{dx} + \beta y = 0$ with $\alpha^2 < 4\beta$ and boundary conditions that $y = 0$ at $x = 0$ and $x = l$ have
 - (a) definite eigenvalues given by $4\beta - \alpha^2 = \frac{4n^2\pi^2}{l^2}$ ($n = 0, 1, \dots$)
 - (b) no definite eigenvalues
 - (c) definite eigenvalues if $\alpha^2 > 4\beta$
 - (d) continuous eigenvalues
6. There are two potentials in one dimension, (i) of finite depth, (ii) of infinite depth. The solutions comprise
 - (a) both the bound and the scattering states in two cases,
 - (b) only the bound states in two cases,
 - (c) the bound and the scattering states in the first case and the bound states alone in the second,
 - (d) only the scattering states in both the cases.

Answers to MCQs

1. (a) 2. (c) 3. (b) 4. (d) 5. (a) 6. (c)

EXERCISE

1. Solve the equation $\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V)\psi = 0$, where $V(x) = 0$ for $-\frac{L}{2} < x < \frac{L}{2}$ and $V(x) \rightarrow \infty$ outside so that you can take $\psi(x) = 0$ at $x = -\frac{L}{2}$ and $x = \frac{L}{2}$. Solve the boundary value problem and find out

the eigenvalues and the eigenfunctions. Repeat the above problem for $V(x) = 0$ for $0 < x < L$ and $V(x) \rightarrow \infty$ outside so that $\psi(x) = 0$ at $x = 0$ and $x = L$.

(This is the infinite-well problem in quantum mechanics. Notice the difference in the eigenfunctions in two cases)

- Convert the Hermite differential equation $\frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + 2ny = 0$ and the Laguerre equation $xy'' + (1-x)y' + ny = 0$ into the self-adjoint forms and verify the self-adjointness of the differential operators pointing out the boundary conditions required to be fulfilled in each case.

[Hint: Proceed in a way similar to that in the case of Legendre differential equation as shown in the section on solved problems].

- Starting from the independent functions $(1, x, x^2, \dots)$ construct a set of orthogonal functions in the interval $(-\infty, \infty)$ with the orthogonality relation given by

$$\int_0^\infty e^{-x^2} f_n f_m dx = \delta_{mn}.$$

[Hint: Proceed in a way shown in section 3.4].

- Find out the Green’s function for the operator $\frac{d^2}{dx^2} + k^2$ for the boundary conditions (a) $(y)_{x=0} = (y)_{x=l} = 0$, (b) $(y)_{x=0} = (y')_{x=l} = 0$ (c) $(y')_{x=0} = (y')_{x=l} = 0$ and for the operator $\left(\frac{d^2}{dx^2} - k^2\right)$ with the boundary condition $(y)_{x=0} = 0$ and $(y)_{x \rightarrow \infty}$ is finite.
- Find out the Green’s function for the above-mentioned cases by expanding the Green’s function in terms of the eigenfunctions of the operator. (see text for the method).
- Convert the following equation to a homogeneous boundary value problem by a proper transformation of the independent variable and then solve by the method of Green’s function.

$$\frac{d^2 y}{dx^2} + y = x, \text{ where } (y)_{x=a} = y_0, (y)_{x=b} = y_1.$$

[Hint: Put $z = \alpha y + \beta$ and adjust α, β so that $(z)_{x=a} = (z)_{x=b} = 0$].

- Solve the equation $\frac{d^2 y}{dx^2} + y = 1$ with $(y)_{x=0} = (y)_{x=l} = 0$ by the Green’s function method and compare it with the solution obtained by the method of the variation of parameters discussed in Chapter 7.

Hints and Answers to Exercise Questions

- (a) $\psi_n = \sqrt{\frac{2}{L}} \sin \frac{2n\pi x}{L}, E_n = \frac{2n^2 \pi^2 \hbar^2}{m L^2}, (n = 1, 2, \dots)$

$$\psi_n = \sqrt{\frac{2}{L}} \cos \frac{(2n+1)\pi x}{L}, E_n = \frac{(2n+1)^2 \pi^2 \hbar^2}{2mL^2} (n = 0, 1, \dots)$$

2. Multiply by e^{-x^2} in the first and e^{-x} in the second.
3. The orthonormalised functions are the Hermite polynomials of different orders.

$$4. (a) G = \frac{\sin K(x' - l)}{K \sin Kl} \sin Kx \text{ for } 0 < x < x' \text{ and}$$

$$= \frac{\sin Kx' \sin K(x - l)}{K \sin Kl} \text{ for } x' < x < l.$$

$$(b) G = -\frac{\cos K(x' - l)}{K \cos Kl} \sin Kx \text{ for } 0 < x < x' \text{ and}$$

$$= -\frac{\sin Kx' \cos K(x - l)}{K \cos Kl} \text{ for } x' < x < l,$$

$$(c) G = -\frac{1}{2K} \left[e^{K(x-x')} - e^{-K(x+x')} \right] \text{ for } 0 < x < x'$$

$$= -\frac{1}{2K} \left[e^{Kx'} - e^{-Kx'} \right] e^{-Kx} \text{ for } x' < x < l.$$

$$5. (a) G(x, x') = \frac{2}{L} \sum_n \frac{\sin \frac{n\pi x}{L} \sin \frac{n\pi x'}{L}}{\frac{n^2 \pi^2}{L^2} - K^2}$$

$$(b) G(x, x') = \frac{2}{L} \sum_n \frac{\sin \frac{(2n+1)\pi x}{2L} \sin \frac{(2n+1)\pi x'}{2L}}{\left\{ \frac{(2n+1)^2 \pi^2}{4L^2} - K^2 \right\}}.$$

Fourier Series and Integral Transforms

CHAPTER

9

9.1 INTRODUCTION, FOURIER EXPANSION, PARSEVAL'S THEOREM

We have seen in the Chapter 3 that a set of orthogonal functions is linearly independent in a certain interval of values of the independent variable. (See Eq. 3.5) and can form the basis of the space of functions in the specified interval. The Legendre, Hermite, Bessel functions can be the appropriate basis functions in suitable intervals. (For Legendre polynomials the interval is $(-1, 1)$, for Hermite polynomials it is $(-\infty, \infty)$ and for Bessel functions it is determined by the zeroes of the Bessel function). Our familiar trigonometric sine and cosine functions are also equally good candidates for the basis functions in the appropriate intervals. The conditions for expanding a function in terms of the infinite series of these functions are the same in all these cases:- (a) the function and its derivatives must be entirely or piecewise continuous in the interval, (By piecewise continuity it is implied that the entire interval can be split up into parts in each of which the function is continuous while suffering a discontinuity in going from one part to another), (b) the function must have a finite number of extrema and also of finite discontinuities, in the interval. These conditions are termed Dirichlet conditions and the series of trigonometric functions used for expanding the function is called the Fourier series.

Thus in the interval $(-L, L)$ the function $f(x)$ can be expanded as, $\frac{1}{2} [f(x_+) + f(x_-)]$
 $= a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L} + \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L}$ meaning thereby that the series on the right

hand side converges to the mean of the right hand and the left hand limits of the function at any point in the interval. At a point where $f(x)$ is continuous the mean becomes the value of the function at the point. Since the series can be extended periodically outside the interval, the function $f(x)$ can be also continued periodically outside the interval meaning that $f(x) = f(x + 2L)$, even when $f(x)$ is originally defined in the range $(-L, L)$.

Thus at a point of continuity of $f(x)$, we can write

$$f(x) = a_o + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L} + \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L} \text{ so that,} \quad (9.1)$$

$$a_o = \frac{1}{2L} \int_{-L}^L f(x) dx, a_n = \frac{1}{L} \int_{-L}^L f(x) \cos \frac{n\pi x}{L} dx \text{ and } b_n = \frac{1}{L} \int_{-L}^L f(x) \sin \frac{n\pi x}{L} dx$$

$$\left[\because \int_{-L}^L \cos \frac{m\pi x}{L} \cos \frac{n\pi x}{L} dx = \int_{-L}^L \sin \frac{m\pi x}{L} \sin \frac{n\pi x}{L} dx = L\delta_{mn} \right]$$

For these coefficients to exist, $f(x)$ must be integrable within the interval. The interval $(-L, L)$ can be altered to any interval (l_1, l_2) by a change of variable given by $y = \frac{l_2 - l_1}{2L} x + \frac{l_2 + l_1}{2}$. Thus for $l_1 = -\pi, l_2 = \pi$,

$$y = \frac{\pi x}{l} \text{ so that } f(y) = a_o + \sum_n a_n \cos ny + \sum_n b_n \sin ny \text{ with } a_o = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) dy, a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos ny dy,$$

and $b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin ny dy$. If $f(x)$ has a periodicity of 2π , the limits of the integral can be changed from

$(-\pi, \pi)$ to $(0, 2\pi)$, or from any x_0 to $x_0 + 2\pi$. If the function has a discontinuity at the end-point [for example in the case of periodic continuation of $f(x) = x$ beyond the interval $(-\pi, \pi)$], the Fourier series converges

there to $\frac{1}{2} [f(-\pi + 0) + f(\pi - 0)]$, where $-\pi$ and π are the end-points and the function is periodically continued beyond the interval. For an even function the a_n s exist but the b_n s vanish and the a_n s are given by

$$a_n = \frac{2}{\pi} \int_0^{\pi} f(y) \cos ny dy, \text{ while for an odd function the } b_n\text{s exist with the } a_n\text{s being zero and the } b_n\text{s given by}$$

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(y) \sin ny dy.$$

A complex Fourier series is defined by $f(x) = \sum_{n=-\infty}^{\infty} C_n e^{inx}$, where the C_n 's are given by $C_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx$. Clearly $C_n = \frac{1}{2} (a_n - ib_n)$. For real $f(x)$, $C_n = C_{-n}^*$.

If a function $f(x)$ is defined only for half of the interval, we can extend it as an odd or an even function to the other half and correspondingly obtain the sine-Fourier or the cosine-Fourier series respectively (Solved problem (2)).

Note: It is to be noted that the Dirichlet conditions are sufficient but not necessary for the expansion of a function in the Fourier series. For example, the function $\log \left| 2 \cos \frac{1}{2} x \right|$ has a logarithmic singularity at $x = \pi$ but it has got a Fourier expansion $\sum_{n=1}^{\infty} (-1)^{n-1} \left(\frac{\cos nx}{n} \right)$ in the interval $(-\pi, \pi)$, (Ref. *A course of Modern Analysis* by Whittaker and Watson).

The Fourier series is absolutely and uniformly convergent everywhere in the interval if $\sum_n |a_n|$ and $\sum_n |b_n|$ converge. (See test of uniform convergence in Chapter 5). However, near the point of discontinuity of the function $f(x)$ the convergence of the series is not uniform. The sum of a finite number of terms $f_n(x)$ overshoots the function $f(x)$ at certain points near discontinuity even as $n \rightarrow \infty$. This is called Gibbs phenomenon. Of course, the series converges in the mean to the function in the interval in the sense that

$$\int_a^b \left| f(x) - \sum_{n=1}^m f_n(x) \right|^2 dx \rightarrow 0 \text{ as } m \rightarrow \infty.$$

In this sense, the Fourier series is said to represent a function in the appropriate interval so that the trigonometric function form a complete set for the expansion of a square-integrable function [i.e., $\int_a^b |f(x)|^2 dx$ is finite.] in that interval.

9.1.1 Parseval's Theorem

For a square integrable function represented by a Fourier series in $(-\pi, \pi)$ we can write,

$$\begin{aligned} \int_{-\pi}^{\pi} [f(x)]^2 dx &= \int_{-\pi}^{\pi} \left[a_o + \sum_n a_n \cos nx + \sum_n b_n \sin nx \right]^2 dx \text{ so that} \\ \int_{-\pi}^{\pi} [f(x)]^2 dx &= 2\pi a_o^2 + \pi \sum_n [a_n^2 + b_n^2] \dots \end{aligned} \quad (9.2)$$

$$\text{because } \int_{-\pi}^{\pi} \cos mx \cos nx dx = \int_{-\pi}^{\pi} \sin mx \sin nx dx = \pi \delta_{mn}$$

$$\text{and } \int_{-\pi}^{\pi} \sin mx \cos nx dx = 0$$

This is actually the mathematical representation of the normalization of the function in terms of the Fourier series used to represent the function in the interval $(-\pi, \pi)$.

The Fourier series finds numerous applications in the problems of physics for example, in acoustics in the decomposition of the vibrations from an oscillating system into different frequencies consisting of the fundamental and its various harmonics, or in electrodynamics in the splitting up of the wave emitted by a bounded system into discrete modes. The Parseval's theorem applied to a vibrating system shows how the energy of an oscillating system is divided into different modes.

Note: Unlike the Taylor series which is the infinite power series expansion of a function about a point where it is single valued, continuous and has an infinite number of derivatives, the Fourier series in an interval of the variable converges to a function that may have a finite discontinuity at the point. A Fourier series, if it be uniformly convergent which is often the case, can be integrated term by term in the interval and also differentiated termwise provided the differentiated series is uniformly convergent.

Also, it must be noted that changing the magnitudes of the function at a set of isolated points does not alter the Fourier coefficients and hence keeps the Fourier series unchanged because the integrals involved in the definitions of the coefficients remain unaffected.

SOLVED PROBLEMS

Problem 1

Find the Fourier-series expansion of the following functions:

(a) $f(x) = 0$, $-\pi < x < 0$ and $= 4 \sin x$ for $0 < x < \pi$ and $f(x) = f(x + 2\pi)$.

(b) $f(t) = c$, for $-\frac{T}{2} < t < 0$ and $= A \sin wt$ for $0 < t < \frac{T}{2}$, $f(t) = f(t + T)$,

(where $T = \frac{2\pi}{w}$ and $A, C = \text{constants}$)

Solution:

1. (a) $a_o = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx = \frac{4}{2\pi} \int_0^{\pi} \sin x dx = \frac{4}{\pi}$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx = \frac{4}{\pi} \int_0^{\pi} \sin x \cos nx dx$$

$$\text{or, } a_n = -\frac{4}{2\pi} \left[\left. \frac{\cos(n+1)x}{(n+1)} \right|_0^{\pi} + \left. \frac{\cos(n-1)x}{1-n} \right|_0^{\pi} \right]$$

$$= -\frac{4}{2\pi} \left[\frac{1}{(n+1)} \{(-1)^{n+1} - 1\} + \frac{1}{(1-n)} \{(-1)^{n-1} - 1\} \right]$$

$$= \frac{4}{2\pi} \left[\left(\frac{1}{n+1} + \frac{1}{1-n} \right) + \left(\frac{(-1)^{n+1}}{1+n} - \frac{(-1)^{n-1}}{1-n} \right) \right]$$

$$= \frac{4}{\pi} \left[\frac{1}{1-n^2} + \frac{(-1)^n \cdot n}{1-n^2} \right] = \frac{4}{\pi(1-n^2)} [1 + n(-1)^n]$$

$$b_n = \frac{4}{\pi} \int_0^{\pi} \sin x \sin nx dx \text{ and is seen to be}$$

$$= 4 \text{ for } n = 1 \text{ and } 0 \text{ otherwise.}$$

$$\text{Thus } f(x) = 4 \left[\frac{1}{\pi} + \sum_{n=1}^{\infty} \frac{1}{(1-n^2)\pi} \{1 + n(-1)^n\} + \sin x \right]$$

(b) $a_o = \frac{1}{T} \left[\int_{-T/2}^0 c dt + \int_0^{T/2} A \sin wt dt \right]$

$$= \frac{1}{T} \left[c \cdot \frac{T}{2} - \frac{A}{w} \cos wt \right]_0^{T/2} = \frac{c}{2} + \frac{2A}{wt}$$

$$\text{or, } a_o = \frac{c}{2} + \frac{A}{\pi}$$

$$a_n = \frac{1}{T} \left[\int_{-T/2}^0 c \cos \frac{2n\pi t}{T} dt + \int_0^{T/2} A \sin wt \cdot \cos \frac{2\pi nt}{T} dt \right],$$

$$b_n = \frac{1}{T} \left[\int_{-T/2}^0 c \sin \frac{2\pi nt}{T} dt + \int_0^{T/2} A \sin wt \sin \frac{2\pi nt}{T} dt \right]$$

Integrating, we see that $a_n = \frac{A}{2\pi(1-n^2)} [1 + (-1)^n]$, and

$$b_n = -\frac{c}{2\pi n} [1 - (-1)^n] + \frac{A}{4} \delta_{n,1}.$$

Thus

$$f(t) = \left(\frac{c}{2} + \frac{A}{\pi} \right) + \sum_n \frac{A}{2\pi(1-n^2)} [1 + (-1)^n] \cos n\omega t - \sum_n \left[\frac{c}{2n\pi} \{1 - (-1)^n\} - \frac{A}{4} \delta_{n,1} \right] \sin n\omega t$$

Problem 2(a)

Expand $f(x) = x$ for $0 < x < \pi$ by (i) a Fourier sine-series and (ii) a Fourier cosine series.

Solution:

- (i) We extend $f(x)$ to the interval $-\pi < x < 0$ by writing $f(x) = x$ for $-\pi < x < 0$ so that we can write $f(x) = x$ in $-\pi < x < \pi$.

Hence $f(x)$ is an odd function throughout the interval $-\pi < x < 0$.

Hence, a_n 's = 0 for all n .

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin nx \, dx = \frac{1}{\pi} \left[-\frac{x \cdot \cos nx}{n} \right]_{-\pi}^{\pi} + \int_{-\pi}^{\pi} \frac{\cos nx}{n} \, dx$$

$$\text{or, } b_n = -\frac{1}{n\pi} [2\pi \cos n\pi] = -\frac{2}{n} (-1)^n$$

$$\text{or, } x = -\sum \frac{2}{n} (-1)^n \sin nx.$$

- (ii) We extend $f(x)$ to the interval $0 < x < \pi$ by writing $f(x) = -x$ so that we can write $f(x) = |x|$ in $-\pi < x < \pi$ and hence $f(x)$ becomes an even function in the interval. The coefficients $b_n = 0$ for all n .

$$a_o = \frac{1}{2\pi} \int_{-\pi}^{\pi} |x| dx = \frac{1}{\pi} \int_o^{\pi} x dx = \frac{\pi}{2}.$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} |x| \cos nx dx = \frac{1}{\pi} \left[- \int_{-\pi}^o x \cos nx dx + \int_o^{\pi} x \cos nx dx \right]$$

$$\begin{aligned} \text{or, } a_n &= \left[\frac{1}{\pi} \left\{ - \frac{x \cdot \sin nx}{n} \right\}_{-\pi}^o + \int_{-\pi}^o \frac{\sin nx}{n} dx \right] + \frac{1}{\pi} \left\{ \frac{x \sin nx}{n} \right\}_o^{\pi} - \int_o^{\pi} \frac{\sin nx}{n} dx \Bigg] \\ &= - \frac{1}{n^2 \pi} [1 - (-1)^n] + \frac{1}{n^2 \pi} [(-1)^n - 1] = \frac{2}{n^2 \pi} [(-1)^n - 1] \end{aligned}$$

$$\text{or, } \dot{x} = \frac{\pi}{2} + \sum_n \frac{2}{n^2 \pi} [(-1)^n - 1] \cos nx.$$

Problem 2(b)

Use Parseval's theorem in (i) to prove that $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$.

Solution:

$$\text{From Parseval's theorem } \int_{-\pi}^{\pi} [f(x)]^2 dx = 2\pi a_o^2 + \pi \sum_n (a_n^2 + b_n^2)$$

For (i), $a_n = 0$ for all n .

$$\text{Hence } \int_{-\pi}^{\pi} x^2 dx = \pi \sum_n b_n^2 = 4\pi \sum_n \frac{1}{n^2}$$

$$\text{or, } \frac{2\pi^3}{3} = 4\pi \sum_n \frac{1}{n^2} \text{ giving } \sum_n \frac{1}{n^2} = \frac{\pi^2}{6}.$$

9.2 MULTIPLE FOURIER SERIES

Since the products of the trigonometric functions $\sin mx, \cos ny \dots$ (like $\sin mx \sin ny, \cos mx \cos ny, \sin mx \cos ny, \cos mx \sin ny$ in two variables) form a series of orthogonal functions in the variables $x, y \dots$ (in the interval $0 \leq x \leq 2\pi, 0 \leq y \leq 2\pi$) we expect to have a uniformly convergent Fourier series expansion in two variables in terms of these as the basis functions. In two dimensions the complex Fourier series can be written as,

$$f(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g_{mn} \exp [i (mx + ny)], \text{ so that using the orthogonality properties of the exponentials}$$

$$\text{we can write, } g_{mn} = \frac{1}{4\pi^2} \int_o^{2\pi} \int_o^{2\pi} dx dy \exp [-i (mx + ny)] f(x, y)$$

The Parseval's theorem extended to two variables becomes,

$$\int_0^{2\pi} \int_0^{2\pi} |f(x, y)|^2 dx dy = \sum_{m, n=-\infty}^{\infty} |g_{mn}|^2 \dots \quad (9.3)$$

The extension to more than two variables is obvious.

9.2.1 Poisson's Sum-Formula

The complex Fourier series $f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$ with $c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iny} dy$ gives, $f(x) = \frac{1}{2\pi} \sum_n \int_{-\pi}^{\pi} f(y) e^{in(x-y)} dy$.

If $f(x)$ is a continuous function of x and $\sum_{n=-\infty}^{\infty} f(2\pi n)$ is an infinite series so that $\sum_{n=-\infty}^{\infty} f(2\pi n + t)$ converges

absolutely and uniformly for all t in the interval $0 \leq t < 2\pi$, then $\sum_{n=-\infty}^{\infty} f(2\pi n + t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n, m=-\infty}^{\infty} f(2\pi n + \tau) e^{im(t-\tau)} d\tau$

$$\text{or, } \sum_{n=-\infty}^{\infty} f(2\pi n + t) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{imt} \sum_{n=-\infty}^{\infty} \int_0^{\pi} f(2\pi n + \tau) e^{-im\tau} d\tau$$

$$\text{or, } \sum_{n=-\infty}^{\infty} f(2\pi n + t) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{imt} \sum_{n=-\infty}^{\infty} \int_{2\pi n}^{\pi(2n+1)} f(\tau) e^{-im\tau} d\tau$$

$$\text{or, } \sum_{n=-\infty}^{\infty} f(2\pi n + t) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{imt} \sum_{n=-\infty}^{\infty} \int_{2\pi n}^{\pi(2n+1)} f(\tau) e^{-im\tau} d\tau$$

$$\text{Putting } t = 0 \text{ we get, } \sum_{n=-\infty}^{\infty} f(2\pi n) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} f(\tau) e^{-im\tau} d\tau \dots \quad (9.4)$$

This is Poisson's formula and is very useful in the theory of numbers.

9.3 FOURIER TRANSFORM

The Fourier series expansion of a periodic function can be generalized to a non-periodic function by making

the period approach infinity. For this purpose we start from the Fourier series expansion $f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/L}$, where

$$c_n = \frac{1}{2l} \int_{-l}^l f(x') e^{-in\pi x'/L} dx' \text{ so that we obtain,}$$

$$f(x) = \frac{1}{2l} \sum_n \int_{-l}^l f(x') \exp \left[\frac{in\pi}{l} (x - x') \right] dx'.$$

Putting $\frac{n\pi}{l} = k$, we get $\Delta k = \frac{\pi}{L}$ and as the period of a non-periodic function tends to infinity, $\Delta k \rightarrow 0$.

We get, $f(x) = \frac{1}{2} \sum \frac{\Delta k}{\pi} \int_{-L}^L f(x') \exp [ik (x - x')] dx'$

or, $f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} c(k) \exp (ikx) dk$, where $c(k) = \int_{-\infty}^{\infty} f(x') e^{-ikx'} dx'$.

[The conditions of continuity, bounded variation, square integrability of the function are required for the existence of the Fourier transform]. The formula can be extended to higher dimensions (in 3-dimensions) for example as,,

$$f(\vec{r}) = \frac{1}{(2\pi)^3} \int g(\vec{k}) \exp [i\vec{k} \cdot \vec{r}] d\vec{k}, \text{ where}$$

$g(\vec{k}) = \int f(\vec{r}') \exp (-i\vec{k} \cdot \vec{r}') d\vec{r}'$, the integrations extending over the entire range of variables of $\vec{r}$ and $\vec{k}$ respectively.

The function $f(\vec{r})$ and $g(\vec{k})$ are said to be the Fourier transforms of each other in the respective $\vec{r}$ and $\vec{k}$ spaces. [It is a convention sometimes to call $g(\vec{k})$ the Fourier transform and $f(\vec{r})$ the inverse Fourier transform].

If $f(x)$ be even, $g(k) = 2 \int_0^{\infty} f(x) \cos kx dx$ and for odd $f(x)$, $g(k) = 2 \int_0^{\infty} f(x) \sin kx dx$. These are called the cosine and the sine-Fourier transforms of $f(x)$ respectively. The inverse transforms are $f(x) = \frac{1}{\pi} \int_0^{\infty} g(k) \cos kx dk$ and $f(x) = \frac{1}{\pi} \int_0^{\infty} g(k) \sin kx dk$ respectively. In quantum mechanics $\psi(\vec{r})$ and $\phi(\vec{k})$ are the wave functions in the coordinate ($\vec{r}$) and momentum ($\vec{k}$) representations respectively. Thus the Fourier transforms play important roles in the theory of representations in quantum mechanics

If we take $\psi(x) = A e^{-\alpha x^2}$, the corresponding wave function in the momentum representation $\phi(p)$ is given by.

$$\begin{aligned} \phi(p) &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \psi(x) e^{-ipx/\hbar} dx \\ &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} A e^{-\alpha x^2} \cdot e^{-ipx/\hbar} dx = \frac{A}{2\pi^{1/2}\hbar} e^{-\frac{p^2}{\hbar^2\alpha}} \end{aligned}$$

Thus, the momentum space wave function also becomes Gaussian similar to that in coordinate space. The spread of the wave function in the coordinate space is $0\left(\frac{1}{\sqrt{\alpha}}\right)$ while that in the momentum space is $0(\hbar \sqrt{\alpha})$

so that the product of the spreads in two spaces is $0(\hbar)$, a result to be expected if Δx and Δp are interpreted as the uncertainties in the position coordinate and the corresponding momentum component respectively.

9.4 FOURIER TRANSFORM OF DERIVATIVES

The Fourier transforms are frequently used for solving differential equations. For that purpose the calculation of the Fourier transforms of derivatives is important.

Thus let $g_1(k) = \frac{1}{2\pi} \int_{-\alpha}^{\alpha} \frac{df}{dx} e^{-ikx} dx$

We obtain, $g_1(k) = \frac{1}{2\pi} \left[e^{-ikx} f(x) \Big|_{-\alpha}^{\alpha} + ik \int_{-\alpha}^{\alpha} e^{-ikx} f(x) dx \right]$

Thus provided $f(x) \rightarrow 0$ as $x \rightarrow \pm \infty$ we get, $g_1(k) = ik g(k)$.

[The vanishing of $f(x)$ as $x \rightarrow \pm \infty$ is not an additional restriction because for the square integrability of the function this is necessary]

Similarly, for the n th order derivative we can write, $g_n(k) = (ik)^n g(k)$

9.4.1 Convolution Integral

The convolution of two functions ϕ and ψ is defined to be

$$\phi * \psi = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(y) \psi(x-y) dy. \quad (9.5)$$

$$\begin{aligned} \text{Or, } \phi * \psi &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(y) \int_{-\infty}^{\infty} g(k) e^{ik(x-y)} dk dy \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} g(k) \left[\int_{-\infty}^{\infty} \phi(y) e^{-iky} dy \right] e^{ikx} dk \\ &= \int_{-\infty}^{\infty} g(k) \cdot h(k) e^{ikx} dk, \text{ where } g(k) \text{ and } h(k) \text{ are the} \end{aligned}$$

Fourier transforms of ψ and ϕ respectively in the k space. (9.6)

Thus the product of the Fourier transforms of two functions is the Fourier transform of the convolution of the original functions. The convolution integral in terms of the Fourier transforms defined by Eq. (9.6) is used in the evaluation of integrals, and also in calculating the inverse Fourier transform of the product of two Fourier transforms.

9.4.2 Dirac Delta Function Revisited

In Chapter 7 we discussed various properties of the delta function and mentioned that this generalized function can be seen as the limiting case of many well-behaved functions. Here, we discuss one such limiting case that will be relevant in the context of Fourier transform.

For example $\frac{\sin Nx}{\pi x}$ in the limit $N \rightarrow \infty$ can be shown to behave as $\delta(x)$. Since $N \rightarrow \infty$, $\int_a^b \frac{\sin Nx}{\pi x} dx$, where the end points a and b include the point $x = 0$, will consist of contributions almost entirely from the

close neighbourhood of $x = 0$ so that $\lim_{N \rightarrow \infty} \int_a^b f(x) \frac{\sin Nx}{\pi x} dx = \lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} f(x) \frac{\sin Nx}{\pi x} dx$

$$= f(0) \lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} \frac{\sin Nx}{\pi x} dx = f(0) \left[\text{because } \int_{-\infty}^{\infty} \frac{\sin Nx}{x} dx = \pi \right]$$

$$\text{or } \lim_{N \rightarrow \infty} \int_0^b f(x) \frac{\sin Nx}{\pi x} dx = f(0) = \int_0^b f(x) \delta(x) dx.$$

[$f(x)$ is here any well-behaved function of x].

Thus in the sense of an integral we can say that

$$\delta(x) = \lim_{k \rightarrow \infty} \frac{\sin kx}{\pi x} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk$$

Clearly, from the definition of Fourier transform we obtain that $\delta(x)$ and 1 are the Fourier transforms of each other. In three dimensions we can write, $\delta(\vec{r}) = \frac{1}{(2\pi)^3} \int e^{i\vec{k} \cdot \vec{r}} d\vec{k}$.

Let us take two functions $f(x)$ and $g(x)$. The integral

$$\int_{-\infty}^{\infty} f(x) g^*(x) dx = \int_{-\infty}^{\infty} \frac{1}{4\pi^2} \left[\int_{-\infty}^{\infty} \bar{f}(k) e^{ikx} dk \int_{-\infty}^{\infty} \bar{g}^*(k') e^{-ik'x} dk' \right] dx$$

Using from the above, $\int_{-\infty}^{\infty} e^{i(k-k')x} dx = 2\pi \delta(k - k')$.

$$\text{Or, } \int_{-\infty}^{\infty} f(x) g^*(x) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{f}(k) \bar{g}^*(k) dk.$$

To bring symmetry in the relations, we sometimes define

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \bar{f}(k) e^{ikx} dk \text{ and } \bar{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx.$$

$$\text{Clearly, then we will obtain, } \int_{-\infty}^{\infty} f(x) g^*(x) dx = \int_{-\infty}^{\infty} \bar{f}(k) \bar{g}^*(k) dk \quad (9.7)$$

This relation can obviously be said to be the generalized Parseval's relation in the sense that when $f = g$, the relation (9.7) becomes the usual Parseval's relation

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |\bar{f}(k)|^2 dk \quad \dots \quad (9.8)$$

Thus the relations (9.7) and (9.8) can be regarded as the independence of the scalar product of two functions and the normalization of any function respectively of any particular representation, [x or k representation in this case]. [The extension to three dimensions is obvious]. This is just an equivalent way of stating that the magnitude of any vector or the scalar product of two vectors when the functions are considered to be the vectors in the space of functions remains the same irrespective of the basis chosen. In quantum mechanics this means that the probability of the system remaining in a state is independent of any particular representations.

9.5 LAPLACE TRANSFORM

Sometimes in electrical circuitry, for example, a signal is switched on at a particular instant of time, so that we can write $f(t) = 0$ for $t < 0$, where $t = 0$ is taken to be the instant when the signal starts. The Laplace transform is of great use in analyzing such a pulse to extract the various frequencies that are present in it.

The Fourier transform of a function $f(x) e^{-\alpha x}$ (the factor α is introduced to take account of the convergence of the Fourier transform) is $F(k) = \int_{-\infty}^{\infty} f(x) e^{-\alpha x} e^{-ikx} dx$ so that

(9.9a)

$$f(x) e^{-\alpha x} = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k) e^{ikx} dk \text{ giving } f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k) \exp [x(\alpha + ik)] dk$$

If now $f(x) = 0$ for $x < 0$ we get,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k) e^{sx} dk, \text{ where } s = \alpha + ik \text{ for } x > 0.$$

Changing the variable to $s = \alpha + ik$ we get, $f(x) = \frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} g(s) e^{sx} ds$, where $g(s) = F\{-i(s - \alpha)\}$

(9.9b)

Then the relation (9.9) (A) gives, $g(s) = \int_0^{\infty} f(x) e^{-sx} dx$

(9.10)

Here, M being the maximum exponential growth rate of $f(x)$ given by $f(x) \leq e^{Mx}$ for all $x > 0$, the real part of s , $\text{Re } s > M$ so that the integral Eq. (9.10) converges and $g(s)$ becomes an analytic function of s .

The Laplace transform $g(s)$ given by Eq. (9.10) satisfies the following properties

(a) *Linearity*: $L[c_1 f_1 + c_2 f_2 + \dots] = c_1 Lf_1 + c_2 Lf_2 + \dots$, where $Lf(x) = g(s) = \int_0^{\infty} f(x) e^{-sx} dx$ and each Laplace transform exists.

$$(b) L[f'(x)] = \int_0^{\infty} e^{-sx} f'(x) dx = \left[e^{-sx} f(x) \right]_0^{\infty} + \int_0^{\infty} s e^{-sx} f(x) dx$$

$$\text{or, } L[f'(x)] = sL[f(x)] - f(0)$$

$$L[f''(x)] = sL[f'(x)] - f'(0) = s^2 Lf(x) - sf(0) - f'(0).$$

This can be continued so that we get,

$$L[f^n(x)] = s^n Lf(x) - s^{n-1} f(0) - s^{n-2} f'(0) - \dots - f^{(n-1)}(0).$$

$$(c) L\left[\int_0^x f(z) dz\right] = \int_0^{\infty} dx e^{-sx} \int_0^x f(z) dz$$

$$= \int_0^{\infty} dz \int_z^{\infty} dx e^{-sx} f(z) = \frac{1}{s} \int_0^{\infty} dz e^{-sz} f(z) = s^{-1} L[f(x)]$$

The inverse Laplace transform given by Eq. 9.9 (a) can also be written as $f(x) = \frac{1}{2\pi i} \int_c g(s) e^{sx} ds$, where the contour c for $x > 0$ (shown in Fig. 9.1)

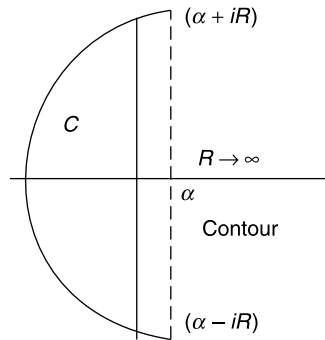


Fig. 9.1

encloses a region that lies to the left of all the singularities of $g(s)$. This ensures that $f(x) = 0$ for $x < 0$ since for $x < 0$ we have to choose a contour that is closed to the right of the line $\text{Re } s = \alpha$, (in order that the contribution of the semicircle goes to zero), thereby making the contour integral go to zero because of Cauchy's theorem of residues. Thus by applying Cauchy's residue theorem we can obtain that $f(x) = \text{sum of the residues at the poles enclosed by the contour } C$ shown in Fig. 9.1.

[If the student finds this approach difficult, he may consult the table of Laplace transforms and its inverse].

9.5.1 Convolution Integral

Defining the convolution of the functions ϕ and ψ by

$$(\phi * \psi)(x) = \int_0^x \phi(y) \psi(x-y) dy \text{ we obtain,}$$

$$\begin{aligned} L[(\phi * \psi)] &= \int_0^\infty dx e^{-sx} \int_0^x \phi(y) \psi(x-y) dy \\ &= \int_0^\infty dx e^{-sx} \int_0^\infty \phi(y) \psi(x-y) dy \\ &= \int_0^\infty e^{-sy} \phi(y) dy \int_0^\infty e^{-s(x-y)} \psi(x-y) dx \end{aligned}$$

$$\text{Putting } x = y + z \text{ we get, } L[\phi * \psi] = \int_0^\infty e^{-sy} \phi(y) dy \int_0^\infty e^{-sz} \psi(z) dz$$

$$\text{or, } L[\phi * \psi] = [L\phi](s) [L\psi](s) \dots \quad (9.11)$$

This is sometimes used to calculate the inverse Laplace transform of the product of two Laplace transform.

SOLVED PROBLEMS

Problem 3

Expand $f(x) = e^x$ for $0 < x < 1$ as a Fourier cosine series of the form $\sum_n a_n \cos n\pi x$.

Solution: We extend the function e^x as an even function of period 2 by writing $f(x) = e^{-x}$ for $-1 < x < 0$.

$$a_0 = \frac{1}{2} \int_{-1}^1 f(x) dx = \frac{1}{2} \left[\int_{-1}^0 e^{-x} dx + \int_0^1 e^x dx \right] = \frac{1}{2} \left[-e^{-x} \Big|_{-1}^0 + e^x \Big|_0^1 \right] = [e - 1]$$

$$a_n = \int_{-1}^1 f(x) \cos n\pi x dx = \int_{-1}^0 e^{-x} \cos n\pi x dx + \int_0^1 e^x \cos n\pi x dx$$

$$\text{or, } a_n = 2 \int_0^1 e^x \cos n\pi x dx = \left(\frac{2n\pi}{1 + n^2 \pi^2} \right) [1 - (-1)^n e]$$

$$\text{or, } e^x = (e - 1) + \sum_{n=1}^{\infty} \left(\frac{2n\pi}{1 + n^2 \pi^2} \right) [1 - (-1)^n e] \cos n\pi x.$$

Problem 4

Find the Fourier transform of

- (a) $f(t) = 1$ for $t > 0$ and $= -1$ for $t < 0$, and vanishes for $|t| > 1$
 (b) $f(t) = v(t) \cos w_0 t$.

Solution:

$$(a) \text{ Fourier transform of } f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-iwt} dt = F(w) \text{ (say)}$$

$$\text{or } F(w) = \frac{1}{2\pi} \left[- \int_{-\infty}^0 e^{-iwt} dt + \int_0^{\infty} e^{-iwt} dt \right] = \frac{-i}{\pi} \int_0^{\infty} \sin wt dt$$

$$\text{or, } F(w) = \frac{-i}{\pi} \int_0^1 \sin wt dt = \frac{i}{\pi w} [\cos w - 1]$$

$$(b) F(w) = \frac{i}{2\pi} \int_{-\infty}^{\infty} v(t) \cos w_0 t \cdot e^{-iwt} dt$$

$$= \frac{1}{4\pi} \int_{-\infty}^{\infty} v(t) [e^{i(w_0 - w)t} + e^{-i(w + w_0)t}] dt$$

$$= \frac{1}{2} [V(w - w_0) + (w + w_0)], \text{ where } V(w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} v(t) e^{-iwt} dt$$

Problem 5Find the Laplace transform of $J_o(x)$ from its integral representation

Solution: $J_o(x) = \frac{1}{\pi} \int_0^{\pi} \cos(x \sin \theta) d\theta$

Laplace transform of $J_o(x) = \frac{1}{\pi} \int_0^{\infty} dx \int_0^{\pi} e^{-sx} \cos(x \sin \theta) d\theta$

or,

$$\begin{aligned}
 L J_o(x) &= \frac{1}{\pi} \int_0^{\pi} d\theta \int_0^{\infty} dx [e^{-sx} \cos(x \sin \theta)] \\
 &= \frac{1}{\pi} \int_0^{\pi} \frac{s}{s^2 + \sin^2 \theta} d\theta \text{ [on carrying out the } x\text{-integral]} \\
 &= \frac{2s}{\pi} \cdot \int_0^{\pi/2} \frac{\sec^2 \theta}{s^2 \sec^2 \theta + \tan^2 \theta} d\theta \text{ and putting } \tan \theta = z \\
 &= \frac{2s}{\pi} \int_0^{\infty} \frac{dz}{z^2 (1 + s^2) + s^2} = \frac{s}{\pi} \cdot \frac{\pi}{s \sqrt{s^2 + 1}} = \frac{1}{\sqrt{s^2 + 1}}.
 \end{aligned}$$

Problem 6Find the inverse Laplace transform of $\frac{(\alpha s^2 + \beta s + \gamma)}{(s^2 + 4)(s^2 + 9)}$, α, β being constants.

Solution: The inverse Laplace transform is the sum of the residues of $e^{st} \left[\frac{\alpha s^2 + \beta s + \gamma}{(s^2 + 4)(s^2 + 9)} \right]$ at

$s = \pm 2i, s = \pm 3i$ [the contour being chosen so that the poles at $s = \pm 2i$ and $s = \pm 3i$ lie to the right of the region enclosed by the contour].

Or, the inverse Laplace transform is thus equal to

$$\begin{aligned}
 &\lim_{s \rightarrow 2i} \left[\frac{(\gamma + 2i\beta - 4\alpha) e^{st}}{(s + 2i)(s^2 + 9)} \right] + \lim_{s \rightarrow -2i} \left[\frac{(\gamma - 2i\beta - 4\alpha) e^{st}}{(s - 2i)(s^2 + 9)} \right] \\
 &+ \lim_{s \rightarrow 3i} \left[\frac{(\gamma + 3i\beta - 9\alpha) e^{st}}{(s^2 + 4)(s + 3i)} \right] + \lim_{s \rightarrow -3i} \left[\frac{(\gamma - 3i\beta - 9\alpha) e^{st}}{(s^2 + 4)(s - 3i)} \right] \\
 &= \left[\frac{(\gamma + 2i\beta - 4\alpha) e^{2it}}{20i} - \frac{(\gamma + 2i\beta - 4\alpha) e^{2it}}{20i} - \frac{(\gamma + 3i\beta - 9\alpha) e^{3it}}{30i} + \frac{(\gamma - 3i\beta - 9\alpha) e^{-3it}}{30i} \right] \\
 &+ \left[\frac{1}{10} (\gamma - 4\alpha) \sin 2t + \frac{\beta \cos 2t}{5} - \frac{(\gamma - 9\alpha)}{15} \sin 3t - \frac{\beta}{5} \cos 3t \right]
 \end{aligned}$$

$$= \left[\frac{(\gamma - 4\alpha)}{10} \sin 2t - \frac{(\gamma - 9\alpha)}{15} \sin 3t + \frac{\beta}{5} (\cos 2t - \cos 3t) \right]$$

[The student can also do this sum by breaking up the function into partial sums and looking at the table].

Problem 7

Solve the differential equation of a series LCR circuit with an externally impressed e. m. f. $E_o \sin wt$ by Laplace transform method. (given that $(q)_{t=0} = (q')_{t=0} = 0$)

Solution: The equation of e. m. f. is,

$$L_o \frac{di}{dt} + R_o i + \frac{q}{c_o} = E_o \sin wt.$$

Putting $i = \frac{dq}{dt}$ we get, $L_o \frac{d^2q}{dt^2} + R_o \frac{dq}{dt} + \frac{q}{c_o} = E_o \sin wt.$

Taking the Laplace transform of both the sides we obtain,

$$L(q) \left(\frac{1}{c_o} + sR_o + L_o s^2 \right) = \frac{E_o w}{w^2 + s^2}, \text{ because the Laplace transform of}$$

$$\begin{aligned} \sin wt &= \int_0^\infty e^{-st} \sin wt \, dt = \int_0^\infty e^{-st} \left[\frac{e^{iwt} - e^{-iwt}}{2i} \right] \\ &= -\frac{1}{2i} \left[\frac{1}{iw - s} + \frac{1}{iw + s} \right] = \frac{w}{w^2 + s^2} \text{ and } (q)_o = (q')_o = 0. \end{aligned}$$

or,

$$L(q) = \frac{E_o w}{(w^2 + s^2) \left(\frac{1}{c_o} + sR_o + s^2 L_o \right)}$$

or,

$$q = \text{Inverse Laplace transform of } \frac{E_o w}{(w^2 + s^2) \left(\frac{1}{c_o} + sR_o + s^2 L_o \right)}.$$

By choosing a suitable contour, as stated in the text, we obtain by calculating the residues that

$$\begin{aligned} q = E_o w \left[\frac{e^{iwt}}{(2iw) \left(\frac{1}{c_o} + iwR_o - w^2 L_o \right)} - \frac{e^{-iwt}}{(2iw) \left(\frac{1}{c_o} - w^2 L_o - iwR_o \right)} + \frac{e^{\alpha_1 t}}{L_o (\alpha_1^2 + w^2) (\alpha_1 - \alpha_2)} \right. \\ \left. + \frac{e^{\alpha_2 t}}{L_o (\alpha_2^2 + w^2) (\alpha_2 - \alpha_1)} \right], \end{aligned}$$

where α_1 and α_2 are the roots of $\alpha^2 + \frac{R_o}{L_o} \alpha + \frac{1}{L_o C_o} = 0$.

The last two terms contribute terms to q that decay as $e - \frac{R_o}{2L_o} t$ and are discarded if we are interested only in the steady state charge and current flowing through the circuit.

Thus at steady state,

$$\begin{aligned}
 q &= \frac{E_o}{2i} \left[\frac{\left(\frac{1}{C_o} - w^2 L_o - iw R_o \right) e^{iwt} - \left(\frac{1}{C} + iw R_o - w^2 L_o \right) e^{-iwt}}{\left(\frac{1}{C} - w^2 L_o \right)^2 + w^2 R_o^2} \right] \\
 &= E_o \left[\frac{\left(\frac{1}{C} - w^2 L_o \right) \sin wt - w R_o \cos wt}{\left(\frac{1}{C} - w^2 L_o \right)^2 + w^2 R_o^2} \right] \\
 &= E_o \frac{\sin (wt - \alpha)}{\left[\left(\frac{1}{C} - w^2 L_o \right)^2 + w^2 R_o^2 \right]^{\frac{1}{2}}}, \text{ where } \tan \alpha = \frac{R_o w}{\left(\frac{1}{C} - w^2 L_o \right)}
 \end{aligned}$$

$$\text{The current } i = \frac{dq}{dt} = \frac{w E_o \cos (wt - \alpha)}{\left[\left(\frac{1}{C} - w^2 L_o \right)^2 + w^2 R_o^2 \right]^{\frac{1}{2}}}$$

[The student can check the result by comparing it with that obtained by the other methods of solving second order linear differential equations with constant coefficients. Also, he can compute the inverse Laplace transform by consulting the table instead of calculating the residues].

SUMMARY

- This chapter deals first with the Fourier series expansion of a function in terms of the orthogonal set of trigonometric function in a suitable interval. The sufficient conditions known as Dirichlet's condition, for expansion are the piece wise continuity, the boundedness and the existence of a finite number of extrema as well as of finite discontinuities of the function in the interval. Thus the infinite series given

by $a_o + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L} + \sum_n b_n \sin \frac{n\pi x}{L}$ converges to $\frac{1}{2} [f(x+o) + f(x-o)]$, which equals $f(x)$ at

the point of continuity for every x in the interval $(-L, L)$. Since the series has a periodicity of $2L$, the function $f(x)$ can be periodically extended beyond the interval meaning thereby that $f(x) = f(x + 2L)$. The

coefficients are given by $a_o = \frac{1}{2l} \int_{-L}^L f(x) dx$, $a_n = \frac{1}{l} \int_{-L}^L f(x) \cos \frac{n\pi x}{l} dx$, $b_n = \frac{1}{l} \int_{-L}^L f(x) \sin \frac{n\pi x}{l} dx$.

Since the period is $2l$, the limits of the integral can be from x_o to $x_o + 2l$, x_o being arbitrary. For the expansion of the function in terms of orthogonal functions of a different period, the limits of integration can be accordingly altered. For a function defined over half of the interval we can extend it to the other half as an odd or an even function.

- The Fourier transform is the extension of the Fourier series to the domain of non-periodic functions. The

Fourier transform of a function $f(x)$ is given by $g(k) = \int_{-\infty}^{\infty} f(x) e^{-ikx} dx$, where $f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(k) e^{-ikx} dk$.

Thus $f(x)$ can be written as $f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x') e^{ik(x-x')} dx' dk$.

Parseval's relation gives $\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |g(k)|^2 dk$, where we symmetrise the definitions of the

Fourier transforms by writing $g(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx$ and $f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g^*(k) e^{-ikx} dk$.

The corresponding relation for the inner product in this case is,

$$\int_{-\infty}^{\infty} f(x) g^*(x) dx = \int_{-\infty}^{\infty} f(k) g^*(k) dk, \left[g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d(k) e^{ikx} dk \right].$$

The corresponding relation in the case of Fourier series is

$$\int_{-\pi}^{\pi} |f(x)|^2 dx = \pi \left[2a_o^2 + \sum_n (a_n^2 + b_n^2) \right].$$

The convolution of two functions ϕ and ψ defined as $\phi * \psi$ is the inverse Fourier transform of the product of the Fourier transforms of the individual functions.

The Laplace transform of a function $f(t)$ which is zero for $t < 0$ is obtained from the formula of the Fourier transform after a slight algebraic manipulation. The Laplace transform $g(s)$ of the function $f(t)$

is given by $g(s) = \int_0^{\infty} e^{-st} f(t) dt$ ($\text{Res} > M$ for convergence where M denotes the maximum exponential

growth rate of $f(t)$ given by $f(t) < e^{Mt}$)

The inverse Laplace transform is given by $f(t) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} g(s) e^{st} ds$, where σ is the real part of s .

The inverse Laplace transform is found out either by contour integration (see text) or from the table of Laplace transforms. The Laplace transform of the convolution of two functions $\phi(t)$ and $\psi(t)$ is the product of their Laplace transforms.

MULTIPLE CHOICE QUESTIONS

1. $\sin mx$ ($m = \text{integer}$) when expanded in a Fourier series has
 - (a) only one Fourier components
 - (b) an infinite number of component
 - (c) the number of components depends on the interval, chosen
 - (d) the number of components does not depend on the interval
2. If $f(x) = a_0 + \sum_n a_n \cos nx + \sum_n b_n \sin nx$ in an interval $(-\pi, \pi)$, the average value of $[f(x)]^2$ is
 - (a) $a_0^2 + \frac{1}{2} \sum_n (a_n^2 + b_n^2)$
 - (b) $\frac{1}{2} \sum_n (a_n^2 + b_n^2)$
 - (c) a_0^2
 - (d) None of these
3. The statement 'tan x cannot be expanded in a Fourier series' is
 - (a) always true
 - (b) always false
 - (c) true if the interval includes the point of infinite discontinuity
 - (d) true if the interval is $(-\pi, \pi)$
4. $\sum_n \frac{1}{(2n+1)^2}$ is equal to
 - (a) $\pi^2/6$
 - (b) $\frac{\pi^2}{8}$
 - (c) $\frac{\pi^2}{3}$
 - (d) none of these

(**Hint:** Take the expansion in the interval $-T < t < T$ of a function $f(t) = -1$ for $-T < t < 0$ and $= 1$ for $0 < t < T$).

5. If $f(t) = f(-t)$, the Fourier series of $f(t)$ contains
 - (a) only constant and the cosine terms
 - (b) only sine terms
 - (c) both the sine and the cosine terms or only cosine terms depending on the interval
 - (d) none of these
6. The Fourier transform of the equation $\frac{d^2 y}{dx^2} + f(x)y = 0$ is,
 - (a) $-k^2 (\tilde{y})_k + \tilde{f}(k) \tilde{y}(k) = 0$,
 - (b) $-k^2 (\tilde{y})_k + \int f(k') \hat{y}(k-k') dk = 0$
 - (c) $-k^2 (\tilde{y})_k + \tilde{f}(k') (\tilde{y})_{k-k'} = 0$
 - (d) $-(k-k')^2 (\tilde{y})_{k-k'} \tilde{f}(k') (\hat{y})_{k'} = 0$
7. The Fourier cosine transform of $e^{-|x|}$ is
 - (a) $\sqrt{\frac{2}{\pi}} \cdot \frac{1}{(1+n)}$
 - (b) $\sqrt{\frac{2}{\pi}} \cdot \left(\frac{1}{1+n^2} \right)$
 - (c) $\sqrt{\frac{2}{\pi}} (1+n)$
 - (d) $\sqrt{\frac{2}{\pi}} (1+n^2)$
8. The convolution of $F(t)$ with $\delta(t-t_0)$ is
 - (a) $F(t_0)$
 - (b) $F(t-t_0)$
 - (c) $F(t+t_0)$
 - (d) $F(t)$

9. The Fourier transform of the convolution of $F(t)$ with $\delta(t - t_0)$ is
 (a) $\bar{F}(w)e^{-iwt_0}$ (b) $\frac{\bar{F}(w)}{e^{-iwt_0}}$ (c) $\bar{F}(w)$ (d) none of these
10. Using Parseval's theorem we can show that $\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx$ is equal to
 (a) $\frac{\pi}{2}$ (b) π (c) 2π (d) $\frac{\pi}{4}$
[Hint:- Take $f(t) = 1$ for $|t| < 1$ and $= 0$ otherwise]
11. The Laplace transform of $f(t)$ being $F(s)$, that of $f(at)$ is
 (a) $F(as)$ (b) $\frac{1}{a}F(s/a)$ (c) $\frac{1}{a}F(s)$ (d) $\frac{1}{a}F(as)$
12. For which of the following functions do the Fourier transforms not exist?
 (a) $e^{-|x|}$ (b) e^{-x}
 (c) e^{-x^2} (d) e^{x^2}
 (e) xe^{-x^2}
13. The Laplace transform of $t^2 \sin at$ is,
 (a) $\frac{2a(3s^2 - a^2)}{(s^2 + a^2)^3}$ (b) $\frac{s^2 a}{s^2 + a^2}$ (c) $\frac{6as^2}{(s^2 + a^2)^3}$ (d) none of these
14. The Laplace transform of $f(t - aT)$ is,
 (a) $e^{-sat} F(s)$ (b) $e^{saT} F(s)$ (c) $\frac{F(s)}{1 - e^{-snT}}$ (d) $F(s - as)$
15. How many Fourier components does $\cos 4x$ contain when expanded in $(-\pi, \pi)$
 (a) Infinite (b) 4 (c) 3 (d) 1
16. The inverse Laplace transform of $\frac{A}{s^2 + 3s + 2}$ ($A = \text{constant}$) is,
 (a) $A(e^{-t} - e^{-2t})$ (b) $A(e^{-t} + 2e^{-2t})$
 (c) $A/3(e^{-t} + e^{-2t})$ (d) $A(2e^{-t} - e^{-2t})$

Answers to MCQs

1. (c)	2. (a)	3. (c)	4. (a)	5. (c)	6. (b)	7. (b)
8. (b)	9. (a)	10. (b)	11. (b)	12. (b), (d)	13. (a)	14. (b)
15. (d)	16. (a).					

EXERCISE

1. Find the Fourier series expansion of
 (a) $f(x) = x^2$ of period $2l$,
 (b) $f(t) = \cos wt$ ($0 < t < 1$) of period $T = 1$, where $wT = 2\pi$. Apply Parseval's theorem in (a) and obtain the result.

2. Expand

(i) $f(x) = \cos \frac{\pi x}{L}$ in $0 < x < L$ as a Fourier sine series

(ii) $f(x) = x$ in $0 < x < L$ as a Fourier cosine series.

3. Find the Fourier cosine transform of $e^{-\alpha|x|}$, $e^{-\alpha x^2}$,

4. (a) Find the Fourier transform of a square pulse $f(x) = 1$ for $-L < x < L$ and $f(x) = 0$ otherwise. Verify the uncertainty relation.

(b) Find the Fourier transform of a triangular pulse $f(x) = 1 + x$ for $-1 < x < 0$, $= 1 - x$ for $0 < x < 1$ and $= 0$ otherwise. Apply Parseval's theorem in this case and see what you obtain.

5. Find the Laplace transforms of $c \cos wt$ ($c = \text{constant}$), $\delta(t^2 - a^2)$ ($a > 0$).

6. Find the inverse Laplace transforms of

(a) $\frac{1}{(s^2 - a^2)(s^2 - b^2)}$, (b) $\frac{1}{(s^2 - 4)(s - 3)}$, (c) $\frac{e^{-\lambda s}}{s - 2}$ ($\lambda = \text{constant}$)

7. Solve the equation $R \frac{dQ}{dt} + \frac{Q}{C} = E_o \sin wt$, given that $(q)_o = 0$, by the method of Laplace transform.

8. Solve the coupled differential equations

$$\ddot{x} = -kx - \alpha y, \quad \ddot{y} = -ky - \beta x \text{ by Laplace transform given that}$$

$$(x)_o = A, (y)_o = 0, \dot{x} = \dot{y} = 0 \text{ at } t = 0.$$

9. Solve by Fourier transform the differential equation of a damped simple harmonic oscillator subjected to an external periodic force $F \sin wt$.

[Hint: The equation of motion is, $m \frac{d^2 x}{dt^2} + k \frac{dx}{dt} + fx = F \sin wt$.

Take the Fourier transform of both the sides, assume x and its derivative to go to zero as $t \rightarrow +\infty$.

The equation will lead to an expression for the Fourier transform of x . When we take the inverse Fourier transform we obtain the steady state solution].

Hints and Answers to Exercise Questions

1. (a) $f(x) = \frac{l^2}{3} + \frac{4l^2}{\pi^2} \sum_n \frac{(-1)^n}{n^2} \cdot \sin \frac{n\pi x}{L} \quad (-L < x < L)$

(b) $f(t) = \frac{\sin w}{w} + \sum_n \left(\frac{2w}{w^2 - 4n^2\pi^2} \right) \cos 2\pi nt + \sum_n \frac{4\pi n(1 - \cos w)}{(4n^2\pi^2 - w^2)} \sin 2\pi nt$

2. (i) $f(x) = \frac{2}{\pi} \sum \left(\frac{n}{n^2 - 1} \right) (1 + (-1)^n) \sin \frac{n\pi x}{L},$

$$(ii) f(x) = \frac{L}{2} + \sum \frac{l}{n^2 \pi^2} \{(-1)^n - 1\} \cos \frac{n\pi x}{L}.$$

$$3. (i) \sqrt{\frac{2}{\pi}} \cdot e^{-\alpha x} (K^2 + \alpha^2)^{-\frac{1}{2}} \sin \left(Kx - \tan^{-1} \frac{\alpha}{K} \right), \quad (ii) \frac{1}{\sqrt{2}} e^{-\frac{K^2}{4\alpha}}$$

$$4. (a) \bar{f}(K) = \frac{\sin Kl}{K\pi}, \Delta K = \frac{2\pi}{L}, \Delta x = 2l$$

$$(b) \frac{2}{\pi} \frac{\sin^2 K}{K^2}, \int_{-\infty}^{\infty} \frac{\sin^4 x}{x^4} dx = \frac{2\pi}{3}$$

$$5. \frac{c}{5}, \frac{\cosh(sa)}{a}.$$

$$6. (a) \frac{1}{(a^2 - b^2)} \left[\frac{\sinh(at)}{a} - \frac{\sinh(bt)}{b} \right]$$

$$(b) \frac{1}{20} (e^{-2t} - 5e^{2t}) + \frac{1}{5} e^{3t}, \quad (c) e^{2(t-\lambda)}$$

$$7. I = 0 \text{ for } t < 0 \text{ and } = \frac{E_0}{\sqrt{R^2 + \frac{1}{w^2 c^2}}} \sin \left(wt - \tan^{-1} \frac{1}{wCR} \right)$$

$$8. \text{ For } \alpha\beta < K^2 \text{ and } w_1^2 = \sqrt{\alpha\beta} + K, w_2^2 = -\sqrt{\alpha\beta} + K,$$

$$x = A \sin w_1 t + B \cos w_1 t + C \sin w_2 t + D \cos w_2 t, \text{ for } t > 0$$

$$y = \{(w_1^2 - K)(A \sin w_1 t + B \cos w_1 t) + (w_2^2 - K)(C \sin w_2 t + D \cos w_2 t)\} / \alpha.$$

$$\text{and } x, y = 0 \text{ for } t < 0.$$

$$\text{or } \alpha\beta < K^2, x = A \sin w_1 t + B \cos w_1 t + Ce^{w_2 t} + De^{-w_2 t} \text{ and}$$

$$y = -\frac{1}{\alpha} (\ddot{X} + Kx) = -\frac{1}{\alpha} [(w_2^2 + K)Ce^{w_2 t} + (w_2^2 + K)De^{-w_2 t}$$

$$+ A(K - w_1^2) \sin w_1 t + B(K - w_2^2) \cos w_1 t] \text{ for } t > 0, \text{ and } x, y = 0 \text{ for } t < 0.$$

Now apply the initial conditions to find out A, B, C, D ,

Partial Differential Equations

CHAPTER

10

10.1 INTRODUCTION

Partial differential equations occur almost everywhere in Physics. This chapter attempts to highlight only a few methods of solving partial differential equations subject to a set of boundary conditions. Though the stress will be primarily on the solutions of second order linear partial differential equations, the first order equations will be briefly touched upon. The method of characteristics along with classification of the partial differential equations based on the types of characteristics will be discussed.

10.2 FIRST ORDER PARTIAL DIFFERENTIAL EQUATIONS

A first order partial differential equation in x and y can be written as,

$$A(x, y) \frac{\partial f}{\partial x} + B(x, y) \frac{\partial f}{\partial y} = C(x, y) \quad (10.1)$$

Let the boundary condition be specified along a boundary curve represented parametrically by $x = x(s)$, $y = y(s)$, s being the arc length measured along the perimeter of the curve.

Thus we can write for the boundary condition $f(x, y) = g(s)$ on the curve so that $\frac{df}{ds} = \frac{dg}{ds}$. Hence we can represent the directional derivative by

$$\frac{\partial f}{\partial x} \cdot \frac{dx}{ds} + \frac{\partial f}{\partial y} \cdot \frac{dy}{ds} = \frac{dg}{ds} \quad (10.2)$$

The determination of $f(x, y)$ at every point within a region bounded by the boundary curve requires finding out uniquely all order partial derivatives everywhere in the enclosed region. For the existence of unique finite values of $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ we obtain from the equations (10.1) and (10.2) that the determinant of the coefficients

$\begin{vmatrix} A & B \\ \frac{dx}{ds} & \frac{dy}{ds} \end{vmatrix}$ must not vanish. The existence of higher order derivatives also requires precisely the same condition. The vanishing of the determinant $\begin{vmatrix} A & B \\ \frac{dx}{ds} & \frac{dy}{ds} \end{vmatrix}$ defines the equation of the characteristic, i.e., $A \frac{dy}{ds} - B \frac{dx}{ds} = 0$ along which the boundary curve must not lie for the existence of the solution.

The solution of Eq. (10.1) for the inhomogeneous term $C = 0$ is clearly $f(x, y) = \text{constant}$, where the constant has to be found out from the integral curve of the equation $\frac{dx}{A} = \frac{dy}{B}$.

[Because if $f(x, y) = \text{constant}$ be the solution of the equation $\frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0$ (10.3), the simultaneous solution of Eqs. (10.1) and (10.3) demands that $\frac{dx}{A} = \frac{dy}{B}$]. If we compare the above equation with the equation of the characteristic, we see that along the characteristics $f(x, y) = \text{constant}$ and, hence, it can be concluded that the characteristics determine the integral curves of the equation. Thus Canchy conditions (specification of $f(x, y)$ and the derivative along the boundary) fix up uniquely the solution.

Example Solve the following equations:

(1) $\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} = 0$.

The equation of the characteristic is $\frac{dt}{1} = \frac{dx}{v}$ or, $x - vt = \text{constant}$. So the solution of the equation is given by $f = \phi(x - vt)$, where ϕ is any arbitrary function to be determined from the initial conditions.

(2) $x^2 \frac{\partial f}{\partial x} - y^2 \frac{\partial f}{\partial y} = 0$ with the boundary condition that for $x = 1$ the function f takes the value $\frac{1}{y} + 3$ for the region between $y = 2$ and $y = 5$.

The equation of the characteristic is, $\frac{dx}{x^2} = -\frac{dy}{y^2}$ giving $\frac{1}{x} + \frac{1}{y} = \text{constant}$ as the solution. The general solution of the equation is thus $f = \phi\left(\frac{1}{x} + \frac{1}{y}\right)$. Since for $x = 1$, $\phi = \frac{1}{y} + 3$ for all y in the specified region, the solution can be written as $f(x, y) = \frac{1}{x} + \frac{1}{y} + 2$.

(3) $xf \frac{\partial f}{\partial x} - yf \frac{\partial f}{\partial y} = y^2 - x^2$ with the boundary condition that the surface passes through the line $x + y = 1$, f being zero in the region of intersection.

The equation of the integral surface is given by $\frac{dx}{xf} = -\frac{dy}{yf} = \frac{df}{y^2 - x^2}$.

The equation can be written in terms of two independent equations $xdy + ydx = 0$ giving $xy = \text{constant}$ and $x dx + y dy + f df = 0$ leading to $x^2 + y^2 + f^2 = \text{constant}$. The general solution of the equation is $\phi(xy, x^2 + y^2 + f^2) = \text{constant}$. Since the integral surface passes through the line $x + y = 1$, f being zero there, we obtain on putting $x = t$, $y = 1 - t$, $c_1 = t(1 - t)$ and $t^2 + (1 - t)^2 = c_2$, c_1 , and c_2 being constants. Thus we obtain $c_2 = 1 - 2c_1$ and hence the equation of the integral surface is $x^2 + y^2 + 2xy + f^2 = 1$.

10.3 SECOND ORDER PARTIAL DIFFERENTIAL EQUATION

Before discussing the method of solving a few selected partial differential equations with given boundary values this section starts with a brief survey of the general theory of second order linear partial differential equations. Here, we will dwell briefly on the role of the characteristics, the classification of different types of second order partial differential equations and the boundary conditions appropriate for the respective equations.

The second-order partial differential equation is of the form

$$A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial y^2} + C \frac{\partial^2 u}{\partial x \partial y} = F\left(x, y, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) \dots \quad (10.3)$$

where A, B, C are, in general, functions of x, y and the right hand side is a function of the arguments inside the bracket. The objective is to find out the solution $u(x, y)$ valid in a certain region satisfying the boundary conditions specified for the problem. The problem of determining $u(x, y)$ in a certain region starting from a specified function on the boundary curve depends on that of finding out the derivatives of all orders in the region. Hence, we deduce below the conditions of the derivatives being finite in the specified region.

Let the function $u(x, y)$ and its derivatives along and normal to the boundary curve at every point be given so that if s be the parameter signifying the curve, we can write,

$\frac{du}{ds} = \frac{\partial u}{\partial x} \frac{dx}{ds} + \frac{\partial u}{\partial y} \frac{dy}{ds}$ and $\frac{du}{dn}(s) = -\frac{\partial u}{\partial x} \cdot \frac{dy}{ds} + \frac{\partial u}{\partial y} \cdot \frac{dx}{ds}$, where the unit normal to the boundary curve has components $-\frac{dy}{ds}$ and $\frac{dx}{ds}$ respectively.

The two equations enable us to find out $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ respectively.

$$\text{Now,} \quad \frac{d}{ds} \left(\frac{\partial u}{\partial x} \right) = \frac{\partial^2 u}{\partial x^2} \cdot \frac{dx}{ds} + \frac{\partial^2 u}{\partial x \partial y} \frac{dy}{ds} \quad (10.4)$$

$$\frac{d}{ds} \left(\frac{\partial u}{\partial y} \right) = \frac{\partial^2 u}{\partial x \partial y} \frac{dx}{ds} + \frac{\partial^2 u}{\partial y^2} \frac{dy}{ds} \quad (10.5)$$

These equations along with the equation (10.4) enable us to find out the three second order derivatives

$\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial y^2}$ and $\frac{\partial^2 u}{\partial x \partial y}$ respectively unless the coefficient determinant given by

$$\begin{vmatrix} \frac{dx}{ds} & \frac{dy}{ds} & 0 \\ 0 & \frac{dx}{ds} & \frac{dy}{ds} \\ A & C & B \end{vmatrix} = 0, \text{ i.e., } A \left(\frac{dy}{ds} \right)^2 + B \left(\frac{dx}{ds} \right)^2 - C \frac{dx}{ds} \cdot \frac{dy}{ds} = 0 \quad (10.6)$$

The curves for which the above equation is satisfied are the characteristic curves or the characteristics of the equation. Hence the Canchy boundary conditions (specification of u and the normal componenty of the derivative of u at every point of the boundary) lead to a solution of the partial differential equation in a certain region unless the boundary curve is tangential to the characteristics anywhere, (similar to that in the first order equations). Continuing the calculations shown above we can conclude that the same condition is required for the existence of all order derivatives of the function $u(x, y)$.

From Eq. (10.7) we see that $\frac{dx}{ds}, \frac{dy}{ds}$ are real for the condition $C^2 > 4AB$. Thus the characteristics are real curves and the corresponding equation is called a hyperbolic differential equation. For $C^2 < 4AB$ the characteristics are imaginary curves the equation being called elliptic. For $C^2 = 4AB$, the characteristics are two straight lines the corresponding equation being called parabolic.

Example

(a) *Wave equation in one dimension*

$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0$, $A = 1$, $B = -C^2$, $C = 0$, thus $C^2 > 4AB$ and hence the equation is a hyperbolic differential equation.

(b) *Laplace's equation in two dimensions*

$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$, $A = 1$, $B = 1$, $C = 0$, i.e., $C^2 < 4AB$ and the equation is elliptic.

(c) *Heat conduction equation in one dimension.*

$\frac{\partial u}{\partial t} = K \frac{\partial^2 u}{\partial x^2}$ ($K = \text{constant}$). Here $A = -K$, $B = C = 0$. Thus $AB = C^2 = 0$

and hence the equation is parabolic.

Let us now discuss the boundary conditions required for the unique solution of each of the three classes of equations specified above.

Let us consider the wave equation $\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0$. The characteristics are found out by solving Eq. (10.7), where $A = 1$, $B = -c^2$, $C = 0$. The characteristics are thus straight lines determined by the equations $x - ct = \text{constant}$, and $x + ct = \text{constant}$. Transforming to the variables $\xi = x - ct$ and $\eta = x + ct$ respectively, we obtain the canonical form $\frac{\partial^2 u}{\partial \xi \partial \eta} = 0$ giving as the solution $u = f_1(\xi) + f_2(\eta)$, where f_1 and f_2 are two functions determined by the boundary conditions. Using Canchy conditions we have, $u(x, t = 0) = f_1(x) + f_2(x)$ and

$\frac{\partial u}{\partial t}(x, t=0) = c(f_2'(x) - f_1'(x))$. Hence we obtain, $f_1(x) = \frac{1}{2}u(x) - \frac{1}{2c} \int \frac{\partial u}{\partial t} dx$ and $f_2(x) = \frac{1}{2}u(x) + \frac{1}{2c} \int \frac{\partial u}{\partial t} dx$.

Specifying $f_1(x)$ along a line that is not tangent to the characteristic $\xi = x - ct = \text{constant}$ enables us to find out $f_1(\xi)$ along all the characteristics, that intersect the line. Similarly, specifying $f_2(x)$ determines $f_2(\eta)$ along the characteristics $\eta = x + ct = \text{constant}$ that intersect the line. Thus, both $f_1(\xi)$ and $f_2(\eta)$ being determined in the region bounded by the characteristics, u is known everywhere within the region. Thus for the hyperbolic differential equation Cauchy's boundary conditions in an open region are sufficient to determine $u(x, t)$ everywhere within the region. For an elliptic differential equation the specification of Dirichlet conditions (i.e., the function) or Neumann boundary conditions (i.e., the normal derivatives) on a closed boundary can be shown to determine uniquely the solution. Cauchy's condition on the boundary will be an over specification. For parabolic differential equations the specification of Dirichlet or Neumann conditions on an open boundary fix up completely the solution in the region enclosed by the boundary.

Note: In three dimensions the wave equation will be of the form $\frac{\partial^2 u}{\partial t^2} - c^2 \nabla^2 u = 0$ and a solution u by analogy with the one-dimensional case will be of the form $u = f_1(\xi) + f_2(\eta)$ where ξ and η are now given by $\xi = \vec{r} \cdot \vec{n} - ct$, $\eta = \vec{r} \cdot \vec{n} + ct$, $\vec{n}$ being the unit vector along the direction of propagation. This is a plane wave solution of the equation, the plane wave front being given by the equation $\vec{r} \cdot \vec{n} = \text{constant}$. However, solutions of other types viz. in the form of spherical or cylindrical waves exist too for this three dimensional wave equation.

10.4 SELECTED PARTIAL DIFFERENTIAL EQUATIONS

In this section we deal with the solution of some selected partial differential equations of the second order by the method of the separation of variables.

1. Waves along a stretched string

When a stretched string tied at both the ends is set into vibration by displacing it perpendicular to the length, transverse waves are set up in the string and being reflected from both the ends give rise to stationary waves. This being the case of the propagation of a wave along the length of the string the equation governing the

propagation of waves is, $\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0$, x denoting the direction of the length of the string. (c = velocity of the wave). (Hyperbolic equation)

Let us attempt a solution of the form $u = X(x) T(t)$, where $X(x)$ and $T(t)$ are hitherto undefined functions of x and t respectively. Substituting for u in the above equation, we get

$$X \frac{d^2 T}{dt^2} = c^2 T \frac{d^2 X}{dx^2}.$$

Dividing by XT , we get
$$\frac{1}{T} \frac{d^2 T}{dt^2} = \frac{c^2}{X} \frac{d^2 X}{dx^2}.$$

Since the left hand side is a function of t while the right hand side depends on x alone each side must separately be equal to a constant. To ensure the wave type solution, X and T must both be oscillatory functions

and hence the constant must be negative. We therefore put $\frac{1}{T} \frac{d^2 T}{dt^2} = \frac{c^2}{X} \frac{d^2 X}{dx^2} = -w^2$, say (w being a constant)

so that we obtain, $T = A \sin wt + B \cos wt$ and $X = C \sin \frac{w}{c} x + D \cos \frac{w}{c} x$, A, B, C and D being constants.

Since the string is tied at $x = 0$ and $x = l$, say, l being the length of the string we obtain, $D = 0$ and $\frac{w}{c} L = n\pi$, $n = 0, \pm 1, \pm 2, \dots$. Thus for a particular n we can write the solution as $u_n = C_n \sin \frac{n\pi x}{L} \left(A_n \sin \frac{n\pi ct}{L} + B_n \cos \frac{n\pi ct}{L} \right)$. Because of the linearity of the equation we can write the general solution using the principle of superposition as, $u = \sum_n \sin \frac{n\pi x}{L} \left(P_n \sin \frac{n\pi ct}{L} + Q_n \cos \frac{n\pi ct}{L} \right)$, where the magnitudes of P_n and Q_n can be fixed up by choosing the initial conditions as shown in the following. (10.7)

Case (a) Suppose the string is excited into vibration by plucking at a point distant a from one end, the amount of displacement being u_o at $x = a$. (Fig. 10.1)

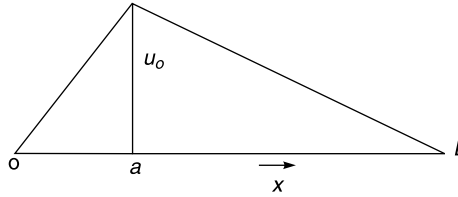


Fig. 10.1

Thus for $0 < x < a$ we can write, $(u)_{t=0} = \frac{u_o x}{a}$ and for $a < x < L$, $(u)_{t=0} = \frac{u_o (L-x)}{L-a}$.

From Eq. (10.8), $(u)_{t=0} = \sum_n Q_n \sin \frac{n\pi x}{L}$ and hence using the orthonormality relation

$$\int_0^L \sin \frac{n\pi x}{L} \sin \frac{m\pi x}{L} dx = \frac{L}{2} \delta_{mn} \text{ we get, } Q_n = \frac{2}{L} \int_0^L (u)_{t=0} \sin \frac{n\pi x}{L} dx.$$

$$\begin{aligned} \text{or, } Q_n &= \frac{2}{L} \left[\int_0^a \frac{u_o x}{a} \sin \frac{n\pi x}{L} dx + \int_a^L \frac{u_o (L-x)}{(L-a)} \sin \frac{n\pi x}{L} dx \right] \\ &= \frac{2u_o L^2}{n^2 \pi^2 a(L-a)} \sin \frac{n\pi a}{L} \dots \end{aligned} \quad (10.8)$$

$$(\dot{u})_{t=0} = \sum_n \frac{n\pi c}{L} P_n \sin \frac{n\pi x}{L} = 0 \text{ for all } x \text{ giving } P_n = 0.$$

$$\text{or, } u = \sum_n Q_n \sin \frac{n\pi x}{L} \cos \frac{n\pi ct}{L}, \text{ where } Q_n \text{ is determined from (10.9)} \quad (10.9)$$

[Of course the initial conditions assumed here are unrealistic. In practice, the plucking is done never at a single point but the region of plucking extends over a certain region. If the extent of the region of plucking be known, we can make a similar calculation with the required modification to obtain $u(x, t)$].

Case (b) Suppose the string is made to vibrate by striking it over a certain region and, thereby, imparted a small initial velocity. Let the initial velocity be given by $(\dot{u})_{t=0} = \frac{p}{2m\epsilon}$, where the region of striking is given by $a - \epsilon \leq x \leq a + \epsilon$, p is the initial impulse or momentum communicated to the string; m the mass per unit length and $(\dot{u})_{t=0} = 0$ for all other values of x .

$(u)_{t=0} = 0$ for all x leading to $Q_n = 0$ for all n .

Thus $u = \sum_n P_n \sin \frac{n\pi x}{L} \cdot \sin \frac{n\pi ct}{L}$ so that $(\dot{u})_{t=0} = \sum_n \frac{n\pi c}{L} P_n \sin \frac{n\pi x}{L}$

$$\text{or, } \frac{n\pi c}{2} P_n = \int_{a-\epsilon}^{a+\epsilon} (\dot{u})_{t=0} \sin \frac{n\pi x}{L} dx \left[\because \int_0^l \sin \frac{n\pi x}{L} \cdot \sin \frac{n\pi x}{L} dx = \frac{L}{2} \delta_{mn} \right]$$

$$\text{or, } P_n = -\frac{2(\dot{u})_{t=0}}{n\pi c} \cdot \frac{L}{n\pi} \left[\cos \frac{n\pi x}{L} \right]_{a-\epsilon}^{a+\epsilon} \dots$$

$$= \frac{2L(\dot{u})_0}{n^2 \pi^2 c} \left[2 \sin \frac{n\pi a}{L} \sin \frac{n\pi \epsilon}{L} \right]$$

$$= \frac{2Lp}{n^2 \pi^2 m c \epsilon} \sin \frac{n\pi a}{L} \sin \frac{n\pi \epsilon}{L} = \frac{2p}{n\pi m c} \sin \frac{n\pi a}{L} \text{ for } \epsilon \rightarrow 0 \quad (10.10)$$

Thus $u = \sum_n P_n \sin \frac{n\pi x}{L} \sin \frac{n\pi ct}{L}$, where P_n is found out from the above (10.11). It can easily be seen that u given by Eqs. (10.9) or (10.10) is of the expected form $u = f_1(ct - x) + f_2(ct + x)$, where f_1 and f_2 are suitable functions as can be determined from u given in the two expressions (10.10) and (10.11). [For example, $\sum_n P_n \sin \frac{n\pi x}{L} \sin \frac{n\pi ct}{L}$ can be written as equal to $\frac{1}{2} \sum_n P_n \left[\cos \frac{n\pi}{L} (x - ct) - \cos \frac{n\pi}{L} (x + ct) \right]$].

2. Vibrations of a membrane

Let us consider the vibrations of a membrane which is free to vibrate at all points except at the boundary, where it is always fixed. The transverse vibrations of a membrane satisfy the two-dimensional wave equation

$\nabla^2 u - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0$. Below we consider the vibrations of a rectangular and a circular membrane respectively.

(i) *Rectangular membrane* The transverse vibrations are determined in this case by the equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0 \quad \dots (10.12),$$

where the boundary conditions are, $u(0, y, t) = u(L_1, y, t) = u(x, 0, t) = u(x, L_2, t) = 0$.

Let $u(x, y, t) = X(x) Y(y) T(t)$ so that we can write,

$$YT \frac{d^2 X}{dx^2} + XT \frac{d^2 Y}{dy^2} - \frac{XY}{c^2} \frac{d^2 T}{dt^2} = 0.$$

Dividing by XYT , we get $\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} = \frac{1}{c^2 T} \frac{d^2 T}{dt^2}$.

Since the l. h. s. is a function of x, y , while the r. h. s. depends on t alone we must put each side equal to a constant so that

$$\frac{1}{c^2 T} \frac{d^2 T}{dt^2} = -\frac{w^2}{c^2}, \quad w \text{ being a constant so that } \frac{d^2 T}{dt^2} + w^2 T = 0 \quad \dots \quad (10.11(A))$$

Similarly separating the x and the y equations we obtain,

$$\frac{1}{X} \frac{d^2 X}{dx^2} = -k_x^2, \quad \frac{1}{Y} \frac{d^2 Y}{dy^2} = -k_y^2, \quad \text{where } k_x^2 = \frac{w^2}{c^2}. \quad (10.11(B))$$

Imposing the boundary condition stated above we can write,

$$X = A \sin k_x x, \quad \text{where } k_x = \frac{n_1 \pi}{L_1}, \quad n_1 = 1, 2, \dots$$

$$Y = B \sin k_y y, \quad \text{where } k_y = \frac{n_2 \pi}{L_2}, \quad n_2 = 1, 2, \dots$$

The solution of Eq. (10.12) gives, $T = C \cos w_{n_1, n_2} t + D \sin w_{n_1, n_2} t$, where $w_{n_1, n_2}^2 = c^2 \pi^2 \left(\frac{n_1^2}{L_1^2} + \frac{n_2^2}{L_2^2} \right)$.

Thus, the general solution combining all the possible modes for the possible values of n_1, n_2 is given by

$$u = \sum_{n_1, n_2} \sin \frac{n_1 \pi}{L_1} x \sin \frac{n_2 \pi}{L_2} y (P_{n_1, n_2} \cos w_{n_1, n_2} t + Q_{n_1, n_2} \sin w_{n_1, n_2} t).$$

Let the initial conditions be $u(x, y, t=0) = f_1(x, y)$ and $\dot{u}(x, y, t=0) = g_1(x, y)$.

We thus obtain, $f_1(x, y) = \sum P_{n_1, n_2} \sin \frac{n_1 \pi}{L_1} x \sin \frac{n_2 \pi}{L_2} y$ and

$$g_1(x, y) = \sum w_{n_1, n_2} Q_{n_1, n_2} \sin \frac{n_1 \pi}{L_1} x \sin \frac{n_2 \pi}{L_2} y$$

or,
$$P_{n_1 n_2} = \frac{4}{L_1 L_2} \int_0^{L_1} \int_0^{L_2} f_1(x, y) \sin \frac{n_1 \pi x}{L_1} \sin \frac{n_2 \pi y}{L_2} dx dy \text{ and}$$

$$Q_{n_1 n_2} = \frac{4}{w_{n_1 n_2} L_1 L_2} \int_0^{L_1} \int_0^{L_2} g_1(x, y) \sin \frac{n_1 \pi x}{L_1} \sin \frac{n_2 \pi y}{L_2} dx dy.$$

The positions of the nodal lines corresponding to the frequency w_{mn} are found out from the zeroes of $\sin \frac{n_1 \pi}{L_1} x$ or $\sin \frac{n_2 \pi}{L_2} y$ respectively.

(ii) *Circular membrane* The transverse vibrations are determined by the same two-dimensional wave equation but using plane polar coordinates for convenience.

Thus, the equation $\nabla^2 u - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0$ becomes

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0$$

Writing $u(r, \theta, t) = X(r) T(\theta)$ we obtain, as before,

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial X}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 X}{\partial \theta^2} + k^2 X = 0, \text{ where } \frac{d^2 T}{dt^2} + k^2 c^2 T = 0$$

Thus, the t -dependence is given by $T = A \cos kct + B \sin kct$.

The (r, θ) equation on writing $X = R(r)$. $F(\theta)$ we get,

$$\frac{1}{rR} \frac{d}{dr} \left(r \frac{dR}{dr} \right) + \frac{1}{r^2 F} \frac{d^2 F}{d\theta^2} = -k^2$$

Thus we can write,
$$\frac{r}{R} \frac{d}{dr} \left(r \frac{dR}{dr} \right) + k^2 r^2 = -\frac{1}{F} \frac{d^2 F}{d\theta^2} = m^2$$

Since the solution is periodic in θ , the separation constant is taken to be positive and m is an integer.

Thus, the r -equation gives,
$$\frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} + \left(k^2 - \frac{m^2}{r^2} \right) R = 0 \dots \quad (10.12)$$

The solution of this differential equation is a linear combination of $J_m(kr)$ and $N_m(kr)$. The latter is discarded because $N_m(kr)$ is singular at $r = 0$. Thus the solution for a particular value of m is given by $u_{mk} = J_m(kr) (A_m \cos m\theta + B_m \sin m\theta) (C_k \cos kct + D_k \sin kct) \dots$ (10.12(A))

Since the circumference of the membrane $r = a$ is permanently fixed, we must have $J_m(ka) = 0$. The frequencies are $w_{mn} = k_{mn} c$. (k_{mn} 's are given by $J_m(k_{mn} a) = 0$).

The general solution obtained by the principle of superposition is given by

$$u(r, \theta, t) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} [\{P_{mn} \cos(k_{mn} ct) + Q_{mn} \sin(k_{mn} ct)\} \cos(m\theta) + \{R_{mn} \cos(k_{mn} ct) + S_{mn} \sin(k_{mn} ct)\} \sin(m\theta)] J_m(k_{mn} r) \quad (10.13)$$

Let the initial conditions be, $u(r, \theta, 0) = f_1(r, \theta)$, $\dot{u}(r, \theta, 0) = f_2(r, \theta)$.
so that we get from Eq. (10.12) (A) for $m, n = 1, 2, \dots$,

$$P_{mn} = \frac{2}{\pi a^2 [J_{m+1}(k_{mn} a)]^2} \int_0^{2\pi} d\theta \int_0^a dr f_1(r, \theta) r J_m(k_{mn} r) \cos m\theta,$$

$$R_{mn} = \frac{2}{\pi a^2 [J_{m+1}(k_{mn} a)]^2} \int_0^{2\pi} d\theta \int_0^a dr f_1(r, \theta) r J_m(k_{mn} r) \sin m\theta,$$

$$Q_{mn} = \frac{2}{\pi a^2 k_{mn} c [J_{m+1}(k_{mn} a)]^2} \int_0^{2\pi} d\theta \int_0^a dr r f_2(r, \theta) J_m(k_{mn} r) \cos m\theta,$$

$$S_{mn} = \frac{2}{\pi a^2 k_{mn} c [J_{m+1}(k_{mn} a)]^2} \int_0^{2\pi} d\theta \int_0^a dr r f_2(r, \theta) J_m(k_{mn} r) \sin m\theta.$$

For $m = 0$, the coefficients are given as above with an additional factor of $1/2$. The orthogonality relation of Bessel functions, viz., $\int_0^a x J_m(\alpha_m x) J_m(\alpha_p x) dx = \delta_{mp} \frac{a^2}{2} [J_{m+1}(\alpha_p a)]^2$ has been used here.

3. Laplace equation

The steady state temperature distribution, variation of the scalar potential as well as the vector potential in a source free region are, to cite a few examples, governed by Laplace equation. The equation is of the form $\Delta^2 \psi = 0$ and thus by nature an elliptic partial differential equation. We will utilise the method of the separation of variables to solve it in different coordinate systems which will be relevant for the solution of boundary value problems under different circumstances.

(i) *Cartesian coordinates* The Laplace equation $\nabla^2 \psi = 0$, written in Cartesian coordinates becomes,

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} = 0,$$

Writing $\psi(x, y, z) = X(x) Y(y) Z(z)$ we obtain,

$$YZ \frac{d^2 X}{dx^2} + XZ \frac{d^2 Y}{dy^2} + XY \frac{d^2 Z}{dz^2} = 0 \text{ and dividing by } XYZ \text{ we get,}$$

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{1}{Z} \frac{d^2 Z}{dz^2} = 0$$

or,

$$-\frac{1}{X} \frac{d^2 X}{dx^2} = \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{1}{Z} \frac{d^2 Z}{dz^2} = k_x^2$$

or,

$$X = A \sin k_x x + B \cos k_x x.$$

Similarly $-\frac{1}{Y} \frac{d^2 Y}{dy^2} = k_y^2, \frac{1}{Z} \frac{d^2 Z}{dz^2} = -k_z^2$ giving $Y = C \sin k_y y + D \cos k_y y$.

and $\frac{1}{Z} \frac{d^2 Z}{dz^2} = k_x^2 + k_y^2$ yielding as the solution,

$$Z = E \exp k_z z + F \exp (-k_z z), \text{ where } k_z = \sqrt{k_x^2 + k_y^2}.$$

Thus the general solution is obtained by taking the product XYZ , superposing the solutions for different k_x, k_y subject to the proper boundary conditions. Let the boundary condition refer to the problem of steady state temperature distribution in a rectangular volume so that $\psi = 0$ at $x = 0, x = a, y = 0, y = b, z = 0$ and $\psi = T_o$ at $Z = c$.

Thus from the expressions of X, Y, Z we obtain that $B = 0, k_x = \frac{m\pi}{a}, D = 0, k_y = \frac{n\pi}{b}, E = -F$ so that the

general solution can be written as, $\psi = \sum_{mn} G_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} \sinh k_z z$,

where $k_z = \pi \sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2}}, m, n = 1, 2, \dots$

Imposing the boundary condition $\psi = T_o$ at $z = c$ we get,

$$T_o = \sum G_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} \sinh k_z c.$$

or,

$$G_{mn} = \frac{4}{ab} \int_0^a \int_0^b T_o \operatorname{cosech} \left[\left(\frac{m^2}{a^2} + \frac{n^2}{b^2} \right)^{\frac{1}{2}} \pi c \right] \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} dx dy$$

Thus inserting the expression for G_{mn} we can find out the steady state temperature T at any point within the region.

(ii) Spherical polar coordinates

Laplace equation becomes, $\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \cdot \frac{\partial^2 \psi}{\partial \phi^2} = 0$

Using the separation of variables technique we write, $\psi = R(r) Z(\theta, \phi)$ so that

we obtain, $\frac{1}{r^2} Z \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{R}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Z}{\partial \theta} \right) + \frac{R}{r^2 \sin^2 \theta} \cdot \frac{\partial^2 Z}{\partial \phi^2} = 0$

Dividing by ZR we obtain,

$$\frac{1}{r^2 R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{Z r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Z}{\partial \theta} \right) + \frac{1}{Z r^2 \sin^2 \theta} \frac{\partial^2 Z}{\partial \phi^2} = 0$$

or,

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = - \frac{1}{Z} \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Z}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \cdot \frac{1}{Z} \frac{\partial^2 Z}{\partial \phi^2} = \lambda,$$

where λ is a constant since the left hand side and the right hand side are functions of r and θ, ϕ respectively. Multiplying by $\sin^2 \theta$ we get,

$$- \frac{\sin \theta}{Z} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Z}{\partial \theta} \right) - \frac{1}{Z} \frac{\partial^2 Z}{\partial \phi^2} = \lambda \sin^2 \theta$$

Writing Z as $F(\theta) G(\phi)$ we get,

$$\frac{\sin \theta}{F} \frac{d}{d\theta} \left(\sin \theta \frac{dF}{d\theta} \right) + \lambda \sin^2 \theta = - \frac{1}{G} \frac{d^2 G}{d\phi^2} = + m^2 \text{ (say).}$$

The separation constant is taken to be positive because the solution is assumed to be periodic in ϕ , which is physically reasonable.

Thus, $G = A \cos m\phi + B \sin m\phi$, where $m = 0, \pm 1, \pm 2, \dots$

The θ -equation becomes,

$$\frac{1}{\sin \theta} \cdot \frac{1}{F} \frac{d}{d\theta} \left(\sin \theta \frac{dF}{d\theta} \right) = \left(\frac{m^2}{\sin^2 \theta} - \lambda \right)$$

Putting $\cos \theta = x$, so that $\frac{dF}{d\theta} = -\sqrt{1-x^2} \frac{dF}{dx}$ and

$$\frac{d}{d\theta} \left(\sin \theta \frac{dF}{d\theta} \right) = \sqrt{1-x^2} \frac{d}{dx} \left((1-x^2) \frac{dF}{dx} \right) = \left[(1-x^2) \frac{d^2 F}{dx^2} - 2x \frac{dF}{dx} \right] \sqrt{1-x^2}$$

so that we get, $(1-x^2) \frac{d^2 F}{dx^2} - 2x \frac{dF}{dx} + \left(\lambda - \frac{m^2}{\sin^2 \theta} \right) F = 0$

This equation has a well-behaved solution at $x = \pm 1$ if we put $\lambda = l(l+1)$, $l =$ positive integer, so that the solution are associated Legendre polynomials given by $P_l^m(x)$, where

$-l \leq m \leq l$. [The other solution $Q_l^m(x)$ is rejected because of finiteness at $x = \pm 1$].

The radial equation gives, $r^2 \frac{d^2 R}{dr^2} + 2r \frac{dR}{dr} - l(l+1) R = 0$

Assuming a series solution of the form $R = \sum_n c_n r^{n+\rho}$ we get,

$$\sum_n (n + \rho)(n + \rho - 1) r^{n+\rho} + 2 \sum_n (n + \rho) c_n r^{n+\rho} - l(l+1) \sum_n c_n r^{n+\rho} = 0$$

Equating the coefficient of r^ρ we get, $c_o [\rho(\rho + 1) - l(l + 1)] = 0$

For $c_o \neq 0$, $\rho = l, -(l + 1)$.

All other c'_n s can be shown to vanish.

Thus the general solution can be written as,

$$\psi = \sum_l \left(A_l r^l + \frac{B_l}{r^{l+1}} \right) \cdot Y_{lm}(\theta, \phi), \text{ where } Y_{lm}(\theta, \phi) = N_{lm} P_l^m(\cos \theta) e^{im\phi},$$

N_{lm} being the proper normalization constant, so that (10.14)

$$\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} Y_{lm}^*(\theta, \phi) Y_{l'm'}(\theta, \phi) \sin \theta d\theta d\phi = \delta_{ll'} \delta_{mm'},$$

For spherical symmetry the problem can be further simplified to give, $\psi = \sum_l \left(A_l r^l + \frac{B_l}{r^{l+1}} \right) P_l(\cos \theta)$ because for $m = 0$,

$$Y_{l0}(\theta, \phi) = \left(\frac{2l+1}{4\pi} \right)^{\frac{1}{2}} P_l(\cos \theta).$$

Example Suppose a solid sphere of radius a has one hemisphere $\left(0 \leq \theta < \frac{\pi}{2}\right)$ at a potential V_1 and the other hemisphere at the potential V_2 . To find out the potential at all points inside the sphere we impose the boundary conditions. Putting $B_l = 0$ (for otherwise the potential would diverge at $r = 0$) we get

$$(\psi)_{r=a} = \sum_l A_l a^l P_l(\cos \theta),$$

Thus $A_l a^l = \left(\frac{2l+1}{2} \right) \int_{-1}^1 (\psi)_{r=a} P_l(\cos \theta) d(\cos \theta)$ by virtue of orthogonality of the Legendre functions.

$$\text{or, } A_l a^l = \left[\int_0^1 V_1 P_l(\cos \theta) d(\cos \theta) + \int_{-1}^0 V_2 P_l(\cos \theta) d(\cos \theta) \right] \left(\frac{2l+1}{2} \right)$$

Using $P'_{l+1}(x) + P'_{l-1}(x) = (2l+1) P_l(x)$, we get

$$\begin{aligned} A_l a^l &= \left[\frac{V_1}{2} \left| P_{l+1}(x) + P_{l-1}(x) \right|_0^1 + \frac{V_2}{2} \left| P_{l+1}(x) + P_{l-1}(x) \right|_{-1}^0 \right] \\ &= \frac{1}{2} [V_1 \{P_{l+1}(1) + P_{l-1}(1) - P_{l+1}(0) - P_{l-1}(0)\} + V_2 \{P_{l+1}(0) + P_{l-1}(0) - P_{l+1}(-1) - P_{l-1}(-1)\}]. \end{aligned}$$

Using $P_l(1) = 1$ for all l , $P_l(0) = 0$ for odd l , $P_l(-x) = (-1)^l P_l(x)$ we get,

$$P_l(0) = \frac{(-1)^{l/2}}{\left(\frac{l}{2}\right)!} \frac{1}{2} \cdot \frac{3}{2} \cdots \frac{(2l-1)}{2} \text{ for even } l.$$

$$A_l a^l = (V_1 + V_2) \text{ for even } l.$$

Similarly, A_l can be determined for odd l .

Thus, $\psi = \left(\sum_l A_l r^l P_l(\cos \theta) \right)$ can be determined for all r inside the sphere.

For finding out ψ for all points outside the sphere we have to use $\psi = \sum_l \frac{B_l}{r^{l+1}} P_l(\cos \theta)$ [otherwise ψ diverges at $r \rightarrow \infty$] ignoring the presence of any field at $r \rightarrow \infty$ for which the term A_l has to be taken into account.

(iii) *Cylindrical coordinate system* Laplace equation in cylindrical coordinate system is,

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \cdot \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \theta^2} + \frac{\partial^2 \psi}{\partial z^2} = 0$$

Substituting $\psi = R(r) F(\theta) Z(z)$ we obtain, after dividing by ψ ,

$$\frac{1}{Rr} \frac{d}{dr} \left(r \frac{dR}{dr} \right) + \frac{1}{r^2 F} \frac{d^2 F}{d\theta^2} + \frac{1}{Z} \frac{d^2 Z}{dz^2} = 0$$

$$\text{or, } \frac{r}{R} \frac{d}{dr} \left(r \frac{dR}{dr} \right) + \frac{r^2}{Z} \frac{d^2 Z}{dz^2} = -\frac{1}{F} \frac{d^2 F}{d\theta^2} = m^2 \text{ in order that the solution is periodic in } \theta.$$

This gives, $F(\theta) = A \cos m \theta + B \sin m \theta$.

We then obtain, $\frac{1}{Rr} \frac{d}{dr} \left(r \frac{dR}{dr} \right) - \frac{m^2}{r^2} = -\frac{1}{Z} \frac{d^2 Z}{dz^2} = -k^2$ so that

$Z = Ce^{kz} + De^{-kz}$. The exponentially varying solution is taken because of the non-periodic behaviour assumed of Z as a function of z . The radial equation then gives

$$\frac{1}{Rr} \frac{d}{dr} \left(r \frac{dR}{dr} \right) + \left(k^2 - \frac{m^2}{r^2} \right) = 0$$

$$\text{or, } \frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} + \left(k^2 - \frac{m^2}{r^2} \right) R = 0, \text{ and putting } kr = x, \text{ we can write the equation as}$$

$$x^2 \frac{d^2 R}{dx^2} + x \frac{dR}{dx} + (x^2 - m^2) R = 0, \text{ which is our known Bessel differential equation.}$$

The solution that is non-singular at $r = 0$ is $R = J_m(kr)$.

Thus, the general solution is $\psi = \sum_{k,m} (A_m \cos m\theta + B_m \sin m\theta) (C_k e^{kz} + D_k e^{-kz}) J_m(kr)$

[This sum excludes $k = 0$ for which $Z = E + Fz$, E and F being constants. The $k = 0$ solution is assumed not to be present by virtue of the boundary conditions imposed. The $m = 0$ term already exists in the sum.]

Example The upper surface of a cylinder of radius a and height h is maintained at a temperature of $T_o^\circ\text{C}$. The curved and the lower flat surfaces are maintained at 0°C . Find the temperature at a point inside the cylinder. (Take $T_o = 100 \cos 2\theta + 200 \sin 2\theta$).

Solution: The solution of Laplace's equation at any point inside the cylinder can be written as $T(r, \theta, z)$
 $= \sum J_m(kr) (A_m \cos m\theta + B_m \sin m\theta) \times (C(k) e^{kz} + D(k) e^{-kz})$, where $A_m, B_m, C(k), D(k)$ are constants and the function $N_m(kr)$ is discarded because of divergence at $r = 0$. ($m = \text{integer}$)

Since $T = 0$ for $z = 0$, we must have $C(k) = -D(k)$ which gives,

$$T(r, \theta, z) = \sum J_m(kr) (A_m \cos m\theta + B_m \sin m\theta) C'(k) \sinh kz, \text{ where } C'(k) = 2 C(k).$$

Again, $T(r, \theta, z) = 0$ at $r = a$, giving.

$$0 = \sum J_m(ka) (A_m \cos m\theta + B_m \sin m\theta) C'(k) \sinh kz \text{ for all } z \text{ and } \theta.$$

This gives $J_m(k_n a) = 0$, where $k_n = \frac{x_{mn}}{a}$, x_{mn} giving then n th zero of the Bessel function of order m .

$$\begin{aligned} \text{or, } T(r, \theta, z) &= \sum_{n=1}^{\infty} \sum_m J_m\left(\frac{x_{mn} r}{a}\right) (A_m \cos m\theta + B_m \sin m\theta) C'(k) \sinh kz. \\ &= \sum_{n=1}^{\infty} \sum_m J_m\left(\frac{x_{mn} r}{a}\right) (A_{mn} \cos m\theta + B_{mn} \sin m\theta) \sinh\left(\frac{x_{mn} z}{a}\right) \dots \end{aligned} \quad (10.15)$$

Now, $T(r, \theta, h) = T_o$ for all r, θ .

$$\text{or, } \sum_n \sum_m J_m\left(\frac{x_{mn} r}{a}\right) (A_{mn} \cos m\theta + B_{mn} \sin m\theta) \sinh\left(\frac{x_{mn} h}{a}\right) = T_o$$

$$\text{or, } \sum_n A_{nn} J_n\left(\frac{x_{nn} r}{a}\right) \sinh\left(\frac{x_{nn} h}{a}\right) \pi = \int_0^{2\pi} T_o \cos m\theta d\theta$$

Putting $T_o = 100 \cos 2\theta + 200 \sin 2\theta$, we get

$$\pi \sum_n J_2\left(\frac{x_{2n} r}{a}\right) \sinh\left(\frac{x_{2n} h}{a}\right) A_{2n} = 100\pi \text{ and other } A_{mn}'\text{'s} = 0. \dots \quad (10.16)$$

$$\text{Similarly, } \pi \sum_n J_2\left(\frac{x_{2n} r}{a}\right) \sinh\left(\frac{x_{2n} h}{a}\right) B_{2n} = 200\pi, \text{ other } B_{mn}'\text{'s} = 0$$

Using the orthogonality relation of Bessel functions, we get

$$\int_0^a r J_p \left(\frac{x_{pm} r}{a} \right) J_p \left(\frac{x_{pn} r}{a} \right) dr = \frac{a^2}{2} [J_{p+1}(x_{pm})]^2 \text{ for } m = n \text{ and} \\ = 0 \text{ otherwise.}$$

Multiplying Eq. (10.17) by $J_m \left(\frac{x_{mq} r}{a} \right)$ and integrating from $r = 0$ to a

$$\text{we get, } \frac{a^2}{2} [J_{m+1}(x_{mq})]^2 A_{2q} \sinh \left(\frac{x_{2q} h}{a} \right) \\ = 100 \int_0^a J_2 \left(\frac{x_{2q} r}{a} \right) r dr, \text{ where } m = 2.$$

$$\text{or, } A_{2q} \sinh \left(\frac{x_{2q} h}{a} \right) = \frac{200}{a^2} \times \frac{1}{[J_3(x_{2q})]^2} \int_0^a J_2 \left(\frac{x_{2q} r}{a} \right) r dr.$$

$$\text{Similarly } B_{2q} \sinh \left(\frac{x_{2q} h}{a} \right) = \frac{400}{a^2 [J_3(x_{2q})]^2} \int_0^a J_2 \left(\frac{x_{2q} r}{a} \right) r dr.$$

Thus using Eq. (10.15) we can write the temperature at any point inside the cylinder as,

$$T(r, \theta, z) = \sum_n J_2 \left(\frac{x_{2n} r}{a} \right) [A_{2n} \cos 2\theta + B_{2n} \sin 2\theta] \sinh \left(\frac{x_{2n} z}{a} \right) \dots, \quad (10.17)$$

the A_{2n} 's and B_{2n} 's being given from the above.

4. Helmholtz equation

Let us consider the heat conduction equation in three dimensions. The equation is, $\frac{\partial T}{\partial t} = h \nabla^2 T$, T being the temperature difference from the equilibrium and h the thermometric conductivity of the material.

Writing $T = F(t) G(\bar{r})$ we obtain, $G \frac{dF}{dt} = h F \nabla^2 G$.

Dividing by T we get, $\frac{1}{F} \frac{dF}{dt} = \frac{h}{G} \nabla^2 G = -k^2 h$, so that we get

$$F(t) = F_0 \exp(-k^2 t).$$

The separation constants is taken to be negative because any deviation from the equilibrium is expected to decay with time. Thus, the spatial evolution of T is given by $\nabla^2 G + k^2 G = 0$. Thus the heat conduction equation which is originally a parabolic differential equation in one dimension is transformed into the Helmholtz equation. $\nabla^2 G + k^2 G = 0$.

Let us solve this equation in the three coordinate systems.

(i) *Cartesian coordinates* The equation becomes, $\frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} + \frac{\partial^2 G}{\partial z^2} = -k^2 G$

Writing $G = X(x) Y(y) Z(z)$ we get by dividing by G ,

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{1}{Z} \frac{d^2 Z}{dz^2} = -k^2.$$

We can thus, write separately three equations involving x , y and z respectively given by,

$$\frac{d^2 X}{dx^2} = -k_x^2 X, \frac{d^2 Y}{dy^2} = -k_y^2 Y, \frac{d^2 Z}{dz^2} = -k_z^2 Z, \text{ where } k_x^2 + k_y^2 + k_z^2 = k^2. \text{ The solutions are given by,}$$

$$X = (A_1 \sin k_x x + B_1 \cos k_x x), Y = (A_2 \sin k_y y + B_2 \cos k_y y), Z = (A_3 \sin k_z z + B_3 \cos k_z z).$$

The boundary conditions are, suppose, $T = 0$ at $x = 0, x = l_x, y = 0, y = l_y$ and $z = 0, z = l_z$ respectively so that the boundary sides of the rectangular volume are kept for all t at the equilibrium values. The separation constants have been taken to be negative because for positive values of the constants X or Y or Z would be a sum of the exponentially growing and decaying solutions and cannot satisfy the boundary condition of vanishing at the walls of the volume.

The boundary conditions give, $X = A_1 \sin \frac{n_1 \pi}{l_1} x$ ($n_1 = \pm 1, \pm 2, \dots$), $Y = A_2 \sin \frac{n_2 \pi}{l_2} y$ ($n_2 = \pm 1, \pm 2, \dots$), $Z = A_3 \sin \frac{n_3 \pi}{l_3} z$ ($n_3 = \pm 1, \pm 2, \dots$)

Thus, the general solution is, $G = \sum_{n_1, n_2, n_3} A_{n_1, n_2, n_3} \sin \frac{n_1 \pi}{l_1} x \sin \frac{n_2 \pi}{l_2} y \sin \frac{n_3 \pi}{l_3} z$.

Clearly, A_{n_1, n_2, n_3} can be known if G be given as a function of x, y, z , which represents the initial value of T .

$$\frac{l_1 l_2 l_3}{8} A_{n_1, n_2, n_3} = \left[\int G(x, y, z) \sin \frac{n_1 \pi}{l_1} x \sin \frac{n_2 \pi}{l_2} y \sin \frac{n_3 \pi}{l_3} z dx dy dz \right]; \text{ where the integration extends}$$

from $x = 0$ to $l_1, y = 0$ to $l_2, z = 0$ to l_3 respectively. Thus finding out A_{n_1, n_2, n_3} we can substitute in T to obtain,

$$T = \sum \left[A_{n_1 n_2 n_3} \sin \frac{n_1 \pi}{l_1} x \sin \frac{n_2 \pi}{l_2} y \sin \frac{n_3 \pi}{l_3} z \right] \exp \left[-h\pi^2 \left(\frac{n_1^2}{l_1^2} + \frac{n_2^2}{l_2^2} + \frac{n_3^2}{l_3^2} \right) t \right] \dots \quad (10.18)$$

(ii) *Spherical polar coordinates* The equation becomes,

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial G}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial G}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 G}{\partial \phi^2} = -k^2 G$$

Writing $G = R(r) X(\theta) Z(\phi)$ we get, after dividing by G ,

$$\frac{1}{r^2 R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + \frac{1}{r^2 X \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial X}{\partial \theta} \right) + \frac{1}{r^2 Z \sin^2 \theta} \frac{\partial^2 Z}{\partial \phi^2} = -k^2$$

$$\text{or, } \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{X \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dX}{d\theta} \right) + \frac{1}{Z \sin^2 \theta} \frac{d^2 Z}{d\phi^2} = -k^2 r^2$$

$$\text{or, } \frac{\sin^2 \theta}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{\sin \theta}{X} \frac{d}{d\theta} \left(\sin \theta \frac{dX}{d\theta} \right) + k^2 r^2 \sin^2 \theta = -\frac{1}{Z} \frac{d^2 Z}{d\phi^2} = -m^2 \text{ (say)}$$

m being an integer to ensure periodicity $Z(\phi) = Z(\phi + 2p\pi)$, $p = \text{integer}$.

Thus $Z = A \cos m\phi + B \sin m\phi$, A and B being constants.

Dividing by $\sin^2 \theta$, we get

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{X \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dX}{d\theta} \right) + k^2 r^2 = \frac{m^2}{\sin^2 \theta}$$

$$\text{or, } \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + k^2 r^2 = -\frac{1}{X \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dX}{d\theta} \right) + \frac{m^2}{\sin^2 \theta} = \lambda \text{ (say)}$$

Thus the θ -equation becomes,

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dX}{d\theta} \right) + \left(\lambda - \frac{m^2}{\sin^2 \theta} \right) X = 0$$

Putting $x = \cos \theta$ we can transform the equation to the form,

$$(1-x^2) \frac{d^2 X}{dx^2} - 2x \frac{dX}{dx} + \left[\lambda - \frac{m^2}{1-x^2} \right] X = 0$$

For the solution at $x = \pm 1$ to stand true, we must have $\lambda = l(l+1)$, where l is a positive integer or zero and $-l \leq m \leq l$. The solution is $X = P_l^m(x)$, the other independent solution $Q_l^m(x)$ being discarded because of the condition of the solution being finite at $x = \pm 1$.

The r -equation becomes

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + [k^2 r^2 - l(l+1)] = 0$$

$$\text{or, } r^2 \frac{d^2 R}{dr^2} + 2r \frac{dR}{dr} + [k^2 r^2 - l(l+1)] R = 0$$

Transforming to $R = \frac{u}{\sqrt{r}}$ we get the equation for u as

$$r^2 \frac{d^2 u}{dr^2} + r \frac{du}{dr} + \left[k^2 r^2 - \left(l + \frac{1}{2} \right)^2 \right] u = 0$$

Changing variable to $\rho = kr$, we obtain $u = J_{l+1/2}(kr)$, the other independent solution $N_{l+1/2}(kr)$ being rejected in order that u is finite at $r = 0$.

Thus, we obtain, $G = \sum_{klm} J_{l+\frac{1}{2}}(kr) P_l^m(\cos \theta) [A_{klm} \cos m\theta + B_{klm} \sin m\theta]$

The general solution $T = \sum_{klm} e^{-k^2 t} J_{l+\frac{1}{2}}(kr) P_l^m(\cos \theta) [A_{klm} \cos m\theta + B_{klm} \sin m\theta] \dots$ (10.19)

For a spherically symmetric problem we can take $m = 0$, so that

$$T = \sum_{kl} e^{-k^2 t} J_{l+\frac{1}{2}}(kr) A_{kl} P_l(\cos \theta)$$

A_{kl} is known by using the orthogonality relations of $P_l(\cos \theta)$ and $J_{l+\frac{1}{2}}(kr)$ if we are given T at $t = 0$, as a function of r and θ .

(iii) *Cylindrical coordinates* The equation becomes,

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial G}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 G}{\partial \theta^2} + \frac{\partial^2 G}{\partial Z^2} = -k^2 G.$$

Writing $G = R(r) X(\theta) Z(z)$ we get, after dividing by G ,

$$\frac{1}{rR} \frac{d}{dr} \left(r \frac{dR}{dr} \right) + \frac{1}{r^2 X} \frac{d^2 X}{d\theta^2} + \frac{1}{Z} \frac{d^2 Z}{dZ^2} = -k^2.$$

or,

$$\frac{1}{rR} \frac{d}{dr} \left(r \frac{dR}{dr} \right) + \frac{1}{r^2 X} \frac{d^2 X}{d\theta^2} + k^2 = -\frac{1}{Z} \frac{d^2 Z}{dZ^2} = -\alpha^2$$

The separation constant is taken to be negative because of the non-periodic behaviour assumed of Z .

$$Z = Ae^{\alpha z} + B e^{-\alpha z}.$$

Multiplying by r^2 we obtain, $\frac{r}{R} \frac{d}{dr} \left(r \frac{dR}{dr} \right) + (k^2 + \alpha^2) r^2 = -\frac{1}{X} \frac{d^2 X}{d\theta^2} = m^2.$

Yielding as a solution $X = C \cos m\theta + D \sin m\theta$. ($m = \text{integer for periodicity}$). The r -equation then becomes,

$$\frac{r}{a} \frac{d}{dr} \left(r \frac{dR}{dr} \right) + [(k^2 + \alpha^2) r^2 - m^2] R = 0$$

Using $\rho = \sqrt{(k^2 + \alpha^2)} r$, the equation becomes the standard Bessel equation so that the solution is, $R = J_m(\rho)$ [the other solution $N_m(\rho)$ is discarded].

Thus the general solution is,

$$T = \sum J_m(\rho) [A_{\alpha km} e^{\alpha z} + B_{\alpha km} e^{-\alpha z}] [C_m \cos m\theta + D_m \sin m\theta] e^{-k^2 t} \dots$$
 (10.20)

[The summation excludes $m = 0$ for which the θ -solution becomes of the form $P\theta + Q$ and $k = 0$ for which the z -solution becomes of the form $Sz + U$, P, Q, S, U being constants. Assuming periodicity in θ ,

$P = 0$, the constants Q being already a term of the sum. The boundary conditions are assumed so that the $k = 0$ term need not be considered]

(In all the calculations involving the cylindrical coordinates the exponent of z has been chosen as discrete. This is because of the assumption of boundary conditions at certain specific values of z which T must satisfy. If there is no such condition required to be satisfied, the exponent can take up a continuous set of values in which case instead of the sum an integration has to be performed over the possible range of values of the exponents. If T at $t = 0$ be specified as a function of z , ρ and θ , it can be determined at any time by using the orthogonality relations.

Note: The heat conduction equation in one dimension (parabolic equation) is discussed briefly in the section 'Appendix'.

10.5 INHOMOGENEOUS PARTIAL DIFFERENTIAL EQUATION

The inhomogeneous partial differential equations occur in many areas of physical science. Instead of a detailed analysis of these equations, attention will be focussed on a particular equation, namely, Poisson equation, which arises in the potential problem with sources in electrostatics or in Newton's gravitational potential problems we will discuss below two methods, (i) Fourier transform method and (ii) Green's function method.

1. Fourier transform method

The equation $\nabla^2 \phi = -\frac{\rho}{\epsilon}$ becomes on writing $\phi = \int \bar{\phi}(\bar{k}) e^{i\bar{k} \cdot \bar{r}} d\bar{k}$ and $\rho = \int \bar{\rho}(\bar{k}) e^{i\bar{k} \cdot \bar{r}} d\bar{k}$, ρ being the charge density at a point in electrostatics or $G\rho_m$ in the gravitational problems, ρ_m being the mass density at a point.

Assuming that $\rho(\bar{r})$ goes to zero as $r \rightarrow \infty$ we get,

$$\bar{\phi} = \frac{\bar{\rho}}{\epsilon_0 k^2} \text{ and hence } \phi(\bar{r}) = \int \frac{\bar{\rho}(k) e^{i\bar{k} \cdot \bar{r}} d\bar{k}}{\epsilon_0 k^2}$$

Writing $\rho(k) = \frac{1}{8\pi^3} \int \rho(\bar{r}') e^{-i\bar{k} \cdot \bar{r}'} d\bar{r}'$, we obtain

$$\phi(\bar{r}) = \frac{1}{\epsilon_0} \int \frac{\rho(\bar{r}') e^{i\bar{k} \cdot (\bar{r} - \bar{r}')}}{8\pi^3 k^2} d\bar{k} d\bar{r}'$$

$$\text{Now, } \int \frac{e^{i\bar{k} \cdot (\bar{r} - \bar{r}')}}{k^2} d\bar{k} = \int e^{i\bar{k} \cdot (\bar{r} - \bar{r}')} \sin\theta d\theta d\phi k^2 dk.$$

$$\text{Or, } \int \frac{e^{i\bar{k} \cdot (\bar{r} - \bar{r}')}}{k^2} d\bar{k} = 4\pi \int_0^\infty \frac{\sin k|\bar{r} - \bar{r}'|}{k|\bar{r} - \bar{r}'|} dk = \frac{2\pi^2}{|\bar{r} - \bar{r}'|}, \text{ using}$$

$$\int_0^\infty \frac{\sin mx}{x} dx = \frac{\pi}{2} \quad (m > 0) \quad [\text{Worked out in the chapter on complex variables}] \text{ and hence}$$

$$\phi(\bar{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\bar{r}')}{|\bar{r} - \bar{r}'|} d\bar{r}'.$$

[The general solution will contain contributions from the homogeneous problem, i. e., the solution of the equation $\nabla^2 \phi_h = 0$ but if the boundary conditions are so adjusted that $\phi_h = 0$ everywhere then the above gives the complete solution.].

2. Green's function method

The Poisson's equation is, $\nabla^2 \phi = -\frac{\rho}{\epsilon_o}$.

The Green's function of the problem satisfies $\nabla^2 G(\vec{r}, \vec{r}') = -\delta(\vec{r} - \vec{r}')$ subject to the prescribed boundary conditions.

The Green's function $G(\vec{r}, \vec{r}') = \frac{1}{4\pi |\vec{r} - \vec{r}'|}$ as can be seen from the equality

$$\nabla^2 \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) = -4\pi \delta(\vec{r} - \vec{r}').$$

Using Green's formula, we obtain $\int_S (\phi \bar{\nabla} G - G \bar{\nabla} \phi) \cdot d\vec{s} = \int_V (\phi \nabla^2 G - G \nabla^2 \phi) dV$

Using $G = \frac{1}{4\pi |\vec{r} - \vec{r}'|}$, we have

$$\phi(\vec{r}) + \frac{1}{4\pi} \int_{V'} \nabla^2 \phi \cdot \frac{1}{|\vec{r} - \vec{r}'|} dV' = - \int_{S'} \left[\phi \nabla \left(\frac{1}{4\pi |\vec{r} - \vec{r}'|} \right) - \frac{1}{4\pi |\vec{r} - \vec{r}'|} \nabla \cdot \phi \right] \cdot d\vec{s}'$$

$$\text{or, } \phi(\vec{r}) = \frac{1}{4\pi \epsilon_o} \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} dV' + \frac{1}{4\pi} \int_{S'} \left[\phi \nabla \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) - \frac{1}{|\vec{r} - \vec{r}'|} \nabla \cdot \phi \right] \cdot d\vec{s}' \quad (10.21)$$

If the potential ϕ and its derivative decrease sufficiently rapidly with distance so that the surface integral vanishes if the surface bounding the volume is taken to an infinitely large distance, we obtain,

$$\phi(\vec{r}) = \frac{1}{4\pi \epsilon_o} \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} dV'.$$

Of course we can add the solution of the homogeneous problem ϕ_h to this expression to obtain the complete solution. If the expression ϕ_h can be made to vanish by proper choice of the boundary condition then the general solution is,

$$\phi = \frac{1}{4\pi \epsilon_o} \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} dV'.$$

In other cases, ϕ_h , which represents the potential at a point due to charges outside the volume, can be adjusted so that along with ϕ given by Eq. (10.22) the first or the second surface integral vanishes. Then the problem reduces to a Dirichlet (specification of ϕ on the boundary) or Neumann (specification of normal derivative of ϕ on the boundary) boundary value problem. (The image problem in electrostatics is an example).

It can be shown that the solution of the potential problem in a region with given boundary conditions on the surface enclosing the volume is unique. The proof is given in any text book on electrostatics and briefly

given here. Let ϕ_1 and ϕ_2 be two solutions so that $\nabla^2 \phi_1 = \frac{\rho}{\epsilon_0}$ and $\nabla^2 \phi_2 = -\frac{\rho}{\epsilon_0}$.

or, $\nabla^2 (\phi_1 - \phi_2) = 0$. If we name $\psi = \phi_1 - \phi_2$ we get, $\nabla^2 \psi = 0$.

$$\text{Now, } \int_V \nabla \cdot (\psi \bar{\nabla} \psi) dV = \int_S \psi \bar{\nabla} \psi \cdot d\bar{s} = \int_V \psi \nabla^2 \psi dV + \int_V [\nabla \psi]^2 dV.$$

For a given boundary condition (Dirichlet or Neumann) the surface integral vanishes and since $\nabla^2 \psi = 0$, $\int_V [\nabla \psi]^2 dV = 0$. Hence $\nabla \psi = 0$ and thus two solutions ϕ_1 and ϕ_2 can differ by trivial constant).

Since both ϕ_1 and ϕ_2 tend to zero on the infinitely distant boundary surface we obtain that $\phi_1 = \phi_2$ and hence the uniqueness is proved.

SUMMARY

A few techniques of solving the partial differential equations that are relevant to the physical problems are discussed.

- First, the first-order partial differential equations are discussed. The idea of characteristics is introduced and used to solve the equations.
- A general theory of linear second-order partial differential equations is discussed briefly. The role of the characteristics in regard to the choice of the boundary conditions for the three types of equations, namely, hyperbolic, elliptic and parabolic, respectively, is discussed. Some selected partial differential equations illustrating each of the above three types, viz., wave equation (hyperbolic), Laplace's equation (elliptic), Helmholtz equation (elliptic) and heat conduction equation in one dimension (parabolic) are discussed. (See the section **Appendix**). The separation of variables technique is employed with prescribed boundary conditions appropriate for the respective equations.
- The Fourier transform method and the Green's function method are briefly discussed in the context of the solution of inhomogeneous partial differential equations.

MULTIPLE CHOICE QUESTIONS

1. The function $u = f(x^2 + ay^2)$ satisfies the equation

$$(a) \quad ay \frac{\partial f}{\partial x} + x \frac{\partial f}{\partial y} = 0 \quad (b) \quad y \frac{\partial f}{\partial x} + ax \frac{\partial f}{\partial y} = 0 \quad (c) \quad ay \frac{\partial f}{\partial x} - x \frac{\partial f}{\partial y} = 0 \quad (d) \quad x \frac{\partial f}{\partial x} - ay \frac{\partial f}{\partial y} = 0$$

2. The function $u = f\left(\frac{x}{y}\right)$ satisfies the equation

$$(a) \quad x \frac{\partial u}{\partial x} - y \frac{\partial u}{\partial y} = 0 \quad (b) \quad x \frac{\partial u}{\partial x} + y \frac{\partial u}{\partial y} = 0 \quad (c) \quad x^2 \frac{\partial u}{\partial x} - y^2 \frac{\partial u}{\partial y} = 0 \quad (d) \quad x^2 \frac{\partial u}{\partial x} + y^2 \frac{\partial u}{\partial y} = 0$$

3. The solution of the equation $\frac{\partial u}{\partial t} + 3 \frac{\partial u}{\partial x} = 4x$ with $u(x, 0) = 6e^{-3t}$ is,
 - (a) $u = \frac{2}{3}x^2 - 6e^{x-3t}$
 - (b) $u = \frac{2}{3}x^2 + 6e^{x-3t}$
 - (c) $u = 2x^2 + 6e^{x-3t}$
 - (d) $u = 2x^2 - 6e^{x-3t}$
4. The solution of the equation $x \frac{\partial u}{\partial x} + y \frac{\partial u}{\partial y} = x^2$ with $u(x=1) = \sin y + 1/2$ is
 - (a) $u = \sin(y/x) + y^2/2$, (b) $u = \sin(y/x) - x^2/2$, (c) $u = \sin(y^2/x^2) + x^2/2$, (d) $u = \sin(y/x) + \frac{x^2}{2}$
5. The solution of the equation $\cos y \frac{\partial u}{\partial x} + \cos x \frac{\partial u}{\partial y} = 0$ with $u = 2 \sin x$ for $y = 0$ is
 - (a) $u = 2(\sin x + \sin^2 y)$
 - (b) $u = 2(\sin x - \sin y)$
 - (c) $u = 2(\sin x + \sin y)$
 - (d) $u = 2(\sin x - \sin^2 y)$
6. The solution of Laplace's equation in problems of axial symmetry involves
 - (a) Associated Legendre polynomials
 - (b) Legendre polynomials
 - (c) Spherical Bessel functions
 - (d) Trigonometric functions.
7. The solution of the equation $\frac{\partial u}{\partial x} + y^2 \frac{\partial u}{\partial y} = 0$ is,
 - (a) $u = A \left(x - \frac{1}{y} \right)^B + C$ ($A, B, C = \text{constants}$)
 - (b) $u = A \left(x + \frac{C}{y} \right)^B$
 - (c) $u = A \left(x + \frac{1}{y} \right)^B + C$
 - (d) $u = A \left(Bx + \frac{C}{y} \right)$
8. The solution of the equation $\frac{\partial^4 u}{\partial x^4} = 0$ is,
 - (a) $u = \phi(y) x^3/6 + \psi(y) x^2/2$
 - (b) $u = \phi(y) x^3/6 + \psi(y) x^2/2 + F(y) x + G(y)$

- (c) $u = \phi(y) x^3 - \psi(y) x^2 + F(y) x + G(y)$
(d) $u = \phi(x) y^3/6 + \psi(y) = x^2/2$, F , G , ϕ and ψ being arbitrary constants
9. The solution of Laplace's equation in cylindrical coordinates involves
(a) cylindrical Bessel function
(b) Spherical Bessel function
(c) Legendre polynomials
(d) Associated Legendre polynomials
10. The type of a partial differential equation of the second order (a) remains unaltered under transformation of variables.
(a) remains unchanged only if the equations is elliptic
(b) remains unchanged only if the equation is hyperbolic
(c) always changes under transformation of variables
11. The equation $\frac{\partial^2 u}{\partial x^2} + 2 \frac{\partial^2 u}{\partial x \partial y} + \alpha \frac{\partial^2 u}{\partial y^2} = 0$ ($\alpha = \text{constant}$) is
(a) elliptic if $\alpha < 1$
(b) hyperbolic if $\alpha^2 < 1$
(c) parabolic if $\alpha^2 = 1$
(d) of a type that cannot be determined straightforwardly

(More than one of the above statements may be true)

Answers to MCQs

1. (c)	2. (b)	3. (b)	4. (d)	5. (b)	6. (b)	7. (c)
8. (b)	9. (a)	10. (a)	11. (b), (c)			

EXERCISE

Solve the following equations:

- $x \frac{\partial u}{\partial x} + x^2 y = u$ and $u = 4y$ on the line $x = 2$.
- $\frac{\partial u}{\partial x} + 2 \frac{\partial u}{\partial y} = 5$, and $u = 2$ on the line $y = x$.
- $y^2 \frac{\partial u}{\partial x} + x^2 \frac{\partial u}{\partial y} = 0$, $u = x^6 + 2$ on the line $y = 0$.

4. $y \frac{\partial u}{\partial x} + 2x \frac{\partial u}{\partial y} = 0, u = 3$ on the hyperbola $y^2 - 2x^2 = 4$

5. $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = x^2 + y^2$ without any boundary condition

6. $\left(\frac{\partial}{\partial x} - a \frac{\partial}{\partial y} \right)^2 u = 0,$

7. $\left(\frac{\partial}{\partial x} - a \frac{\partial}{\partial y} \right) \left(\frac{\partial}{\partial x} - b \frac{\partial}{\partial y} \right) u(x, y) = 0$

[No boundary condition in (6) and (7)].

8. Solve $u_t = a^2 u_{xx}$, where $u(x, 0) = x$ for $0 < x < l$, $u(0, t) = u(l, t) = 0$ for all t .
9. Solve $\nabla^2 \phi = 0$ in a plane such that $\phi = \phi_0$ for $0 < \theta < \pi$ and $\phi = -\phi_0$ for $-\pi < \theta < 0$ for $r = a$.
10. A region between two concentric spherical shells is maintained at a steady state so that the two spherical surfaces are maintained at temperatures θ_1 and θ_2 respectively. (The surfaces are of radii r_1 and r_2 respectively). Find out the temperature at any point between the spheres.
11. A cylinder is cut into two equal halves so that one half ($\cos \phi > 0$) and the other half ($\cos \phi < 0$) are at the potential $\pm V_0$ respectively. Find out the potential at any point inside the cylinder.
12. A string is excited by plucking over a region so that at $t = 0$ the shape of the string looks like that in the figure (Fig. 10.2).

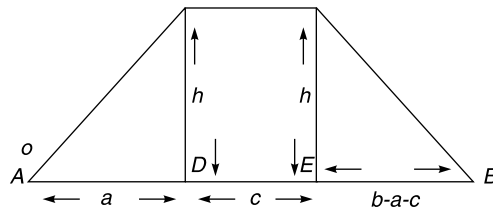


Fig. 10.2

Find out the displacement u as a function of x and t where x denotes the direction of the length of the string.

13. Solve Schrodinger equation for a free particle in three dimension in spherical polar coordinates and find out $\psi(r, \theta, \phi, t)$. (ψ is the wave function).
14. Solve Schrodinger equation for a particle localized in a box with sides at $x = 0, x = a, y = 0, y = b$ and $z = 0, z = c$ respectively. The initial condition is, $\psi = xyz$ at $t = 0$

(Schrodinger equation is, $i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi$. Solve the problems (12) and (13) by the method of the separation of variables).

Hints and Answers to Exercise Questions

1. $u = 4xy - x^2y,$

2. $u = 5(y - x) + 2,$

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

4. $u = y^2 - 2x^2 - 1,$

5. $u = \frac{1}{12} (x^4 + y^4) + f(z) + \phi(\bar{z}),$ where $z = x + iy$, where f and ϕ are two arbitrary functions of z and $\bar{z}$ respectively.

6. $u = f(z_1) - \frac{z_2}{2a} \phi(z_1),$ where $z_1 = y + ax, z_2 = y - ax$ and f, ϕ are two arbitrary functions of z_1 .

[**Hint:** Use two variables $z_1 = y + ax, z_2 = y - ax$ and convert the derivatives into those involving z_1 and z_2 respectively].

7. $u = f(y + ax) + g(y + bx), f$ and g being any two arbitrary functions.

8. $u = -\frac{2l}{\pi} \sum_n \frac{(-1)^n}{n} \sin \frac{n\pi x}{l} e^{-\frac{a^2\pi^2}{l^2} n^2 t}.$

[**Hint:** Use the method of the separation of variables].

9. $\phi = \sum_m \left(A_m r^m + \frac{B_m}{r^m} \right) D_m \sin m\theta,$

where $D_m \left(A_m a^m + \frac{B_m}{a^m} \right) = \frac{2\phi_0}{m\pi} [1 - (-1)^m].$

13. $\psi(r, \theta, \phi, t) = \sum_{lm} A_{lm} j_l(Kr) P_m^l(\cos \theta) e^{im\phi} e^{-iE_K t/\hbar}, E_K = \frac{\hbar^2 K^2}{2m}$

for a fixed energy. The coefficients A_{lm} are found out by solving the initial values prescribed for the problem. (for a given $l, -l \leq m \leq l$)

14. $\psi = \sum_n A_{mnp} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} \sin \frac{p\pi z}{c} \quad (m = \pm 1, \dots, n = \pm 1, \dots, p = \pm 1, \dots),$

where $A_{mnp} = \frac{8abc}{mnp} \cdot \frac{(-1)^{m+n+p}}{\pi^3}.$

APPENDIX

A Parabolic Equation

The heat conduction equation in one dimension is given by $\frac{\partial T}{\partial t} = h \frac{\partial^2 T}{\partial x^2}$, the symbols having their usual meanings. This is a parabolic equation.

The general solution with the boundary condition $T = 0$ at $x = 0$ and $x = l$ respectively for all t is a particular case of (10.19), namely,

$$T(x, t) = \sum_m A_n \sin \frac{n\pi x}{L} e^{-\frac{h\pi^2}{L^2} n^2 t}$$

Putting $T(x, 0) = f(x)$, where $f(x)$ is any given function we obtain,

$$A_n = \frac{2}{L} \int_0^l f(x) \sin \frac{n\pi x}{l} dx$$

Note: No separate section on “Solved Problems” has been provided in this chapter because the sufficient number of examples given in the chapter will, I believe, serve the same purpose.

Probability and Statistics

CHAPTER

11

11.1 INTRODUCTION

Often certain phenomena that occur are influenced by diverse factors, all of which are virtually impossible to take care of by any theoretical calculation. As a result the outcomes of such phenomena cannot be predicted with certainty. The tossing of coins, the game of dice are two such examples in this context. However, if we perform a set of identical experiments and focus our attention on a particular outcome, we will often observe a certain regularity. It is seen that the ratio of the number of such outcomes (also called events) to the total number of experiments tends to a fixed limit as the number of experiments tends to infinity. This is called the probability of the event. For example, the probability of an unbiased coin falling on the ground with its head up (or otherwise) is $\frac{1}{2}$. If we name two events A and B and try to see the

outcomes of these two events there can be the following possibilities:

- (a) Event A but not B occurs.
- (b) Event B but not A occurs.
- (c) Events A and B , both, occur.
- (d) Neither A nor B occurs.

The probability of any occurrence involving A and B can be seen to belong to one of the above categories or a combination of them. The probabilities of the various events can be represented, as explained below, by the Venn diagrams. Figure 11.1(A) and 11.1(B).

The area enclosed in the rectangular region represents the sample space, which is the outcome of all the experiments performed. The region enclosed by the closed curve A is a measure of the number of events A and that by B a measure of the number of events B . In Fig. (11.1(A)) the shaded region of intersection of two curves represents the class (c) listed above. The region within A to the left of the shaded region represents (a) while that to the right of the shaded region within B represents (b). The region within the rectangle excluding the two curves represents the type (d).

The set of events stated by (a), (b), (c) and (d) is symbolised by $A - B$, $B - A$, $A \cap B$ and $\overline{A + B}$ respectively ($\overline{A}$ is the set complement to A). Thus if N_1, N_2, N_3 and N_4 denote the number of events of the four types stated above, we can write with $N = \sum_{i=1}^4 N_i$ the total number of events.

$$P(A) = \text{Probability that the events } A \text{ occurs} = \frac{N_1 + N_3}{N},$$

$$P(B) = \text{Probability that } B \text{ occurs} = \frac{N_2 + N_3}{N},$$

$$P(A \cap B) = \text{Probability that both } A \text{ and } B \text{ occur} = \frac{N_3}{N},$$

$$P(\overline{A + B}) = \text{Probability that neither } A \text{ nor } B \text{ occurs} = \frac{N_4}{N},$$

$$\text{Thus, } P(A \cup B) = \text{Probability that either } A \text{ or } B \text{ or both occur} = \frac{N_1 + N_2 + N_3}{N}$$

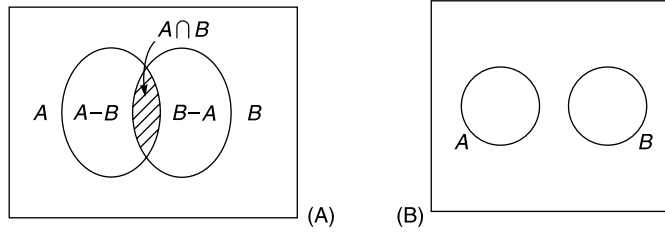


Fig. 11.1 (Venn diagram)

The same definitions can be extended to more than two events. Thus, the commutativity, associativity and the distributivity relations given below always hold good.

Commutativity, $A \cap B = B \cap A$, $A \cup B = B \cup A$.

Associativity, $(A \cap B) \cap C = A \cap (B \cap C)$, $(A \cup B) \cup C = A \cup (B \cup C)$.

Distributivity, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

11.2 JOINT PROBABILITY, CONDITIONAL PROBABILITY ETCETRA

The probability that a set of events happen together is called their joint probability. If two events are independent, the joint probability is the product of the individual probabilities so that we can write, $P(A \cap B) = P(A) \cdot P(B)$. (The same rule is valid for any number of events happening together). The problem of tossing of two coins is an example of this type. Thus the probability that two coins fall with (i) both the heads up

(ii) both the heads down, (iii) one head up and the other down. (iv) the up and the down interchanged is $\frac{1}{4}$

each. If the two events can never happen together, i.e., are mutually exclusive, the joint probability vanishes. Thus the simultaneous precise measurement of the two canonically conjugate variables (like x and p_x of a particle) in a single experiment are mutually exclusive phenomena. In such a case $P(A \cap B) = 0$. (Fig. 11.1 (B)).

The conditional probability in an experiment involving two events A and B is defined as the probability that A occurs if B occurs or alternatively as the probability that B happens provide A does so. This is denoted by $P(A/B)$ or $P(B/A)$ respectively. Thus we can write,

$$\begin{aligned} P(A \cap B) &= P(B) \cdot P(A/B) = P(A) P(B/A) \\ P(A \cup B) &= P(A) + P(B) - P(A \cap B). \end{aligned} \quad (11.1)$$

For example, the probability of obtaining two aces when two cards are drawn successively from a pack of 52 cards. $= \frac{4}{52} \times \frac{3}{51} = \frac{1}{221}$.

(This follows by writing $P(2 \text{ aces}) = P(1 \text{ ace}) P(2\text{nd. ace}/1\text{st ace})$).

From the relation (11.1) we can write, $P(A/B) = \frac{P(A) P(B/A)}{P(B)}$ (11.2)

If instead of one single event A , there are a number of mutually exclusive events A_1, A_2 , covering the entire space, we can apply Eq. (11.1) to get,

$$P(A/B) = \frac{P(A) P(B/A)}{\sum_i P(A_i) P(B/A_i)}.$$

If the sample space involves three events A, B , and C , we can apply Eq. (11.2) to obtain,

$$\frac{P(B/A)}{P(C/A)} = \frac{P(B) P(A/B)}{P(C) P(A/C)} \quad (11.3)$$

The relation (11.3) is called Baye's theorem.

Equation (11.1) can be applied to the one-dimensional motion of the molecules in a gas and undergoing collision. The probability of a collision in the interval between x and $x + dx$ can be written as the probability of a molecules being at x multiplied by the probability of a collision between x and $x + dx$ so that we can write, $-dn$ = decrease in the number of molecules due to the collision in the distance $dx = n dx/\lambda$, where λ is the mean free path.

On integrating we get, $n = n_0 e^{-x/\lambda}$, a relation giving the number of molecules remaining at x .

If each individual occurrence in an experiment is accorded equal *a-priori* probability, the total probability of the event is proportional to the number of ways in which the event can occur. For example, if out of a pack of 52 cards we pick up cards in succession, the net probability of getting any colour (heart, for example) is $\frac{13}{4}$ times that of getting an ace because the former can be obtained in 13 ways while the latter in 4 ways, though each occurrence has the same probability in a single draw.

The problem of finding out the number of ways in which a particular event occurs belongs in general to the topic of combinatorials and involves the knowledge of permutations and combinations that you are already familiar with. Suppose the event, we consider here, is the distribution of a set of objects among certain boxes following a set of prescriptions. Clearly, the answer will depend upon what type of distribution we want to

have. In the following paras we consider the arrangement of a certain number of non-interacting identical particles among the various states satisfying definite conditions. These arrangements are important in the context of statistical mechanics.

- (a) Suppose, we distribute n_i particles among g_i states so that a state contains at most one particle, where the particles are considered to be indistinguishable. The number of such arrangements is equal to $\frac{1}{n_i!} [g_i (g_i - 1) \dots (g_i - n_i + 1)]$ or, $g_{iC_{n_i}}$.

$$\text{Thus } W_i = \text{no. of such arrangements} = \frac{g_i!}{n_i! (g_i - n_i)!}.$$

[The factor $n_i!$ in the denominator comes due to distributions of n_i particles among themselves which because of their indistinguishability lead to a single arrangement]. Thus the total number of such arrangements $W = \prod_i W_i$. The equilibrium distribution corresponds to the most probable distribution,

i.e., $W = W_{\max}$, subject to the constraints that the total number of particles $\sum_i n_i$ and the total energy $\sum_i E_i n_i$ remain fixed. The result of the constrained maximum problem is the Fermi-Dirac distribution.

- (b) Suppose in the same distribution problem we impose the condition that a state has no restriction on the number of particles it can contain. Here, the number of distinct arrangements = $\frac{(g_i + n_i - 1)!}{(g_i - 1)! n_i!}$, where

$(g_i + n_i - 1)!$ is the number of all possible distributions of $(g_i - 1)$ partitions separating the g_i states from one another and n_i particles among themselves. The $(g_i - 1)$ imaginary partitions and n_i particles being indistinguishable; we have to divide by $(g_i - 1)! n_i!$ to obtain the number of distinct arrangements. The remaining procedure is the same as in the previous case, namely, the maximisation of $\prod_i \frac{(g_i + n_i - 1)!}{(g_i - 1)! n_i!}$ with the condition that $\sum_i n_i$ and $\sum_i E_i n_i$ remain fixed. The result is Bose-Einstein distribution.

- (c) For distinguishable particles, the number of possible distributions $W = \frac{N!}{\prod_i n_i!} \prod_i w_i$, where W_i is the

number of ways in which n_i distinguishable particles can be arranged among g_i states and given by $W_i = g_i^{n_i}$, so that $W = N! \prod_i \frac{g_i^{n_i}}{n_i!}$. The remaining procedure is the same as in the earlier two cases and

leads to Maxwell-Boltzman distribution. The exact forms of these distributions and the other results following from these are not discussed here. They are given in standard text books on statistical physics.

11.3 RANDOM VARIABLES, DISTRIBUTION FUNCTION ETCETRA

We have earlier learnt the basics of calculating the probabilities of various events. The moment we have stepped into the world of probabilities the questions that naturally arise are, “what is the expected outcome of the result of a large number of independent trials? What is the extent of the spread of the results about the average?” The remaining part of this chapter attempts to answer these questions which are of widespread

relevance in the various disciplines of science (including social science also) by introducing the concepts of random variable, probability distribution function, etc., and performing mathematical operations with them.

The outcomes of a set of independent experiments being subject, in general, to a ‘law of chance’ are called *random variables*. The random variables may be discrete or continuous. The number of times a coin falls on the ground in a given number of tosses with its head up or the number of occurrences of a given signal in a fixed period of time is an example of a *discrete* random variable. When the experiments are performed a large number of times we can define the probability of the occurrence of an event as,

$P(x) = p_i$ if the random variable $x = x_i$ or $= 0$ otherwise. Thus the probability $P(x)$ plotted against x shows a number of peaks at certain isolated points and vanishes elsewhere.

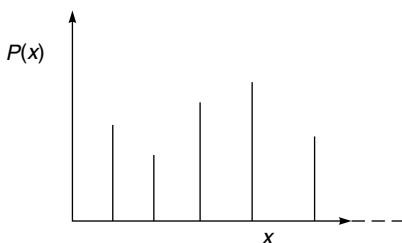


Fig. 11.1 (Typical probability distribution function of a discrete random variable).

$P(x)$ obtained by dividing the number of times that we get a certain value of x by the total number of trials has to satisfy the normalisation condition $\sum_i P(x_i) = 1$.

The time-period of a pendulum, the volume of liquid flowing out through a tube in a given amount of time etc., are examples of random *continuous* variables because in principle there is no restriction to the values of these variables. The probability distribution function is a continuous function of the variable.

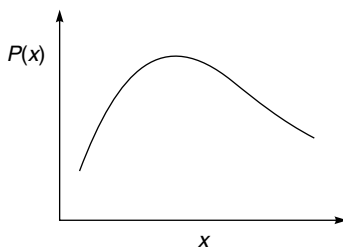


Fig. 11.2 (Typical probability distribution function of a continuous random variable)

Thus, the probability that x lies between x_1 and x_2 is given by $\int_{x_1}^{x_2} P(x) dx$, where $\int_0^{\infty} P(x) dx = 1$ from normalisation (Assume that $x > 0$ here).

11.2.1 Mean, Mode and Median

For any significant analysis of the data obtained from a set of experiments a knowledge of the three parameters (i) mean, (ii) mode and (iii) median of the probability distribution of the random variable is extremely important.

- (i) The mean of any discrete distribution is given by $m = \sum_{i=1}^n x_i P(x_i)$, where n is the number of independent trials, x_i , $i = 1, 2, \dots, n$ are the individual results and $P(x_i)$ the corresponding probabilities. If n_i be the number of events in which x_i occurs we can write, $P(x_i) = \frac{n_i}{n}$ and hence $m = \sum_{i=1}^n \frac{n_i x_i}{n}$. If the number of trials be very large, the mean m is the sum of an almost infinite series and finite if the series converges. The sum can be determined once we can fit the data to any theoretical distribution function. [Some of the well-known universally accepted distribution functions will be discussed later]. For a continuous distribution the mean m is found out from the relation $m = \int_{-\infty}^{\infty} x P(x) dx$ and is finite provided the integral converges. (It is of course assumed here that $\sum_i P(x_i) = 1$ and $\int_{-\infty}^{\infty} P(x) dx = 1$ so that the distribution functions, discrete or continuous are properly normalised.) The theoretically calculated mean value of any random variable coincides with the expected value of the average obtained from a large number of independent trials if the theoretical distribution function gives a good fit to the experimental data. For any experiment involving more than one random variable, the mean of any random variable x is found out from the relation $m_x = \sum_{i=1}^n x_i P(x_i, y_i, \dots)$ or theoretically from the formula

$$m_x = \int \dots \int x P(x, y, z, \dots) dx dy dz \dots,$$

where the distribution function is properly normalised.

- (ii) The mode of any distribution function is the value of the random variable for which the probability is maximum. If there be more than one values of x for which the probability is maximum then each one corresponds to the mode. For certain distribution functions (normal distribution function, for example, to be discussed later), the mode coincides with the mean.
- (iii) The median of a distribution function is the magnitude of the random variable x_0 for which $\int_{-\infty}^{x_0} P(x) dx = \frac{1}{2}$ or $\sum_{i=1}^n P(x_i)$, where n is such that $x_n = x_0$ equals $\frac{1}{2}$. [The integral or the sum is called the cumulative probability distribution].

Variance - Apart from the mean, mode and the median, another important characteristic of a distribution function is the variance that determines the spread of the data points about the average. It is defined by $\text{Var}[x] = \text{mean of } [x - \bar{x}]^2$, where $\bar{x}$ is the average. Thus $\text{Var}[x] = \sum_k (x_k - \bar{x})^2 P_k$, where P_k is the probability of obtaining a particular x_k . Using $\sum_k P_k = 1$ and $\bar{x} = \sum_k x_k P_k$ we get, $\text{Var}[x] = \sum_k x_k^2 P_k - \bar{x}^2 = \overline{x^2} - (\bar{x})^2$.

For a continuous distribution we define $\text{Var}[x] = \int_{-\infty}^{\infty} (x - \bar{x})^2 P(x) dx$. Using $\int_{-\infty}^{\infty} P(x) dx = 1$, $\int_{-\infty}^{\infty} x P(x) dx = \bar{x}$ and $\int_{-\infty}^{\infty} x^2 P(x) dx = \overline{x^2}$ we get, $\text{Var}[x] = \overline{x^2} - (\bar{x})^2$. [we can make use of Schwartz inequality (see Chapter 3) to prove that $\overline{x^2} - (\bar{x})^2 > 0$ both for the continuous and the discrete distribution functions]. Another quantity

which is important in the analysis of data, called the standard deviation σ , is given by $\sigma = \sqrt{\text{Var}(x)}$. In the next section we will discuss a few standard distribution functions that have proved to be useful in the statistical analysis of data obtained in diverse random phenomena. Before going into those we take a note of the following points that deserve mention.

- (1) The calculation of variance, as shown above, is valid also for a multivariate distribution with the modification that the integrals have to be evaluated in several dimensions over all possible values of the random variables concerned.
- (2) The probability density of a function $y(x)$ can be known given that of x .

Thus (i) for a discrete distribution we can write the density $F(y)$ as $F(y) = \sum_i P(x_i)$ if x_i 's are the values of x for which $y_i = f(x_i)$ for a fixed i and $F(y) = 0$ otherwise and (ii) for a continuous distribution we can write, $F(y) = P(x(y)) \left| \frac{dx}{dy} \right|$,

We get equivalent relations for a multivariate distribution with the difference that $\left| \frac{dx}{dy} \right|$ is to be replaced by the Jacobian $\left| \frac{\partial(x_1 \dots x_n)}{\partial(y_1 \dots y_n)} \right|$.

- (3) Chebyshev's inequality $\left(P(|x - m| \geq t) < \frac{\sigma^2}{t^2} \right)$ shows the concentrations of the data about the mean.
(Proof is omitted)

11.4 SOME STANDARD DISTRIBUTIONS

11.4.1 Binomial Distributions

Suppose we perform n independent trials in which the outcomes can be of two types—success and failure, the former referring to the situation when a particular event occurs and the latter when the event does not occur, the probability of success in each trial being same. These types of trials are called Bernoulli trials, the relevant distribution in this context being the binomial distribution. The repeated tossing of a coin in which there can be two results, the coin falling with the head at the top (success) and the head below (failure) with the probability of success equal to $\frac{1}{2}$ in each trial, provides an example of a Bernoulli trial. Thus if there be n independent trials, the probability of x successes, the probability of each success in a trial being p = Number of ways in which we get x out of a set of $n \times$ probability of x successes $\times$ probability of $(n - x)$ failures yielding $P(x) = nC_x p^x q^{n-x}$, where $q = 1 - p$ is the probability of a failure. These are nothing but the successive terms of the binomial expansion $(p + q)^n$ with $x = 0, 1, \dots, n$. (Fig. 11.3)

Thus $\sum P(x) = 1$.

$$\begin{aligned} \text{Mean } x = \langle x \rangle &= \sum_{x=0}^n x P(x) = np \sum \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x} \\ &= np(p + q)^{n-1} = np. \end{aligned}$$

$$\langle x^2 \rangle = \text{mean } x^2 = \sum_{x=0}^n x^2 P(x) \text{ and can be seen to be equal to}$$

$$np + n(n-1)p^2 = npq + n^2 p^2$$

Thus the variance $\text{Var}(x) = \langle x^2 \rangle - \langle x \rangle^2 = npq$.

An application of binomial distributions to physics is in the random walk problem leading to the Brownian motion. (The random walk problem is briefly discussed later)

11.4.2 Hypergeometric Distribution

In deducing the binomial distribution formula we assumed that the trials were independent and the probability of success remained the same throughout the trials. This requires that the initial conditions remain always unaltered. But often we encounter situations in which these conditions are not satisfied. For example, suppose we draw a number of balls out of a set of balls, perform some trial and do not replace the balls. The initial conditions change leading to the successive trials being dependent on one another. The corresponding distribution being a variant of the binomial distribution is called the hypergeometric distribution.

Suppose, out of a collection of N balls in any trial p are white and q black. The number of ways in which in a draw of n balls x are white and the remaining ones black is equal to $p_{C_x} N - p_{C_{n-x}}$. The total number of ways in which n balls can be chosen out of a set of N balls = N_{C_n} .

Thus, the probability of obtaining x white balls in n trials = $\frac{p_{C_x} N - p_{C_{n-x}}}{N_{C_n}}$

The mean and the variance of a hypergeometric distribution are given respectively by $\langle x \rangle = n \frac{P}{N}$ and $\text{Var}(x) = \left(\frac{N-n}{N-1} \right) np'(1-p')$

[Left as an exercise for the students], where $p' = \frac{p}{N}$.

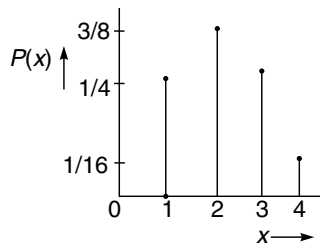


Fig. 11.3 A typical binomial distribution for $n = 4$, $p = 0.5$

11.4.3 Poisson Distribution

The Poisson distribution occurs as a limit of the binomial distribution when the number of independent trials is infinitely large, i.e., $n \rightarrow \infty$ and the probability of success p in each trial $p \rightarrow 0$ such np is finite. The phenomenon of radioactive decay from a finite mass of radioactive material or number of wrong telephone numbers dialled in a day gives an example of the Poisson distribution.

$$P(x) \text{ according to the binomial distribution} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$\lim_{n \rightarrow \infty} n_{C_x} = \lim_{n \rightarrow \infty} n(n-1) \dots (n-x+1) x! = n^x x!$$

$$\text{or, } P(x) \rightarrow \frac{(np)^x (1-p)^{n-x}}{x!} \rightarrow \frac{\lambda^x (1-p)^{n-x}}{x!} \text{ where } \lambda = \lim_{\substack{p \rightarrow 0 \\ n \rightarrow \infty}} np.$$

$$\text{Now } (1-p)^{n-x} \rightarrow \frac{(1-p)^n}{(1-p)^x} \rightarrow \left(1 - \frac{\lambda}{n}\right)^n \text{ or, } \rightarrow e^{-\lambda} \text{ as } n \rightarrow \infty, p \rightarrow 0.$$

$$\text{or, } P(x) \rightarrow \frac{\lambda^x e^{-\lambda}}{x!} \text{ which is the celebrated Poisson distribution.}$$

The Poisson distribution can also be deduced alternatively in the following way.

Let $P_x(t)$ be the probability of having x radioactive decays in a time interval t . If λ be the probability of a disintegration per unit time, the probability of having x such decays in an interval of time $t + \Delta t$ is given by

$$P_x(t + \Delta t) = P_x(t) (1 - \lambda \Delta t) + P_{x-1}(t) \cdot \lambda \Delta t.$$

$$\text{Or, } \frac{P_x(t + \Delta t) - P_x(t)}{\Delta t} = \lambda P_{x-1}(t) - \lambda P_x(t).$$

$$\text{In the limit } \Delta t \rightarrow 0 \text{ we get, } \frac{dP_x(t)}{dt} + \lambda P_x(t) = \lambda P_{x-1}(t) \quad (11.4)$$

$$\text{Clearly putting } x = 0 \text{ we get, } \frac{dP_0(t)}{dt} = -\lambda P_0(t)$$

$$\text{having the solution } P_0(t) = P_0(0) e^{-\lambda t}$$

$$\text{Putting in Eq. (11.1) we get, } \frac{dP_1(t)}{dt} + \lambda P_1 = \lambda P_0(0) e^{-\lambda t}$$

$$\text{or, } P_1 e^{\lambda t} = \lambda P_0(0) t + \text{constant.}$$

$$\text{Since at } t = 0, P_1 = 0 \text{ we have, } P_1 = \lambda t P_0(0) e^{-\lambda t},$$

$$\text{Proceeding in this way we obtain, } P_x(t) = \frac{(\lambda t)^x}{x!} e^{-\lambda t}, \text{ where } P_0(0) \text{ has been put equal to 1.}$$

$$\text{The mean } \langle x \rangle = \sum \frac{(\lambda t)^x x}{(x-1)!} e^{-\lambda t} = \lambda t \left[\sum_{n=0} \frac{(\lambda t)^n}{n!} = e^{\lambda t} \right]$$

$$\langle x^2 \rangle = \sum \frac{(\lambda t)^x x}{(x-1)!} e^{-\lambda t} = \sum \left[\frac{(\lambda t)^x}{(x-2)!} + \frac{(\lambda t)^x}{(x-1)!} \right] e^{-\lambda t} = \lambda^2 t^2 + \lambda t$$

Thus $\text{Var}(x) = \langle x^2 \rangle - \langle x \rangle^2 = \lambda t$.

For $t = 1$ the probability of obtaining x decays in unit time is given by $P_x(1) = \frac{\lambda^x}{x!} e^{-\lambda}$. A typical Poisson distribution for $\lambda = 1$ is shown.

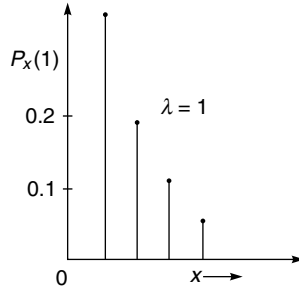


Fig. 11.4 A typical Poisson distribution for $\lambda = 1$.

11.4.4 Gaussian or Normal Distribution

It is a very popular continuous distribution utilized to analyse the data in various experiments involving random variables which form a continuous set. In the following we show that this distribution is obtained as the limit of the binomial and the Poisson distributions discussed above.

Gaussian from Binomial

The Gaussian distribution follows from the binomial distribution in the limit $n \rightarrow \infty$, $p = \text{finite}$ so that $np \rightarrow \infty$.

For the binomial distribution $P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$.

In the limit as $n \rightarrow \infty$ and $x \rightarrow \infty$ we get, putting $n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$, the Stirling approximation for large n .

We obtain, $P(x) = \frac{1}{\sqrt{2\pi n}} \left\{ \left(\frac{n}{e}\right)^n / \left[\left(\frac{x}{e}\right)^x \left(\frac{n-x}{e}\right)^{(n-x)} \right] \right\} \cdot p^x (1-p)^{n-x} \sqrt{x} \cdot \sqrt{n-x}$

Or, $P(x) = \frac{1}{\sqrt{2\pi n}} \left(\frac{x}{n}\right)^{-n-\frac{1}{2}} \left(\frac{n-x}{n}\right)^{-n+x-\frac{1}{2}} p^x (1-p)^{n-x}$

$$= \frac{1}{\sqrt{2\pi n}} \exp \left[- \left(x + \frac{1}{2} \right) \ln \frac{x}{n} - \left(n - x + \frac{1}{2} \right) \ln \frac{n-x}{n} + x \ln p + (n-x) \ln (1-p) \right]$$

Putting $1 \ll x - np \ll np$ and keeping the leading terms, we get

$$P(x) \approx \frac{1}{\sqrt{2\pi n}} \cdot \frac{1}{\sqrt{p(1-p)}} \exp \left[- \frac{1}{2} \frac{(x - np)^2}{np(1-p)} \right].$$

$$\text{Mean } x = \langle x \rangle = \frac{1}{\sqrt{2\pi n}} \cdot \frac{1}{\sqrt{p(1-p)}} \int_{-\infty}^{\infty} x \exp \left[-\frac{1}{2} \frac{(x-np)^2}{np(1-p)} \right] dx.$$

Using $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$ we get $\langle x \rangle = np$ and similarly

$$\text{Var}(x) = \langle x^2 \rangle - \langle x \rangle^2 = np(1-p).$$

The Gaussian distribution is symmetric about the mean $x = np$.

Gaussian from Poisson

Since the Gaussian distribution follows from the binomial distribution as the mean $np \rightarrow \infty$ it is expected that the Gaussian distribution can be obtained from Poisson distribution $P_x = \frac{\lambda^x}{x!} e^{-\lambda}$ ($t=1$) as $\lambda \rightarrow \infty$.

From $P_x = e^{-\lambda} \frac{\lambda^x}{x!}$ we get, $\ln P_x = -\lambda + x \ln \lambda - \ln x!$

Using Stirling approximation, $x! = \sqrt{2\pi x} \left(\frac{x}{e}\right)^x$ we get,

$$\ln P_x = -\lambda + x \ln \lambda - (x \ln x - x) - \ln \sqrt{2\pi x}.$$

Since the Poisson distribution function peaks at $x = \lambda$ we get by putting $x = \lambda + \epsilon$

$$\ln P_x = -\lambda + (\lambda + \epsilon) \ln \lambda + (\lambda + \epsilon) - (\lambda + \epsilon) \ln (\lambda + \epsilon) - \ln \sqrt{2\pi (\lambda + \epsilon)}$$

Using $\ln(1+x) = x - \frac{x^2}{2} + \dots$ we obtain,

$$\ln P_x = -\frac{\epsilon^2}{2\lambda} - \ln \sqrt{2\pi \lambda}, \text{ keeping terms to } 0 (\epsilon^2).$$

$$\text{or, } P_x = \frac{1}{\sqrt{2\pi \lambda}} \exp \left[\frac{-\epsilon^2}{2\lambda} \right] = \frac{1}{\sqrt{2\pi \lambda}} \exp \left[-\frac{(x-\lambda)^2}{2\lambda} \right], \text{ which is the Gaussian distribution function.}$$

11.5 APPLICATION

11.5.1 Fokker-Planck and Diffusion Equation

Apart from the applications to the analysis of data in the experimental science the study of probability is of immense importance in theoretical science, statistical physics being one of the important disciplines for example. The Bose-Einstein and the Fermi-Dirac distributions (Section 11.2) form the foundations of quantum statistics. One of the other applications is the problem of diffusion in statistical physics. The problem of diffusion is the outcome of a random walk phenomenon, which a molecule undergoes due to collisions with the other molecules in its motion in the surrounding medium.

Let, p , in the one-dimensional random walk problem, be the probability of a step towards the right (increasing x) and $q = 1 - p$ be that of a step to the left respectively. Suppose the length of a step covered in time $t_0 = l$.

The probability that in a total of n steps m steps are taken to the right and $(n - m)$ to the left is $P(n, m) = n C_m p^m q^{n-m}$ (binomial distribution) so that we can then write, $P(x, t + t_0) = p P(x - l, t) + q P(x + l, t)$, where $P(x, t)$ stands for the probability that a molecule covers a path length x in time t .

$$\text{or,} \quad P(x, t + t_0) - P(x, t) = p[P(x - l, t) - P(x, t)] + q[P(x + l, t) - P(x, t)]$$

In the limit $t_0 \ll t$ we can write,

$$t_0 \frac{\partial P}{\partial t} = p \left[-l \frac{\partial P}{\partial x} + \frac{l^2}{2} \frac{\partial^2 P}{\partial x^2} \right] + q \left[l \frac{\partial P}{\partial x} + \frac{l^2}{2} \frac{\partial^2 P}{\partial x^2} \right]$$

$$\text{or,} \quad \frac{\partial P}{\partial t} = - \frac{(p - q)l}{t_0} \frac{\partial P}{\partial x} + \frac{1}{2} \frac{l^2}{t_0} \frac{\partial^2 P}{\partial x^2}$$

In the limit $t_0, l \rightarrow 0$ so that $\frac{l^2}{2t_0} \rightarrow D$, a finite constant and $\frac{(p - q)l}{t_0} \rightarrow \alpha$, a constant, we obtain

$$\frac{\partial P}{\partial t} = - \alpha \frac{\partial P}{\partial x} + D \frac{\partial^2 P}{\partial x^2}, \text{ the Fokker-Planck equation.}$$

If we set $p = q = 1$ we get, $\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial x^2}$, the diffusion equation.

The extension to three dimensions is obvious, where $P = P(\vec{r}, t)$ giving $\frac{\partial P}{\partial t} = D \nabla^2 P$.

11.6 THEORY OF ERRORS

A student, asked to perform a certain experiment designed to measure certain physical variables takes a lot of data and records the mean as the magnitude of the variable. His experiment is always associated with random errors resulting from the fluctuations of background conditions or unavoidable human errors even when other sources of error, such as those coming from uses of faulty instruments, defective calibrations or wrong theoretical formulae etc. have been eliminated. The experimental data may thus be taken as the random variables and since the actual value of the variable under measurement is unknown it is pertinent to ask "How reliable is the estimation of the variable?" The present section attempts to address this question.

Suppose we do a very large number of experiments and choose a representative sample as the n -member

set consisting of $x_1, x_2, \dots, x_n$. The average value, also called the sample mean is given by $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

The deviation of the m th. reading is given by $d_m = x_m - \bar{x}$.

The error in the m th. reading is defined as $e_m = x_m - x_0$, x_0 being the true reading. The standard error of this

distribution in the sample space is given by $\sigma = \left[\frac{1}{n} \sum_i e_i^2 \right]^{\frac{1}{2}}$.

The standard deviation of the sample σ_s is given by $\sigma_s = \left[\frac{1}{n} \sum_m d_m^2 \right]^{\frac{1}{2}}$

Now, $d_m = x_m - \bar{x} = (e_m + x_0) - (e_0 + x_0)$, ($e_0 = \bar{x} - x_0$ represents the error of the mean) $= e_m - e_0$

$$\text{Thus } \sigma_s^2 = \frac{1}{n} \sum_m d_m^2 = \frac{1}{n} \sum (e_m - e_0)^2 = \frac{1}{n} \sum e_m^2 + e_0^2 - \frac{2e_0}{n} \sum e_m.$$

$$\text{or, } \sigma_s^2 = \frac{1}{n} \sum e_m^2 - e_0^2, \text{ where } e_0 = \frac{1}{n} \sum (x_m - x_0) = \frac{1}{n} \sum e_m$$

$$\text{Now, } e_0^2 = \frac{1}{n^2} \sum e_m^2 + \frac{1}{n^2} \sum \sum_{i \neq j} e_i e_j$$

$$\text{or, } \langle e_0^2 \rangle_{av} = \frac{1}{n^2} \left(\sum e_m^2 \right)_{av} + \frac{1}{n^2} \left(\sum \sum_{\substack{i, j \\ i \neq j}} e_i e_j \right)_{av}$$

and because e_i 's can be equally +ve or -ve, the second term on the r.h.s. Vanishes and hence

$$\langle e_0^2 \rangle_{av} = \frac{\langle e^2 \rangle_{av}}{n} \quad (11.5)$$

Putting $\langle e_0^2 \rangle_{av} = \sigma_0^2$ and $\langle e^2 \rangle_{av} = \sigma^2$ we get, $\sigma_0 = \frac{\sigma}{\sqrt{n}}$, where

$\sigma_0^2 = \langle e_0^2 \rangle_{av}$ gives the mean of the squares of the individual errors over all the sets. Since the time value x_0 of the variable is unknown, the relation (11.5) is not of much practical use. Now, we can write

$$(\sigma_s^2)_{av.} = \frac{1}{n} \left(\sum e_m^2 \right)_{av.} - \langle e_0^2 \rangle_{av.} = \sigma^2 - \sigma_0^2.$$

Using $\sigma_0 = \frac{\sigma}{\sqrt{n}}$ we obtain, $\sigma^2 = \frac{n}{n-1} (\sigma_s^2)_{av.}$

$$\text{or, } \sigma = \sqrt{\frac{n}{n-1} (\sigma_s^2)_{av.}} = \sqrt{\frac{\sum (x_m - \bar{x})^2}{n-1}}$$

$$\text{or, } \sigma_0 = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sum (x_m - \bar{x})^2}{n(n-1)}}, \text{ where } \sigma_s^2 \text{ has to be taken to be almost equal to } (\sigma_s^2)_{av.} \quad (11.6)$$

Thus, the standard error of the mean is calculated from Eq. (11.6) in terms of the experimentally measurable parameters on the right hand side.

We can similarly write, $d_0 = \frac{1}{n} \sum d_m$, giving $d_0^2 = \frac{1}{n^2} \sum d_m^2$

[because the cross terms cancel out as before].

$$\text{Thus } (d_0^2)_{av} = \left[\frac{1}{n^2} \left(\sum d_m^2 \right) \right]_{av} = \frac{(d^2)_{av}}{n} = \frac{1}{n} \sigma_s^2$$

$$\text{Hence we obtain, } \sigma_{\text{mean}} = \sqrt{(d_0^2)_{av}} = \sqrt{\frac{\sigma_s^2}{n}} \quad (11.7)$$

The results deduced hitherto are valid independently of any specific type of distribution, used to fit the data. But a knowledge of the distribution function is essential for the determination of the various moments of the distribution, (i.e., expressions of the form

$$\int_{-\infty}^{\infty} (x - \bar{x})^r f(x) dx, f(x) \text{ being the distribution function, } r \text{ and } \bar{x} \text{ the order of the moment and the mean}$$

respectively) and the estimation of the reliability of the data, which is needed to ascertain which are the relevant readings to be retained for any subsequent analysis.

For example for a Gaussian distribution given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{(x - \bar{x})^2}{2\sigma^2} \right] \text{ we obtain by numerical integration that } P(-\sigma < |\bar{x} - x| < \sigma) = 0.68,$$

$$P(-2\sigma < |\bar{x} - x| < 2\sigma) = 0.954, P(-3\sigma < |\bar{x} - x| < 3\sigma) = 0.997.$$

Thus it is almost cent per cent probable that the data points will lie within an interval of length 3σ centred about $x = \bar{x}$.

It will be hence quite safe to reject the data outside the interval for any subsequent calculation.

Now it can be shown (the proof is omitted here) that independent of the form of the scatter of the initial data in an infinitely extended ensemble the sample mean, i.e., the mean of a finite but large collection of points taken out of the original set follows a Gaussian distribution. (This is called the central limit theorem). Thus the fact that the features of a Gaussian distribution have been well-studied enables one to make reasonable statistical predictions about the observations.

11.6.1 Propagation of Errors

Often the physical quantity of interest is determined in terms of several independent variables which are assumed to be measured a large number of times with their characteristic standard deviations. The problem is how to evaluate the resulting standard deviation in the measurement of the desired physical variable in terms of the deviations of the others on which depends the required variable.

Let y be the desired variable which depends on the independent variables $x_1, x_2, \dots, x_n$ so that we can write, $y = F(x_1, x_2, \dots, x_n)$. Let the mean values of $x_1, \dots, x_n$ be denoted by $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$ so that $y = F(\bar{x}_1 + \xi_1, \bar{x}_2 + \xi_2, \dots, \bar{x}_n + \xi_n)$

$$\text{or, } y = y_0 + \sum_i \xi_i \left. \frac{\partial F}{\partial x_i} \right|_{\bar{x}_i} + \frac{1}{2} \sum_{i,j} \xi_i \xi_j \left. \frac{\partial^2 F}{\partial x_i \partial x_j} \right|_{\bar{x}_i}, \text{ where}$$

$$y_0 = F(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n).$$

Hence $\langle y \rangle = y_0 + \frac{1}{2} \sum_{i,j} \langle \xi_i \xi_j \rangle \left[\frac{\partial^2 F}{\partial x_i \partial x_j} \right]_{\bar{x}_i}$, Since $\langle \xi_i \rangle = \langle x_i - \bar{x}_i \rangle = 0$

The error in putting $y = y_0$ is defined in the lowest order by $\langle (y - y_0)^2 \rangle = \sum_{i,j} \delta_i \delta_j \langle \xi_i \xi_j \rangle$, where $\delta_i = \left. \frac{\partial F}{\partial x_i} \right|_{\bar{x}_i}$
If the variables $x_i, x_j, \dots$ are independent of one another,

$$\langle \xi_i \xi_j \rangle = \langle \xi_i \rangle \langle \xi_j \rangle = 0 \text{ for } i \neq j \text{ and } = \langle \xi_i \rangle^2 \text{ for } i = j.$$

Thus, $\sigma_y^2 = \langle (y - y_0)^2 \rangle = \sum_i \delta_i^2 \langle \xi_i \rangle^2 = \sum_i \delta_i^2 \sigma_{x_i}^2$, where $\sigma_{x_i}^2$ denotes the variance in the measurement of x_i .

This is the principle of propagation of random errors. For example in the determination of the coefficient of viscosity by Poissenllie's method the relevant formula is:

$$\eta = \frac{\pi P r^4}{8 l l}. \text{ Thus } \eta = K \frac{P r^4}{8}, \text{ where } K = \frac{\pi}{8 l} \text{ is assumed to be supplied. } P_0 \text{ is the pressure difference}$$

between two ends of the tube whose radius is r and the volume of liquid flowing out through it per unit time is v .

$$\text{Thus } \sigma_\eta^2 = \frac{\pi}{8 l} \left[\left\langle \frac{r^4}{V} \right\rangle \sigma_P^2 + \left\langle \frac{4 P r^3}{V} \right\rangle \sigma_r^2 + \left\langle \frac{P r^4}{V^2} \right\rangle \sigma_V^2 \right], \text{ where the quantities within the brackets}$$

$$\left[\left\langle \frac{r^4}{V} \right\rangle, \left\langle \frac{P r^3}{V} \right\rangle, \left\langle \frac{P r^4}{V^2} \right\rangle \right] \text{ denote their mean values.}$$

11.6.2 Graph Drawing, Least Square Fit

A graph is an elegant way of displaying the results of observations. Suppose we are required to draw a curve showing the dependent variable as a function of the random variable x and a number of parameters $a, b, \dots$

Let the theoretical formula we want to fit the data be given by $y_{0i} = f(x_i, a, b, \dots)$, and let $y = y_i \pm \sigma_i$ for all i .

$$\text{The quantities } y_i \text{ have Gaussian distributions given by } f_i(y_i) = (2\pi\sigma_i^2)^{-1/2} \exp \left[- \left(\frac{y_i - y_{0i}}{\sqrt{2}\sigma_i} \right)^2 \right].$$

The best estimate for the parameters $a, b, \dots$ corresponds to the maximisation of $\prod_{i=1}^n f_i(y_i)$ and thus to the minimum of $\sum_{j=1}^n \left(\frac{y_j - y_{0j}}{\sigma_j} \right)^2$. If σ_j 's are all equal this condition reduces to the least value of $\sum_i (y_i - y_{0j})^2$.

If there be two parameters a and b only, the minimisation of $\sum_i (y_i - y_{0j})^2$ with respect to the variations of a and b respectively gives, where $y_{0j} = a + bx_j$,

$$\frac{\partial}{\partial a} \left[\sum_i (y_i - y_{0j})^2 \right] = \frac{\partial}{\partial b} \left[\sum_j (y_j - y_{0j})^2 \right] = 0$$

Let us assume, $y_{0j} = a + bx_j$ for all j .

$$\begin{aligned}\text{Now, } \sum_j (y_j - y_{0j})^2 &= \sum_j (y_j - a - bx_j)^2 \\ &= \sum_j y_j^2 + na^2 + b^2 \sum_j x_j^2 - 2b \sum_j x_j y_j + 2ab \sum_j x_j - 2a \sum_j y_j\end{aligned}$$

$$\text{Putting } S = \sum_j (y_j - y_{0j})^2, \left| \frac{\partial S}{\partial a} \right|_{\bar{a}, \bar{b}} = 2n\bar{a} + 2\bar{b} \sum x_j - 2 \sum y_j = 0$$

$$\left| \frac{\partial S}{\partial b} \right|_{\bar{a}, \bar{b}} = 2\bar{b} \sum x_j^2 - 2 \sum x_j y_j + 2\bar{a} \sum x_j = 0$$

Solving the two equations we get,

$$\begin{aligned}\bar{a} &= \frac{\sum x_j^2 \sum y_j - \sum x_j \sum x_j y_j}{n \sum x_j^2 - (\sum x_j)^2}, \\ \bar{b} &= \frac{n \sum x_j y_j - \sum x_j \sum y_j}{n \sum x_j^2 - (\sum x_j)^2}\end{aligned}$$

Thus knowing a and b we can draw the best curve fitting the data with a linear relation. The above method works for any relation in y and x involving the parameters $a, b, c, \dots$ linearly.

(for example, $y = a + bx + cx^2$). In many cases, a relation between y and x involving the parameters $a, b, c, \dots$ in a non-linear way can by a transformation of variables be changed to a relation which involves linear functions of the parameters, so that the above procedure can be applied.

Only a least square fit with the experimental data does not imply that the assumed distribution is correct for the explanation of the data. There is a popular test called the χ^2 -test, briefly introduced in the appendix, that can give a clue to the correctness of our hypothesis.

SOLVED PROBLEMS

Problem 1

One of the two boys can solve 80% of the problems of a book while the other can do 70% of them. What is the probability that a problem selected at random can be solved if both the boys try?

Solution: Let A and B be two events corresponding to the individual successes of them.

Then $P(A \cup B)$, the desired probability $= P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A) \cdot P(B)$ since two events being independent $P(A \cap B) = P(A) \cdot P(B)$.

Thus $P(A \cup B) = 0.80 + 0.70 - 0.80 \times 0.70 = 0.94$

Problem 2

A box has 12 red balls and 8 black balls. A ball is selected at random from the box.

If the ball is red it is replaced while if it be black it is not returned. Find the probability that a second ball from the box is (a) red, (b) black.

Solution: Let A_i and B_i be the respective events corresponding to the i th-ball drawn being red and black respectively.

The required probability in (a) is $P(A_2) = P(A_1) P(A_2/A_1) + P(B_1) P(A_2/B_1)$

$$\text{Or, } P(A_2) = \frac{12}{20} \times \frac{12}{20} + \frac{8}{20} \times \frac{12}{19} = \frac{9}{25} + \frac{24}{95} = 0.61$$

The probability in (b) is $P(B_2) = P(A_1) P(B_2/A_1) + P(B_1) P(B_2/B_1)$

$$\text{or, } P(B_2) = \frac{12}{20} \times \frac{8}{20} + \frac{8}{20} \times \frac{7}{19} = \frac{6}{25} + \frac{14}{95} = 0.38$$

Problem 3

Three boxes contain silver and copper coins. The percentages of the total number

of coins are respectively 20, 30 and 50, the fractions of silver coins in the box being 0.2, 0.3 and 0.5 respectively. If a coin be taken at random from anyone box, what is the probability that it is a silver coin?

Solution: $P(S) = \sum_{n=1}^3 P(B_n) P(S/B_n)$, where S and B_n stand for silver and the n th. box respectively.

$$\text{or, } P(S) = 0.2 \times 0.2 + 0.3 \times 0.3 + 0.5 \times 0.5 = 0.38.$$

Problem 4

A target is hit independently by two rifles. The probability of the first one hitting

the target is 0.7 and that of the second 0.4. If the target is destroyed by a single shot, determine the probability that it is destroyed by the first and the second rifles respectively.

Solution: There can be four types of events.

(a) both the rifles failed, (b) both succeeded, (c) the first one succeeded but the second one failed, (d) the first one failed but the second one succeeded.

$$P(a) = (1 - 0.7) (1 - 0.4) = 0.18, P(b) = 0.7 \times 0.4 = 0.28$$

$$P(c) = 0.7(1 - 0.4) = 0.42, P(d) = (1 - 0.7) 0.4 = 0.12.$$

$$\text{Let the event } A \text{ be the destruction of the target by a single hit. } P(c/A) = \frac{0.42 \times 1}{0.18 + 0.28 + 0.42 + 0.12} = \frac{0.42}{1} = 0.42$$

$[P(c/A)$ is the probability of destruction by the first]

$$P(d/A) = \text{Prob. of destruction by the second one} = \frac{0.12}{1} = 0.12.$$

Problem 5

The distribution functions of two random variables X and Y independent of each

other are given by $Af(x)$ and $Bg(y)$ respectively. Calculate the probability that (a) X lies between a and b , (b) Y lies in the interval $c < Y < d$, (c) X lies between x_1 and x_2 . and Y between y_2 and y_1 [A and B are constants].

Solution: Normalising the total probability distribution function we get, $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} AB f(x) g(y) dx dy = 1$,
 which gives, $AB = \frac{1}{\int_{-\infty}^{\infty} f(x) dx \int_{-\infty}^{\infty} g(y) dy}$

(a) Probability that x lies between a and b

$$= AB \int_a^b f(x) dx \int_{-\infty}^{\infty} g(y) dy = \frac{\int_a^b f(x) dx}{\int_{-\infty}^{\infty} f(x) dx}$$

(b) Probability that y lies between c and d

$$= AB \int_{-\infty}^{\infty} f(x) dx \int_c^d g(y) dy = \frac{\int_c^d g(y) dy}{\int_{-\infty}^{\infty} g(y) dy}$$

(c) Probability that X lies between x_1 and x_2 and Y between y_1 and y_2

$$= AB \int_{x_1}^{x_2} f(x) dx \int_{y_1}^{y_2} g(y) dy = \frac{\int_{x_1}^{x_2} f(x) dx \int_{y_1}^{y_2} g(y) dy}{\int_{-\infty}^{\infty} f(x) dx \int_{-\infty}^{\infty} g(y) dy}$$

Problem 6

The position distribution function of a particle is given by $A r e^{-r/a}$, A and a being constants. Find out (i) the mean position. (ii) the variance of this distribution function, (iii) the mode of the function.

Solution:

(i) Normalising the distribution function we get,

$$\int_0^{\infty} A r e^{-r/a} 4\pi r^2 dr = 1, \text{ yielding } 4\pi A a^4 \Gamma(4) = 1, \text{ or, } A = \frac{1}{24\pi a^4}.$$

The mean position of the particle is given by $\int_0^{\infty} 4\pi A r^4 e^{-r/a} dr$, which is equal to $4\pi A a^5 \Gamma(5)$ or, $4a$.

(ii) The mean r^2 is given by $\langle r^2 \rangle = \int_0^\infty 4\pi A r^5 e^{-r/a} dr$ or $20 a^2$.

Thus variance $= \langle r^2 \rangle - \langle r \rangle^2 = 4a^2$.

(iii) The mode of the function is given by $\frac{d}{dr} [r e^{-r/a}] = 0$.

Thus the mode is at $e^{-r/a}(1 - r/a) = 0$ or $r = a$.

[The second derivative is minimum here implying a maximum]

Problem 7

(a) The probability function $f(x) = \frac{1}{a} e^{-x/a}$ [$0 < x < \infty$]. If $y = x^2$,

Calculate the probability of y lying between y_1 and y_2 .

(b) Calculate $\text{Var}(aX + bY)$ from $\text{Var}(X)$ and $\text{Var}(Y)$ respectively if X and Y are two independent random variables.

Solution:

(a) If the probability function in y be given by $g(y)$ we obtain, $g(y) dy = f(x) dx$, so that $g(y) = \frac{f(x)}{2x}$
[Since $dy = 2x dx$]

$$\text{The probability that } y \text{ lies between } y_1 \text{ and } y_2 = \int_{x_1}^{x_2} f(x) dx = \int_{\sqrt{y_1}}^{\sqrt{y_2}} f(x) dx = \int_{y_1}^{y_2} g(y) dy$$

Or, the required probability is given by

$$P = \frac{1}{a} \int_{\sqrt{y_1}}^{\sqrt{y_2}} e^{-x/a} dx = - \left[e^{-x/a} \right]_{\sqrt{y_1}}^{\sqrt{y_2}} = \left[\exp \left(-\frac{1}{a} \sqrt{y_1} \right) - \exp \left(-\frac{1}{a} \sqrt{y_2} \right) \right]$$

(b) $\text{Var}(aX + bY) = \langle (aX + bY)^2 \rangle - \langle aX + bY \rangle^2$
 $= a^2 [\langle X^2 \rangle - \langle X \rangle^2] + b^2 [\langle Y^2 \rangle - \langle Y \rangle^2] + 2ab [\langle XY \rangle - \langle X \rangle \langle Y \rangle]$
 $= a^2 \text{Var}(X) + b^2 \text{Var}(Y)$; the last term vanishes due to mutual independence of X and Y .

Problem 8

(a) A coin is tossed n times. Find out the probability of obtaining all possible values of the heads.

(b) There are p typographical errors in a book of n pages. Assuming the errors to be randomly distributed, what is the probability that (i) a given page will have more than 3 errors, (ii) q pages will have no error?

Solution:

(8) (a) This is an example of binomial distribution. The probability of getting a head in a toss $= \frac{1}{2}$.
Probability of getting heads in r throws

$$= n_{cr} \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{n-r}, \text{ where } r = 0, 1, \dots$$

- (b) This is an example of Poisson distribution since the number of events is large and the probability of each event is negligibly small.

$$P = \frac{\lambda^x}{x!} e^{-\lambda}, \text{ where } \lambda = \lim_{\substack{p \rightarrow 0 \\ n \rightarrow \infty}} np$$

$$(i) P(x > 3) = 1 - P(x = 0) - P(x = 1) - P(x = 2) - P(x = 3),$$

$$= 1 - \left[e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda^2}{2!} e^{-\lambda} + \frac{\lambda^3}{3!} e^{-\lambda} \right] = 1 - e^{-\lambda} \left[1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} \right]$$

- (ii) $P(x = 0) = e^{-\lambda}$ gives the probability that a page will have no errors. Thus the probability that q page will be free from errors $= e^{-q\lambda}$.

SUMMARY

- The chapter introduces the basic concepts of probability and statistics which are used in the various disciplines of human knowledge. The probability of an event is defined as the number of times the event occurs divided by the total number of experiments involving the event. The probability of an experiment involving more than one events depends on how these events are mutually correlated. For example, the probability of two events occurring simultaneously is zero if they are mutually exclusive while it is the product of two independent probabilities if they are independent. The various outcomes of an experiment can be illustrated well by Venn diagrams shown in the text.
- Each individual arrangement having the same probability, the net probability is proportional to the number of ways in which the distribution is obtained. The number of possible arrangements is calculated in the text for a set of identical particles satisfying Bose-Einstein, Fermi-Dirac and Maxwell-Boltzmann distributions under different constraints as explained in the text and these results are extensively used in Statistical Mechanics. A few distribution relevant in the context of various experiments, namely, Binomial, Poisson, Hypergeometric and Gauss (or normal) distributions involving random variables, have been discussed in the text.
- The probability distributions referred to earlier can be applied to the statistical analysis of the experimental data. If we are given a huge collection of data regarding a random variable we can calculate the average value and the variance, given by $[(x^2)_{av.} - (x)_{av.}^2]$, x being the magnitude of the random variable in a single trial, which shows how the results are distributed about the average. If we can, however, fit the data with a theoretical distribution we can determine the mean, variance (or standard deviation which is $\sqrt{\text{Variance}}$), mode, (magnitude of the variable corresponding to the maximum probability) etc. from the distribution itself by straightforward calculation. The mean (also called the expectation value or average) and the variance have been calculated in the text for the standard distributions stated earlier. The Gauss distribution deserves special mention in this regard because the mean of a number of data follows the Gaussian distribution for a large number of independent experiments involving the random variable irrespective of the distribution of the initial set of data with standard deviation equal

to $\frac{\sigma}{\sqrt{n}}$, where n is the number of data points over which the average has been taken and σ refers to

the standard deviation computed for the initial set. The chapter concludes with (i) a brief discussion of the propagation of the errors that shows how the error in the calculation of a function of the random variables is related to errors of the determination of the individual variables, (ii) reliability of the data which tells us which data are to be retained and which may safely be rejected, (iii) the correct way of drawing a curve that serves as the best fit with the experimental data.

MULTIPLE CHOICE QUESTIONS

- When two dice are thrown simultaneously the probability of getting the sum 9 or 10 is
 (a) $\frac{8}{36}$, (b) $\frac{7}{36}$, (c) $\frac{5}{36}$, (d) $\frac{11}{36}$
- The probability of having at least one head in four throws of a coin is
 (a) $\frac{15}{16}$, (b) $\frac{1}{16}$, (c) $\frac{1}{4}$, (d) $\frac{1}{2}$
- An unbiased coin is tossed thrice. The probability that the head turns up in exactly two such events is
 (a) $\frac{1}{8}$, (b) $\frac{1}{16}$, (c) $\frac{1}{4}$, (d) $\frac{3}{8}$
- If $P(A) = 0.3$, $P(A \cup B) = 0.6$ and A, B are independent, then $P(B)$ is
 (a) $\frac{1}{3}$, (b) $\frac{1}{5}$, (c) $\frac{3}{7}$, (d) $\frac{1}{7}$
- If for two events E_1 and E_2 , $P_r(E_1) = \frac{1}{2}$, $P_r(E_2) = \frac{1}{4}$ and the probability of E_1 and E_2 both happening together is $\frac{1}{5}$. Which of the following statements is true?
 (a) $P_r(E_1 \text{ or } E_2) = \frac{2}{3}$, (b) The two events are independent,
 (c) The two events are not independent, (d) $P(E_1/E_2) = \frac{3}{4}$
- A committee of 6 persons is formed from 9 gentlemen and 4 ladies. What is the probability that the committee consists of (i) exactly 3 ladies (ii) at least one lady?
 (a) $\frac{28}{143}$, $\frac{136}{143}$ (b) $\frac{115}{143}$, $\frac{136}{143}$ (c) $\frac{15}{143}$, $\frac{7}{143}$ (d) None of these.
- Two dice are tossed n times. What is the probability that the first p tosses give a sum of 6 and the last q tosses a sum of 8.
 (a) $\frac{6^p \cdot 8^q}{2^n}$, (b) $\frac{5^{p+q}}{36^n}$, (c) $\frac{6^p \cdot 8^q}{36^n}$, (d) $\frac{5^{p+q}}{2^n}$.

8. The probability of getting an ace or a king from a pack of cards when (i) the first card is replaced after the draw (ii) the first card is not replaced.
- (a) $\frac{4}{169}, \frac{4}{169}$ (b) $\frac{1}{169}, \frac{1}{221}$ (c) $\frac{4}{169}, \frac{14}{663}$ (d) $\frac{1}{169}, \frac{14}{663}$
9. Items out of a set of 10 mangoes are defective. What is the probability that 3 out of any set of 4 mangoes will prove to be good?
- (a) $\frac{1}{3}$, (b) $\frac{4}{8}$, (c) $\frac{8}{21}$, (d) $\frac{1}{2}$
10. The probability of destroying a target in one hit is equal to $\frac{2}{3}$. The number of hits required to destroy the target with probability greater than or equal to $\frac{8}{9}$ is
- (a) 9 (b) 3 (c) 1 (d) 2
11. In a machine consisting of 3 parts the probabilities of defects in the three parts are p_1, p_2 and p_3 respectively. The probability that the entire machine will be free from any defect is
- (a) $(1 - p_1)(1 - p_2)(1 - p_3)$
 (b) $(1 - p_1)(1 - p_2) + (1 - p_1)(1 - p_3) + (1 - p_2)(1 - p_3)$
 (c) $1 - \frac{1}{3}(p_1 + p_2 + p_3)$
 (d) $\left(1 - \frac{1}{3}p_1\right)\left(1 - \frac{1}{3}p_2\right)\left(1 - \frac{1}{3}p_3\right)$
12. The probability that an event will occur twice (i) in two trials, (ii) in ten trials if the probability of occurrence of the event in each trial is 0.3. [Hint:- Think of binomial distribution].
- (a) 0.09, 0.23 (b) 0.09, $(0.3)^{10}$ (c) 0.6, $(0.3)^{10}$ (d) None of these.
13. There are p typographical errors in a book of n pages. The probability that a given page will have more than two errors is,
- (a) λ^2 (b) $1 - (1 - \lambda)^2$ (c) $1 - e^{-\lambda} \left(1 + \lambda + \frac{\lambda^2}{2}\right)$ (d) None of these.
- where $\lambda = p/n$.
14. The mean and the variance of a binomial distribution $(p + q)^n = \sum_{r=0}^n {}^nC_r p^r q^{n-r}$ are given by
- (a) $\frac{p+q}{2}, npq$ (b) $np, \sqrt{npq}$ (c) np, npq (d) $np, \sqrt{npq}$
15. The length and the radius of a cylinder are l and r respectively with probable errors $\pm e_1$ and $\pm e_2$ respectively. The probable errors in the measurement of the curved surface area and the volume are
- (a) $\pm \sqrt{l^2 e_1^2 + r^2 e_2^2}, \pm \pi r \sqrt{e_1^4 + 4l^2 e_2^2}$, (b) $\pm \sqrt{l^2 e_2^2 + r^2 e_1^2}, \pm \pi r \sqrt{4e_1^2 r^2 + l^2 e_2^2}$
 (c) $\pm 2\pi \sqrt{l^2 e_2^2 + r^2 e_1^2}, \pm \pi r \sqrt{4r^2 e_1^2 + l^2 e_2^2}$ (d) $\pm 2\pi \sqrt{l^2 e_2^2 + r^2 e_1^2}, \pm \pi r \sqrt{4l^2 e_2^2 + r^2 e_1^2}$

16. The binomial distribution with mean 20 and variance 12 is
 (a) $\left(\frac{2}{5} + \frac{3}{5}\right)^{400}$ (b) $\left(\frac{2}{5} + \frac{3}{5}\right)^{50}$ (c) $\left(\frac{2}{5} + \frac{3}{5}\right)^{20}$ (d) $\left(\frac{3}{5} + \frac{2}{5}\right)^{12}$
17. The mean and the standard deviation of a binomial distribution are 10 and 2 respectively. The value of the probability of success in each trial is,
 (a) 0.6, (b) 0.4, (c) 0.2, (d) None of these.
18. A box contains p white and q black balls. The balls are randomly picked up consecutively from the box, without replacement. The probability for both the balls being white is,
 (a) $\frac{2}{p}$ (b) $\frac{2}{p+q}$ (c) $\frac{p(p-1)}{(p+q)(p+q-1)}$ (d) $\frac{p^2}{p+q}$
19. The variance of a random variable which takes the values 1, 2, 3 and 4 with probabilities 0.1, 0.2, 0.3 and 0.4 respectively, is
 (a) 1.25 (b) 1.0 (c) 1.12 (d) None of these

Answers to MCQs

1. (b)	2. (a)	3. (d)	4. (c)	5. (c)	6. (a)	7. (b)
8. (c)	9. (c)	10. (d)	11. (d)	12. (a)	13. (c)	14. (c)
15. (d)	16. (b)	17. (a)	18. (c)	19. (b)		

EXERCISE

1. (i) Two dice are thrown simultaneously. Calculate the probability that a sum of 6 will turn up.
 (ii) A box containing n articles has p defectives while another one with m articles has q defectives. If the articles in two boxes are mixed, and one is taken out what is the probability that (a) the article is defective and (b) the defective article is from the second box.
- (iii) A dice is thrown n times. What is the probability that a particular digit is obtained (i) at least once, (ii) in all the throws?
- (iv) Calculate the probability of having three kings in three consecutive draws from a pack of cards when (i) the drawn card is replaced after each draw, (ii) the card is not replaced.
- (v) The probability that out of a total of 5 machines one is in the operating condition is p . Calculate the probability that (i) two machines are in order, (ii) at least one machine is in the working condition.
- (vi) The probability that a shot hits the target is 0.5. What is the probability that a total of three shots will result in (i) three hits, (ii) no hit, (iii) at least one hit?
- (vii) A can solve 80% of the problems while B and C can solve 75% and 70% of the problems of a problem book. What is the probability that a problem selected at random will be solved when A , B and C try simultaneously?
[Hint: Use the formula for $P(A \cup B \cup C)$].
- (viii) Out of a total of n machines, p are of grade 1, q of grade 2, r of grade 3 and s of grade 4. The probabilities of obtaining correct machines are p_1 , q_1 , r_1 and s_1 respectively in the four grades. What is the probability that any one machine is free of defects?

- (ix) Two cards are drawn randomly from a deck. Calculate the probability of drawing either a king or a heart.
- (x) From a box containing a copper and b silver coins two coins are drawn successively with replacement and c coins of the type drawn being added after each draw. Find out the probability that the (a) two coins are of copper, (b) the two are of different types.
- (xi) The events $A_1, A_2, \dots, A_n$ occur with probabilities $p_1, p_2, \dots, p_n$ respectively. Find out the probability that (i) no event occurs, (ii) at least one of the events occurs. (iii) only one of these events occurs and (iv) all these events occur. (Take the events to be independent).
- (xii) Let a probability distribution function be given by $f(x) = Kx^3$, $0 \leq x \leq 2$, K being a constant and 0 otherwise. Find out K . Find out c_1 and c_2 such that the probability of $x \leq c_1 = 0.2$ and $x \leq c_2 = 0.8$.
- (xiii) Calculate the average and the standard deviation of the following probability distribution functions (i) $f(x) = \alpha x + \beta x^2$ for $0 \leq x \leq 1$ and zero elsewhere, (ii) $f(x) = Ax^2 \exp(-ax^2)$ for $0 \leq x \leq \infty$ and $a \geq 0$.
- (xiv) A radioactive source emits 1200 α -particles during an experiment lasting 10 hours. What is the probability of detecting (i) no α -particle, (ii) 5 α -particles. Justify the probability distribution function you use in this problem.
- (xv) Find out the average and the variance of $3x + 4$ if you know those of x .
- (xvi) The expectation and the variance of a random variable obeying the Gaussian (or normal) distribution are 20 and 100 respectively. Find out the probability that the variable lies between 10 and 40.
- (xvii) A G.M. counter records an average of 30 counts per minute. Find out the probability that there will be 2 counts in a period of 5 seconds.
- (xviii) A particle is confined to a three dimensional box with sides a, b and c respectively. The probability distribution is given according to quantum mechanics by $\sin^2 \frac{n_1 \pi x}{a} \sin^2 \frac{n_2 \pi y}{b} \sin^2 \frac{n_3 \pi z}{c}$ (n_1, n_2, n_3 are integers) for a particular state. Find out the expectation and the variance of the particle's position coordinate x (or y or z) according to both classical and quantum mechanics.
- [Hint: Probability of the position vector $\vec{r}$ lying between $\vec{r}$ and $\vec{r} + d\vec{r}$ is proportional to the time spent in the integral d interval $\vec{r}$ in classical mechanics].
- (xix) The random variables x and y have a joint P.D.F. proportional to $x^2 y^2 (x + y)^2$ for $0 \leq x \leq 1$ and $0 \leq y \leq 2$. Find the distribution function for x and y only (also called the marginal distribution).
- (xx) Suppose a random variable x has the distribution $f(x) = \alpha e^{-x}$ ($0 \leq x < \infty$), α being a constant.
- (a) Find $\langle x \rangle$, (b) Find $\langle x_1^2 + x_2^2 \rangle$, $\langle x_1 x_2 \rangle$ for two independent values x_1 and x_2 .
- [Hint: $f(x_1, x_2) = f(x_1) f(x_2)$ for independent x_1, x_2].
- (xxi) (a) A boy chooses randomly one of the three coins with probability of a head $\frac{3}{4}$, $\frac{1}{4}$ and $\frac{1}{2}$ respectively. What is the probability of (i) getting two heads, (ii) getting exactly one head?
- (b) From a basket containing x white and y black balls, n balls are drawn successively. After each draw z balls of the colour drawn are added. Find out the probability that the first m draws will give only white and the remaining draws will result in $(n - m)$ black balls.

Hints and Answers to Exercise Questions

1. (i) $\frac{5}{36}$, (ii) $\frac{p+q}{m+n}, \frac{q}{m+n}$,
- (iii) $1 - \left(\frac{5}{6}\right)^n, \left(\frac{1}{6}\right)^n$, (iv) $\frac{1}{(13)^3}, \frac{1}{5525}$,
- (v) $5c_2p^2(1-p)^3, 1 - (1-p)^5$. (vi) $\frac{1}{8}, \frac{1}{8}, \frac{7}{8}$
- (vii) 98.5%, (viii) $\frac{1}{n}(p_1p + q_1q + r_1r + s_1s)$,
- (ix) $\frac{4}{13}$, (x) $\left(\frac{a}{a+b}\right) \cdot \left(\frac{a+c}{a+b+c}\right)$,
- (xi) $\prod_{i=1}^n (1-p_i), 1 - \prod_{i=1}^n (1-p_i), p_i \prod_{j \neq i}^n (1-p_j), p_1p_2 \dots p_n$.
- (xii) $K = \frac{3}{8}, 0.001, 0.064$.
- (xiii) (i) $\langle x \rangle_{av} = \frac{\alpha}{4} + \frac{\beta}{5}, \sigma = \sqrt{\left(\frac{\alpha}{4} + \frac{\beta}{5}\right) - \left(\frac{\alpha}{3} + \frac{\beta}{4}\right)^2}$.
- (ii) $\langle x \rangle_{av} = \frac{A}{2\alpha^2}, \langle x^2 \rangle = \frac{3A\sqrt{\pi}}{8\alpha^{5/2}}, \sigma = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$,
- where $A = \frac{4}{\sqrt{\pi}} \alpha^{3/2}$
- (xiv) $\langle 3x + 4 \rangle = 3\langle x \rangle + 4, \text{Var}(3x + 4) = 9\text{Var}(x)$.
- (xv) $\langle x \rangle_{cl} = \langle x \rangle_q + \frac{a}{2}, \langle y \rangle_{cl} = \langle y \rangle_q = \frac{b}{2}, \langle z \rangle_{cl} = \langle z \rangle_q = \frac{c}{2}$
- $(\text{Var}(x))_{cl} = (\text{Var}(x))_q = \frac{a^2}{12}, (\text{Var}(y))_{cl} = (\text{Var}(y))_q = \frac{b^2}{12}$,
- $(\text{Var}(z))_{cl} = (\text{Var}(z))_q = \frac{c^2}{12}$ (for a particular n_1, n_2, n_3).

$$(xvi) \ A \left(\frac{8}{3} x^2 + 8x + \frac{32}{5} \right) x^2 dx, \ A \frac{y^2}{10} (5y^2 + 5y + 2), \ A = \frac{3}{14}$$

$$(xvii)(a) \ \langle x \rangle = 1,$$

$$(b) \ \langle x_1^2 + x_2^2 \rangle = 2.$$

$$(c) \ \langle x_1 x_2 \rangle = 1$$

$$(xviii)(a) \ (i) \ \frac{13}{24}$$

$$(ii) \ \frac{13}{96},$$

$$(b) \ \frac{x}{(x+y)} \cdot \frac{(x+z)}{(x+y+z)} \cdots \frac{x+(m-1)z}{\{x+y+(m-1)z\}} \cdot \frac{y}{(x+y+mz)} \cdots \frac{y+(n-m-1)z}{x+y+(n-1)z}$$

The answers to equations (xiv), (xvi) and (xvii) are not given because students can obtain these results by simply applying the formulae given in the text and consulting the table of functions and integrals.

APPENDIX

Introduction to χ^2 -test

Let N be the number of observation points sufficiently large and p_j , the probability of the data lying in a certain interval I_j , theoretically expected according to the assumed distribution. The expected number of data

in the interval $= Np_j = E_j$, say, and let us calculate $\chi_0^2 = \sum_{j=1}^n \frac{(E_j - 0_j)^2}{E_j}$, where n is the number of such slots and

0_j the experimental points in the interval, each slot being supposed to contain a large number of data points. Clearly the lower is χ_0^2 , the greater is the chance that the assumed hypothesis is good in the present context. From the table showing the probabilities for different χ^2 we can find out the probability that χ^2 is as large as χ_0^2 . Setting up a confidence limit ($c\%$, say) we can accept or reject a hypothesis in regard to the explanation of our experimental results depending on whether the probability is less or greater than $c\%$.

Group Theory

CHAPTER

12

12.1 INTRODUCTION

The topic of group theory finds immense applications in diverse areas of physics, crystallography, molecular physics, quantum mechanics for example. In this chapter we try to familiarise the students with the basic ideas of group theory without going into the depth of the subject. Detailed derivations of many important theorems will be omitted keeping in view the simplicity of presentation and the general level of the readers.

Definition - A group is a set of elements along with a definite rule of combination of any two elements belonging to the set. The result of this composition, called the product of two elements, obeys the following properties.

- (i) The product $a \cdot b$ belongs to the set.
- (ii) The product is associative, i.e., $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for any three elements a, b, c of the set.
- (iii) There is a unique inverse given by $a^{-1} \cdot a = a \cdot a^{-1} = I$ (I = identity element), proving the existence of a unique identity element of the group. Further, if $a \cdot b = b \cdot a$, the group is said to be commutative or Abelian. A group with a finite number of elements is called a finite group.

The number of such elements is called the order of the group.

Examples

- (1) The set of real or complex numbers forms a group with respect to addition, zero being the identity element. The group is Abelian and infinite.
- (2) The four numbers $1, -1, i, -i$ form a group under multiplication. All the elements of this group are obtained by raising i to different powers with $i^4 = 1$. Such a group which is Abelian with the different elements obtained by raising an element A of the group to different powers with $A^n = I$, I being the identity element, is called a cyclic group of order n . Thus the order of the cyclic group above is 4.

- (3) *Permutation group* All possible rearrangements of n objects among themselves form a group with $n!$ elements. That the rearrangements form a group can be understood by realising that the combination of two arrangements is again another possible arrangement, where the identity element corresponds to the objects remaining unaltered in their respective positions and the inverse is an arrangement opposite to the original arrangement. The group is, in general, non-Abelian and called the permutation or the symmetric group designated by S_n .
- (4) *Symmetry operations on an equilateral triangle* (taken to be in xy -plane). The symmetry operations on an equilateral triangle described below form a group called the dihedral group (D_3). The elements of the group are: (i) No operation or rotation through an angle of 360° about any axis (I), (ii) rotation by $\frac{2\pi}{3}$ in the anti-clockwise sense about the z -axis (A) (iii) rotation through an angle $\frac{4\pi}{3}$ about the same axis (B) in the same direction (iv) reflections about the axes 1, 2 and 3 forming the three elements C , D and F respectively. (The three axes 1, 2 and 3 are the three medians of the triangle). That these six elements form a group can easily be verified and this will be further demonstrated by the group multiplication table illustrated below. Though we will discuss finite groups mainly, we give a brief introduction to the infinite group in the Section 12.6.

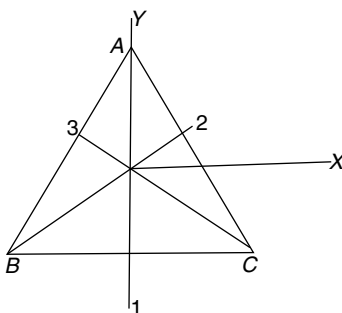


Fig. 12.1

12.2 MULTIPLICATION TABLE, ISOMORPHISM

The group multiplication table displays in a tabular form the product of any two group elements. Thus the group multiplication table of the group D_3 is shown in Table 12.1. The elements are placed along the rows and the columns of the table, while the product element ij occurs at the intersection of the i th row and the j th. column respectively. For the group D_3 , let I, A, B, C, D, F represent the six elements of D_3 in the same order as stated in the text. (I is the identity element here).

Table 12.1

	I	A	B	C	D	F
I	I	A	B	C	D	F
A	A	B	I	F	C	D
B	B	I	A	D	F	C
C	C	D	F	I	A	B
D	D	F	C	B	I	A
F	F	C	D	A	B	I

(The elements along the row and the column need not be in the same order as shown here in the table).

Two groups G and G' are said to be homomorphic to each other if there is a correspondence between the elements of the two groups that is not necessarily one to one. If the elements G_i, G_j of the group G can be mapped homomorphically to the elements G'_i, G'_j of the group G' respectively then the product $G_i G_j = G_k$ of G will correspond to the element $G'_i G'_j = G'_k$ of the group G' . Isomorphism is a special case of homomorphism when the mapping is one to one. For example, the six elements I, A, B, C, D, E of the group D_3 represent six possible rearrangements of the vertices of the equilateral triangle among themselves and can be said to have one to one correspondence with the six possible distributions of positions of the three objects among themselves. Thus there is isomorphism between the dihedral group D_3 and the permutation group S_3 . Two groups isomorphic with each other have the same multiplication table. Regarding other examples of isomorphic groups we can say that (i) the cyclic group $(A, A^2, A^3, A^4 = I)$ is isomorphic to the group $(i, -1, -i, 1)$, (ii) the group $(A, A^2, A^3 = I)$ is isomorphic to the group $(w, w^2, w^3 = 1)$. The group multiplication tables are shown below (Tables 12.2 (A) and 12.2 (B)).

Table 12.2 (A)

	A	A^2	A^3	$A^4 = I$
A	A^2	A^3	I	A
A^2	A^3	I	A	A^2
A^3	I	A	A^2	A^3
$A^4 = I$	A	A^2	A^3	I

Table- 12.2 (B)

	i	-1	$-i$	1
i	-1	$-i$	1	i
-1	$-i$	1	i	-1
$-i$	1	i	-1	$-i$
1	i	-1	$-i$	1

The examples of homomorphic mapping are:

- Mapping of the group $(1, -1, i, -i)$ to the cyclic group $(A, A^2 = I)$ where the elements $(1, -1)$ and $(i, -i)$ correspond to I and A respectively.
- The mapping of the group $(\pm I, \pm \sigma_1, \pm i\sigma_2, \pm \sigma_3)$ to (± 1) , where $\sigma_1, \sigma_2, \sigma_3$ are the Pauli matrices given by,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ and the elements } (\pm I, \pm \sigma_1) \text{ and } (\pm i\sigma_2, \pm \sigma_3) \text{ correspond to}$$

1 and -1 respectively.

The characteristic of a group multiplication table is that every element must occur only once in any row or column of the table. This is known as the Rearrangement Lemme. The proof is obvious because if $AB = C$, and $AD = C$, then $AB = AD$, and multiplying by A^{-1} we obtain $B = D$. But two elements B and D being different, we obtain the equality $B = D$, which is absurd. Hence $AB \neq AD$ for any A, B, D . Thus the group multiplication involves only rearrangement of the elements among different positions.

12.3 SECTION

- Subgroup** – It is a part of the main group, the elements of which constitute a group under the same rule of multiplication as in the original group. The identity element and the entire group are trivial subgroups of the original group. Among the nontrivial subgroups we may mention (i) the group of real numbers which forms a subgroup of the original group of complex numbers, (ii) the group of even

integers including zero which is a subgroup of the group of integers. (The nontrivial subgroups are also called proper subgroups).

- (2) *Cosets* – If H is a subgroup of the group G and $a, b, c, \dots$ are the elements of G that do not belong to H , the set consisting of the elements $(aH, bH, \dots)$ forms the left coset of H . Similarly, the set of elements $(Ha, Hb, \dots)$ forms the right coset of H . If all the left cosets of H are the same as the right cosets of H , the subgroup H is called the normal or the invariant subgroup of H . It is easily seen that the cosets are not subgroups because the identity element is not included in the cosets.

Example

If out of the original group $G (= \pm I, \pm \sigma_1, \pm i\sigma_2, \pm \sigma_3)$ we take out a subgroup (I, σ_3) , (i) the left cosets are $\pm \sigma_1(I, \sigma_3) = (\pm \sigma_1, \mp i\sigma_2) \pm i\sigma_2(I, \sigma_3) = (\pm i\sigma_2, \mp \sigma_1)$. (ii) the right cosets are $\pm (I, \sigma_3) \sigma_1 = (\pm \sigma_1, \pm i\sigma_2) \pm (I, \sigma_3) i\sigma_2 = (\pm i\sigma_2, \pm \sigma_1)$. The invariant subgroups of the group G are $(\pm I)$, $(\pm I, \pm \sigma_1)$, $(\pm I, \pm i\sigma_2)$, $(\pm I, \pm \sigma_3)$. [When one says that $aH = Ha$, it means that $aH_i = H_j a$, where i is not necessarily equal to j , i.e., H_i and H_j may be any two members of H].

Two short theorems are relevant in this context.

- (a) Two cosets of a subgroup are either completely identical or have no common element.

To prove this let us suppose that $AH_i = BH_j$, where A and B do not belong to H . We get, $B^{-1}A = H_j H_i^{-1}$ and, hence, it belongs to H . Thus $B^{-1}AH = H$ (by rearrangement lemma). We get therefore, $AH = BH$. Thus the two cosets are entirely identical. Similar is the case for two right hand cosets.

- (b) The order of a finite group G is a multiple of that of any finite subgroup H .

This is actually a corollary of (a). Taking the elements of H and those obtained by multiplying the subgroup H by the elements $A, B, \dots$ belonging not to H but to the group, we can obtain the total group G . Thus the order of the entire group = order of H + the order of the subgroup multiplied by the number of such elements $A, B, \dots$. Hence the order of the group G is a multiple of that of the subgroup.

- (3) *Factor group* – An invariant subgroup, together with its cosets can be shown to form a group called the factor group (or quotient group) of G by H , H being the invariant subgroup of the group G and denoted by G/H . The order of the group is g/h .

Proof:

(i) $(AH)(BH) = A(HB)H = A(BH)H = ABHH = ABH$ and is thus another coset of H .

(ii) The identity element in the group = $IH = H$.

(iii) $(A^{-1}H)(AH) = A^{-1}(HA)H = A^{-1}AH = H$.

$(AH)(A^{-1}H) = A(HA^{-1})H = A(A^{-1}H)H = AA^{-1}HH = H$, proving that the inverse of the coset $AH = A^{-1}H$.

Thus all the group properties are satisfied.

Example

The factor group of $G = (\pm I, \pm \sigma_1, \pm i\sigma_2, \pm \sigma_3)$ with respect to the invariant subgroup $H = (\pm I, \pm \sigma_1)$ comprises H and the coset $(\pm \sigma_3, \pm i\sigma_2)$.

- (4) *Class* – Two elements a and b are defined to be conjugate to each other (denoted by $a \sim b$) if $b = g^{-1}ag$ for any element g of the group. This type of conjugacy relation satisfies the properties (i) $a \sim a$ (ii) $a \sim b$ implies $b \sim a$ (iii) $a \sim b$ and $b \sim c$ implies $a \sim c$. Such a relation existing between the elements of a group is called an equivalence relation. The equivalence relation divides the entire group into mutually disjoint sets of elements having a relation $a \sim b$ between any pair and said to be in the same class. In an

Abelian group every element a satisfies a relation $ag = ga$ and hence $g^{-1}ag = a$ so that it forms a class by itself. It can be shown that no two classes have any element common between them and thus a group comprises sets of distinct classes. (proofs are given in the section “solved problems”)

Examples (1) The identity element of each group forms a class by itself. (2) The dihedral group D_3 consists of three distinct classes (I), (A , B) and (C , D , E) as is evident from Table 12.1.

12.4 REPRESENTATION OF GROUPS

12.4.1 Definition

The group representation resembles that of vectors and operators in a certain basis, where we regard groups as linear operators acting on an underlying vector space and accordingly give representations of groups in terms of matrices. The matrices obey the same multiplication rules as the group elements to which they correspond. Thus if $G_i G_j = G_k$, we have $M(G_i) M(G_j) = M(G_k)$, $M(G_i^{-1}) = [M(G_i)]^{-1}$,

$M(I) = e$, e being the identity operator of the matrix representation. The same group may be represented by matrices of different dimensions. The dimension of the matrix is called the dimension of the representation. The action of the matrix operator acting on the basis vectors of the vector space is represented by $M(G)e_i =$

$\sum_{j=1}^n e_j M(G)_{ij}$, n being the dimension of the representation. If there be a one-to-one correspondence between the group elements and the corresponding matrices the representation is said to be faithful.

Examples

- (i) For every group there exist trivial representations of all dimensions corresponding to each element given by $M(G) = I$, I being the identity element. The representation is clearly not faithful.
- (ii) For the dihedral group D_3 we have a faithful representation of two-dimensions given by

$$D(I) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, D(A) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, D(B) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix},$$

$$D(C) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, D(D) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, D(F) = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

- (iii) The group (I , P , T , PT), where P , T represent the parity and the time reversal operations respectively having a four-dimensional representation given by

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, P = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, PT = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

- (iv) The cyclic group of order four has the 1D representation given by $M^{(1)}(I) = 1, M^{(1)}(A) = i, M^{(1)}(A^2) = -1, M^{(1)}(A^3) = -i$ or, $M^{(1)}(I) = 1, M^{(1)}(A) = -i, M^{(1)}(A^2) = -1, M^{(1)}(A^3) = i$.

12.4.2 Reducible and Irreducible Representation

Recalling our earlier discussion in Chapter 3 we can say that a matrix is said to have an irreducible representation if no proper subspace of the entire vector space is left invariant by the action of the matrix. (A proper subspace means that it is neither the complete space nor the null space).

On the contrary, if a subspace of the entire vector space is left invariant by the action of the representative matrix, it is said to be in a reducible form.

If the matrix representation $G \rightarrow M(G)$ be completely reducible, then the matrix $M(G)$ can by a similarity transformation be decomposed in the simplest case into the block-diagonal form $M(G) = \begin{pmatrix} M^{(1)}(G) & 0 \\ 0 & M^{(2)}(G) \end{pmatrix}$,

where $M_1(G)$ and $M_2(G)$ keep the subspaces V_1 and V_2 of the entire vector space V invariant so that the matrix $M(G)$ as well as the vector space V on which it operates can be written as $M(G) = M^{(1)}(G) \oplus M^{(2)}(G)$ and $V = V_1 \oplus V_2$ respectively; V_1 and V_2 being non-overlapping. (The symbol $\oplus$ means the direct sum). Of course $M(G)$ can have more complicated decomposition like $M(G) = \sum_i a_i M^{(i)}(G) = a_1 M^{(1)}(G) \oplus a_2 M^{(2)}(G) + \dots$,

where $a_1, a_2, \dots$ are the number of times the representative matrices $M^{(1)}(G), M^{(2)}(G) \dots$ appear in $M(G)$. Each matrix $M^{(i)}(G)$ is in an irreducible representation.

If a single similarity transformation can transform all the matrices representing the group elements to the block-diagonal forms the representation is said to be reducible.

For example, (i) the group (I, P, T, PT) shown in the previous section can be regarded as the direct sum of four one-dimensional irreducible representations. (ii) the 2D representation of the group D_3 given in the previous section is irreducible because no such transformation exists. (iii) the 2-D representation of the cyclic

group C_4 represented by $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ can be shown to be reducible. (Students are instructed to find out the similarity transformation).

We, now, state a few theorems that are relevant in the context of the theory of representation. (The proofs are omitted here).

- (1) Every finite dimensional representation is equivalent to unitary representation, where the matrices are unitary matrices acting on complex vector spaces.
- (2) If in an irreducible representation M of a group G , there is a matrix operator A such that $M(g) A = A M(g)$ for all the elements g belonging to G , then A is a multiple of the unit matrix. This is called Schur's Lemma.

As a corollary of this lemma it can be proved that any irreducible representation of a finite Abelian group must be of one-dimension. [The corollary can be proved easily as follows. For an Abelian group. $M(g) M(g') = M(g') M(g)$ for all g, g' belonging to G . From Schur's lemma, $M(g)$ for an irreducible representation = $K(g) \cdot I$, $K(g)$ being a constant and I the identity matrix. Hence, $M(g)$ is one dimensional because for higher dimension $M(g)$ becomes necessarily reducible.]

The irreducible representations are the basic units of building up any representation of a group because any finite representation of a group can be expressed as a linear combination of the irreducible representations. Below we will study briefly, omitting the proofs of some important theorems that are available in standard

textbooks on group theory, the condition of a representation being irreducible, the decomposition of a representation into irreducible ones, etc.

First, we introduce the idea of the character of a representation.

Character – We have already defined classes of a group in Section 12.3. Since a class is formed by the elements conjugate to one another the matrices representing the elements are related by similarity transformation. Thus if A and B be two conjugate elements related by $B = S^{-1} A S$, S being an element of the group we can write $M(B) = [M(S)]^{-1} M(A) M(S)$, $M(A)$, $M(B)$ and $M(S)$ being the respective matrix representations. The character of a class is defined as the trace of the matrix representing the element in the class.

The elements belonging to a class have the same character since the trace remains invariant under similarity transformation. Hence a class is characterised uniquely by its character for a particular representation.

The different classes of a group and the characters of the classes in the possible irreducible representations can be displayed in the character table to be discussed later. We cite below an important theorem, called the orthogonality theorem, that is useful in the construction of the character table. as well as in ascertaining if a representation is irreducible or in breaking up a reducible representation into irreducible ones.

(i) Let G be a group of order g and $M^r (r = 1, 2, \dots)$ represent the different irreducible representations.

We then have, $\sum_{n=1}^g M_{ij}^r(G_n) M_{kl}^s(G_n^{-1}) = \frac{g}{n_r} \delta_{rs} \delta_{il} \delta_{kj}$, n_r being the dimension of M^r . For the unitary representation it follows that $\sum_{n=1}^g M_{ij}^r(G_n) M_{lk}^s(G_n)^* = \frac{g}{n_r} \delta_{rs} \delta_{il} \delta_{kj}$. (12.1)

[We do not prove the theorem but refer the reader to the text books on group theory for its proof].

For the same representation $r = s$ and we get,

$$\sum_{n=1}^g M_{ij}^r(G_n) M_{lk}^r(G_n)^* = \frac{g}{n_r} \delta_{il} \delta_{kj}.$$

Comparison with the orthogonality relation of vector leads one to interpret Eq. (12.1) as the orthogonality relation in group theory where the terms $M_{ij}^r(G_k)$ can be regarded as the components of g -dimensional complex vectors.

Putting $i = j$, $k = l$ in Eq. (12.1) we get, $\sum_{n=1}^g x^r(G_n) x^s(G_n)^* = g \delta_{rs}$, where we sum over i , k and $X^r(G_n) = \text{tr } M^r(G_n)$ denotes the trace of the matrix representing the element G_n . If n_i be the number of elements in the class c_i we have $\sum_{t=1}^c n_t x_t^r x_t^{s*} = g \delta_{rs}$, $x_t^{(r)}$ denoting the character of the class t . (12.2)

If M be a reducible representation that can be written in the fully reducible form as a direct sum of the irreducible representations M^i , given by $M = \sum_{i=1}^{n_i} a_i M^i$, where a_i is the number of times the irreducible representation M^i occurs in M , (here Σ represents the direct sum) we have $x_t = \sum_{i=1}^{n_i} a_i x_t^i$ (12.3)

Combining Eqs. (12.2) and (12.3) we get,

$$a_i = \frac{1}{g} \sum_t n_t x_t^{(i)*} x_t = \frac{1}{g} \sum_G x^{i*}(G) x(G). \quad (12.4)$$

Using (12.3) we get, $\sum_t n_t |x_t|^2 = g \sum_i a_i^2$.

$$\text{Or,} \quad \sum_{i=1}^{n_i} a_i^2 = \frac{1}{g} \sum_{t=1}^c n_t |x_t|^2 = \frac{1}{g} \sum_G |x(G)|^2 \quad (12.5)$$

Clearly if $\sum_{i=1}^n a_i^2 = 1$ or, $\sum_G |x(G)|^2 = g$, there can be only a single non-vanishing a_i and hence the

representation is irreducible. In the general case the a_i 's can be found out from (12.4) if we know the characters of the classes present in the group, (the method of finding out the characters will be discussed in a subsequent section). Below we will discuss a representation called the regular representation that is reducible and for which the number of irreducible representations can be easily determined.

The regular representation:

The regular representation of a group is constructed in two steps.

- The group multiplication table is rearranged in such a way that the elements along the row are the reciprocals of those at the corresponding positions along the columns. The identity element thus appears at the diagonal positions of the table.
- The matrix representation of the group is formed by putting unity where the element occurs and zero otherwise. Thus we can write

$$M_{ij}(G_k) = 0 \text{ if } G_i^{-1} G_j \neq G_k \text{ and } = 1 \text{ otherwise.}$$

Obviously the trace of the matrix representing the identity element is the dimension of the representation and zero for the other elements.

$$\text{Thus } [M(G_p) M(G_q)]_{ij} = \sum_k M_{ik}(G_p) M_{kj}(G_q)$$

$$M_{ik}(G_p) = 1 \text{ if } G_i^{-1} G_k = G_p \text{ and } M_{kj}(G_q) = 1 \text{ if } G_k^{-1} G_j = G_q.$$

Hence $[M(G_p) M(G_q)]_{ij} = 1$ only if $G_i G_p = G_j G_q^{-1}$ or if $G_p G_q = G_i^{-1} G_j$ and zero otherwise. We thus obtain, $[M(G_p) M(G_q)]_{ij} = [M(G_p G_q)]_{ij}$. Thus the matrices $M(G_i)$ form a representation (All other properties of a group representation can be seen to hold good).

That the regular representation is reducible can be proved by writing,

$$M^{\text{reg}}(G_i) = \sum_{\alpha} a_{\alpha} M^{\alpha}(G_i), \text{ where } M^{\text{reg}}(G_i) \text{ is a regular representation.}$$

$$\text{Now from (12.4) we obtain, } a_i = \frac{1}{g} \sum_G x^{i*}(G_i) x^{\text{reg}}(G)$$

$$\text{or, } a_i = \frac{1}{g} \cdot g x^{i*}(I) \text{ since } x^{\text{reg}}(G_i) = g \text{ if } G_i = I, \text{ the identity element and } = 0 \text{ for the other elements.}$$

or, $a_i = x^{i*}(I) = d_i$, d_i being the dimension of the representation. Thus a regular representation contains the irreducible representation $M^i(G_i)$ as many times as its dimension.

The rearranged group multiplication table for the group D_3 so as to form the regular representation is shown below. (see Table 12.1)

	I	A	B	C	D	F
I^{-1}	I	A	B	C	D	F
A^{-1}	B	I	A	D	F	C
B^{-1}	A	B	I	F	C	D
C^{-1}	C	D	F	I	A	B
D^{-1}	D	F	C	B	I	A
F^{-1}	F	C	D	A	B	I

Thus the regular representation corresponding to the element I is the matrix

1	0	0	0	0	0
0	1	0	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
0	0	0	0	1	0
0	0	0	0	0	1

The character $\chi^{\text{reg}}(I) = 6 = \text{number of elements of the group}$.

The regular representation corresponding to the element A is the matrix

0	1	0	0	0	0
0	0	1	0	0	0
1	0	0	0	0	0
0	0	0	0	1	0
0	0	0	0	0	1
0	0	0	1	0	0

The character $\chi^{\text{reg}}(A) = 0$. Students can verify that $\chi^{\text{reg}}(G_i)$ for every G_i except $I = 0$.

Direct Product Representation

If we have two groups of operators A_i and B_j ($A_i = A_1, \dots, I, \dots, A_m$; $B_j = B_1, \dots, I, \dots, B_n$) acting on vector spaces such that all the operators A_i corresponding to the group G_A commute with all B_j corresponding to G_B , we can form a representation of the product group $G_A \otimes G_B$ by taking the direct product of the elements of the groups G_A and G_B [See Appendix of Chapter 3 for the definition of the direct product], respectively. The order of representation of the product group is thus equal to the product of the individual orders. The character of the representation matrix of the product group = Product of the individual characters. That the direct product of two groups gives rise to another group with the number of elements equal to the product of the number of elements of each group can be verified by following the multiplication rules of the matrices. Also, it can be shown that all the irreducible representations of the product group can be generated by taking the direct product of the irreducible representations of the individual groups. (Both of these are worked out in the section on “Solved examples”).

We can form higher order representations of a single group by taking the direct product of two lower order representations. If the lower order representations be irreducible the higher order representation is, in general, reducible and can be decomposed into a sum of the irreducible representations. Thus, for example, the irreducible representations of the rotation group acting on the spaces spanned by the angular momenta j_1 and j_2 are of dimensions $2j_1 + 1$ and $2j_2 + 1$ respectively. By counting the two angular momenta we obtain the angular momentum of the coupled state J given by $|j_1 - j_2| \leq J \leq j_1 + j_2$.

Thus the rotation matrix for the coupled state is given by the product representation $M^J = M^{(j_1)} \otimes M^{(j_2)}$ and also as $M^{(J)} = M^{|j_1 - j_2|} \oplus M^{|j_1 - j_2| + 1} \oplus \dots \oplus M^{(j_1 + j_2)}$

[$\otimes$ is the symbol for the direct product and $\oplus$ the direct sum]. Thus the product representation is reducible and expressed as the direct sum of the irreducible representations.

12.5 CHARACTER TABLE

The necessary pieces of information regarding a group viz., the different classes, the distinct irreducible representations, the characters of the classes in the different representations can all be displayed through the character table of a group. Now we proceed to construct the character tables of a few groups. The orthogonality theorem referred to in (12.4.2) and a few results derived from this theorem are relevant in this context.

The useful results can be summed up as follows.

- (1) $\sum_{t=1}^c n_t x_t^r x_t^{s*} = g \delta_{rs}$, a result deduced in the section (12.4.2) from the orthogonality relation, (eqn. (12.2))
- (2) $\sum_{r=1}^{n_r} n_t x_t^r x_u^{r*} = g \delta_{tu}$, a relation deducible from the orthogonality theorem and derived in the appendix. (12.6)
- (3) The number of inequivalent irreducible representations equals that of the classes of the group. [By the term inequivalent we mean that the matrices of the representations are not related by similarity transformation with one another.]

This can be proved by putting $\bar{x}_t^r = \left(\frac{n_t}{g}\right)^{\frac{1}{2}} x_t^r$, $\bar{x}_t^s = \left(\frac{n_t}{g}\right)^{\frac{1}{2}} x_t^s$ in (12.2)

we obtain, $\sum_{t=1}^c \bar{x}_t^r \bar{x}_t^{s*} = \delta_{rs}$. This relation shows that the vectors $\bar{x}_t^r$ ($r = 1, 2, \dots, n_r$, where n_r is the number of dimensions) form an orthonormal set in a c -dimensional vector space. Since the number of vectors cannot exceed the dimension of the vector space we obtain that $n_r \leq c$. (12.7)

Similarly from (12.6) we can write, $\sum_{r=1}^{n_r} \bar{x}_t^r \bar{x}_u^{r*} = \delta_{tu}$, which reflects the fact that the vectors

$\bar{x}_t^r$ ($t = 1, 2, \dots, c$) form an orthonormal set in the n_r -dimensional vector space. Hence by the same logic as followed in (12.7) we get, $c \leq n_r$. (12.8)

Combining (12.7) and (12.8) we obtain, $n_r = c$.

- (4) We have already seen that a regular representation contains an irreducible representation a number of times equal to the dimension of the representation. If d_i be the dimension of the i th-irreducible

representation, the total dimension of the regular representation is equal to $\sum_i d_i^2$ and clearly equal to the order of the group. Thus we obtain, $\sum_i d_i^2 = g$.

The steps for the construction of the character table can be summarized as follows:

- (1) The number of classes in a group is determined by studying the group multiplication table and finding out the conjugate elements. The number of inequivalent irreducible representations that equals the number of classes is thus obtained.
- (2) The dimensions of the irreducible representations are calculated from the relation $\sum_i d_i^2 = g$ leading to a unique set of values of d_i 's in most of the cases.
- (3) The character table is an array of elements which are the characters of the classes in different irreducible representations, where the classes are placed along the rows and the representations along the columns. The identity element of a group is a class by itself and generally occupies the first position of the row. (Because $g_i^{-1} I g_i = I$ for all g_i) The first column representing the characters of this one-member class in different dimensional representations which are equal to the respective dimensions is easily filled up.
- (4) Every group has a trivial one-dimensional representation in which each element is represented by unity. The first row can thus be filled up in which every character is unity.
- (5) The orthogonality relations (12.2) and (12.6) respectively are now used to complete the character table. [If the group be Abelian, the only finite dimensional irreducible representation is one-dimensional and hence the character table can be written very simply without going through all the above mentioned steps]

Examples

- (1) Cyclic group - $(I, A, A^2, \dots, A^{n-1} = I)$.

This group is Abelian and hence has only one-dimensional irreducible representation, represented by $M^{(r)}$.

Clearly $x^{(r)}(A^n) = 1 = [x^{(r)}(A)]^n$.

The n roots of this relation yielding, $x^{(r)}(A) = \exp\left[\frac{2\pi i r}{n}\right]$, where $r = 0, 1, \dots, n-1$, are the n characters corresponding to n elements,

- (2) The group D_3 (the group of symmetry operations of the equilateral triangle).

The group has three classes as can be demonstrated from the character table,

- (i) One-element class c_1 containing the identity element I ,
- (ii) Two-member class c_2 consisting of A, B .
- (iii) Three-member class c_3 comprising C, D, F .

From $\sum_i d_i^2 = 6$, we have the unique solution $d_1 = d_2 = 1, d_3 = 2$.

The first row and the first column are filled up as prescribed in (3) and (4).

Since for the one-dimensional representation $x(c^2) = 1 = [x(c)]^2$ for $c = c_2$ or $c = c_3$ we obtain, $x(c) = \pm 1$. Thus using Eq. (12.2) the characters of the classes corresponding to two one-dimensional representations are easily written. For the remaining two-dimensional representation the orthogonality relations (12.2) is used.

We get, $2 + 2x_{c_2}(M^{(3)}) + 3x_{c_3}(M^{(3)}) = 0$ and

$2 + 2x_{c_2}(M^{(3)}) - 3x_{c_3}(M^{(3)}) = 0$, so that, we get,

$x_{c_3}(M^{(3)}) = 0, x_{c_2}(M^{(3)}) = -1$.

Thus the entire character table is shown below

	$C_1(I)$	$C_2(A, B)$	$C_3(C, D, F)$
$M^{(1)}$	1	1	1
$M^{(2)}$	1	1	-1
$M^{(3)}$	2	-1	0

12.6 INFINITE GROUPS

The number of elements in an infinite group being infinite we cannot construct the multiplication table or the character table in this case. The basic conditions for the formation of a group must however be satisfied also for the infinite groups.

Infinite groups may be discrete or continuous. The set of integers (including zero) under addition provides an example of discrete groups. In the case of infinite continuous groups the elements of the group vary continuously as a function of the parameters of the group. The group is compact or non-compact according as the parameter space is bounded and closed or not. (There exists an alternative definition that is mathematically more rigorous but this will not be given here). The rotation group (where the angle of rotation can be taken to vary between 0 and 2π) is compact, whereas the Lorentz group is non-compact.

Taking the identity element to refer to the case when the group parameters vanish, we can write for an infinitesimal change in the neighbourhood of the identity, (summation over k is implied), $g = I + Y_k da_k$, where a_k represents the k th. parameter, Y_k is called the corresponding generator and I the identity element. If the group elements correspond to the angles of rotation about the different axes for a rotation group, we can write for the rotation about the i th. axis through angles θ_i and $\theta_i + d\theta_i$ respectively,

$$g(\theta_i + d\theta_i) = g(\theta_i) g(d\theta_i) = g(\theta_i) + Y_i g d\theta_i$$

or, $dg = Y_i g d\theta_i = \frac{\partial g}{\partial \theta_i} d\theta_i$, giving $g = \exp(Y_i \theta_i) c(\theta_j's, j \neq i)$, c being the constant of integration depending

on the angles of rotation about the other axes. Considering a rotation about any, arbitrary axis in 3D Cartesian system that can be written as a combination of three successive rotations about three coordinate axes we can

write, $g \sim \exp\left(\sum_{k=1}^3 Y_k \theta_k\right)$, the Y_k 's being the generators for the rotations about the corresponding axes. For

rotation about the z -axis for example, $dg = g(x + dx, y + dy) - g(x, y)$

$$\text{Or, } dg = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy$$

For rotation through an infinitesimally small angle $d\theta$ about the z -axis, $dx = y d\theta$, $dy = -x d\theta$ (rotation being anticlockwise).

or, $dg = \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y}\right) g d\theta = \frac{iL_z}{\hbar} g d\theta$, where $L_z = (\vec{r} \times \vec{p})_z$ is given by $L_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\right)$, the

z -component of orbital angular momentum in Cartesian coordinates.

Thus in this case $Y = -\frac{iL_z}{\hbar}$ [often, instead of Y , L_z is called the generator of rotation about the z -axis]. For rotation about an arbitrary axis having the unit vector $\vec{n}$ we can write the group element as $g(\theta) = \exp\left[\frac{-i}{\hbar}(\vec{L} \cdot \vec{n})\theta\right]$, where $g(0)$ = identity element.

A group whose elements vary continuously as functions of the parameters and which may be written as $I + Y_k da_k$, I being the identity element, a_k the group parameter and Y_k the corresponding generators (or proportional to them) for an infinitesimal change in the neighbourhood of the identity is called the Lie-group. Thus the groups of rotations about any axis, or the general linear transformation of the form $x' = ax + by + c$, $y' = dx + ey + f$, a, b, c, d, e, f being constants are examples of Lie-groups. Below, we will briefly discuss a few essential features of some important Lie groups.

(1) $O(n)$ or the orthogonal groups of order n . The condition $O^T O = I$ gives $\frac{n(n+1)}{2}$ independent

equations. The number of independent parameters $= n^2 - \frac{n(n+1)}{2} = \frac{1}{2}n(n-1)$ and $\det O = \pm 1$. The

only invariant under the orthogonal transformation is $\sum_{i=1}^n x_i^2$, i.e., the norm of the vector of the vector

space on which the matrix operates. The group is composed of two disjoint pieces, one corresponding to pure rotation with determinant 1 and symbolised by $SO(n)$, while the other has determinant -1 and is related to inversion. The group $SO(n)$ has the same number of independent parameters, i.e.,

$\frac{1}{2}n(n-1)$. In three dimensions the group $SO(3)$ has three independent parameters, namely, the three

angles about the coordinate axes. Any arbitrary element of the group $SO(3)$ can be written as $\exp(-iJ_z\alpha)\exp(-iJ_y\beta)\exp(-iJ_x\gamma)$ for any general rotation about any axis which is an ordered product of three rotations about three coordinate axes through Euler angles γ, β, α respectively. [J_x, J_y and J_z , the generators of rotation about the x, y and the z axes are the x, y and the z components of the total angular momentum, respectively. In the absence of spin they become L_x, L_y and L_z respectively as stated earlier]

The group $SO(3)$ is compact, the angles of rotation being limited between 0 and 2π . A variant of the group

$SO(n)$ is $SO(m, n)$, where the invariant is $\sum_{i=1}^m x_i^2 - \sum_{i=m+1}^n x_i^2$. An example of this pseudo-orthogonal group is

$SO(3, 1)$, the proper Lorentz-group, which is non-compact and where the invariant is $\sum_{i=1}^3 x_i^2 - x_4^2$ ($x_4 = ct$).

Unitary group – The unitary representations of an infinite group are given by unitary matrices $U(n)$, the determinants of which have unit moduli. The set of all unitary matrices of dimension n with determinant $+1$ forms a group called $SU(n)$, which is a subgroup of $U(n)$. The number of independent parameters of the group

$SU(n)$ is $n^2 - 1$. A general $SU(2)$ matrix is represented by $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where the condition of unitarity gives, $|a|^2 + |b|^2 = 1$, $|c|^2 + |d|^2 = 1$, $ad - bc = 1$, $a^*c + b^*d = 0$

we get, $a^* = (a^*a + b^*b)d = d$ and $b^* = -(a^*a + b^*b)c = -c$.

Thus, the general $SU(2)$ matrix is written as, $\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}$, with $|a|^2 + |b|^2 = 1$.

A useful parametrisation of $SU(2)$ is obtained by putting $a = \cos \theta e^{i\phi}$, $b = \sin \theta e^{i\xi}$.

We express a traceless two-dimensional matrix as a linear combination of Pauli spin matrices $\sigma_1, \sigma_2, \sigma_3$ given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ as } X = \sum_{i=1}^3 x_i \sigma_i$$

or, $X = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}$. If O be the general unitary matrix represented by $\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}$ we have

$X' = O X O^\dagger = O X O^{-1}$ and $X'^\dagger X' = X^\dagger X$ so that we get, $x_1'^2 + x_2'^2 + x_3'^2 = x_1^2 + x_2^2 + x_3^2$.

Thus, the two dimensional matrix O representing the $SU(2)$ group induces an orthogonal transformation from (x_1, x_2, x_3) to (x_1', x_2', x_3') given by straightforward calculation as

$$\begin{aligned} x_1' &= \frac{1}{2} (a^2 + a^{*2} - b^2 - b^{*2}) x_1 + i/2 (a^{*2} + b^{*2} - a^2 - b^2) x_2 - (ab + a^* b^*) x_3 \\ x_2' &= i/2 (a^2 - a^{*2} + b^2 - b^{*2}) x_1 + \frac{1}{2} (a^2 + a^{*2} + b^2 + b^{*2}) x_2 + i(a^* b^* - ab) x_3, \\ x_3' &= (a^* b + ab^*) x_1 + i(a^* b - ab^*) x_2 + (aa^* - bb^*) x_3 \end{aligned}$$

If the matrix of the orthogonal matrix be denoted by $R(u)$ we can write $\vec{r}' = R(u) \vec{r}$.

Taking $O_1(\alpha) = \begin{pmatrix} e^{i\alpha/2} & O \\ O & e^{-i\alpha/2} \end{pmatrix}$ we have, $R_1(\alpha) = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$ so that $x_1' = x_1 \cos \alpha + x_2 \sin \alpha$,

$x_2' = -x_1 \sin \alpha + x_2 \cos \alpha$, $x_3' = x_3$ representing a rotation about the z -axis. Similarly, taking $O_2(\beta) =$

$\begin{pmatrix} \cos \beta/2 & \sin \beta/2 \\ -\sin \beta/2 & \cos \beta/2 \end{pmatrix}$ we have, $R_2(\beta) = \begin{pmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{pmatrix}$ defining a rotation about the y -axis.

A transformation from a given set of orthogonal axes to another similar set can be effected by three successive rotations through the Euler angles α, β, γ about the x, y and z axes respectively. We get, $O(\alpha, \beta, \gamma)$

$$= O_1(\gamma) O_2(\beta) O_1(\alpha) = \begin{pmatrix} e^{i(\alpha+\gamma)/2} \cos \beta/2 & e^{i(\gamma-\alpha)/2} \sin \beta/2 \\ -e^{-i(\gamma-\alpha)/2} \sin \beta/2 & e^{-i(\alpha+\gamma)/2} \cos \beta/2 \end{pmatrix}$$

Thus a and b are identical with those given earlier, where $\theta = \frac{\beta}{2}$, $\phi = \frac{(\alpha + \gamma)}{2}$, $\xi = \frac{(\gamma + \alpha)}{2}$.

Thus an appropriate matrix from the $SU(2)$ group can generate an arbitrary rotation which is represented by an orthogonal matrix that is formed by the product of three orthogonal matrices $R_1(\alpha)$, $R_2(\beta)$ and $R_1(\gamma)$ respectively. Clearly, O and $-O$ formed by changing α, β, γ by 2π each give rise to the same orthogonal matrix in the $SO(3)$ representation. Thus there is a two to one homomorphism between the group of 2×2 matrices with determinant +1 and the three-dimensional orthogonal matrices representing pure rotation in the $SO(3)$ representation.

If $\begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ represents a two-dimensional vector that gets transformed to $\begin{pmatrix} V'_1 \\ V'_2 \end{pmatrix}$ under the action of the $SU(2)$

matrix operator we obtain, $V'_1 = aV_1 + bV_2$, $V'_2 = -b^*V_1 + a^*V_2$.

Defining the set of functions $F_m^j = \frac{V_1^{j+m} V_2^{j-m}}{\sqrt{(j+m)!(j-m)!}}$ we obtain that F_m^j changes to

$$F_m^{j'} = \frac{(aV_1 + bV_2)^{j+m} (a^*V_2 - b^*V_1)^{j-m}}{[(j+m)!(j-m)!]^{\frac{1}{2}}} \text{ for } -j \leq m \leq j \text{ and integer } j.$$

Applying binomial theorem we can write after straightforward algebraic calculations,

$$F_m^{j'} = \sum_{m'} F_{m'}^j U_{m'm}^j(a, b), \text{ where}$$

$$U_{m'm}^j(a, b) = \sum_p \frac{[(j+m)!(j-m)!(j+m')!(j-m')!]^{\frac{1}{2}}}{p!(j+m-p)!(j-m'-p)!(m'-m+p)!} a^{j-m-p} a^{*j-m'-p} b^p (-b^*)^{m'-m+p}$$

The matrix elements $U_{m'm}^j(a, b)$ can be shown to form a $(2j+1)$ dimensional irreducible representation of $SU(2)$. (Proof is omitted and can be obtained in any text book on group theory, Tinkham's book for example).

The expressions of the matrix elements calculated for $j = \text{odd}$ and even integral multiples of $\frac{1}{2}$ show after

substituting $a = \frac{\cos \beta}{2} e^{i(\alpha+\gamma)/2}$ and $b = e^{i\left(\frac{\gamma-\alpha}{2}\right)} \sin \frac{\beta}{2}$ respectively that for $j = \text{odd}$ multiple of $\frac{1}{2}$,

$U_{m'm}^j(\alpha, \beta, \gamma) = -U_{m'm}^j(\alpha, \beta + 2\pi, \gamma)$ and for $j = \text{integer}$, $U_{m'm}^j(\alpha, \beta, \gamma) = U_{m'm}^j(\alpha, \beta + 2\pi, \gamma)$. The elements of the $SU(2)$ group thus exhibit a double-valued representation for $j = \text{odd}$ multiple of $\frac{1}{2}$ and become single-valued for $j = \text{integer}$, though the representation of the $SO(3)$ group remains single-valued for all j .

Generators The generators of a group are found out (a) by operating the group elements in the neighbourhood of the identity on the coordinates and determining the resulting changes of a function of the coordinates, (b) by ascertaining directly the changes in the matrix elements as a consequence of the infinitesimal changes of the parameters. The following examples are given as illustrations.

(1) Let us discuss the $SO(3)$ matrices given by

$$\begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & \sin \psi \\ 0 & -\sin \psi & \cos \psi \end{pmatrix}. \text{ Corresponding to rotations about the } z, y \text{ and } x\text{-axes respectively.}$$

The matrices acting on the coordinates yield

- (i) $x' = x \cos \theta + y \sin \theta, y' = -x \sin \theta + y \cos \theta, z' = z,$
- (ii) $x' = x \cos \phi + z \sin \phi, y' = y, z' = -x \sin \phi + z \cos \phi,$
- (iii) $x' = x, y' = y \cos \psi + z \sin \psi, z' = -y \sin \psi + z \cos \psi.$

Considering (a) we define the generator in (i) as X_θ , where $X_\theta f = \frac{f(x', y', z') - f(x, y, z)}{\Delta \theta}$ as $\Delta \theta \rightarrow 0$,

where $\Delta \theta$ represents the change in θ from its magnitude corresponding to the identity element. For infinitesimally small change near the identity, we obtain, in

- (i) $x' = x + y\theta, y' = -x\theta + y, z' = z$, where $\theta = 0$ corresponds to the identity element. Thus,

$$X_\theta f = \left[\frac{f(x + y\theta, y - x\theta, z) - f(x, y, z)}{\theta} \right]_{\theta \rightarrow 0}$$

$$\text{or, } X_\theta f = y \frac{\partial f}{\partial x} - x \frac{\partial f}{\partial y} = \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) f$$

$$\text{or, } X_\theta = y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} = -\frac{iL_z}{\hbar}, \text{ where } L_z = z\text{-component of the orbital angular momentum operator } (\vec{r} \times \vec{p}).$$

- (ii) $y' = y, x' = x - z\phi, z' = z + x\phi$ so that

$$X_\phi f = \left[\frac{f(x - z\phi, y, z + x\phi) - f(x, y, z)}{\phi} \right]_{\phi \rightarrow 0}$$

$$= -\left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) f \text{ so that}$$

$$X_\phi = -\left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) = -\frac{iL_y}{\hbar}, \text{ and similarly, } X_\psi = -\frac{iL_x}{\hbar}$$

The generators satisfy the commutation relation,

$$(X_\theta, X_\phi) = -\frac{1}{\hbar^2} (L_z, L_y) = -X_\psi \text{ or, } (X_\phi, X_\theta) = X_\psi,$$

$$\text{Similarly, } (X_\psi, X_\phi) = X_\theta, (X_\theta, X_\psi) = -X_\phi \text{ or, } (X_\psi, X_\theta) = X_\phi.$$

- (b) The changes in the matrix elements due to infinitesimal changes in the parameters θ, ϕ and ψ respectively

$$\text{are } \begin{pmatrix} 0 & \theta & 0 \\ -\theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -\phi \\ 0 & 0 & 0 \\ \phi & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \psi \\ 0 & -\psi & 0 \end{pmatrix} \text{ [changes are found out with respect to the identity elements}$$

which are $\theta = 0, \phi = 0$ and $\psi = 0$ respectively.

The matrix generators are

$$X'_\theta = \left\{ \begin{bmatrix} 0 & \theta & 0 \\ -\theta & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} / \theta = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\text{and similarly } X'_\phi = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, X'_\psi = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \text{ respectively,}$$

$$\text{Thus, } (X'_\psi, X'_\phi) = -X'_\theta, (X'_\psi, X'_\theta) = -X'_\phi, (X'_\phi, X'_\theta) = -X'_\psi$$

The commutation relations are the negatives of those in the case (a).

Lie algebra – The Lie algebra is a vector space comprising the generators as elements where the product of the two elements implies their commutator. Thus we can write for any two elements X_i, X_j the commutator (X_i, X_j) (also called the L_{ie} bracket) given by $(X_i, X_j) = \sum C_{ijk} X_k$, is a linear combination of the generators and hence lies in the vector space.

The constants C_{ijk} are called the structure constants of the Lie group. Thus for the SO_3 group the structure constants are each equal to unity.

Since $[[X_i, X_j], X_k] + [[X_j, X_k], X_i] + [[X_k, X_i], X_j] = 0$ according to Jacobi's identity we have the following relation between the structure constants.

$$c_{ijl} c_{lkp} + c_{jkl} c_{lip} + c_{kil} c_{ljp} = 0.$$

An operator that commutes with all the generators is called the Casimir operator. The $SO(3)$ group has only one Casimir operator which is L^2 .

Lorentz group – The set of Lorentz transformations represented by, $x' = \gamma(x - Vt), y' = y, z' = z, t' = \gamma\left(t - \frac{Vx}{c^2}\right)$

is restricted to the velocity of initial frame along the x -direction. The general Lorentz transformations given by $x' = \Lambda x + x_0$, where Λ is the 4×4 matrix and x, x', x_0 are the four-dimensional vectors form a group called the inhomogeneous Lorentz group. The group property can be easily established by using the Lorentz transformation equation. The Lorentz invariant given by $x_4^2 - x_1^2 - x_2^2 - x_3^2$, where $x_4 = ct, x_1 = x, x_2 = y, x_3 = z$

leads to the relation $\Lambda^T g \Lambda = g$, g being the matrix
$$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$
 The inhomogeneous Lorentz group

including translations and reflections in space time, also called the Poincare group, is a ten-parameter group denoted by $O(3, 1)$. The group is noncompact because the parameter of transformation v lies in the open interval $(-c, c)$. The parameter space is thus bounded but not closed. The Lorentz group plays a significant role in physics because of the requirement of Lorentz invariance of the equations of relativistic quantum mechanics and quantum field theory under Lorentz transformations.

12.7 APPLICATIONS

1. Symmetry and degeneracy in quantum mechanics

An operation is said to be a symmetry operation for a system if the state of the system does not change its behaviour due to the action of the operator. For example, if the Hamiltonian does not change due to some symmetry operation the time-evolution of the state will remain essentially unaltered. For such a symmetry operation O we can write $O^{-1}HO = H$.

Thus the operator O commutes with the Hamiltonian. Since the Hamiltonian does not change, the action of the operator on a state of the system gives rise to a combination of the states of the system degenerate with respect to energy. If the degeneracy be n -fold the dimension of the degenerate subspace is n . The action of the operator on the state of the system will lead to a state that is a linear combination of all the n degenerate states. We can write $O|\psi_p^E\rangle = \sum_{m=1}^n O_{pm} \psi_m^E$ (the superscript E denotes that the states belong to the same energy

eigenvalue E). The matrix elements O_{pm} thus correspond to a n -dimensional irreducible representation of the symmetry operator O . For a particle in a central potential $v(r)$ the operator corresponding to a rotation about any axis does not change the Hamiltonian of the system and hence is a symmetry operator of the system. Thus for a state of given angular momentum j the degeneracy is $(2j + 1)$ fold. Hence we can write,

$$D^j|jm\rangle = \sum_{m'=-j}^j D_{m'm}^j |jm'\rangle. \text{ The matrix elements } D_{m'm}^j \text{ correspond thus to a } (2j + 1) \text{ dimensional irreducible}$$

representation of the rotation group. That the matrices defined as above form a representation can be proved as follows.

We define for any operator O as, $O_{XY} \psi_n = O_Y O_X \psi_n$

$$\text{Or, } D_{XY} \psi_n = D_Y \sum (D_X)_{nn'} \psi_{n'} = \sum (D_X)_{nn'} (D_Y)_{n'n''} \psi_{n''}$$

$$\text{Or, } D_{XY} \psi_n = \sum (D_X D_Y)_{nn''} \psi_{n''} = \sum (D_{XY})_{nn''} \psi_{n''}$$

$$\text{Or, } (D_{XY})_{nn''} = (D_X D_Y)_{nn''} \text{ for any } n, n'' \text{ proving thereby that the } D\text{-matrices form a representation.}$$

For an orthonormal basis the representation can be shown to be unitary. The degenerate subspace spanned by the eigenstates for a particular value of the energy is not always exhausted by the symmetry operations usually found in the physical problems. There occur degeneracies arising from the particular form of the potential called accidental degeneracies. The l -degeneracy for a particular energy in the hydrogen atom problem offers one such example. This occurs from some hidden symmetry in the problem as a consequence of the constancy of the Runge-Lenz vector for this coulomb potential problem. Thus, if we exclude the accidental degeneracies, the dimensions of the representations determine the degeneracies.

2. Bloch's theorem, Wave-function in a periodic potential

Let us consider a particle placed in a one-dimensional crystal with the distance between the two consecutive identical units equal to a . This is the problem of restricted translational invariance of the Hamiltonian of the system where H remains unchanged due to change of x by a . If we apply periodic boundary conditions at two walls n such units apart we get $na = L$, where L is the distance separating two walls. Thus the state returns to its original magnitude after a distance L so that $\psi(x) = \psi(x + L)$. If we consider a cyclic group of order n with elements $A, A^2, \dots, A^n (= I)$, we can understand a basic similarity between the original problem and that of a cyclic group. The cyclic group is a finite dimensional Abelian group and so its only irreducible representation is one dimensional. Thus the one-dimensional representation satisfies $[\Gamma(A)]^n = \Gamma(I) = 1$,

and hence has the one-dimensional matrix element $\Gamma(A) = e^{2\pi i m/n}$, where $m = 1, 2, \dots, n$. We can say that the eigenfunctions of the one-dimensional Hamiltonian with the periodicity of the crystal must transform according to the irreducible representation of the one-dimensional cyclic group. Thus we can write,

$$\psi(x+a) = \Gamma(A) \psi(x) = e^{2\pi i m/n} \psi(x) = e^{\frac{2\pi i m a}{L}} \psi(x).$$

Thus putting $k = \frac{2\pi m}{L}$, we can write, $\psi_k(x+a) = \psi_k(x) e^{ika}$.

Thus any function satisfying the above relation can be represented as, $\psi_k(x) = u_k(x) e^{ikx}$, where $u_k(x) = u_k(x+a)$.

The extension to three dimensions is obvious.

3. Splitting of spectral lines

We have earlier seen that the dimension of the irreducible representation of a symmetry group is equal to the degree of degeneracy of a level. If the environment of the system be changed, for example, by subjecting the system to an external field, the symmetry group is changed. The irreducible representation of the original group is now no longer irreducible for the new symmetry group, in general, and breaks up into irreducible representations of the new group according to the relation (12.4). For example, if an electron is subjected to a central potential in an atom, the Hamiltonian is spherically symmetric and the symmetry group is $SO(3)$. If the system is placed in an external magnetic field along the z -axis, say, the $SO(3)$ symmetry is broken and the new symmetry group is $SO(2)$, corresponding to rotation about the z -axis. The irreducible representation of the Abelian $SO(2)$ group is one dimensional. Hence, the original energy level now becomes split up into a number of levels equal to the degeneracy of the original level.

4. Symmetries in elementary particle physics – an introduction

The symmetries in elementary particle physics are of two types - (i) Space-time symmetry, (ii) internal symmetry. The space time symmetry group is the Poincare group which is a general inhomogeneous Lorentz group and the internal symmetry group the Lie group usually. The operator $p^i (i = 1, 2, 3, 4)$, the energy-momentum 4-vector, the generator of the Poincare group and the generators of the internal symmetry group G commute. The operation of the group G , therefore, keeps the norm of the energy-momentum 4-vector, i.e., the rest mass invariant. Hence the different states belonging to the representation of the group G correspond to the multiplets with the same mass. The simplest such group, the $SU(2)$ group, reflects the symmetry in the space of a quantum number called the isospin. The lowest dimensional representation, which is two, representing the isospin doublet corresponds to the states of the proton and the neutron which have isospin quantum no. $I = 1/2$ and $I_3 = \pm 1/2$, respectively. I_3 is thus related to the charge of the nucleons by the relation

$Q = \frac{1}{2} (1 + 2I_3)$. We can also have higher dimensional representations of the $SU(2)$ group by taking the direct

product of two $SU(2)$ groups in their lowest dimensions. Using $D(2) \otimes D(2) = D(3) \oplus D(1)$ we can have the isospin triplet corresponding to the $D(3)$ representation. The isospin triplet corresponding to $D(3)$ comprises π^+ and π^0 , the three pions having isospin quantum number $I = 1$ and $I_3 = \pm 1$ and 0 respectively. $I = 0$ corresponds to the Λ and the η -particles. These particles behave identically in respect of strong interactions. Other symmetry groups have been introduced to explain other particles with additional quantum numbers but these are not discussed here.

SOLVED PROBLEMS

Problem 1

- (a) Prove that a group is Abelian if and only if the elements of the group are isomorphic to their inverses.
 (b) Prove that all the elements in the same class of a cyclic group have the same order.

Solution:

- 1 (a) If there exist isomorphisms between each of the pairs (A, A^{-1}) and (B, B^{-1}) then AB should be isomorphic to $A^{-1} B^{-1}$ or $(BA)^{-1}$. Since according to the definition of isomorphism AB must be isomorphic to $(AB)^{-1}$, $AB = BA$ or the group is Abelian.
 (b) For a cyclic group, $A^n = I$ where n is the order of A . For A and B to be in the same class $B = G^{-1} A G$, G being any other element of the group. Now, $B^n = (G^{-1} A G)^n = G^{-1} A^n G = G^{-1} I G = I$. Hence the elements in a cyclic group have the same order if they belong to the same class.

Problem 2

- (a) Prove that the direct product of two matrices generally of different dimensions representing two groups forms another representation.
 (b) Prove that all the irreducible representations of the product group can be generated by taking the direct product of the irreducible representations of the individual groups.

Solution:

$$\begin{aligned}
 2(a) \quad D^{(\mu \times \nu)}(G_i H_j) D^{(\mu \times \nu)}(G_{i'} H_{j'}) &= [D^{(\mu)}(G_i) \times D^{(\nu)}(H_j)] [D^{(\mu)}(G_{i'}) \times D^{(\nu)}(H_{j'})] \\
 \text{or, } D^{(\mu \times \nu)}(G_i \cdot H_j) D^{(\mu \times \nu)}(G_{i'} H_{j'}) &= [D^{(\mu)}(G_i) D^{(\mu)}(G_{i'})] \times [D^{(\nu)}(H_j) D^{(\nu)}(H_{j'})] \\
 &= [D^{(\mu)}(G_i G_{i'})] \times [D^{(\nu)}(H_j H_{j'})] \\
 &= D^{(\mu \times \nu)}(G_i G_{i'} H_j H_{j'}) \\
 &= D_{ij i' j'}(GH).
 \end{aligned}$$

Thus the direct product of two groups gives a representation of the product group. It can be shown (proof is given in the standard texts on group theory, for example Wigner's book) that the direct product of two irreducible representations of two groups gives an irreducible representation of the product group. The direct product of two irreducible representations of the same group is, in general, reducible.

- (b) We have, $\sum_i (d_i^{G_1})^2 = g_{G_1}$ and $\sum_j (d_j^{G_2})^2 = g_{G_2}$, the symbols having the meanings followed earlier in the text.

$$\sum_{i,j} d_{ij}^2 = \sum_{i,j} (d_i^{G_1} d_j^{G_2})^2 = \sum_i (d_i^{G_1})^2 \sum_j (d_j^{G_2})^2 = g_{G_1} g_{G_2} = g_{G_1 \otimes G_2}$$

where $g_{G_1 \otimes G_2}$ denotes the number of elements or the order of the product group. The condition

$\sum_{i,j} d_{ij}^2 = g_{G_1 \otimes G_2}$ implies the irreducibility of the product group.

Problem 3

- (a) Prove that the elements common to two intersecting subgroups H_1 and H_2 form a subgroup of the group G .
- (b) Prove that a group, where order is a prime number must be a cyclic group.

Solution:

- 3 (a) (i) The product of the elements common to both H_1 and H_2 belongs to both of them and hence to the intersection subset H_3 .
- (ii) The associative law is valid for H_3
- (iii) The identity element and the inverse of any element belong both to H_1 and H_2 and, hence, to H_3 . Thus the elements belonging to both H_1 and H_2 form a subgroup of H .
- 3 (b) Since the order of a subgroup is a divisor of that of the entire group, the group with its order equal to a prime number cannot have any proper subgroup with order less than that of the group. (The one-element group containing only the identity element is excluded from our consideration). Since a group must contain apart from any element its various powers, a group with the order equal to the prime number must be a cyclic group with the order of an element equal to that of the group.

Problem 4

Prove that

- (a) The elements of the two different classes are all distinct.
- (b) A group comprises entire classes.

Solution:

- 4 (a) Let $(G_1, G_2, G_3, \dots)$ and $(G_1, H_2, H_3, \dots)$ be two classes so that the element G_1 is common between them. Now clearly $G_1 = g G_2 g^{-1}$, where g is any element of the group G . Since G_1 belongs also to the second class, $G_1 = g' H_2 g'^{-1}$, where g' is another element of G . We, thus, have, $g G_2 g^{-1} = g' H_2 g'^{-1}$.
- or $(g'^{-1} g) G_2 (g^{-1} g') = H_2$. Since $g'^{-1} g$ and $g^{-1} g'$ are two members of G , G_2 and H_2 are conjugate to each other and hence should belong to the same class. Similarly we can show that the elements of two classes are identical. But this contradicts the assumption that the two classes are distinct. Hence there can be no element common between two different classes.
- (b) The identity element forms a one-member class. We take any other element and find out from the group multiplication table the elements conjugate to it and thus combine them to form a class. Taking other elements in turn and ascertaining the elements conjugate to it we go on forming different classes till we exhaust the entire group. Thus the group becomes split up into a set of distinct classes.

Problem 5

Find out the symmetry group of a square and construct the character table.

Solution: The symmetry operations of a square are assuming the square to be in xy -plane, (i) rotations by $\frac{n\pi}{2}$ ($n = 0, 1, 2, 3$) about the z -axis, (ii) reflections with respect to the two axes passing through the centre and parallel to the sides of the square respectively (taken to be parallel to the x and the y -axes respectively), (iii) reflections about two diagonals.

We recall the matrix corresponding to rotation about the z -axes through an angle θ in the anticlockwise sense; given by $\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$. Thus, the four rotations stated in (i) are represented by $\pm I, \pm i\sigma_2$, where I is the identity matrix and $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. The reflections in the x and the y -axes corresponding to the transformations $x \rightarrow x, y \rightarrow -y$ and $x \rightarrow -x, y \rightarrow y$. The relevant matrix representations are given by $\pm \sigma_3$. The reflections with respect to the diagonals are described by the relations $x \rightarrow y, y \rightarrow x$ and $x \rightarrow -y, y \rightarrow -x$ respectively. They are represented by the matrices $\pm \sigma_1$. Thus the symmetry group is the eight-element group consisting of $\pm I, \pm \sigma_1, \pm i\sigma_2, \pm \sigma_3$.

The group comprises five classes formed by $(I), (I), (\pm \sigma_1), (\pm \sigma_3), (\pm i\sigma_2)$

From the relation $\sum_{\mu} n_{\mu}^2 = 8$, we get $n_1 = n_2 = n_3 = n_4 = 1, n_5 = 2$.

The relations $\sigma_1^2 = \sigma_3^2 = 1$ show that the characters of the one-dimensional representations given by σ_1 and σ_3 are ± 1 respectively. Again the relation $(\pm \sigma_3)(\pm \sigma_1) = i\sigma_2$ shows that the product of the characters of the classes in the left hand side $(\pm \sigma_3, \pm \sigma_1) = \text{character of the class } (\pm i\sigma_2)$. Since $(\pm i\sigma_2)^2 = -I$, the character of the second class $(-I) = 1$.

With the help of the orthogonality relations we can write the complete character table as

	I	$-I$	$(\pm \sigma_1)$	$(\pm \sigma_2)$	$(\pm i\sigma_2)$
	C_1	C_2	C_3	C_4	C_5
$D^{(1)}$	1	1	1	1	1
$D^{(2)}$	1	1	-1	-1	1
$D^{(3)}$	1	1	-1	1	-1
$D^{(4)}$	1	1	1	-1	-1
$D^{(5)}$	2	-2	0	0	0

SUMMARY

- A group comprises a set of elements along with the rule of composition which satisfies certain constraints, such as, (i) closure under composition, (ii) existence of the inverse and the identity elements, (iii) associativity and distributivity. These have been illustrated in the text with specific examples. A group with a finite number of elements is called a finite group with its order equal to that number. The cyclic group and the permutation group are two relevant examples in this context and have been explained in the text.
- A subgroup is a part of the original group, the members of the former forming a group under the same law of composition as the complete group. Apart from the identity element and the complete group all subgroups are called proper subgroups. The sets $aH, bH, \dots$, where $a, b, \dots$ are elements of the group other than the subgroup H are called the left cosets. Similarly the sets $Ha, Hb, \dots$ are called the right cosets. If all the left cosets are equal the right cosets the subgroup H is called an invariant subgroup.

The group multiplication table displays the elements obtained by the multiplication of the elements of the group and hence can show if the group contains any proper subgroup.

- The representation of a group means associating a matrix with each element of the group. such that the representative matrices satisfy the same multiplication rule as the original elements. The mapping from the elements to the matrices is isomorphic or homomorphic according as there is one-to-one or many to one relation between the two sets. The representation is completely reducible if all the matrices corresponding to the elements can be transformed by a single similarity transformation to the

form $\begin{bmatrix} D^{(1)} & O \\ O & D^{(2)} \dots \end{bmatrix}$, where $D^{(1)}, D^{(2)}$ are matrices of lower dimensions. Clearly, in this case there

are spaces of dimensions less than the original one that are left invariant by the actions of $D^{(1)}, D^{(2)} \dots$. If no similarity transformation can reduce the matrices to the block-diagonal forms the representation is irreducible. The irreducible representation is important because it serves as the building unit of all representations so that any representation D can be written as $D = a_1 D^{(1)} + a_2 D^{(2)} + \dots$, where $a_1, a_2, \dots$ are the number of times that the irreducible representations occur in the expression of D . The various properties of irreducible representations have been described in the text.

- A set of elements a, b , such that $g^{-1}ag = b$, g being any element of the group are said to form a class. The matrices corresponding to the elements of a class are related by similarity transformations to one another. The trace of the matrix representing any element of the class, also called the character, is the same for all the elements in the class and can be said to characterise a class. A table can be constructed showing the characters of the different classes in a group in various possible irreducible representations of different dimensions, in general. The different relations useful for the construction of the character tables are given in the text (though all of them have not been proved here). The character table enables one to know the different classes in a group, the various irreducible representations and the characters in different representations.
- Groups comprising an infinite number of elements are called infinite groups. The infinite groups may be discrete (for example, the group of integers including zero) or continuous (the rotation groups for example) in which the group elements vary continuously as functions of a set of parameters (the angles of rotation in the case of the rotation groups). The continuous groups are compact if the space of parameters is bounded and closed. (There are, of course, more rigorous definitions of compactness but they are not given here). The groups in which the group elements vary continuously as functions of the parameters are called Lie groups. The groups of linear transformation or the groups of rotations are examples of Lie groups. Expanding the group elements in the neighbourhood of the identity element we can define the generators of the group. The unitary or the orthogonal groups represented by $SU(n)$ or $SO(n)$ respectively, where n can be 2, 3, ..., are compact while the Lorentz groups are non-compact Lie groups. These groups have been briefly described in the text. The chapter concludes with a short discussion of a few applications of group theory to quantum mechanics.

EXERCISE

1. Find out which of the following are groups (also the conditions if necessary):
 - (a) The set of finite dimensional matrices under addition
 - (b) The set of the above matrices under multiplication.
 - (c) The set of all Hermitian $(n \times n)$ matrices under multiplication

- (d) The set of complex numbers satisfying $z^n = a$ ($a = \text{real}$) under multiplication
- (e) 1, 2, 3, 4, under multiplication (mod. 5)
(mod. n means dividing by n and taking the remainder as the net magnitude of the product).
- (f) The set of odd integers and zero under addition.
- (g) The set of complex numbers including zero, under multiplication.
- (h) The set of matrices of the type $\begin{pmatrix} \lambda & 0 & 0 \\ a & \lambda & 0 \\ b & c & \lambda \end{pmatrix}$, a, b, c and λ being constants.

Also, find out the nature of the group in each case. (Abelian or not)

2. Prove that for a group, the identity and the inverse are unique.
3. (a) Prove that $c_i c_j = \sum_k \alpha_{ijk} c_k$, where c_p, c_j and c_v denote the classes of a group and α_{ijk} 's are constants.
(b) Find out the one-dimensional representations of a dihedral group D_s .

[**Hint:** Follow the group multiplication table given in Table 12.1].

4. Find out the symmetry group of
(a) A rectangle,
(b) A square without considering reflections (b) an equilateral triangle if you do not consider reflections about the medians. Construct the multiplication tables.
5. The group elements satisfy $X^2 = Y^2 = Z^2 = XYZ = I$. (I is the identity element). Construct the multiplication and the character tables.
6. Do the following form Lie groups?
(a) The Lorentz transformation given by

$$x' = \frac{x - Vt}{\sqrt{1 - \frac{V^2}{c^2}}}, y' = y, z' = z, t' = \frac{t - \frac{Vx}{c^2}}{\sqrt{1 - \frac{V^2}{c^2}}}$$

- (b) $x' = ax + by + c, y' = dx + ey + f$ (a, b, c, d, e, f are constants) If so, find out the generators.
7. Prove that for a finite group the sum of the characters of all the elements in any irreducible representation of the group G except for the identity representation = 0.
[Use the relation (12.2) and take either x^r or x^s as the character corresponding to the identity representation, where x^r or $x^s = 1$]
8. (a) Prove that if a subgroup contains half of the elements of a group, it must be a normal subgroup.
(b) Find out the example of permutation symmetry in problems of quantum mechanics and explain its consequences.

[**Hint:** Look for the examples in the chapter on Identical Particles]

Hints and Answers to Exercise Questions

1. (a) Yes (A), (b) Yes (N.A),
 (c) They must commute (N.A) (d) Yes (A),
 (e) Yes (A), (f) No,
 (g) No, (h) Yes (N.A)
 (A $\Rightarrow$ Abelian, N.A $\Rightarrow$ Non-Abelian)
3. (b) $D(E) = 1, D(A) = D(B) = 1, D(c) = D(D) = D(F) = 1$
 and $D(A) = D(B) = D(E) = 1, D(C) = D(D) = D(F) = -1,$
6. (a) Yes, Generator is $-t \frac{\partial}{\partial x} - \frac{x}{c^2} \frac{\partial}{\partial t}$ (when acting as transformation operator on functions) and
- $$\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{c^2} & 0 & 0 & 0 \end{bmatrix} \text{ when acting directly on the coordinates.}$$
- (b) $x \frac{\partial}{\partial x}, y \frac{\partial}{\partial x}, \frac{\partial}{\partial x}, x \frac{\partial}{\partial y}, y \frac{\partial}{\partial y}, \frac{\partial}{\partial y}$. Students should try to get the other form.

Linear Integral Equations

CHAPTER

13

13.1 INTRODUCTION

An integral equation involves the function to be determined inside an integral. We will address the problems of linear integral equations that are expressed in a general form given by $y(x) = g(x) + \lambda \int_{\alpha}^{\beta} K(x, z)y(z)dz$, where the unknown function $y(x)$ appears under the integral sign. λ is a constant and $K(x, z)$ the kernel of the integral equation.

If $f(x) = 0$ we have a Fredholm integral equation of the first kind while if $f(x) = 1$ we obtain Fredholm's equation of the second kind. If $g(x) = 0$ the equation is called homogeneous while if $g(x) \neq 0$ we obtain an inhomogeneous integral equation. If $K(x, z) = 0$ for $z > x$ we have a Volterra integral equation. The Fourier transform

relations given by $f(x) = \int_{-\infty}^{\infty} g(K)e^{ikx} dk$ or $g(K) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)e^{-ikx} dx$ are examples

of integral equations, (Fredholm type of the first kind) where e^{ikx} or e^{-ikx} plays the role of the kernel. The Green's function method of solution also furnishes such an example of an integral equation, where the Green's function plays the role of the kernel.

Relation with a Differential Equation

We can show that a differential equation can after imposition of suitable boundary conditions be transformed to an integral equation. This is illustrated below by a simple example

Suppose, the equation is $\frac{d^2 y}{dx^2} = F(x, y)$ with the condition that $(y)_{x=0} = a$ and $(y')_{x=0} = b$.

We obtain after integrating once, $\frac{dy}{dx} = \int_0^x f\{z, y(z)\}dz + c$, c being an arbitrary constant.

This can be further integrated to yield,

$$y = \int_0^x \left[\int_0^\xi f\{z, y(z)\}dz d\xi \right] + cx + d,$$

where d is another constant.

By suitably changing the order of integration and properly adjusting the limits so that the region of integration is not altered we can write,

$$y = \int_0^x \left(\int_z^x d\xi \right) f\{z, y(z)\}dz + cx + d$$

$$\text{or, } y = \int_0^x (x-z)f\{z, y(z)\}dz + cx + d \quad (13.1)$$

To illustrate the above we take $y'' + y = 0$ with $y(0) = 0$, $y'(0) = 1$. We immediately obtain, from (13.1)

$$y = \int_0^x (z-x)y(z)dz + cx + d$$

Putting $y(0) = 0$ and $y'(0) = 1$ we get, $y(x) = x + \int_0^x (z-x)y(z)dz$

The equation (13.1) is a Volterra type integral equation. Similarly imposing the conditions on $y(x)$ at two values of x we can transform the differential equation into an integral equation of Fredholm type. Thus substituting in (13.1) the conditions $y(x = 0) = \alpha$ and $y(x = 1) = \beta$ we obtain, $d = \alpha$ and

$$\beta = \int_0^1 (1-z)f\{z, y(z)\}dz + c + \alpha \text{ giving } c = (\beta - \alpha) - \int_0^1 (1-z)f\{z, y(z)\}dz$$

Thus putting the values of c and d in (13.1) we get,

$$y(x) = F(x) + \int_0^1 K(x, z)f\{z, y(z)\}dz, \text{ where}$$

$F(x) = d + (\beta - \alpha)x$ while the kernel $K(x, z) = z(x-1)$ for $x \geq z$ and $= x(z-1)$ for $x \leq z$ respectively.

In the following we discuss briefly a few commonly used methods mainly by way of examples. Keeping the level of the readers in mind we will leave out the detailed theoretical discussions regarding integral equations.

A linear integral equation can be regarded as a transformation from $y(x)$ to $g(x)$ such that if corresponding to the functions $y_1(x), y_2(x), \dots$ there exist functions $g_1(x), g_2(x), \dots$ then $c_1y_1 + c_2y_2 + \dots$ will give rise to $c_1g_1 + c_2g_2 + \dots$. Similarly if $y_1, y_2, \dots$ are the solutions of the homogeneous integral equation then $c_1y_1 + c_2y_2 + \dots$ is also a solution of the same equation.

13.2 DEGENERATE KERNEL

A kernel $K(x, y)$ expressible as $K(x, y) = \sum_{i=1}^n \xi_i(x) \eta_i(y)$ is called separable for degenerate. The integral equation can then be written as,

$$y(x) = g(x) + \lambda \int \left(\sum_i \xi_i(x) \eta_i(z) \right) y(z) dz.$$

Thus we can write $y(x) = g(x) + \lambda \sum_i A_i \xi_i(x)$, where $A_i = \int \eta_i(z) y(z) dz$ and is a constant for Fredholm

integral equations where the integration is between the fixed limits. Putting the expressions for $y(x)$ in A_i we get a set of algebraic equations relating the different A_i 's. The A_i 's, found out by solving these equations, lead to the solution of the integral equation provided the denominators occurring in the expression for A_i 's do not vanish.

Example Let us take the equation $f(x) = x + \lambda \int_0^1 (x^2 y + xy^2) f(y) dy$

or,
$$f(x) = x + \lambda x^2 \int_0^1 y f(y) dy + \lambda x \int_0^1 y^2 f(y) dy \quad (13.1)(A)$$

or,
$$f(x) = x + \lambda A x^2 + \lambda B x, \text{ where } A = \int_0^1 y f(y) dy, B = \int_0^1 y^2 f(y) dy \quad (13.1)(B)$$

Substituting in (13.1)(B) we obtain,

$$A = \int_0^1 y(y + \lambda A y^2 + \lambda B y) dy = \frac{1}{3} + \frac{\lambda A}{4} + \frac{\lambda B}{3}$$

or, $A \left(1 - \frac{\lambda}{4} \right) - \frac{\lambda B}{3} = \frac{1}{3}$ yielding $A(12 - 3\lambda) - 4\lambda B = 4$ (13.2)(A)

and
$$B = \int_0^1 y^2 (y + \lambda A y^2 + \lambda B y) dy = \frac{1}{4} + \frac{\lambda A}{5} + \frac{\lambda B}{4}$$

as $\frac{\lambda A}{5} + B \left(\frac{\lambda}{4} - 1 \right) = \frac{-1}{4}$ giving $4A\lambda + B(5\lambda - 20) = -5$ (13.2)(B)

Solving (13.2)(A) and (13.2)(B) we get,

$$A = \frac{80}{\lambda^2 + 120\lambda - 240} \text{ and } B = -\frac{(\lambda + 60)}{\lambda^2 + 120\lambda - 240} \text{ respectively.}$$

Putting A and B in the original equation we obtain $f(x)$ if the denominator $\lambda^2 + 120\lambda - 240$ does not vanish. Clearly the values of λ when $\lambda^2 + 120\lambda - 240 = 0$ correspond to the situation when the corresponding

homogeneous integral equation $f(x) = \lambda \int_0^1 (x^2 y + xy^2) f(y) dy$ has non-trivial solutions. These values of λ are

called the eigenvalues of the integral equation.

A few statements given below can be made (proof is omitted) regarding the solutions of integral equations with degenerate kernels, that are bounded.

- (1) Either the inhomogeneous equation $y(x) = g(x) + \lambda \int_a^b K(x, z)y(z)dz$ possesses a unique finite solution

or the corresponding homogeneous equation has at least one non-trivial solution called the eigenfunction for the particular value of λ called the eigenvalue of the integral equation. Also if λ is an eigenvalue of the integral equation it also corresponds to an eigenvalue of the transposed equation

$$y(x) = \lambda \int_a^b K(z, x)y(z)dz.$$

- (2) For an eigenvalue λ the inhomogeneous equation has a solution if and only if $\int_a^b F(x)g(x)dx = 0$, where

$$F(x) \text{ satisfies the equation } F(x) = \lambda \int_a^b K(z, x)F(z)dz.$$

13.3 ITERATIVE SOLUTION AND OTHER METHODS

A straightforward solution of the integral equation $y(x) = g(x) + \lambda \int_a^b K(x, z)y(z)dz$ when λ is small can be obtained by the method of iteration, also called the method of Neumann series.

We start with the zeroth order approximation $y_0(x) = g(x)$.

The successive approximants are given by $y_1(x) = g(x) + \lambda \int_a^b K(x, z)g(z)dz$,

$$y_2(x) = g(x) + \lambda \int_a^b K(x, z)g(z)dz + \lambda^2 \int_a^b dz \int_a^b K(x, z)K(z_1, z)g(z_1)dz \text{ and so on.}$$

For the n th order approximant we can write

$$y_n(x) = g(x) + \sum_{r=1}^n \lambda^r \int_a^b K_r(x, z)g(z)dz, \text{ where}$$

$$K_0(x, z) = K(x, z); K_1(x, z) = \int_a^b K(x, z_1)K(z_1, z)dz_1, \dots, K_n(x, z) = \int_a^b K(x, z_1)K_{n-1}(z_1, z)dz_1$$

$$\text{The series can be shown to converge if } |\lambda|^2 \int_a^b dx \int_a^b |K(x, z)|^2 dz < 1 \quad (13.3)$$

As an example in physics we consider the scattering state solutions of schroedinger equation

$$\nabla^2 \psi + \frac{2m}{\hbar^2} E \psi = \frac{2m}{\hbar^2} V \psi.$$

Thus for small V we can write the different order solutions as, $\psi_0 = e^{ikz}$, corresponding to a plane wave with wave number k moving along the z -direction,

$\psi_1 = e^{ikz} + \alpha \int G(\bar{r}, \bar{r}') V(\bar{r}') \psi_0(\bar{r}') d\bar{r}'$, where $G(\bar{r}, \bar{r}')$, the Green's function given by $\frac{e^{ik|\bar{r}, \bar{r}'|}}{4\pi|\bar{r}, \bar{r}'|}$, plays the role of the kernel and $\alpha = -\frac{2m}{4\pi\hbar^2}$. We can continue the expansion to get,

$$\begin{aligned}
 \psi(\bar{r}) &= e^{ikz} - \frac{2m}{4\pi\hbar^2} \int G(\bar{r}, \bar{r}') V(\bar{r}') \psi_0(\bar{r}') d\bar{r}' \\
 &+ \left(\frac{2m}{4\pi\hbar^2} \right)^2 \iint G(\bar{r}, \bar{r}') G(\bar{r}, \bar{r}'') V(\bar{r}') V(\bar{r}'') \psi_0(\bar{r}'') d\bar{r}' d\bar{r}'' + \dots
 \end{aligned}$$

This is called the Born series in the scattering theory where the different terms in the expansion represent multiple scattering events at different points of the scatterer and the integrals extend over the region occupied by the scatterer.

Fredholm obtained the solution in the form of a series by subdividing the interval $a < x < b$ and replacing the definite integral by the limit of a sum as the number of subdivisions tends to infinity. The resulting algebraic equations are solved and lead to the solution given by $y(x) = g(x) + \lambda \int_a^b R(x, z, \lambda) g(z) dz$, where $R(x, z, \lambda)$, called the resolvent kernel is given by $R(x, z, \lambda) = \frac{D(x, z, \lambda)}{D(\lambda)}$,

$$\begin{aligned}
 D(\lambda) &= 1 - \lambda \int du K(u, u) + \frac{\lambda^2}{2!} \int du du' \begin{vmatrix} K(u, u) & K(u, u') \\ K(u', u) & K(u', u') \end{vmatrix} - \dots, \\
 D(x, z, \lambda) &= K(x, z) - \lambda \int du \begin{vmatrix} K(u, z) & K(u, u) \\ K(u, z) & K(u, u) \end{vmatrix} \\
 &+ \frac{\lambda^2}{2!} \int du du' \begin{vmatrix} K(x, z) & K(x, u) & K(x, u') \\ K(u, z) & K(u, u) & K(u, u') \\ K(u', z) & K(u', u) & K(u', u') \end{vmatrix} \quad (13.4)
 \end{aligned}$$

The power series (13.4) converges for all λ unlike in the case of Neumann series where λ has to satisfy the condition (13.3).

Other Methods: Apart from the methods described above, there exist a few other techniques that enable us to arrive at the solution in a comparatively easy way for some specific problems outlined below.

(a) Let us take a Volterra equation of the form $y(x) = x^2 + \int_0^x x^2 z y(z) dz$

or, $y(x) = x^2 + x^2 f(x)$, where $f(x) = \int_0^x z y(z) dz$

or, $f'(x) = x y(x) = x(x^2 + x^2 f(x))$

or, $f'(x) - x^3 f(x) = x^3$ yielding the solution,

$$f(x) e^{\frac{-x^4}{4}} = \int x^3 e^{\frac{-x^4}{4}} dx + c, \text{ } c \text{ being any constant.}$$

$$\text{This gives, } f(x) = c e^{\frac{x^4}{4}} - 1$$

$$\text{or, } y(x) = c x^2 e^{\frac{x^4}{4}}$$

Putting in the original integral equation we get,

$$y(x) = x^2 + x^2 C \int_0^x z^3 e^{\frac{z^4}{4}} dz \text{ giving } C = 1.$$

$$\text{The solution is therefore, } y(x) = x^2 e^{\frac{x^4}{4}}$$

(b) Let the kernel $K(x, y)$ be of the form $K(x, y) = K(x - y)$. and the integral equation be $y(x) = g(x) + 1 \int_{-\infty}^{\infty} K(x - z) y(z) dz$.

Taking the Fourier transform we obtain, $\bar{y}(K) = \bar{g}(K) + \lambda \bar{K}(K) \bar{y}(K)$, (see the chapter on Fourier transform) so that we get, $\bar{y}(K) = \frac{\bar{g}(K)}{1 - \lambda \bar{K}(K)}$

Taking the inverse Fourier transform we obtain $y(x)$. If $K(x, z)$ vanishes for $z < x$ and $y(x) = 0$ for $x < 0$, we have to perform a Laplace transform in order to arrive at the solution.

13.4 SCHMIDT-HILBERT THEORY

The Schmidt-Hilbert theory of integral equations resembles closely Sturm-Lionville theory of differential equations. We introduce below briefly the essential features of the theory.

Similar to the Hermitian operators in Sturm-Lionville theory the kernel $K(x, y)$ playing the role of an integral operator is said to be Hermitian if $K(y, x) = K^*(x, y)$. By following the same procedure as in the case

of Sturm-Lionville theory one can prove that the homogeneous integral equation $y(x) = \lambda \int_a^b K(x, z) y(z) dz$,

or written symbolically as $y = \lambda K y$, has all the eigenvalues real and the eigenfunctions belonging to different eigenvalues mutually orthogonal.

[Hint: One can write as a definition of the Hermitian Kernel, $(y_2, K y_1) = (K y_2, y_1)$, where y_1 and y_2 are two functions on which the integral operator K operates. Thus, if they correspond to different eigenvalues λ_1 and

λ_2 respectively one can easily prove that $(y_2, y_1) = \int_a^b y_2^* y_1 dx = 0$ for $\lambda_1 \neq \lambda_2$ and λ 's are real]

If the eigenvalues are degenerate the eigenfunctions can be orthogonalised by Gram-Schmidt method similar to that in the Sturm-Lionville theory.

For the inhomogeneous case we write, $y(x) = g(x) + \lambda \int_a^b K(x, z) y(z) dz$.

If the y_i 's be the eigenfunctions of the homogeneous equation $y_i = \lambda_i K y_i$, we obtain, by writing $y = g + \sum_i c_i y_i$, the c_i 's being the appropriate expansion coefficients,

$$g + \sum_i c_i y_i = g + \lambda K g + \lambda \sum_i \frac{c_i y_i}{\lambda_i}, \text{ where } K g = \int_a^b K(x, z) g(z) dz.$$

Taking the inner product with y_j , using the Hermitian property of the Kernel and the orthonormality of the corresponding eigenfunctions we can write, $y(x) = g(x) + \lambda \sum_i \left(\frac{\int_a^b y_i^*(x) g(x) dx}{\lambda_i - \lambda} \right) y_i$ (13.5)

Clearly if $\lambda = \lambda_j$, the integral equation has a finite solution only if $\int_a^b y_j^*(x) g(x) dx = 0$, i.e. the eigenfunction is orthogonal to $g(x)$

SOLVED PROBLEMS

Problem 1

$$1. \text{ Solve :- } u(x) = \sin x + \lambda \int_0^1 x^2 t u(t) dt$$

Solution:

Putting $A = \int_0^1 t u(t) dt$ we get, $u(x) = \sin x + \lambda x^2 A$

or, $A = \int_0^1 t [\sin t + \lambda t^2 A] dt = (\sin 1 - \cos 1) + \frac{\lambda A}{4}$, so that

$$A = \left(\frac{\sin 1 - \cos 1}{1 - \frac{\lambda}{4}} \right).$$

Thus, $u(x) = \sin x + \left(\frac{\lambda}{1 - \frac{\lambda}{4}} \right) x^2 (\sin 1 - \cos 1)$.

The solution exists for all λ except $\lambda = 4$ which is seen to be the eigenvalue of the homogeneous equation

$$u = \lambda \int_0^1 x^2 t u(t) dt.$$

Problem 2

Solve by Neumann series method to terms $O(\lambda^2)$,

$$f(x) = x^3 + \lambda \int_0^1 x^2 y^2 f(y) dy.$$

Solution: First approximation: $f(x) = x^3$.

Second approximation: $f(x) = x^3 + \lambda \int_0^1 x^2 y^5 dy = x^3 + \frac{\lambda x^2}{6}$

Third approximation: $f(x) = x^3 + \lambda \int_0^1 x^2 y^2 \left(y^3 + \frac{\lambda y^2}{6} \right) dy$

or,
$$f(x) = x^3 + \frac{\lambda x^2}{6} + \frac{\lambda^2 x^2}{30}$$

Problem 3

Solve:- $f(x) = x^2 + \int_0^x f(y) dy$

Solution: Differentiating with respect to x we get,

$$\frac{df}{dx} = 2x + f(x), \text{ yielding the solution}$$

$$f e^{-x} = 2 \int x e^{-x} dx = -2[e^{-x} x e^{-x}] + c$$

or, $f = -2[1 + x] + c e^x$

Putting in the original equation we obtain, $c = 2$. Thus, the solution is, $f(x) = 2(e^x - x - 1)$.

Problem 4

Solve by Schmidt-Hilbert method:- $f(x) = e^x + \lambda \int_0^1 e^{x+y} f(y) dy$

Solution: The eigenfunction of the homogeneous equation is given by, $f(x) = \lambda \int_0^1 e^{x+y} f(y) dy = \lambda A e^x$,

where $A = \int_0^1 e^y f(y) dy$

or, $A = \int_0^1 e^y f(y) dy = \lambda A \int_0^1 e^{2y} dy = \frac{\lambda A}{2} (e^2 - 1)$ giving, $\lambda = \frac{2}{e^2 - 1}$

The normalized eigenfunction is given by $f(x) = \sqrt{\frac{2}{e^2 - 1}} \cdot e^x$

Substituting in the equation (13.5) we obtain,

$$\begin{aligned} y(x) &= e^x + \frac{\lambda}{\left(\frac{2}{e^2-1} - \lambda\right)} \left(\int_0^1 \sqrt{\frac{2}{e^2-1}} e^{2x} dx \right) \sqrt{\frac{2}{e^2-1}} \cdot e^x \\ &= e^x + \frac{\lambda}{\left(\frac{2}{e^2-1} - \lambda\right)} e^x = \frac{2e^x}{2 - \lambda(e^2-1)} \end{aligned}$$

SUMMARY

The chapter introduces briefly the integral equations in which the unknown function appears within the integral as outlined in the following:

- Differential equations with the boundary conditions included, give rise to integral equation.
- There are two types of integral equation,

(a) Fredholm type, such as, $y(x) = g(x) + \lambda \int_a^b K(x, z)y(z)dz$, and (b) Volterra type, such as, $y(x) = \lambda \int_a^x K(x, z)y(z)dz + h(x)$, $K(x, z)$ being called the kernel and g, h the given functions.

- (c) Some standard methods of solving integral equations, stated below, have been discussed with examples.
- The integral equations with separable kernels.
 - The Neumann series method which involves an infinite series expansion in powers of a parameter.
 - The Fredholm determinant method which also yield, the solution as a ratio of two infinite series each of which involves determinants of different orders.
 - Volterra type equations that admits of a solutions by converting the integral equation into a differential equation by differentiation.
 - The method of integral transforms that yields solutions of integral equations with some special types of kernels.
 - The Schmidt-Hilbert method, which discusses some essential features of integral equations.

EXERCISE

- Solve the following problems involving separable kernels, determining the eigenvalues and the eigenfunctions of the homogeneous problem.

(a) $u(x) = \lambda \int_a^b e^{x+y} u(y) dy,$

$$(b) \quad u(x) = x^2 + \lambda \int_0^1 x u(t) dt,$$

$$(c) \quad u(x) = x + \lambda \int_0^\pi \cos(x-t) u(t) dt.$$

2. Solve the following equations by (i) the method of Neumann series, (ii) the method of Fredholm determinant.

$$(a) \quad y(x) = x^2 + \lambda \int_0^1 x^2 z^2 y(z) dz,$$

$$(b) \quad y(x) = e^{-x} + \lambda \int_a^b e^{-x-z} y(z) dz$$

3. Solve the following Volterra equations

$$(a) \quad f(x) = x^2 + 2 \int_0^x y f(y) dy$$

$$(b) \quad y(x) = x^2 - \int_0^x x^2 z^2 y(z) dz$$

4. Solve the following equations by Schmidt-Hilbert method

$$(a) \quad y(x) = x^2 + \lambda \int_0^1 x^2 z^2 y(z) dz,$$

$$(b) \quad y(x) = 2 \cos x + \lambda \int_0^\pi \sin(x-z) y(z) dz$$

Hints and Answers to Exercise Questions

$$1. (a) \quad u(x) = \frac{2e^x}{(e^{2b} - e^{2a})}, \quad (b) \quad u(x) = x^2 + \frac{2\lambda x}{3(1-\lambda)}.$$

$$(c) \quad u(x) = x + 1 (A \cos x + B \sin x), \text{ where } A = \frac{4}{\lambda-2}, B = \frac{2\pi}{2-\lambda}.$$

$$2. (a) \quad y(x) = x^2 \left(1 + \frac{\lambda}{5} + \frac{\lambda^2}{15} \right) + \dots$$

$$(b) \quad y(x) = e^x \left[1 - \frac{\lambda}{2} (e^{-2b} - e^{-2a}) - \frac{\lambda^2}{2} (b-a)(e^{-2b} - e^{-2a}) + \dots \right]$$

$$3. (a) \quad f(x) = e^{x^2} - 1.$$

$$(b) \quad y(x) = x^2 e^{-x^{5/5}}$$

4. (a) $y = \frac{5x^2}{5-\lambda}.$

(b) $y = 2 \cos x + \lambda A\pi \left[\frac{A(\sin x + \cos x)}{\left(\frac{2}{\pi} - \lambda\right)} - \frac{1}{\left(\frac{2}{\pi} + \lambda\right)} (\sin x - \cos x) \right],$

where $A^2 = \pi/4$.

Bibliography

CHAPTER 1

1. M.L.Boas-Mathematical methods in physical sciences-Wiley Student Edition (3rd Edition, 2011),
2. K.F.Riley, M.P.Hobson and S.J.Bence- mathematical methods for Physics and Engineering-Cambridge University Press (3rd. Editions), 2002.
3. B.M.Budak and S.V.Fomin-Multiple integrals, field theory and series - Mir publishers.Moscow (2nd. Edition), 1978.
4. W. Kaplan-Advanced calculus- Addition-Wesley publishing company, Inc., (Ist. Edition), 1962.
5. H.Margenean and G.M.Murphy- The mathematics of Physics and Chemistry- D. Van.Nostrand company, Inc. (1956).

CHAPTER 2

1. H.Lass- Vector and Tensor Analysis- International student Edition, McGraw Hill, Kogakusha Limited.
2. G.B.Arflen and H.J.Weber- Mathematical methods for physicists (6th. Edition), Academic Press, 2005.
3. Tulsi Dass and Satish Sharma- Mathematical methods in classical and Quantum physics- Universities Press (India) Private Limited., 2009.
4. K.F.Riley, M.P. Hobson and S. J. Bence- Mathematical methods for physics and Engineering-Cambridge University Press (3rd Edition) 2002.
5. B.M. Budak and S.V. Fomin-Multiple integrals, field theory and series- Mir publishers, Moscow (2nd Edition), 1978.

B.2 Fundamentals of Mathematical Physics

6. M.L. Boas- Mathematical methods in physical Sciences -Wiley Student Edition (3rd. Edition), 2011.
7. M.R.Spiegel- Theory and Problems of Vector Analysis- Schanm's Outline Series- Mc-Graw Hill Book Company, 1974.

CHAPTER 3

1. R.Courant and D.Hilbert- Methods of mathematical physics (vol-1) Interscience, New York (1953).
2. P.R.Halmos- Finite Dimensional Vector spaces, D. Van Nostrand and Company, Inc. Princeton, N.J. (1958).
3. Tulsi Dass and Satish Sharma- Mathematical methods in classical and quantum physics- Universities Press (India) Pvt. Ltd, (2nd. Edition), 2009.
4. Mathews and Walker- Mathematical methods of physics, (2nd Edition)/Pearson Education.
5. J.J.Sakurai- Modern quantum mechanics. Pearson Education (4th. Impression), 2009, (Particularly for application to the problems of quantum mechanics).

CHAPTER 4

1. H.Lass- vector and Tensor Analysis- International Student Edition- Mc.Graw Hill, Kogakusha Ltd.
2. A.W.Joshi- Matrices and Tensors in physics. Wiley Eastern, New Delhi (1973).
3. G.B.Arftken and H.J.Weber- Mathematical methods for physicists (6th-Edition), Academic Press, 2005.
4. H.Margenean and G.M.Murphy- The Mathematics of Physics and Chemistry- D.Van Nostrand company, Inc. (1956).
5. General relativity and cosmology - Sriranjana Banerjee and Asit Banerjee, Elsevier, a division of reed Elsevier India Pvt. Ltd. (2007).

CHAPTER 5

1. B.M.Budak and S.V. Fomin- Multiple integrals, field theory and series- Mir publishers (Moscow) (2nd Edition), 1978.
2. N.Pishkunov- Differential and Integral calculus- volume II, Mir publishers, Moscow (Student Edition), 1974.
3. G.B.Arftken and H.J. Weber. Mathematical methods for physicists, (6th. Edition) Academic press, 2005.
4. M.L.Boas- Mathematical methods in physical sciences- Wildy student Edition (3rd Edition), 2011.
5. E.T.Whittaker and G.N.Watson- A course of modern analysis- Cambridge University Press (1946).

CHAPTER 6

1. R.V.Churchill, J.W.Brown and R.F.Verley- complex Variables and Applications-McGraw Hill International student Edition, 1979.
2. E.T.Copson- An introduction to the functions of a complex variable- Oxford University Press. (1962).
3. P.Dennery and A.Krzywicki- Mathematics for physicists- Dover Publications (1995).

4. P.K.Chattopadhyay- Mathematical physics (2nd. Edition), New Age International Pvt. Ltd, 2013.
5. F.W.Byron and R.W.Fueller-Mathematical methods of classical and quantum physics- Dover Publications. Inc., 1992.
6. M.L.Boas - Mathematical methods in physical sciences- Wiley student Edition (3rd. Edition), 2011.

CHAPTER 7

1. G.F.Simmons- Differential equations with applications and Historical Notes- Tata McGraw Hill, New Delhi (1979).
2. Mathews and Walker- Mathematical methods of physics- Pearson Education (2nd. Edition),
3. P.M.Morse and H.Feshbach- Methods of theoretical physics, McGraw Hill, New York, 1953.
4. P.Dennery and A.Krzywicki- Mathematics for physicists - Dover publications (1995).
5. Tulsi Dass and Satish Sharma- Mathematical methods in classical and Quantum physics - Universities Press (India) Pvt. Ltd. (2nd. Edition), 2009.
6. M.L.Boas- Mathematical methods in physical sciences- Wiley student Edition (3rd. Edition), 2011.

CHAPTER 8

1. R.Courant and D.Hilbert- Methods of mathematical physics. (Vol.1) Interscience, New York (1953).
2. P.Dennery and A.Krzywicki- Mathematics for physicists- Dover publications (1995).
3. J.Mathews and R.L.Walker- Mathematical methods of physics. Pearson Education (2nd. Edition).
4. P.M.Morse and H.Feshbach- Methods of theoretical physics. McGraw Hill, New York, 1953.
5. Tulsi Dass and Satish Sharma- Mathematical methods in classical and Quantum physics - Universities Press (India) Pvt. Ltd., (2nd. Edition) 2009.

CHAPTER 9

1. P.Denery and A.Krzywicki- Mathematics for physicists - Dover publications. 1995.
2. G.B.Arffen and H.J.Weber- Mathematical methods for physicists -Academic press (6th. Edition), 2005.
3. M.L.Boas- Mathematical methods in physical sciences- Wiley student Edition (3rd. Edition), 2011.
4. E.T.Whittaker and G.N.Watson- A course of modern analysis- Cambridge University Press (1946).
5. R.Courant and D.Hilbert- Methods of mathematical physics - (Vol.1) Interscience, New York (1953).

CHAPTER 10

1. J.Mathews and R.L.Walker- Mathematical methods of Physics, Pearson Education (2nd. Edition).
2. P.K.Chattopadhyay- Mathematical physics (2nd. Edition)- New Age International Pvt. Ltd., 2013.

B.4 Fundamentals of Mathematical Physics

3. G.B.Arffen and H.J.Weber- Mathematical methods for physicists Academic Press (6th. Edition), 2005.
4. I.Sneddon- Elementary partial differential Equations- Mc.Graw Hill International Edition, 1957.

CHAPTER 11

1. J. Mathews and R.L.Walker - Mathematical methods of Physics, Pearson Education (2nd. Edition) 2005.
2. T.K.Chandra and D.Chatterjee - A First course in Probability-Narosa Publishing House Pvt. Ltd. (3rd. Edition) 2005.
3. K.F.Riley, M.P.Hobson and S.J.Bence - Mathematical Methods for Physics and Engineering - Cambridge University Press (2nd. Edition) 2005.
4. Tulsi Dass and Satish Sharma - Mathematical Methods in classical and Quantum Physics - Universities Press (India) Pvt. Ltd. 2009.
5. M.L.Boas - Mathematical Methods in Physical Sciences - Wiley Student Edition (3rd. Edition) 2011.

CHAPTER 12

1. M. Tinkham - Group Theory and Quantum mechanics-McGraw Hill Book Bompany, New York (1964).
2. A.W.Joshi- Elements of Group Theory for physicists, Wiley Eastern, New Delhi (1973).
3. J.Mathews and R.L.Walker- Mathematical Methods of Physics, Pearson Education (2nd. Edition).

CHAPTER 13

1. J.Mthews and R.L.Walker- Mathematical methods of Physics, Pearson Education (2nd. Edition). 2005.
2. Mathematical methods for Physics and Engineering - V.F.Riley, M.P.Hobson and S.J.Bence. Cambridge University Press (2nd. Edition) 2005.
3. P.M.Morse and H.Feshbach-Methods of theoretical Physics- McGraw Hill, New York, 1953.
4. H.Margenean and G.M.Murphy-The Mathematics of Physics and Chemistry- D.Van Nostrand Company, Inc. (1956).