M.Sc Previous Year
Physics
MP 01

MATHEMATICAL PHYSICS

मध्यप्रदेश भोज (मुक्त) विश्वविद्यालय – भोपाल
MADHYA PRADESH BHOJ (OPEN) UNIVERSITY - BHOPAL

Reviewer Committee

1. Dr. P.K. Agrawal
   Director
   IEHE College, Bhopal (M.P.)
2. Dr. S.K.Vijay
   Professor Govt. MLB College,
   Bhopal (MP)
3. Dr. Seema Hardikar
   Professor Sarojini Naidu Govt Girls PG
   (Autonomous) College, Bhopal (MP)

Advisory Committee

1. Dr. Jayant Sonwalkar
   Hon'ble Vice Chancellor
   Madhya Pradesh Bhoj (Open) University,
   Bhopal
2. Dr. L.S. Solanki
   Registrar
   Madhya Pradesh Bhoj (Open) University,
   Bhopal
3. Dr. Neelam Wasnik
   Asst. Director
   Madhya Pradesh Bhoj (Open) University,
   Bhopal
4. Dr. P.K. Agrawal,
   Director
   IEHE College,
   Bhopal (M.P.)
5. Dr. S.K.Vijay
   Professor
   Govt. MLB College, Bhopal (MP)
6. Dr. Seema Hardikar
   Professor
   Sarojini Naidu Govt Girls PG
   (Autonomous) College, Bhopal (MP)

COURSE WRITERS

B D Gupta, Former Reader & Head Department of Mathematics, Vardhaman College BIJNOR (UP) (Rohilkhand University)
Units (1.0-1.5, 1.8-1.16, 2.0-2.1, 2.6, 2.7, 2.8, 2.10, 2.13-2.17, 3.0-3.6, 3.8-3.14, 4.0-4.5, 4.10-4.14, 5.0-5.1, 5.3, 5.8-5.14)

V K Khanna and S K Bhambri, Retd. Associate Professor, Deptt. of Mathematics, Kirori Mal College, University of Delhi
Units (1.7)

KA Lakshminarayanan, Senior Faculty, Deptt. of Mathematics and Computer Applications, PSG College of Technology, Coimbatore
V Sundaram, Senior Faculty, Deptt. of Mathematics and Computer Applications, PSG College of Technology, Coimbatore
R Balasubramanian, Senior Faculty, Deptt. of Mathematics and Computer Applications, PSG College of Technology, Coimbatore
Units (5.2, 5.5, 5.6)

Dr Samiksha Sirakwar, Assistant Professor, Integrated Basic Science School for Physical and Decision Science, Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow.
Dr Bhupendra Pratap Singh, Former Faculty, Deptt of Physics, University of Lucknow
Units (1.6, 2.2-2.5, 2.7.1, 2.9, 2.11, 2.12, 3.7, 4.6-4.9, 5.4, 5.7)

Published by Registrar, MP Bhoj (open) University, Bhopal in 2020

VIKAS®

Vikas® is the registered trademark of Vikas® Publishing House Pvt. Ltd.

VIKAS® PUBLISHING HOUSE PVT. LTD.
E-28, Sector-8, Noida - 201301 (UP)
Phone: 0120-4078900 • Fax: 0120-4078999
Regd. Office: A-27, 2nd Floor, Mohan Co-operative Industrial Estate, New Delhi 1100 44
• Website: www.vikaspublishing.com • Email: helpline@vikaspublishing.com

SYLLABI-BOOK MAPPING TABLE

Mathematical Physics

Syllabi	Mapping in Book
UNIT - I: Vectors, Matrices and Tensors Curvilinear coordinates, Orthogonal curvilinear coordinate system, Derivation of gradient, divergence and curl in polar, spherical and cylindrical coordinate systems. Eigenvalue problem, Cayley-Hamilton theorem, Function of Matrix, Kronecker sum and product of matrices. Definition of Tensor, Coordinate transformation, Controvariant , Covariant and Mixed tensors. Addition, subtraction, multiplication and contraction operations with tensor. Quotient law. Christoffel's symbols.	**Unit-1:** Vectors, Matrices, And Tensors **(Pages 3-80)**
UNIT - II: Partial Differential Equations and Group Theory Solutions of the following partial differential equations with boundary and initial conditions. Replace equation·, Poisson equation, wave equation. Heat conduction equation and its application, to rectangular bar with finite and infinite length. Definition of group, subgroups, classes, invariant subgroups, factor group complexes, Isomorphism and Isomorphism group representation. Direct sum and product, Reducible and irreducible representations, Schur's lemmas and orthogonality theorem, character of a representation some applications of group theory in physics: classification of states and elementary particles splitting of energy levels, Matrix elements and selection rules.	**Unit-2:** Partial Differential Equations and Group Theory **(Pages 81-162)**
UNIT - III: Functions of Complex Variable Definition, Argand diagram, function of a complex variable, Derivatives, Analyticity of complex function. Cauchy-Reimann conditions, Cauchy's theorem, Cauchy's intergral formula, poles, residue, Cauchy's Residue theorem, Contour Integration.	**Unit-3:** Functions of Complex Variable **(Pages 163-218)**
UNIT - IV: Special Functions and Spherical Harmonics Legendre, Bessel; Hermite and laugurre functions. Their generating functions, Recursions, relations properties. Spherical Harmonics, Series solutions of Hermite and Laugurre polynomials, their generating functions, orthogonality, Associated laugurre polynomials. Hypergeometric functions, representation of Bessel, Laguerre and Legendre functions in terms of hypergeometric functions.	**Unit-4:** Special Functions and Spherical Harmonics **(Pages 219-278)**
UNIT - V: Integral Transform Fourier transform and its properties, Application of Fourier transform to Dirac delta function and potential problems. Replace transform and its properties. Applications to potential and oscillatory problems. Evaluation of Simple integrals using Fourier and Laplace transforms.	**Unit-5:** Integral Transform **(Pages 279-318)**

CONTENTS

INTRODUCTION

NOTES

Mathematical physics seeks to apply rigorous mathematical ideas to problems in physics, or problems inspired by physics. As such, it is a remarkably broad subject. The *Journal of Mathematical Physics* defines the field as 'the application of mathematics to problems in physics and the development of mathematical methods suitable for such applications and for the formulation of physical theories'. Traditionally mathematical physics has been quite closely associated to ideas in calculus, particularly those of differential equations. In recent years, however, in part due to the rise of superstring theory, there has been a great enlargement of branches of mathematics which can now be categorized as part of mathematical physics.

Mathematical physics emphasizes tools and techniques of particular use to physicists and engineers. It focuses on vector spaces, matrix algebra, differential equations (especially for boundary value problems), integral equations, integral transforms, infinite series, and complex variables. Its approach can be tailored to applications in electromagnetism, classical mechanics, and quantum mechanics.

This book, *Mathematical Physics* is divided into five units that follow the self-instruction mode with each unit beginning with an Introduction to the unit, followed by an outline of the Objectives. The detailed content is then presented in a simple but structured manner interspersed with Check Your Progress Questions to test the student's understanding of the topic. A Summary along with a list of Key Terms and a set of Self-Assessment Questions and Exercises is also provided at the end of each unit for recapitulation.

UNIT 1 VECTORS, MATRICES, AND TENSORS

NOTES

Structure

1.0 INTRODUCTION

It is generally observed that there exist two types of physical measurements in applied mathematics, physics and mechanics: one involving *only magnitude and no direction* in the space of three dimensions, such as volume, mass, length, speed, temperature, potential, electric charge, etc., and the other involving *a definite direction in space associated with their magnitudes* such as velocity, acceleration, momentum, force, electric or magnetic field intensities, etc., the former being called *scalar quantities* or simply *scalars* and the latter, *vector quantities* or simply *Vectors*. The complete characterisation of a vector quantity requires length, support and sense, *i.e.*, a specified unit, a number stating how many times that unit is contained in that quantity and the statement of the direction. In this unit, you will learn about gradient, divergence and curl along with matrix and its different types. You will also learn various operations that can be performed with matrices and how the inverse of a matrix is computed by elementary transformations. The matrix inversion process is explained for finding matrix B that satisfies the prior equation for a given invertible matrix A. In linear algebra, an n-by-n square matrix A is called invertible or non-singular, if there exists an n-by-n matrix B such that, $AB = BA = In,$ where In denotes the n-by-n identity matrix and the multiplication used is ordinary matrix multiplication. In this case, matrix B is uniquely determined by A and is called the inverse of A, denoted by A^{-1}. The rank of a matrix is the maximum number of independent rows or the maximum number of independent columns. In fact, any matrix can be reduced to Echelon form by elementary row

NOTES

operations. A system of equations is called consistent if and only if there is a common solution to all of them, otherwise it is called inconsistent. You will learn how an eigenvector of a matrix cannot correspond to two different eigenvalues and how an eigenvalue of a matrix can and will correspond to different eigenvectors. In this unit you will also learn the method to verify the Cayley-Hamilton theorem to express the equations in matrix form and the nature of characteristic roots of a diagonal. Tensor Analysis forms that part of study which is rather suitable for the mathematical formulation of natural laws in forms which are invariant with respect to underlying frames of reference. In brief, tensors are quantities obeying certain transformation laws. In wider sense a tensor formulation is very compact and good deal of clarity in its use. The tensor formulation was originated by G. Ricci and it became rather popular when Albert Einstein used it as a natural tool for the description of his general theory of relativity.

In mathematics, a tensor is a geometric object that maps in a multi-linear manner geometric vectors, scalars, and other tensors to a resulting tensor. Vectors and scalars which are often used in elementary physics and engineering applications, are considered as the simplest tensors. Vectors from the dual space of the vector space, which supplies the geometric vectors, are also included as tensors. Geometric in this context is chiefly meant to emphasize independence of any selection of a coordinate system. This unit introduces you to the concept of tensor. You will also study the algebra of tensors.

1.1 OBJECTIVES

After going through this unit, you will be able to:

- Explain the meaning of gradient, divergence and curl
- Describe the use of matrix and its various types
- Explain various operations performed on matrices
- Describe the symmetric, skew-symmetric, Hermitian, skew-Hermitian, orthogonal, unitary and adjoint matrices
- Discuss the method of finding inverse and rank of a matrix
- Explain various elementary transformations performed on matrices
- Describe the meaning of tensors
- Discuss tensors as classification of transformation laws
- Explain symmetric and anti-symmetric tensors
- State the rules which govern tensor analysis

1.2 CURVILINEAR COORDINATES

We know that the equations

$$u = f_1(x, y, z);\ v = f_2(x, y, z);\ w = f_3(x, y, z)$$

where u, v, w are parameters, represent three families of surfaces when expressed in the form

$$u = \text{const.},\ v = \text{const.},\ w = \text{const.}, \quad \text{... (1.1)}$$

where u, v, w are continuously differentiable functions defined in any region R of space.

NOTES

Suppose that the three surfaces u = const., v = const., w = const., intersect in a point P of the region R. The values of u, v, w for the three surfaces intersecting at P are called the *curvilinear co-ordinates* of the point P. The three surfaces are then known as *co-ordinate surfaces.* The three surfaces intersect pairwise in three curves known as *co-ordinate curves.* Only one co-ordinate is variable on each of the co-ordinate curves. The curve on which u varies known as u-curve and similarly v-curve and w-curve are those on which v and w respectively vary. One variable is constant on each of the co-ordinate surfaces. The surface on which u is constant is known as u-surface and similarly v-surface and w-surface are those on which v and w respectively are constant.

Using the equation (1.1) the rectangular co-ordinates (x, y, z) and therefore the position vector **r** of any point in the region of space may be expressed in terms of curvilinear co-ordinates. Since there is a one to one correspondence between x, y, z and u, v, w the position vector **r** is a vector function of u, v, w.

Note. The loci of $u = C_1$, $v = C_2$, $w = C_3$; C_1, C_2, C_3 being constants represent the co-ordinate surfaces and the equations of the co-ordinate curves then are

$$v = C_2,\ w = C_3;\ w = C_3,\ u = C_1;\ u = C_1,\ v = C_2.$$

1.3 ORTHOGONAL CURVILINEAR COORDINATES SYSTEM

A system of orthogonal curvilinear co-ordinates is one which corresponds to the points of intersection of a triply orthogonal system of three families of surfaces

$$u\,(x, y, z) = \text{const.},\ v\,(x, y, z) = \text{const.},\ w\,(x, y, z) = \text{const.}$$

which are such that, through each point P in any region R of space passes one and only one member of each family, each of the three surfaces cutting the other two orthogonally. In short the curvilinear co-ordinates u, v, w are said to be orthogonal if the co-ordinate curves are mutually perpendicular at every point P (x, y, z) of space.

Let us suppose that $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ form a right handed system of unit vectors tangent to the co-ordinate curves u, v, w respectively at P and directed towards increasing u, v, w. Then we have

$$\left.\begin{aligned} \mathbf{e}_1 = \mathbf{e}_2 \times \mathbf{e}_3,\ \mathbf{e}_2 = \mathbf{e}_3 \times \mathbf{e}_1 \text{ and } \mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2 \\ \mathbf{e}_1 \cdot \mathbf{e}_2 = \mathbf{e}_2 \cdot \mathbf{e}_3 = \mathbf{e}_3 \cdot \mathbf{e}_1 = 0 \end{aligned}\right\} \quad \text{... (1.2)}$$

Let the arc lengths measured along the co-ordinate curves in the positive directions of u, v, w be respectively s_1, s_2, s_3. Now consider an infinitesimal parallelopiped whose diagonal is the element of arc ds along a curve tangent to PQ at P and faces coincide with planes u, v or w and length of edges are ds_1, ds_2, ds_3. Therefore,

$$ds^2 = ds_1^{\ 2} + ds_2^{\ 2} + ds_3^{\ 3}. \quad \text{... (1.3)}$$

NOTES

Let us now introduce the three numbers h_1, h_2, h_3 known as *metrical coefficients* with the property

$$\frac{ds_1}{du} = h_1, \frac{ds_2}{dv} = h_2, \frac{ds_3}{dw} = h_3$$

i.e., $$ds_1 = h_1 du,\ ds_2 = h_2 dv,\ ds_3 = h_3 dw. \qquad \text{... (1.4)}$$

Substituting, the values of ds_1, ds_2, ds_3 from (1.4) in (1.3), we get

$$ds^2 = h_1^2 du^2 + h_2^2 dv^2 + h_3^2 dw^2. \qquad \text{... (1.5)}$$

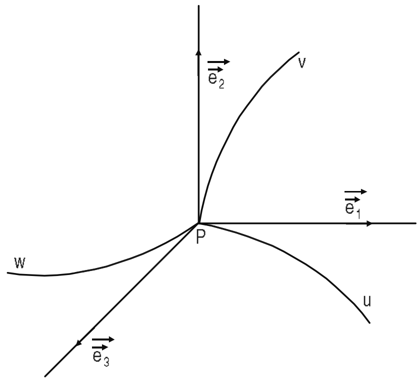

Fig 1.1

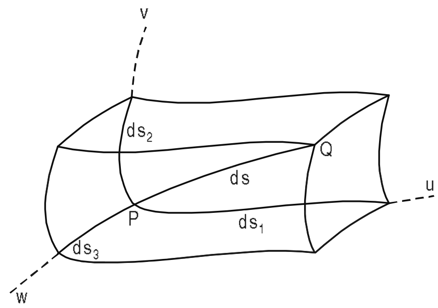

Fig. 1.2

Now if **r** be the position vector of P, referred to the origin of a rectangular co-ordinate system, the tangents to the co-ordinate curves at P are parallel to the directions of $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and have the magnitudes h_1, h_2, h_3 respectively. Therefore,

$$\frac{\partial \mathbf{r}}{\partial u} = h_1\, \mathbf{e}_1, \quad \frac{\partial \mathbf{r}}{\partial v} = h_2\, \mathbf{e}_2, \quad \frac{\partial \mathbf{r}}{\partial w} = h_3\, \mathbf{e}_3. \qquad \text{... (1.6)}$$

These give, $$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = h_1 h_2\, \mathbf{e}_1 \times \mathbf{e}_2$$

$$= h_1 h_2\, \mathbf{e}_3 \qquad \text{from (1.2)}$$

i.e., $$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = \frac{h_1 h_2}{h_3} \frac{\partial \mathbf{r}}{\partial w} \qquad \text{from (1.6)}$$

Similarly $$\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w} = \frac{h_2 h_3}{h_1} \frac{\partial \mathbf{r}}{\partial u} \qquad \text{from (1.6)} \qquad \text{... (1.7)}$$

and $$\frac{\partial \mathbf{r}}{\partial w} \times \frac{\partial \mathbf{r}}{\partial u} = \frac{h_3 h_1}{h_2} \frac{\partial \mathbf{r}}{\partial v} \qquad \text{from (1.6)}$$

Also $$\left[\frac{\partial \mathbf{r}}{\partial u}\, \frac{\partial \mathbf{r}}{\partial v}\, \frac{\partial \mathbf{r}}{\partial w}\right] = \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) \cdot \frac{\partial \mathbf{r}}{\partial w} = h_1 h_2 \mathbf{e}_3 \cdot h_3 \mathbf{e}_3 = h_1 h_2 h_3.$$

... (1.8)

Check Your Progress

1. What are curvilinear co-ordinates?
2. What do you mean by co-ordinate curves?

1.4 GRADIENT, DIVERGENCE, AND CURL

In this section, we will study about gradient, divergence and curl.

NOTES

The Gradient of A Scalar Field

Consider a scalar function *i.e.*, a function whose value depends upon the values of coordinates (x, y, z). Being a scalar its value is constant at a fixed point in space.

The gradient of any scalar function ϕ is defined as

$$\text{grad } \phi = \mathbf{i}\frac{\partial \phi}{\partial x} + \mathbf{j}\frac{\partial \phi}{\partial y} + \mathbf{k}\frac{\partial \phi}{\partial z} = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial x}\right)\phi \quad \text{... (1.9)}$$

$$= \nabla \phi$$

where operator ∇ is generally known as '*del*' or '*nabla*' operator and read as 'gradient' or 'grad' in short. We have already mentioned that a scalar field is the region in which the scalar point function specifies the scalar physical quantity like temperature, electric potential, density, etc. It is represented by a continuous scalar function giving the value of the quantity at each point. In scalar field all the points having same value of ϕ can be connected by means of surfaces, which are called *equal* or *level* surfaces.

Consider a co-ordinate system with axes such that any level surface lies in *x*-*y* plane while *z*-axis is along the normal to that level surface. Since the value of ϕ does not change along the level surface, *i.e.*,

$$\frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial y} = 0,$$

therefore $$\text{grad } \phi = \mathbf{k}\frac{\partial \phi}{\partial z}. \quad \text{... (1.10)}$$

Clearly grad ϕ is directed along *z*-axis, *i.e.*, along the normal to the level surface. Therefore equation (9) may be written as

$$\text{grad } \phi = \frac{\partial \phi}{\partial n}\mathbf{n} \quad \text{... (1.11)}$$

where **n** is unit vector along the normal to the level surface at any point.

From equation (1.11) we may state, "*The magnitude of grad ϕ at any point is rate of change of function ϕ with distance along the normal to the level surface at the point and is directed along unit vector* **n**."

Note. It is to be noted that gradient of any scalar quantity is a vector.

Problem 1.1. *Prove* $\nabla r^n = nr^{n-2}\mathbf{r}$.

$$\text{L.H.S.} = \nabla r^n = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)(r^n)$$

$$= \mathbf{i}\frac{\partial r^n}{\partial x} + \mathbf{j}\frac{\partial r^n}{\partial y} + \mathbf{k}\frac{\partial r^n}{\partial z}$$

$$= \mathbf{i}\, nr^{n-1}\frac{\partial r}{\partial x} + \mathbf{j}\, nr^{n-1}\frac{\partial r}{\partial y} + \mathbf{k}\, nr^{n-1}\frac{\partial r}{\partial z}$$

$$= nr^{n-1}\left[\mathbf{i}\frac{\partial r}{\partial x} + \mathbf{j}\frac{\partial r}{\partial y} + \mathbf{k}\frac{\partial r}{\partial z}\right]$$

NOTES

since $r^2 = x^2 + y^2 + z^2$; $\therefore \quad \frac{\partial r}{\partial x} = \frac{x}{r}, \frac{\partial r}{\partial y} = \frac{y}{r}, \frac{\partial r}{\partial x} = \frac{z}{r}$

$$= nr^{n-2}\,(\mathbf{i}\,x + \mathbf{j}\,y + \mathbf{k}\,z) = nr^{n-2}\,\mathbf{r}.$$

Problem 1.2. *If* **r** *is the position vector of a point, deduce the value of grad (1/r).*

As given, $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

So that $\frac{1}{r} = \frac{1}{(x^2+y^2+z^2)^{1/2}}$.

$\therefore$ grad $(1/r) = \nabla\,(1/r)$

$$= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)\left\{\frac{1}{(x^2+y^2+z^2)^{1/2}}\right\}$$

$$= \mathbf{i}\frac{\partial}{\partial x}\left[\frac{1}{(x^2+y^2+z^2)^{1/2}}\right] + \mathbf{j}\frac{\partial}{\partial y}\left[\frac{1}{(x^2+y^2+z^2)^{1/2}}\right] + \mathbf{k}\frac{\partial}{\partial z}\left[\frac{1}{(x^2+y^2+z^2)^{1/2}}\right]$$

$$= \mathbf{i}\left[-\frac{1}{2}\cdot\frac{2x}{(x^2+y^2+z^2)^{3/2}}\right] + \mathbf{j}\left[-\frac{1}{2}\cdot\frac{2y}{(x^2+y^2+z^2)^{3/2}}\right] + \mathbf{k}\left[-\frac{1}{2}\cdot\frac{2z}{(x^2+y^2+z^2)^{3/2}}\right]$$

$$= -\frac{x\mathbf{i}+y\mathbf{j}+z\mathbf{k}}{(x^2+y^2+z^2)^{3/2}} = -\frac{\mathbf{r}}{(r^2)^{3/2}} = -\frac{\mathbf{r}}{r^3}.$$

The Gradient of a Scalar-Point Function

If $\phi\,(x, y, z)$ be defined and differentiable at each point (x, y, z) in a certain region of space specified as a scalar field, we have

$$= \left(l\frac{\partial\phi}{\partial x} + m\frac{\partial\phi}{\partial y} + n\frac{\partial\phi}{\partial z}\right) = (l\mathbf{i} + m\mathbf{j} + n\mathbf{k})\cdot\left(\mathbf{i}\frac{\partial\phi}{\partial x} + \mathbf{j}\frac{\partial\phi}{\partial y} + \mathbf{k}\frac{\partial\phi}{\partial z}\right) \quad \text{... (1.12)}$$

Its R.H.S. is the scalar product of two vectors $(l\mathbf{i} + m\mathbf{j} + n\mathbf{k})$ and $\left(\mathbf{i}\frac{\partial\phi}{\partial x} + \mathbf{j}\frac{\partial\phi}{\partial y} + \mathbf{k}\frac{\partial\phi}{\partial z}\right)$, where the vector $(l\mathbf{i} + m\mathbf{j} + n\mathbf{k})$ is a unit vector along a line whose direction cosines are l, m, n and the second vector depends only on the point (x, y, z) and not on any direction. Thus we conclude that directional derivative along any line can be obtained by multiplying the vector

$\mathbf{i}\frac{\partial\phi}{\partial x} + \mathbf{j}\frac{\partial\phi}{\partial y} + \mathbf{k}\frac{\partial\phi}{\partial z}$ scalarly with the unit vector $l\mathbf{i} + m\mathbf{j} + n\mathbf{k}$.

The vector function $\mathbf{i}\frac{\partial\phi}{\partial x}+\mathbf{j}\frac{\partial\phi}{\partial y}+\mathbf{k}\frac{\partial\phi}{\partial z}$ is called *the gradient of a scalar-point function* ϕ and is written as grad ϕ or $\nabla\phi$. Thus,

NOTES

$$\nabla\phi = \text{grad } \phi \equiv \mathbf{i}\frac{\partial\phi}{\partial x}+\mathbf{j}\frac{\partial\phi}{\partial y}+\mathbf{k}\frac{\partial\phi}{\partial z} \equiv \frac{\partial\phi}{\partial x}\mathbf{i}+\frac{\partial\phi}{\partial y}\mathbf{j}+\frac{\partial\phi}{\partial z}\mathbf{k}.$$

It is clear that the gradient of a scalar-point function is a vector.

In case, ϕ is a constant, grad $\phi = 0$ since $\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}$, all will be zero in that case. Its converse is also true.

The Gradient or Sum of Two Scalar-Point Functions

If u and v are two differentiable scalar functions of x, y, z, then the gradient of their sum is given by

$$\begin{aligned}
\nabla(u+v) &= \left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial x}+\mathbf{k}\frac{\partial}{\partial x}\right)(u+v) \\
&= \mathbf{i}\frac{\partial}{\partial x}(u+v)+\mathbf{j}\frac{\partial}{\partial y}(u+v)+\mathbf{k}\frac{\partial}{\partial z}(u+v) \\
&= \mathbf{i}\frac{\partial u}{\partial x}+\mathbf{i}\frac{\partial v}{\partial x}+\mathbf{j}\frac{\partial u}{\partial y}+\mathbf{j}\frac{\partial v}{\partial y}+\mathbf{k}\frac{\partial u}{\partial z}\mathbf{k}\frac{\partial v}{\partial z} \\
&= \left(\mathbf{i}\frac{\partial u}{\partial x}+\mathbf{j}\frac{\partial u}{\partial y}+\mathbf{k}\frac{\partial u}{\partial z}\right)+\left(\mathbf{i}\frac{\partial v}{\partial x}+\mathbf{j}\frac{\partial v}{\partial y}+\mathbf{k}\frac{\partial v}{\partial z}\right) \\
&= \left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)u+\left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)v \\
&= \nabla u+\nabla v.
\end{aligned}$$

Showing that the gradient of sum of two scalar-point functions is equal to the sum of their gradients.

This rule may be generalised for any number of scalar-point functions.

The Gradient of product of Two Scalar-Point Functions

If u and v be two differentiable scalar-point functions of x, y, z, then the gradient of their product is given by

$$\begin{aligned}
\nabla(uv) &= \left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)(uv) \\
&= \mathbf{i}\frac{\partial}{\partial x}(uv)+\mathbf{j}\frac{\partial}{\partial y}(uv)+\mathbf{k}\frac{\partial}{\partial z}(uv) \\
&= \mathbf{i}\left(u\frac{\partial v}{\partial x}+v\frac{\partial u}{\partial x}\right)+\mathbf{j}\left(u\frac{\partial v}{\partial y}+v\frac{\partial u}{\partial y}\right)+\mathbf{k}\left(u\frac{\partial v}{\partial z}+v\frac{\partial u}{\partial z}\right) \\
&= u\left[\mathbf{i}\frac{\partial v}{\partial x}+\mathbf{j}\frac{\partial v}{\partial y}+\mathbf{k}\frac{\partial v}{\partial z}\right]+v\left[\mathbf{i}\frac{\partial u}{\partial x}+\mathbf{j}\frac{\partial u}{\partial y}+\mathbf{k}\frac{\partial u}{\partial z}\right] \\
&= u\left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)v+v\left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)u \\
&= u\nabla v+v\nabla u.
\end{aligned}$$

Showing that the gradient of the product of two scalar-point functions is obtained by the same rule as is valid for derivatives of the algebraic functions.

NOTES

The Divergence of a Vector-Point Function

If $\mathbf{V}(x, y, z) = V_1\mathbf{i} + V_2\mathbf{j} + V_3\mathbf{k}$ *be a continuous differentiable vector-point function specified in a vector field, then the divergence of* **V** *is defined as:*

$$\mathbf{i}\cdot\frac{\partial \mathbf{V}}{\partial x}+\mathbf{j}\cdot\frac{\partial \mathbf{V}}{\partial y}+\mathbf{k}\cdot\frac{\partial \mathbf{V}}{\partial z}$$

and is written as $\nabla \cdot \mathbf{V}$ *or div* **V** *and read as divergence* **V**.

$$\therefore \qquad \nabla\cdot\mathbf{V} = \left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)\cdot\left(V_1\mathbf{i}+V_2\mathbf{j}+V_3\mathbf{k}\right)$$

$$=\frac{\partial V_1}{\partial x}+\frac{\partial V_2}{\partial y}+\frac{\partial V_3}{\partial z}$$

which is clearly a scalar quantity.

Note. If $\nabla \cdot \mathbf{V} = 0$ then **V** is known as *Solenoidal Vector.*

The Divergence of Sum of two Vector Functions

If **U** and **V** be two vector-point functions expressed as

$$\mathbf{U} = U_1\mathbf{i} + U_2\mathbf{j} + U_3\mathbf{k}$$
$$\mathbf{V} = V_1\mathbf{i} + V_2\mathbf{j} + V_3\mathbf{k}.$$

Then

$$\nabla\cdot(\mathbf{U}+\mathbf{V}) = \left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)\cdot[(U_1+V_1)\,\mathbf{i}+(U_2+V_2)\,\mathbf{j}+(U_3+U_3)\,\mathbf{k}]$$

$$= \frac{\partial}{\partial x}[U_1+V_1]+\frac{\partial}{\partial y}[U_2+V_2]+\frac{\partial}{\partial z}[U_3+V_3]$$

$$= \left(\frac{\partial U_1}{\partial x}+\frac{\partial U_2}{\partial y}+\frac{\partial U_3}{\partial z}\right)+\left(\frac{\partial V_1}{\partial x}+\frac{\partial V_2}{\partial y}+\frac{\partial V_3}{\partial z}\right)$$

$$= \left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)\times(U_1\mathbf{i}+U_2\mathbf{j}+U_3\mathbf{k})$$
$$+\left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)\cdot(V_1\mathbf{i}+V_2\mathbf{j}+V_3\mathbf{k})$$
$$= \nabla\cdot\mathbf{U}+\nabla\cdot\mathbf{V}.$$
$$= \text{div}\,\mathbf{U}+\text{div}\,\mathbf{V}.$$

Showing that the divergence of the sum of two vector functions is equal to the sum of their divergences.

This rule may be generalised for any number of vector functions.

The Divergence of Product

If the vector point function **U** is expressed as

$\mathbf{U} = U_1\mathbf{i} + U_2\mathbf{j} + U_3\mathbf{k}$ and V is a scalar point-function.

$$\text{Then} \nabla \cdot (\mathbf{U}V) = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \cdot [(U_1\mathbf{i} + U_2\mathbf{j} + U_3\mathbf{k})V]$$

$$= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \cdot [VU_1\mathbf{i} + VU_2\mathbf{j} + VU_3\mathbf{k}]$$

$$= \frac{\partial}{\partial x}(VU_1) + \frac{\partial}{\partial y}(VU_2) + \frac{\partial}{\partial z}(VU_3)$$

$$= U_1\frac{\partial V}{\partial x} + V\frac{\partial U_1}{\partial x} + U_2\frac{\partial V}{\partial y} + V\frac{\partial U_2}{\partial y} + U_3\frac{\partial V}{\partial z} + V\frac{\partial U_3}{\partial z}$$

$$= \left(U_1\frac{\partial V}{\partial x} + U_2\frac{\partial V}{\partial y} + U_3\frac{\partial V}{\partial z}\right) + V\left(\frac{\partial U_1}{\partial x} + \frac{\partial U_2}{\partial y} + \frac{\partial U_3}{\partial z}\right)$$

$$= \left(\mathbf{i}\frac{\partial V}{\partial x} + \mathbf{j}\frac{\partial V}{\partial y} + \mathbf{k}\frac{\partial V}{\partial z}\right) \cdot (U_1\mathbf{i} + U_2\mathbf{j} + U_3\mathbf{k})$$

$$+ V\left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \cdot (U_1\mathbf{i} + U_2\mathbf{j} + U_3\mathbf{k})$$

$$= (\nabla V) \cdot \mathbf{U} + V(\nabla \cdot \mathbf{U})$$

i.e., $\quad \text{div}(\mathbf{U}V) = (\text{grad } V) \cdot \mathbf{U} + V \text{ div } \mathbf{U}.$

The curl or Rotation of a Vector Point Function

Let $\mathbf{f}(x, y, z) = f_1\mathbf{i} + f_2\mathbf{j} + f_3\mathbf{k}$ be a continuous differentiable vector-point function; then the curl of **f** or rotation of **f** is given by

$$\left(\mathbf{i} \times \frac{\partial \mathbf{f}}{\partial x} + \mathbf{j} \times \frac{\partial \mathbf{f}}{\partial y} + \mathbf{k} \times \frac{\partial \mathbf{f}}{\partial z}\right)$$

and is written as curl **f** or $\nabla \times \mathbf{f}$ or rot **f**.

i.e., $$\nabla \times \mathbf{f} = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \times (f_1\mathbf{i} + f_2\mathbf{j} + f_3\mathbf{k})$$

$$= \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial z}\right)\mathbf{j} + \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right)\mathbf{k}$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & f_3 \end{vmatrix}$$

It is clear that **curl f** or **rotatioin f** is a vector quantity and read as **del cross f**.

Note. If curl = 0, f is known as *Irrotatinal Vector.*

Interpretation of the curl f. *If a rigid body is in motion, the curl of its linear velocity at any point gives twice its angular velocity.*

Consider the motion of a rigid body rotating with angular velocity ω about an axis *OA*; *O*, being a fixed point in the body. Let **r** be the position vector of any point *P* of the body. Draw *PQ* perpedicular from *P* to the axis *OA*. Then,

NOTES

Linear velocity V of P due to circular motion

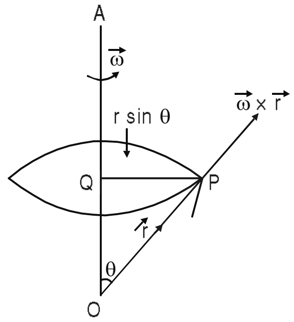

Fig. 1.3

$$= |\mathbf{V}|$$

$$= \omega QP = \omega r \sin\theta = |\vec{\omega} \times \mathbf{r}|$$

i.e., $\mathbf{V} = \vec{\omega} \times \mathbf{r}$

where, $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$

and $\vec{\omega} = \omega_1\mathbf{i} + \omega_2\mathbf{j} + \omega_3\mathbf{k}$.

But we know that curl $\mathbf{V}$ = $\nabla \times \mathbf{V} = \nabla \times (\vec{\omega} \times \mathbf{r})$

$$= \nabla \cdot \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \omega_1 & \omega_2 & \omega_3 \\ x & y & z \end{vmatrix}$$

$$= \nabla \times [(\omega_2 z - \omega_3 y)\,\mathbf{i} + (\omega_3 x - \omega_1 z)\,\mathbf{j} + (\omega_1 y - \omega_2 x)\,\mathbf{k}]$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \omega_2 z - \omega_3 y & \omega_3 x - \omega_1 z & \omega_1 z - \omega_2 x \end{vmatrix}$$

$$= 2\,[\omega_1\mathbf{i} + \omega_2\mathbf{j} + \omega_3\mathbf{k}]$$

$$= 2\vec{\omega}$$ which proves the proposition.

Curl of the Sum of Two Vector-Point Functions

If **u** and **v** be two vector-point functions given by

$$\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$$

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$

then $\nabla \times (\mathbf{u} + \mathbf{v}) = \nabla \times [(u_1 + v_1)]\mathbf{i} + (u_2 + v_2)\,\mathbf{j} + (u_3 + v_3)\,\mathbf{k}]$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ u_1 + v_1 & u_2 + v_2 & u_3 + v_3 \end{vmatrix}$$

$$= \mathbf{i}\left[\frac{\partial}{\partial y}(u_3 + v_3) - \frac{\partial}{\partial z}(u_2 + v_2)\right] + \mathbf{j}\left[\frac{\partial}{\partial z}(u_1 + v_1) - \frac{\partial}{\partial x}(u_3 + v_3)\right]$$

$$+ \mathbf{k}\left[\frac{\partial}{\partial x}(u_2 + v_2) - \frac{\partial}{\partial y}(u_1 + v_1)\right]$$

NOTES

$$= \mathbf{i}\left[\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}\right] + \mathbf{j}\left[\frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}\right] + \mathbf{k}\left[\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}\right]$$

$$+ \mathbf{i}\left[\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right] + \mathbf{j}\left[\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right] + \mathbf{k}\left[\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right]$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u_1 & u_2 & u_3 \end{vmatrix} + \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix}$$

$$= \nabla \times \mathbf{u} + \nabla \times \mathbf{v}. \qquad \text{(by the definition)}$$

i.e., curl $(u + v)$ = curl **u** + curl **v**.

Hence curl of sum of two vector point functions is equal to the sum of their curls.

The result may be generalised for any number of vector-point functions.

Note. If **r** is the position vector of a variable point with respect to a fixed origin such that $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ then curl $\mathbf{r} = \mathbf{0}$.

$$\text{Since curl } \mathbf{r} = \left(\mathbf{i}\times\frac{\partial}{\partial x} + \mathbf{j}\times\frac{\partial}{\partial y} + \mathbf{k}\times\frac{\partial}{\partial z}\right) \times (x\mathbf{i} + y\mathbf{i} + z\mathbf{k})$$

$$= \left[\frac{\partial}{\partial y}(z) - \frac{\partial}{\partial z}(y)\right] + \mathbf{j}\left[\frac{\partial}{\partial z}(x) - \frac{\partial}{\partial x}(z)\right] + \mathbf{k}\left[\frac{\partial}{\partial x}(y) - \frac{\partial}{\partial y}(x)\right]$$

$$= \mathbf{0}.$$

Curl of the Product of two Vector-Point Functions

We have to consider the curl of the forms $u\mathbf{v}$ and $\mathbf{u} \times \mathbf{v}$ where u is a scalar and **u**, **v** vector point functions.

Suppose, $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$

$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$

and u is a scalar point function.

Then, curl $(u\mathbf{v}) = \nabla \times (u\mathbf{v}) = \left(\mathbf{i}\times\frac{\partial}{\partial x} + \mathbf{j}\times\frac{\partial}{\partial y} + \mathbf{k}\times\frac{\partial}{\partial z}\right)(uv_1\mathbf{i} + uv_2\mathbf{j} + uv_3\mathbf{k})$

$$= \nabla \times (uv_1\mathbf{i} + uv_2\mathbf{j} + uv_3\mathbf{k})$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ uv_1 & uv_2 & uv_3 \end{vmatrix}$$

$$= \left[\frac{\partial}{\partial y}(uv_3) - \frac{\partial}{\partial z}(uv_2)\right]\mathbf{i} + \left[\frac{\partial}{\partial z}(uv_1) - \frac{\partial}{\partial x}(uv_3)\right]\mathbf{j} + \left[\frac{\partial}{\partial x}(uv_2) - \frac{\partial}{\partial y}(uv_1)\right]\mathbf{k}$$

$$= \left[u\frac{\partial v_3}{\partial y} + v_3\frac{\partial u}{\partial y} - u\frac{\partial v_3}{\partial z} - v_2\frac{\partial u}{\partial z}\right]\mathbf{i} + \left[u\frac{\partial v_1}{\partial z} + v_1\frac{\partial u}{\partial z} - u\frac{\partial v_3}{\partial x} - v_3\frac{\partial u}{\partial x}\right]\mathbf{j}$$

$$+ \left[u\frac{\partial v_2}{\partial x} + v_2\frac{\partial u}{\partial x} - u\frac{\partial v_1}{\partial y} - v_1\frac{\partial u}{\partial y} \right]\mathbf{k}$$

NOTES

$$= u\left[\left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial v_1}{\partial z} - \frac{\partial y_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\mathbf{k}\right]$$

$$+ \left[\left(\frac{\partial u}{\partial y}v_3 - \frac{\partial u}{\partial z}v_2\right)\mathbf{i} + \left(\frac{\partial u}{\partial z}v_1 - \frac{\partial u}{\partial x}v_3\right)\mathbf{j} + \left(\frac{\partial u}{\partial x}v_2 - \frac{\partial u}{\partial y}v_1\right)\mathbf{k}\right]$$

$$= u\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix} + \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix} = u \text{ curl } \mathbf{v} + (\text{grad } u) \times \mathbf{v}$$

i.e., $\nabla \times (uv) = u\,\nabla \times \mathbf{v} + (\nabla u) \times \mathbf{v}$.

Again $\text{curl}(\mathbf{u} \times \mathbf{v}) = \nabla \times (\mathbf{u} \times \mathbf{v})$

$$= \nabla \times \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

$$= \nabla \times [(v_3u_2 - v_2u_3)\,\mathbf{i} + (v_1u_3 - v_3u_1)\,\mathbf{j} + (v_2u_1 - v_1u_2)\,\mathbf{k}]$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (v_3u_2 - v_2u_3) & (v_1u_3 - v_3u_1) & (v_2u_1 - v_1u_2) \end{vmatrix}$$

$$= \mathbf{i}\left[\frac{\partial}{\partial y}(v_2u_1 - v_1u_2) - \frac{\partial}{\partial z}(v_1u_3 - u_1v_3)\right] + \mathbf{j}\left[\frac{\partial}{\partial z}(v_3u_2 - v_2u_3) - \frac{\partial}{\partial x}(v_2u_1 - v_1u_2)\right]$$

$$+ \mathbf{k}\left[\frac{\partial}{\partial x}(v_1u_3 - v_3u_1) - \frac{\partial}{\partial y}(v_3u_2 - v_2u_3)\right]$$

$$= \mathbf{i}\left[u_1\frac{\partial v_2}{\partial y} + v_2\frac{\partial u_1}{\partial y} - u_2\frac{\partial v_1}{\partial y} - v_1\frac{\partial u_2}{\partial y} - u_3\frac{\partial v_1}{\partial z} - v_1\frac{\partial u_3}{\partial z} + u_1\frac{\partial v_3}{\partial z} + v_3\frac{\partial u_1}{\partial z}\right]$$

$$+ \mathbf{j}\left[u_2\frac{\partial v_3}{\partial z} + v_3\frac{\partial u_2}{\partial z} - u_3\frac{\partial v_2}{\partial z} - v_2\frac{\partial u_3}{\partial z} - u_1\frac{\partial v_2}{\partial x} - v_2\frac{\partial u_1}{\partial x} + v_1\frac{\partial v_2}{\partial x} + u_2\frac{\partial v_1}{\partial x}\right]$$

$$+ \mathbf{k}\left[u_3\frac{\partial v_1}{\partial x} + v_1\frac{\partial u_3}{\partial x} - v_3\frac{\partial u_1}{\partial x} - u_1\frac{\partial v_3}{\partial x} - u_2\frac{\partial v_3}{\partial y} - v_3\frac{\partial u_2}{\partial y} + v_2\frac{\partial u_3}{\partial y} + u_3\frac{\partial v_2}{\partial y}\right]$$

$$= (u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k})\left[\left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)\cdot(v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}\right]$$

$$- (v_1\mathbf{i} + v_2\mathbf{j} + v_3\ )\left[\left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)\cdot(u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}\right]$$

NOTES

$$+\left[(v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k})\cdot\left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}+\right)\right](u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k})$$

$$-\left[(u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k})\cdot\left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}+\right)\right](v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k})$$

$$= \mathbf{u}\ \text{div}\ \mathbf{v} - \mathbf{v}\ \text{div}\ \mathbf{u} + (\mathbf{v}\cdot\nabla)\ \mathbf{u} - (\mathbf{u}\cdot\nabla)\ \mathbf{v}.$$

Aliter $$\text{curl}\ \mathbf{f} = \mathbf{i}\times\frac{\partial\mathbf{f}}{\partial x}+\mathbf{j}\times\frac{\partial\mathbf{f}}{\partial y}+\mathbf{k}\times\frac{\partial\mathbf{f}}{\partial z}=\Sigma\,\mathbf{i}\times\frac{\partial\mathbf{f}}{\partial x}.$$

Now $$\text{curl}\ (\mathbf{u}\times\mathbf{v}) = \Sigma\,\mathbf{i}\times\frac{\partial[\mathbf{u}\times\mathbf{v}]}{\partial x}$$

$$=\Sigma\,\mathbf{i}\times\left[\mathbf{u}\times\frac{\partial\mathbf{v}}{\partial x}+\mathbf{v}\times\frac{\partial\mathbf{u}}{\partial x}\right]$$

$$=\Sigma\,\mathbf{i}\times\left(\mathbf{u}\times\frac{\partial\mathbf{v}}{\partial x}\right)+\Sigma\,\mathbf{i}\times\left(\mathbf{v}\times\frac{\partial\mathbf{u}}{\partial x}\right)$$

$$=\Sigma\left(\mathbf{i}\cdot\frac{\partial\mathbf{v}}{\partial x}\right)\mathbf{u}-\Sigma(\mathbf{i}\cdot\mathbf{u})\frac{\partial\mathbf{v}}{\partial x}+\Sigma(\mathbf{i}\cdot\mathbf{v})\frac{\partial\mathbf{u}}{\partial x}-\Sigma\left(\mathbf{i}\cdot\frac{\partial\mathbf{u}}{\partial x}\right)\mathbf{v}$$

[by vector triple product]

$$=\left[\left(\Sigma\,\mathbf{i}\cdot\frac{\partial\mathbf{v}}{\partial x}\right)\right]\mathbf{u}-[\Sigma(\mathbf{i}\cdot\mathbf{u})]\frac{\partial\mathbf{v}}{\partial x}+[\Sigma(\mathbf{i}\cdot\mathbf{v})]\frac{\partial\mathbf{u}}{\partial x}-\left[\Sigma\left(\mathbf{i}\cdot\frac{\partial\mathbf{u}}{\partial x}\right)\right]\mathbf{v}$$

$$= \mathbf{u}\ \text{div}\ \mathbf{v} - \mathbf{v}\ \text{div}\ \mathbf{u} + (\mathbf{v}\cdot\nabla)\ \mathbf{u} - (\mathbf{u}\cdot\nabla)\ \mathbf{v}.$$

To express Gradient of Scalar Product in Terms of Curl

We have to show that

grad $(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\times\text{curl}\ \mathbf{v}+\mathbf{v}\times\text{curl}\ \mathbf{u}+(\mathbf{u}\cdot\nabla)\ \mathbf{v}+(\mathbf{v}\cdot\nabla)\ \mathbf{u}$

We know that

$$\text{grad}\ (\mathbf{u}\cdot\mathbf{v}) = \Sigma\,\mathbf{i}\frac{\partial}{\partial x}\ (\mathbf{u}\cdot\mathbf{v})$$

$$=\Sigma\,\mathbf{i}\left[\mathbf{u}\cdot\frac{\partial\mathbf{v}}{\partial x}+\mathbf{v}\cdot\frac{\partial\mathbf{u}}{\partial x}\right]$$

$$=\Sigma\,\mathbf{i}\left(\mathbf{u}\cdot\frac{\partial\mathbf{v}}{\partial x}\right)+\Sigma\,\mathbf{i}\left(\mathbf{v}\cdot\frac{\partial\mathbf{u}}{\partial x}\right) \quad \text{... (1.13)}$$

And $$\mathbf{u}\times\left(\mathbf{i}\times\frac{\partial\mathbf{v}}{\partial x}\right) = \left(\mathbf{u}\cdot\frac{\partial\mathbf{v}}{\partial x}\right)\mathbf{i}-(\mathbf{u}\cdot\mathbf{i})\frac{\partial\mathbf{v}}{\partial x}$$

or $$\left(\mathbf{u}\cdot\frac{\partial\mathbf{v}}{\partial x}\right)\mathbf{i}=\left[\mathbf{u}\times\left(\mathbf{i}\times\frac{\partial\mathbf{v}}{\partial x}\right)\right]+(\mathbf{u}\cdot\mathbf{i})\frac{\partial\mathbf{v}}{\partial x}$$

$\therefore$ $$\Sigma\left(\mathbf{u}\cdot\frac{\partial\mathbf{v}}{\partial x}\right)\mathbf{i} = \Sigma\left[\mathbf{u}\times\left(\mathbf{i}\times\frac{\partial\mathbf{v}}{\partial x}\right)\right]+\Sigma\,(\mathbf{u}\cdot\mathbf{i})\frac{\partial\mathbf{v}}{\partial x}$$

$$= \mathbf{u} \times \text{curl}\, \mathbf{u} + (\mathbf{u} \cdot \nabla)\, \mathbf{v} \qquad \text{... (1.14)}$$

Similarly $\Sigma \left(\mathbf{v} \cdot \dfrac{\partial \mathbf{v}}{\partial x} \right) \mathbf{i} = \mathbf{v} \times \text{curl}\, \mathbf{u} + (\mathbf{v} \cdot \nabla)\, \mathbf{u}$ ... (1.15)

NOTES

Substituting values of (1.14) and (1.15) in (1.13) we find

$$\text{grad}\, (\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \times \text{curl}\, \mathbf{v} + (\mathbf{u} \cdot \nabla)\, \mathbf{v} + \mathbf{v} \times \text{curl}\, \mathbf{u} + (\mathbf{v} \cdot \nabla)\, \mathbf{u}$$
$$= \mathbf{u} \times \text{curl}\, \mathbf{v} + \mathbf{v} \times \text{curl}\, \mathbf{u} + (\mathbf{u} \cdot \nabla)\, \mathbf{v} + (\mathbf{v} \cdot \nabla)\, \mathbf{u}.$$

To express divergence of vector product in terms of curl

We have to show that div ($\mathbf{u} \times \mathbf{v}$) = curl $\mathbf{u} \cdot \mathbf{v}$ – curl $\mathbf{v} \cdot \mathbf{u}$.

We know that,

$$\text{div}\, (\mathbf{u} \times \mathbf{v}) = \Sigma\, \mathbf{i} \cdot \frac{\partial}{\partial x}\, (\mathbf{u} \times \mathbf{v})$$
$$= \Sigma\, \mathbf{i} \cdot \left[\frac{\partial \mathbf{u}}{\partial x} \times \mathbf{v} + \mathbf{u} \times \frac{\partial \mathbf{v}}{\partial x} \right]$$
$$= \left[\Sigma\, \mathbf{i} \cdot \left(\frac{\partial \mathbf{u}}{\partial x} \times \mathbf{v} \right) + \Sigma\, \mathbf{i} \cdot \left(\mathbf{u} \times \frac{\partial \mathbf{v}}{\partial x} \right) \right]$$
$$= \Sigma\, \mathbf{i} \times \frac{\partial \mathbf{u}}{\partial x} \cdot \mathbf{v} + \Sigma\, \mathbf{i} \cdot \mathbf{u} \times \frac{\partial \mathbf{v}}{\partial x}$$

(interchanging dot and cross)

$$= \left(\Sigma\, \mathbf{i} \times \frac{\partial \mathbf{u}}{\partial x} \right) \cdot \mathbf{v} + \left(\Sigma\, \mathbf{i} \times \frac{\partial \mathbf{v}}{\partial x} \right) \cdot \mathbf{u}$$
$$= \text{curl}\, \mathbf{u} \cdot \mathbf{v} - \text{curl}\, \mathbf{v} \cdot \mathbf{u}.$$

Check Your Progress

3. What is a scalar field?
4. What are level surfaces in a scalar field?

1.5 EIGENVALUE PROBLEM

To discuss some relative eigen value problem, we first summarise the useful results.

A **matrix polynomial** of degree n is an expression such as

$$\mathbf{P}(\lambda) = \mathbf{A}_0 + \mathbf{A}_1\lambda + \mathbf{A}_2\lambda^2 + \ldots\ldots.. + \mathbf{A}_m\lambda^m;\ \mathbf{A}_m \neq 0 \qquad \text{...(1.16)}$$

where all the square matrices $\mathbf{A}_0$, $\mathbf{A}_1$, $\mathbf{A}_2$,....., $\mathbf{A}_m$ are of the same order, say, n, *i.e.,* the matrix polynomial is n-rowed.

Thus if **A** be an n-rowed square matrix and **I** be an r-rowed unit matrix, then $\mathbf{A}-\lambda\mathbf{I}$ is the matrix polynomial of the first degree, and known as the **characteristic** or **proper** or **latent** or **eigen** or **varient matrix** of the matrix **A** and its determinant $|\mathbf{A} - \lambda\mathbf{I}|$ is known as the **characteristic polynomial** of the matrix **A** whereas $|\mathbf{A} - \lambda\mathbf{I}| = \mathbf{O}$ is the **characteristic equation** of the matrix **A**.

According to **Caylay-Hamilton Theorem** every square matrix satisfies its own characteristic equation, *i.e.*, if

$$|\mathbf{A}-\lambda\mathbf{I}| = a_0 + a_1\lambda + a_2\lambda^2 + ... + a_n\lambda^n = 0 \qquad ...(1.17)$$

$$\text{then } a_0\mathbf{I} + a_1\mathbf{A} + a_2\mathbf{A}^2 + ... + a_n\mathbf{A}^n = 0 \qquad ...(1.18)$$

NOTES

In the case when there exists a scalar λ s.t.

$$\mathbf{AX} = \lambda\mathbf{X},\ \mathbf{X} \neq 0 \qquad ...(1.19)$$

then the non-zero vector **X** is termed as the **characteristic vector** or **eigen vector** of the matrix **A** and λ is known as an **eigen-value** or **eigen-root** or **characteristic root** coresponding to the characteristic vector **X**.

Now, **I** being a unit matrix,

$$\mathbf{AX} = \lambda\mathbf{X} = \lambda\mathbf{IX} \Rightarrow (\mathbf{A}-\lambda\mathbf{I})\mathbf{X} = \mathbf{O} \qquad ...(1.20)$$

$$(5) \Rightarrow |\mathbf{A}-\lambda\mathbf{I}| = \text{O as } \mathbf{X} \neq \mathbf{O} \text{ so that } (\mathbf{A}-\lambda\mathbf{I}) \text{ is singular} \qquad ...(1.21)$$

$\Rightarrow$ every eigen root **X** of a matrix **A** is a root of its eigen equation $|\mathbf{A}-\lambda\mathbf{I}| = \mathbf{O}$

Conversely, if λ is a root of the eigen equation $|\mathbf{A}-\lambda\mathbf{I}| = \mathbf{O}$, then the matrix-equation $(\mathbf{A}-\lambda\mathbf{I})\mathbf{X} = \mathbf{O}$ possesses a non-trivial solution for **X** and as such ∃ a vector $\mathbf{X} \neq \mathbf{O}$ s.t.

$$\mathbf{AX} = \lambda\mathbf{IX} = \lambda\mathbf{X} \qquad ...(1.22)$$

implying that every root of the eigen equation of a matrix A is an eigen root or value of the matrix.

Thus if $\mathbf{A} = [a_{ij}]$ be an n-rowed square matrix and λ is an indeterminate, then

$$|\mathbf{A}-\lambda\mathbf{I}| = \mathbf{O} \Rightarrow \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13}\cdots a_{1n} \\ a_{21} & a_{22}-\lambda & a_{23}\cdots a_{2n} \\ a_{31} & a_{32} & a_{33}-\lambda\cdots a_{3n} \\ \hline a_{n1} & a_{n2} & a_{n3}\cdots a_{nn}-\lambda \end{vmatrix} = \mathbf{O} \qquad ...(1.23)$$

$\Rightarrow$ ∃ n values, *i.e., n* eigen values or eigen roots of this equation

Also ∃ an eigen vector corresponding to every eigen root (value), as

$$[\mathbf{A}-\lambda\mathbf{I}]\mathbf{X} = \mathbf{O} = \begin{bmatrix} a_{11-\lambda} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22-\lambda} & \cdots & a_{2n} \\ \hline a_{n1} & a_{n2} & & a_{nn-\lambda} \end{bmatrix} \mathbf{X} = \mathbf{O} \qquad ...(1.24)$$

The sum of eigen roots = the sum of diagonal elements

$$i.e. \sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n}a_{ii} \text{ and } \prod_{i=1}^{n}\lambda_i = |\mathbf{A}_i| \qquad ...(1.25)$$

[A] The Power Method or Iterative Method for Dominant Eigen Roots

Consider a matrix $\mathbf{A} = [a_{ij}]_{n\times n}$ having n eigen values, say, $\lambda_1, \lambda_2, ... \lambda_n$. s.t. $|\lambda_1| \geq |\lambda_2| \geq ...... \geq |\lambda_n|$ so that λ_1 is the **largest** or **dominant** root and take a vector $\mathbf{Y}_0$ which is a linear combination of eigen vectors $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ s.t.

$$\mathbf{Y}_0 = x_1\mathbf{X}_1 + x_2\mathbf{X}_2 + ... + x_n\mathbf{X}_n \qquad ...(1.26)$$

where $x_1, x_2, ..., x_n$ are scalars.

NOTES

Now, if we operate $\mathbf{A}$ repeatedly on $\mathbf{Y}_0$ s.t.

$$\left.\begin{aligned}\mathbf{Y}_1 = \mathbf{A}\mathbf{Y}_0 &= x_1\mathbf{A}\mathbf{X}_1 + x_2\mathbf{A}\mathbf{X}_2 + ... + x_n\mathbf{A}\mathbf{X}_n \\ &= x_1\lambda_1\mathbf{X}_1 + x_2\lambda_2\mathbf{X}_2 + ... + x_n\lambda_n\mathbf{X}_n\end{aligned}\right\} \qquad ...(1.27)$$

Similarly,

$$\mathbf{Y}_2 = \mathbf{A}\mathbf{Y}_1 = x_1\lambda_1^2\mathbf{X}_1 + x_2\lambda_2^2\mathbf{X}_2 + ... + x_n\lambda_n^2\mathbf{X}_n ...(1.28)$$

$$\mathbf{Y}_p = \mathbf{A}\mathbf{Y}_{p-1} = x_1\lambda_1^p\mathbf{X}_1 + x_2\lambda_2^p\mathbf{X}_2 + ... + x_n\lambda_n^p\mathbf{X}_n$$

$$= \lambda_1^p\left[x_1\mathbf{X}_1 + x_2\left(\frac{\lambda_2}{\lambda_1}\right)^p\mathbf{X}_2 + ... + x_n\left(\frac{\lambda_n}{\lambda_1}\right)^p\mathbf{X}_n\right] \qquad ...(1.29)$$

For sufficiently large value of p, the vector

$$x_1\mathbf{X}_1 + x_2\left(\frac{\lambda_2}{\lambda_1}\right)^p\mathbf{X}_2 + ... + x_n\left(\frac{\lambda_n}{\lambda_1}\right)^n\mathbf{X}_n \; i.e. \; x_1\mathbf{X}_1 + \sum_{i=2}^{n} x_i\left(\frac{\lambda_i}{\lambda_1}\right)^p\mathbf{X}_i$$

converges towards $x_1\mathbf{X}_1$ which is the eigen vector of λ_1, since

$$\operatorname*{Lin}_{p\to\infty}\left(\frac{\lambda_i}{\lambda_1}\right)^p \to 0 \text{ as } \lambda_i < \lambda_1 \; i.e. \; \frac{\lambda_i}{\lambda_1} < 1 \qquad ...(1.30)$$

Conclusively,

$\mathbf{Y}_p = x_1\lambda_i^p\mathbf{X}_1$, a scalar multiple of eigen vector $\mathbf{X}_1$ ...(1.31)

from which it follows that $\mathbf{Y}_p$ is also an eigen vector corresponding to the largest eigen value λ_1.

In practice, if we take

$\mathbf{Y}_1 = \mathbf{A}\mathbf{Y}_0 = k_1\mathbf{Z}_1 \Rightarrow \mathbf{Z}_1 = \frac{1}{k_1}\mathbf{Y}_1$, k_1 being the largest element of $\mathbf{Y}_1$ and the process be repeated upto convenient choice of $\mathbf{Y}_0$ s.t.

$$\mathbf{Y}_{p+1} = \mathbf{A}\mathbf{Z}_p = \mathbf{A}^{p+1}\mathbf{Y}_0 \qquad ...(1.32)$$

If we finally take

$$\mathbf{Y}_{p+1} = k_{p+1}\mathbf{Z}_{p+1}$$

then k_{p+1} is the eigen value and $\mathbf{Z}_{p+1}$ is the eigen vector.

It is **notable** that after evaluating the largest eigen value λ_1, of the matrix A, it is possible to find the smallest eigen value on modifying the matrix in the form $(\mathbf{A} - \lambda, \mathbf{I})$ having the eigen values

$$\lambda'_p = (\lambda_p - \lambda_1), p = 1, 2, ...n \qquad ...(1.33)$$

where λ_p are the eigen values of $\mathbf{A}$. But for λ_p to be smallest, λ_p' must be the largest normed (magnitude) eigen values of $(\mathbf{A} - \lambda_1\mathbf{I})$. In case $\mathbf{X}_p'$ is the corresponding eigen vector, then

$$(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{X}'_p = (\lambda_p - \lambda_1)\mathbf{X}_p' \Rightarrow \mathbf{A}\mathbf{X}_p' = \lambda_p\mathbf{X}_p' \qquad ...(1.34)$$

NOTES

implying that $\mathbf{X}_p'$ is also the eigen vector of **A** corresponding to the smallest eigen value of **A**.

Thus, for evaluation of the eigen values, if we diminish (or increase) the eigen values of **A** by a constant, say, μ by setting $\mathbf{B} = \mathbf{A} - \mu\mathbf{I}$, where λ is an eigen value of **A** with eigen vector **X** s.t. $\mathbf{AX} = \lambda\mathbf{X}$, then we can write

$$\mathbf{BX} = (\mathbf{A} - \mu\mathbf{I})\mathbf{X} = \mathbf{AX} - \mu\mathbf{IX} = \lambda\mathbf{X} - \mu\mathbf{X} = (\lambda - \mu)\mathbf{X} \qquad ...(1.35)$$

which shows that latent roots of **B** are merely the latent roots of **A** diminished by a constant μ with the same eigen vectors.

ILLUSTRATIVE EXAMPLES

Illustration 1.1 *Using the iterative method find the largest eigen value and the corresponding eigen vector of the matrix.*

$$\mathbf{A} = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \qquad [R.U.\ 1988]$$

Take the initial vector $\mathbf{Y}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, so that

$$\mathbf{Y}_1 = \mathbf{AY}_0 = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \Rightarrow k_1 = 1,\ \mathbf{Z}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{Y}_2 = \mathbf{AZ}_1 = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \\ 0 \end{bmatrix} = 7\begin{bmatrix} 1 \\ 0.4 \\ 0 \end{bmatrix} = 7\mathbf{Z}_2 \text{ (say)}$$

$$\mathbf{Y}_3 = \mathbf{AZ}_2 = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0.4 \\ 0 \end{bmatrix} = \begin{bmatrix} 3.4 \\ 1.8 \\ 0 \end{bmatrix} = 3.4\begin{bmatrix} 1 \\ 0.53 \\ 0 \end{bmatrix} = 3.4\mathbf{Z}_3$$

$$\mathbf{Y}_4 = \mathbf{AZ}_3 = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0.53 \\ 0 \end{bmatrix} = \begin{bmatrix} 4.18 \\ 2.06 \\ 0 \end{bmatrix} = 4.18\begin{bmatrix} 1 \\ 0.49 \\ 0 \end{bmatrix} = 4.18\mathbf{Z}_4$$

$$\mathbf{Y}_5 = \mathbf{AZ}_4 = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0.49 \\ 0 \end{bmatrix} = \begin{bmatrix} 3.94 \\ 1.98 \\ 0 \end{bmatrix} = 3.94\begin{bmatrix} 1 \\ 0.5 \\ 0 \end{bmatrix} = 3.94\mathbf{Z}_5$$

$$\mathbf{Y}_6 = \mathbf{AZ}_5 = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0.5 \\ 0 \end{bmatrix} = 4\mathbf{Z}_6$$

Obviously $\mathbf{Z}_5 \approx \mathbf{Z}_6$, therefore the largest eigen root is 4 and the corresponding eigen vector is $\begin{bmatrix} 1 \\ 0.5 \\ 0 \end{bmatrix}$.

NOTES

Illustration 1.2 *Find all the eigen roots of* $\mathbf{A} = \begin{bmatrix} 1 & 6 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ *and point out the smallest eigen value.*

Here $\mathbf{A} = \begin{bmatrix} 1 & 6 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ whose largest eigen root is 4 by illustration 1.1.

Now, consider

$$\mathbf{B} = \mathbf{A} - 4\mathbf{I} = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix} - 4\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -3 & 6 & 1 \\ 1 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

then, starting with $\mathbf{Y}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, we have the following iterations:

$$\mathbf{Y}_1 = \mathbf{B}\mathbf{Y}_0 = \begin{bmatrix} -3 & 6 & 1 \\ 1 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} = -3\begin{bmatrix} 1 \\ -0.33 \\ 0 \end{bmatrix} = -3\mathbf{Z}_1 \text{ (say)}$$

$$\mathbf{Y}_2 = \mathbf{B}\mathbf{Z}_1 = \begin{bmatrix} -3 & 6 & 1 \\ 1 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ -0.33 \\ 0 \end{bmatrix} = \begin{bmatrix} -4.98 \\ 1.66 \\ 0 \end{bmatrix} = -4.98\begin{bmatrix} 1 \\ -0.33 \\ 0 \end{bmatrix} = -4.98\,\mathbf{Z}_2$$

$$\mathbf{Y}_3 = \mathbf{B}\mathbf{Z}_3 = \begin{bmatrix} -3 & 6 & 1 \\ 1 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ -0.33 \\ 0 \end{bmatrix} = \begin{bmatrix} -4.98 \\ 1.66 \\ 0 \end{bmatrix} = -4.98\begin{bmatrix} 1 \\ -0.33 \\ 0 \end{bmatrix}$$

implying that the largest eigen root of **B** is – 4.98 or –5 approximately and the corresponding eigen vector is $\begin{bmatrix} 1 \\ -0.33 \\ 0 \end{bmatrix}$.

Thus the smallest eigen root is $-5 + 4 = -1$

Ultimately using (10), *i.e.* $\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}$, the remaining third root, say, λ_3 is given by

$$4 - 1 + \lambda_3 = 1 + 2 + 3 \Rightarrow \lambda_3 = 3.$$

[B] Jacobi's Method

All the eigen values of a real symmetric matrix are real. To show that a real orthogonal matrix **O** (say) s.t. $\mathbf{O}^{-1}\mathbf{A}\mathbf{O}$ is a diagonal matrix;

$$\mathbf{O}^{-1}\mathbf{A}\mathbf{O} = \mathbf{D} \text{ (a diagonal matrix)}$$

we take $\lambda_1, \lambda_2, \ldots\ldots, \lambda_n$ as n distinct eigen values of **A** and $\mathbf{X}_1, \mathbf{X}_2, \ldots\ldots \mathbf{X}_n$ as the corresponding eigen vectors.

Consider a column vector

$$\mathbf{X}_i = \begin{bmatrix} \mathbf{X}_{1i} \\ \mathbf{X}_{2i} \\ \vdots \\ \mathbf{X}_{ni} \end{bmatrix} \qquad ...(1.36)$$

and set **O** consisting of column vectors as the n eigen vectors s.t.

$$\mathbf{O} = [\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n] = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \cdots \mathbf{X}_{1n} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \cdots \mathbf{X}_{2n} \\ \hline \mathbf{X}_{n1} & \mathbf{X}_{n2} \cdots \mathbf{X}_{nn} \end{bmatrix} = [\mathbf{X}_{ij}] \qquad ...(1.37)$$

$$\text{Assuming} \quad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 \cdots 0 \\ 0 & \lambda_2 \cdots 0 \\ \hline 0 & 0 \cdots \lambda_n \end{bmatrix} = \text{diag}\,[\lambda_1, \lambda_2, ..., \lambda_n] \qquad ...(1.38)$$

we have

$$\mathbf{OD} = \begin{bmatrix} \lambda_1 \mathbf{X}_{11} & \lambda_2 \mathbf{X}_{12} \cdots \lambda_n \mathbf{X}_{1n} \\ \lambda_1 \mathbf{X}_{21} & \lambda_2 \mathbf{X}_{22} \cdots \lambda_n \mathbf{X}_{2n} \\ \hline \lambda_1 \mathbf{X}_{n1} & \lambda_2 \mathbf{X}_{n2} \cdots \lambda_n \mathbf{X}_{nn} \end{bmatrix} = [\lambda_j\, \mathbf{X}_{ij}] \qquad ...(1.39)$$

(no summation over j)

$$\begin{aligned} &= [\mathbf{AX}_1, \mathbf{AX}_2, ..., \mathbf{AX}_n] \text{ on expressing matrix as vectors} \\ &= \mathbf{A}\,[\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n] \text{ as } \mathbf{AX} = \lambda \mathbf{X} \qquad ...(1.40) \\ &= \mathbf{AO} \end{aligned}$$

If **O** is non-singular, then premultiplying (25) by $\mathbf{O}^{-1}$, we have

$$\mathbf{O}^{-1}\mathbf{OD} = \mathbf{O}^{-1}\mathbf{AO} \text{ when } \mathbf{O}^{-1}\mathbf{O} = \mathbf{I},\ \mathbf{O}^{-1}\mathbf{OD} = \mathbf{ID} = \mathbf{D} \text{ so that}$$

$$\mathbf{O}^{-1}\mathbf{AO} = \mathbf{D} \qquad ...(1.41)$$

As such we face the problem of diagonalizing a real symmetric matrix and so try to find an orthogonal matrix expressible as the product of very special orthogonal matrices.

Let us choose the numerically largest off-diagonal element, say, a_{ij}, s.t. $|a_{ij}|$ = Max. not belonging to the diagonal, but following a 2×2 sub matrix $\begin{bmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix}$ with $a_{ij} = a_{ji}$ so that the matrix being symmetrical, can be easily diagonalized.

Take an orthogonal matrix

$$\mathbf{O} = \begin{bmatrix} \cos\phi & -\sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \text{ s.t. } |\mathbf{O}| = \cos^2\phi + \sin^2\phi = 1 \qquad ...(1.42)$$

yielding

$$\mathbf{O}^{-1} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \qquad ...(1.43)$$

$$\text{Now, if} \qquad \mathbf{D} = \begin{bmatrix} d_{ii} & d_{ij} \\ d_{ji} & d_{jj} \end{bmatrix} \qquad ...(1.44)$$

NOTES

then

$$\mathbf{D} = \mathbf{O}^{-1}\mathbf{AO} \Rightarrow \begin{bmatrix} d_{ii} & d_{ij} \\ d_{ji} & d_{jj} \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} d_{ii} & d_{ij} \\ d_{ji} & d_{jj} \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_{ii}\cos\phi + a_{ij}\sin\phi & -a_{ii}\sin\phi + a_{ij}\cos\phi \\ a_{ji}\cos\phi + a_{jj}\sin\phi & -a_{ji}\sin\phi + a_{jj}\cos\phi \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} d_{ii} & d_{ij} \\ d_{ji} & d_{jj} \end{bmatrix}$$

$$= \begin{bmatrix} a_{ii}\cos^2\phi + a_{ij}\sin 2\phi + a_{jj}\sin^2\phi & -\frac{1}{2}\left(a_{ii} - a_{jj}\right)\sin 2\phi + a_{ij}\cos 2\phi \\ -\frac{1}{2}\left(a_{ii} - a_{ji}\right)\sin 2\phi + a_{ji}\cos 2\phi & a_{ii}\sin^2\phi - a_{ji}\sin 2\phi + a_{jj}\cos^2\phi \end{bmatrix} \quad ...(1.45)$$

$$\Rightarrow \left.\begin{aligned} d_{ii} &= a_{ii}\cos^2\phi + a_{ij}\sin 2\phi + a_{ji}\sin^2\phi \\ d_{ji} = d_{ij} &= -\frac{1}{2}\left(a_{ii} - a_{jj}\right)\sin 2\phi + a_{ij}\cos 2\phi \\ d_{jj} &= a_{ii}\sin^2\phi - a_{ij}\sin 2\phi + a_{jj}\cos^2\phi \end{aligned}\right\}$$

For $\mathbf{D} = \begin{bmatrix} d_{ii} & d_{ij} \\ d_{ji} & d_{jj} \end{bmatrix}$ to be a diagonal matrix, we have

$$d_{ij} = d_{ji} = \mathbf{O} \Rightarrow -\frac{1}{2}\left(a_{ii} - a_{jj}\right)\sin 2\phi + a_{ij}\cos 2\phi = 0,$$

$$\Rightarrow \tan 2\phi = \frac{2a_{ij}}{a_{ii} - a_{jj}} \quad ...(1.46)$$

We can find four different values of ϕ from (32), but we choose $-\frac{\pi}{4} \leq \phi \leq \frac{\pi}{4}$ for getting the least possible relations.

Setting $\mathbf{R} = \sqrt{\left(a_{ii} - a_{jj}\right)^2 + 4a_{ij}^2}$ and $\sigma = \begin{cases} 1, \text{ when } a_{ii} \geq a_{jj} \\ -1, \text{ when } a_{ii} < a_{jj} \end{cases}$ equation (1.47) yields

$$\frac{\sin 2\phi}{2a_{ij}} = \frac{\cos 2\phi}{a_{ii} - a_{jj}} = \frac{\sqrt{\sin^2 2\phi + \cos^2 2\phi}}{\sqrt{(a_{ii} - a_{jj})^2 + 4a^2_{ij}}} = \frac{\pm 1}{\mathbf{R}} = \frac{\sigma}{\mathbf{R}} \quad ...(1.48)$$

giving $\sin 2\phi = \frac{2\sigma a_{ij}}{\mathbf{R}}$ and $\cos 2\phi = \frac{\sigma\left(a_{ij} - a_{jj}\right)}{\mathbf{R}}$ ...(1.49)

$$\text{Again (1.47)} \Rightarrow \begin{cases} \text{if } \tan 2\phi > 0, 2\phi \text{ lies in first quadrant, } i.e., 0 \leq 2\phi \leq \frac{\pi}{2} \\ \text{if } \tan 2\phi < 0, 2\phi \text{ lies in fourth quadrant, } i.e., -\frac{\pi}{2} \leq 2\phi \leq 0 \end{cases}$$

and thus writing (1.47) as

$$\phi = \frac{1}{2}\tan^{-1}\frac{2a_{ij}}{a_{ii} - a_{jj}} \text{ for } a_{ii} \neq a_{jj},$$

we conclude that

$$\phi = \begin{cases} \dfrac{\pi}{4} & \text{when } a_{ij} > 0 \\ -\dfrac{\pi}{4} & \text{when } a_{ij} < 0 \end{cases} \quad \text{where } a_{ii} \neq a_{jj}$$

As such, choosing $-\dfrac{\pi}{2} < \tan^{-1}\dfrac{2a_{ij}}{a_{ii} - a_{jj}} < \dfrac{\pi}{2}$, we conclude from (1.46),

$$d_{ii} = \frac{1}{2}\left[a_{ii}(1+\cos 2\phi) + 2a_{ij}\sin 2\phi + a_{jj}(1-\cos 2\phi)\right]$$

$$= \frac{1}{2}\left[a_{ii} + a_{jj} + \frac{\sigma}{R}\cdot R^2\right]$$

$$= \frac{1}{2}\left[a_{ii} + a_{jj} + \sigma R\right] \qquad \text{...(1.50)}$$

Similarly $\quad d_{jj} = \dfrac{1}{2}\left[a_{ii} + a_{jj} - \sigma R\right]$ with $d_{ij} = d_{ji}$ ...(1.51)

Now (1.50) and (1.51) $\Rightarrow d_{ii} + d_{jj} = a_{ii} + a_{jj}$ and $d_{ii}\, d_{jj} = a_{ii}a_{jj} - a_{ij}^2$ ...(1.52)

Hence we determine a diagonal matrix $\begin{bmatrix} d_{ii} & \mathbf{O} \\ \mathbf{O} & d_{jj} \end{bmatrix}$.

Performing a series of such rotations, we shall finally approach a diagonal matrix whose diagonal elements will constitute the eigen values of the matrix **A** under consideration. By a series of such two-dimensional rotations, the annihilation of an element a_{ij} gives the transformation matrix a position to have the above given form **O** in the element (i, i), (i, j), (j, i) and (j, j) and identical with the unit matrix else where, but in each rotation, we have to choose such values i and j that $|a_{ij}|$ = Max. After such p rotations, say,

$$\mathbf{P}_p = \mathbf{O}_1\,\mathbf{O}_2.....\mathbf{O}_p$$

the matrix $\mathbf{B} = \mathbf{P}_p^{-1}\,\mathbf{AP}_p$ for increasing p approaches a diagonal matrix **D** with diagonal elements as the eigen values of **A**. Moreover, the eigen vectors appear as the corresponding columns of $\mathbf{O} = \lim\limits_{p\to\infty}\mathbf{P}_p$, as $\mathbf{O}^{-1}\mathbf{AO} = \mathbf{D} \Rightarrow \mathbf{AO} = \mathbf{OD}$.

Hence if $\mathbf{X}_j$ be the jth column vector of **O** and λ_j is the jth diagonal element of **D**, then

$$\mathbf{AX}_j = \lambda_j\,\mathbf{X}_j \qquad \text{...(1.53)}$$

Note that the diagonalized matrix of a real symmetric matrix is orthogonal, since if $\lambda_1, \lambda_2, ..., \lambda_n$ are eigen roots of a real symmetric matrix **A** and **D** its diagonalized matrix, then

$$\mathbf{D}^{-1}\mathbf{AD} = \text{diag}\,[\lambda_1, \lambda_2, ..., \lambda_n]$$

On taking transpose of either side, we get

$$[\mathbf{D}^{-1}\mathbf{AD}]' = \text{diag}\,[\lambda_1, \lambda_2, ..., \lambda_n]'$$

NOTES

$$\Rightarrow \mathbf{D}'\mathbf{A}'(\mathbf{D}^{-1})' = \text{diag}\,[\lambda_1, \lambda_2, \ldots\ldots, \lambda_n] = \mathbf{D}^{-1}\mathbf{A}\mathbf{D}$$

$\Rightarrow \mathbf{D}' = \mathbf{D}$ showing that $\mathbf{D}$ is an orthogonal matrix.

As an illustration, if we consider a real symmetric matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \text{ (say),}$$

then the off-diagonal elements are either 1 or 0 of which 1 is numerically largest. It is also clear that

$$a_{12} = 1 = a_{21},\ a_{23} = 1 = a_{32} \text{ and } a_{13} = a_{31} = 0.$$

For the first rotation, consider $a_{ij} = a_{23} = 1$, where $i = 2, j = 3$, so that $a_{ii} = a_{22} = 4$ and $a_{jj} = a_{33} = 4$ in the 2×2 submatrix $\begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$.

$$\therefore \qquad \tan 2\phi = \frac{2a_{ij}}{a_{ii} - a_{jj}} = \frac{2 \times 2}{4 - 4} = \infty \Rightarrow 2\phi = \frac{\pi}{2} \Rightarrow \phi = \frac{\pi}{4}.$$

$$\text{Thus,} \quad \mathbf{O}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.7071 & -0.7071 \\ 0 & 0.7071 & 0.7071 \end{bmatrix}$$

$$\text{as} \qquad \sin\frac{\pi}{4} = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}} = 0.7071 \text{ approx.}$$

$$\text{so that} \qquad \mathbf{O}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.7071 & 0.7071 \\ 0 & -0.7071 & 0.7071 \end{bmatrix}$$

giving new value of $\mathbf{A}$, say, $\mathbf{A}_1$, as

$$\mathbf{A}_1 = \mathbf{O}_1^{-1}\mathbf{A}\mathbf{O}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.7071 & 0.7071 \\ 0 & -0.7071 & 0.7071 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.7071 & -0.7071 \\ 0 & 0.7071 & 0.7071 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.7071 & 0.7071 \\ 0 & -0.7071 & 0.7071 \end{bmatrix} \begin{bmatrix} 2 & 0.7071 & -0.7071 \\ 1 & 3.5355 & -2.1213 \\ 0 & 3.5355 & 2.1213 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0.7071 & -0.7071 \\ 0.707 & 4.9999 & 0 \\ -0.7071 & 0 & 2.9999 \end{bmatrix}$$

For the second rotation, consider $a_{ij} = a_{12} = .7071$, whence $a_{ii} = a_{11}$, $a_{jj} = a_{22}$

$$\therefore \qquad \tan 2\phi = \frac{2a_{ij}}{a_{ii} - a_{jj}} = \frac{2 \times 7.701}{2 - 4.9999}$$

NOTES

$\Rightarrow \qquad \phi = -27°22'$ by trigonometric tables.

$\Rightarrow \qquad \cos\phi = 0.8881$ and $\sin\phi = -0.4596$

so that
$$\mathbf{O}_2 = \begin{bmatrix} \cos\phi & -\sin\phi & 1 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.8881 & 0.4596 & 0 \\ -0.4596 & 0.8881 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus
$$\mathbf{O}_2^{-1} = \begin{bmatrix} 0.8881 & 0.4596 & 0 \\ -0.4596 & 0.8881 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

giving new value of **A**, say, $\mathbf{A}_2$ as

$$\mathbf{A}_2 = \mathbf{O}_2^{-1}\mathbf{A}_1\mathbf{O}_2 = \begin{bmatrix} 0.8881 & 0.4596 & 0 \\ -0.4596 & 0.8881 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0.7071 & -0.7071 \\ 0.7071 & 4.9999 & 0 \\ -0.7071 & 0 & 2.9999 \end{bmatrix}\begin{bmatrix} 0.8881 & 0.4596 & 0 \\ -0.4596 & 0.8881 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.8881 & -0.4596 & 0 \\ 0.4596 & 0.8881 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1.4512 & 1.5471 & -0.7071 \\ -1.6699 & -4.7653 & 0 \\ -6.279 & .3249 & 2.9999 \end{bmatrix}$$

$$= \begin{bmatrix} 2.0562 & -0.8186 & -0.6279 \\ -0.8160 & 4.9431 & -0.3249 \\ -0.6279 & -0.3249 & 0 \end{bmatrix}$$

Continuing this process, after 12 rotations, we find

$$\mathbf{A}_{12} = \begin{bmatrix} 1.52 & 0 & 0 \\ 0 & 5.17 & 0 \\ 0 & 0 & 3.31 \end{bmatrix}$$

giving eigen values of **A** as 1.52, 5.17, 3.31 with sum 10 and product 26.0113 in good agreement with the exact eigen-equation

$$\lambda^3 - 10\lambda^2 + 30\lambda - 26 = 0$$

[C] Complex Eigen-Values

If the eigen values of a matrix **A** are complex number, then their corresponding eigen vectors will also be complex quantities. Consider a square matrix **A** of order n and **X** a vector comprising complex elements. Separating into real and imaginary parts the equation $(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{O}$ will render $2n$ equations in $2n$ unknowns.

Taking the eigen root $\lambda = \alpha + i\beta$ and its corresponding eigen vector as

$\mathbf{X} = \begin{bmatrix} x_1 + iy_1 \\ x_2 + iy_2 \\ \cdots \\ x_n + iy_n \end{bmatrix}$, the equation $(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{O}$ leads

NOTES

$$\left[\begin{array}{cccc} a_{11}-\alpha & a_{12}\cdots a_{1n} & \beta & 0\cdots 0 \\ a_{21} & a_{22}-\alpha\cdots a_{2n} & 0 & \beta\cdots 0 \\ \hline a_{n1} & a_{n2}\cdots a_{nn}-\alpha & 0 & 0\cdots\beta \\ -\beta & 0\cdots 0 & a_{11}-\alpha & a_{12}\cdots a_{1n} \\ 0 & -\beta\cdots 0 & a_{21} & a_{22}-\alpha.......a_{2n} \\ \hline 0 \quad 0 & -\beta & a_{n1} & a_{n2}.......a_{nn}-\alpha \end{array}\right] = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad ...(1.54)$$

which being equivalent to $2n$ equations in $2n$ unknowns can give on solving $x_i's$ and $y_i's$ and so **X** corresponding to $\lambda = \alpha + i\beta$. But complex roots always occur in pairs, therefore $a - i\beta$ will also be its eigen values corresponding to an eigen vector conjugate to $\lambda = \alpha + i\beta$. We thus require only one eigen vector for each pair of complex eigen values and the other will be found by conjugation.

For example, if $\mathbf{A} = \begin{bmatrix} -1 & -5 \\ 1 & 3 \end{bmatrix}$, then its eigen values are given by

$$\begin{bmatrix} -1-\lambda & -5 \\ 1 & 3-\lambda \end{bmatrix} = 0 \Rightarrow -(1+\lambda)(3+\lambda) + 5 = 0 \Rightarrow \lambda^2 - 2\lambda + 2 = 0$$

$$\Rightarrow \quad \lambda = \frac{2 \pm \sqrt{4-8}}{2} = 1 \pm i, \text{ say } \alpha = 1, \beta = 1$$

If **X** be the eigen vector corresponding to $\lambda = 1 + i$ s.t. $\mathbf{X} = \begin{bmatrix} x_1 + iy_1 \\ x_2 + iy_2 \end{bmatrix}$,

then by (1.54), $[\mathbf{A} - \lambda\mathbf{I}]\mathbf{X} = \mathbf{O} \Rightarrow \begin{bmatrix} -1-1 & -5 & 1 & 0 \\ 1 & 3-1 & 0 & 1 \\ -1 & 0 & -1-1 & -5 \\ 0 & -1 & 1 & 3-1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$$\Rightarrow \quad \begin{bmatrix} -2 & -5 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ -1 & 0 & -2 & -5 \\ 0 & -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad \begin{bmatrix} 1 & 2 & 0 & 1 \\ -2 & -5 & 1 & 0 \\ -1 & 0 & -2 & -5 \\ 0 & -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ by } R_{12}$$

$$\Rightarrow \quad \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 2 & -2 & -4 \\ 0 & -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ by } R_{21(2)}, R_{31(1)} \text{ and then } R_{2(-1)}$$

$$\Rightarrow \quad \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ by } R_{33(-2)}, R_{42(1)}$$

$$\Rightarrow \quad x_1 + 2x_2 + y_2 = 0 \text{ and } x_2 - y_1 - 2y_2 = 0$$

There being only two equations in four unknowns, the number of independent solutions = $n - r = 4 - 2 = 2$

Choosing $x_1 = 0, y_1 = 1$, we get $x_3 = 0.2, y_2 = -0.4$, so that

$$\mathbf{X} = \begin{bmatrix} i \\ 0\cdot 2 - 0\cdot 4i \end{bmatrix} \text{ for } \lambda = 1 + i$$

whereas for $\lambda = 1 - i$, the corresponding eigen vector is $\begin{bmatrix} i \\ 0.2 + 0.4i \end{bmatrix}$ on replacing i by $-i$ in the value of **X**.

[D] The General Eigen Value Problem

Consider two symmetric matrices **A** and **B** s.t. **B** is positive definite. Then the eigen value problem is

$$|\mathbf{A} - \lambda \mathbf{B}| = \mathbf{O}$$

Since **B** can be splitted up as **B** = **LL′**, **L** having lower triangular matrix, therefore

$$\mathbf{A} - \lambda\mathbf{B} = \mathbf{A} - \lambda\mathbf{LL}'$$
$$= \mathbf{L}\left[\mathbf{L}^{-1}\mathbf{A}(\mathbf{L}')^{-1} - \lambda\mathbf{I}\right]\mathbf{L}'$$

Setting $\mathbf{P} = \mathbf{L}^{-1}\mathbf{A}(\mathbf{L}')^{-1}$ so that $\mathbf{P}' = \mathbf{P}$, we have

$$|\mathbf{A} - \lambda\mathbf{B}| = |\mathbf{L}|^2 \cdot (\mathbf{P} - \lambda\mathbf{I})$$

which can be used to determine eigen values.

1.6 CAYLEY–HAMILTON THEOREM

Every square matrix satisfies its own characteristic equation.

The characteristic equation of a square matrix A is $|A - \lambda I| = 0$. This can be written as

$$p_0\lambda^n + p_1\lambda^{n-1} + p_2\lambda^{n-2} + ... + p_n = 0 \quad ...(1.55)$$

We have to prove that,

$$p_0A^n + p_1A^{n-1} + p_2A^{n-2} + ... + p_nI = 0 \quad ...(1.56)$$

Consider the matrix $B = Adj(A - \lambda I)$. The elements of B are polynomials in λ of degree $(n - 1)$ or less. Therefore, B can be written in the form,

$$B = B_0\lambda^{n-1} + B_1\lambda^{n-2} + ... + B_{n-2}\lambda + B_{n-1} \quad ...(1.57)$$

Where, $B_0, B_1, B_{n-2}, ..., B_{n-1}$ are matrices of order n and whose elements are polynomials of the elements of A. It is known that for any matrix A.

$$A\ Adj\ A = |A|I$$

Using this property for the matrix $(A - \lambda I)$, we get,

$$(A - \lambda I)\ Adj\ (A - \lambda I) = |A - \lambda I|I$$
$$(A - \lambda I)\ B = |A - \lambda I|I$$

NOTES

Using Equations (3.56) and (1.57) we get,

$$(A - \lambda I)\,(B_0\lambda^{n-1} + B_1\lambda^{n-2} + ... + B_{n-2}\lambda + B_{n-1}) = (p_0\lambda^n + p_1\lambda^{n-1} + ... + p_n)I$$

On equating the coefficients of λ^n, λ^{n-1} we get,

$$-B_0 = p_0 I$$
$$AB_0 - B_1 = p_1 I$$
$$AB_1 - B_2 = p_2 I$$
$$\vdots$$
$$AB_{n-1} = p_n I$$

Premultiplying these equations by $A^n, A^{n-1}......, A, I$ we get,

$$-A^n B_0 = p_0 A^n$$
$$A^n B_0 - A^{n-1} B_1 = p_1 A^{n-1}$$
$$A^{n-1} B_1 - A^{n-2} B_2 = p_2 A^{n-2}$$
$$\vdots$$
$$A^2 B_{n-2} - AB_{n-1} = p_{n-1} I$$
$$AB_{n-1} = p_n I$$

Adding these equations we get,

$0 = p_0 A^n + p_1 A^{n-1} + ... + p_n I$, which proves the theorem.

Example 1.1: Verify Cayley-Hamilton theorem for the matrix $\begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}$ and hence find A^{-1} and A^4.

Solution: Let, $A = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}$

$$\beta_1 = 6,\ \beta_2 = 11,\ \beta_3 = |A| = 6$$

Characteristic equation is $\lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0$

We have to prove that $A^3 - 6A^2 + 11A - 6I = 0$ ....(1.58)

$$A^2 = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 & -4 \\ 0 & 4 & 0 \\ -4 & 0 & 5 \end{bmatrix} \text{ and}$$

NOTES

$$A^3 = \begin{bmatrix} 14 & 0 & -13 \\ 0 & 8 & 0 \\ -13 & 0 & 14 \end{bmatrix}$$

Using them in Equation (1.58), we find that the equation is satisfied.

$\therefore \quad A^3 - 6A^2 + 11A - 6I = 0$

Premultiplying each term by A^{-1}, we get $A^2 - 6A + 11I - 6A^{-1} = 0$

$$\therefore \quad A^{-1} = \frac{1}{6}[A^2 - 6A + 11I] = \frac{1}{6}\begin{bmatrix} 4 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix}$$

Premultiplying each term of Equation (1.58) by A, we get,

$$A^4 = 6A^3 - 11A^2 + 6A = \begin{bmatrix} 41 & 0 & -40 \\ 0 & 16 & 0 \\ -40 & 0 & 41 \end{bmatrix}$$

1.7 FUNCTION OF MATRIX

Let F be a an array and n, m be two integers ≥ 1. An array of elements F of the type

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

is called a matrix F. We denote this matrix by (a_{ij}), $i = 1, ..., m$ and $j = 1, ..., n$. We say that it is a $m \times n$ matrix (or matrix of order $m \times n$). It has m rows and n columns. For example, the first row is $(a_{11}\ a_{12}\ ..., a_{1n})$ and first column is,

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

Also, a_{ij} denotes the element of the matrix (a_{ij}) lying in ith row and jth column and we call this element as the (i, j)th element of the matrix.

For example, in the matrix,

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

$a_{11} = 1$, $a_{12} = 2$, $a_{32} = 8$, i.e., (1, 1)th element is 1
(1, 2)th element is 2
(3, 2)th element is 8

NOTES

Notes:

1. Unless otherwise stated, we shall consider matrices over the field C of complex numbers only.
2. A matrix is simply an arrangement of elements and has no numerical value.

Example 1.2: If $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 0 & 1 & 2 \end{pmatrix}$, find $a_{11}, a_{22}, a_{33}, a_{31}, a_{41}$.

Solution: a_{11} = Element of A in first row and first column = 1

a_{22} = Element of A in second row and second column = 5

a_{33} = Element of A in third row and third column = 9

a_{31} = Element of A in third row and first column = 7

a_{41} = Element of A in fourth row and first column = 0

Types of Matrices

1. ***Row Matrix:*** A matrix which has exactly one row is called a *row matrix*.
   For example, (1 2 3 4) is a row matrix.
2. ***Column Matrix:*** A matrix which has exactly one column is called a *column matrix*.

   For example, $\begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}$ is a column matrix.
3. ***Square Matrix:*** A matrix in which the number of rows is equal to the number of columns is called a *square matrix*.

   For example, $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ is a 2×2 square matrix.
4. ***Null or Zero Matrix:*** A matrix each of whose elements is zero is called a *null matrix* or *zero matrix*.

   For example, $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is a 2×3 Null matrix.
5. ***Diagonal Matrix:*** The elements a_{ij} are called diagonal elements of a *square matrix* (a_{ij}). For example, $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ in matrix,

   the diagonal elements are $a_{11} = 1$, $a_{22} = 5$, $a_{33} = 9$

   A square matrix whose every element other than diagonal elements is zero, is called a *diagonal matrix*. For example,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \text{ is a diagonal matrix.}$$

NOTES

Note that, the diagonal elements in a diagonal matrix may also be zero. For example,

$$\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ are also diagonal matrices.}$$

6. ***Scalar Matrix:*** A diagonal matrix whose diagonal elements are equal, is called a *scalar matrix*. For example,

$$\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ are scalar matrices.}$$

7. ***Identity Matrix:*** A diagonal matrix whose diagonal elements are all equal to 1 (unity) is called *identity matrix or (unit matrix)*. For example,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ is an identity matrix.}$$

8. ***Triangular Matrix:*** A square matrix (a_{ij}), whose elements $a_{ij} = 0$ when $i < j$ is called a *lower triangular matrix.*

   Similarly, a square matrix (a_{ij}) whose elements $a_{ij} = 0$ whenever $i > j$ is called an *upper triangular matrix.*

   For example,

$$\begin{pmatrix} 1 & 0 & 0 \\ 4 & 5 & 0 \\ 7 & 8 & 9 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \text{ are lower triangular matrices}$$

$$\text{and } \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \text{ are upper triangular matrices.}$$

Algebra of Matrices

Equality

Two matrices A and B are said to be equal if,

(*i*) A and B are of same order.

(*ii*) Corresponding elements in A and B are same. For example, the following two matrices are equal.

$$\begin{pmatrix} 3 & 4 & 9 \\ 16 & 25 & 64 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 9 \\ 16 & 25 & 64 \end{pmatrix}$$

NOTES

But the following two matrices are not equal.

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

As matrix on left is of order 2 × 3, while on right it is of order 3 × 3

The following two matrices are also not equal.

$$\begin{pmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 8 & 9 \end{pmatrix}$$

As (2, 1)th element in LHS matrix is 7 while in RHS matrix it is 4

Addition of Matrices

If A and B are two matrices of the same order then addition of A and B is defined to be the matrix obtained by adding the corresponding elements of A and B.

For example, if

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, B = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{pmatrix}$$

$$\text{Then, } A + B = \begin{pmatrix} 1+2 & 2+3 & 3+4 \\ 4+5 & 5+6 & 6+7 \end{pmatrix} = \begin{pmatrix} 3 & 5 & 7 \\ 9 & 11 & 13 \end{pmatrix}$$

$$\text{Also,} \quad A - B = \begin{pmatrix} 1-2 & 2-3 & 3-4 \\ 4-5 & 5-6 & 6-7 \end{pmatrix} = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

Note that addition (or subtraction) of two matrices is defined only when A and B are of the same order.

Properties of Matrix Addition

(*i*) Matrix addition is commutative.

i.e., $A + B = B + A$

For, (i, j)th element of $A + B$ is $(a_{ij} + b_{ij})$ and of $B + A$ is $(b_{ij} + a_{ij})$, and they are same as a_{ij} and b_{ij} are real numbers.

(*ii*) Matrix addition is associative,

i.e., $A + (B + C) = (A + B) + C$

For, (i, j)th element of $A + (B + C)$ is $a_{ij} + (b_{ij} + c_{ij})$ and of $(A + B) + C$ is $(a_{ij} + b_{ij}) + c_{ij}$ which are same.

(*iii*) If O denotes null matrix of the same order as that of A then,

$A + O = A = O + A$

For (i, j)th element of $A + O$ is $a_{ij} + O$ which is same as (i, j)th element of A.

(*iv*) To each matrix A there corresponds a matrix B such that,

$A + B = O = B + A$.

For, let (i, j)th element of B be $-a_{ij}$. Then (i, j)th element of $A + B$ is, $a_{ij} - a_{ij} = 0$.

Thus, the set of $m \times n$ matrices forms an abelian group under the composition of matrix addition.

NOTES

Example 1.3: If $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix}$

Verify $A + B = B + A$.

Solution: $A + B = \begin{pmatrix} 1+0 & 2+1 & 3+2 \\ 4+3 & 5+4 & 6+5 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \end{pmatrix}$

$$B + A = \begin{pmatrix} 0+1 & 1+2 & 2+3 \\ 3+4 & 4+5 & 5+6 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \end{pmatrix}$$

So, $A + B = B + A$

Example 1.4: If A and B are matrices as in Example 1.3.

and $C = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix}$, verify $(A + B) + C = A + (B + C)$.

Solution: Now $A + B = \begin{pmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \end{pmatrix}$

So, $(A + B) + C = \begin{pmatrix} 1-1 & 3+0 & 5+1 \\ 7+1 & 9+2 & 11+3 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 6 \\ 8 & 11 & 14 \end{pmatrix}$

Again, $B + C = \begin{pmatrix} 0-1 & 1+0 & 2+1 \\ 3+1 & 4+2 & 5+3 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 3 \\ 4 & 6 & 8 \end{pmatrix}$

So, $A + (B + C) = \begin{pmatrix} 1-1 & 2+1 & 3+3 \\ 4+4 & 5+6 & 6+8 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 6 \\ 8 & 11 & 14 \end{pmatrix}$

Therefore, $(A + B) + C = A + (B + C)$

Example 1.5: If $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, find a matrix B such that $A + B = 0$.

Solution: Let, $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$

Then, $A + B = \begin{pmatrix} 1+b_{11} & 2+b_{12} \\ 3+b_{21} & 4+b_{22} \\ 5+b_{31} & 6+b_{32} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$

It implies, $b_{11} = -1$, $b_{12} = -2$, $b_{21} = -3$, $b_{22} = -4$,

$b_{31} = -5$, $b_{32} = -6$

Therefore required $B = \begin{pmatrix} -1 & -2 \\ -3 & -4 \\ -5 & -6 \end{pmatrix}$

NOTES

Multiplication of Matrices

The product AB of two matrices A and B is defined only when the number of columns of A is same as the number of rows in B and by definition the product AB is a matrix G of order $m \times p$ if A and B were of order $m \times n$ and $n \times p$, respectively. The following example will give the rule to multiply two matrices:

Let, $$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \quad B = \begin{pmatrix} d_1 & e_1 \\ d_2 & e_2 \\ d_3 & e_3 \end{pmatrix}$$

Order of $A = 2 \times 3$, Order of $B = 3 \times 2$

So, AB is defined as,

$$G = AB = \begin{pmatrix} a_1d_1 + b_1d_2 + c_1d_3 & a_1e_1 + b_1e_2 + c_1e_3 \\ a_2d_1 + b_2d_2 + c_2d_3 & a_2e_1 + b_2e_2 + c_2e_3 \end{pmatrix}$$

$$= \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$$

g_{11} : Multiply elements of the first row of A with corresponding elements of the first column of B and add.

g_{12} : Multiply elements of the first row of A with corresponding elements of the second column of B and add.

g_{21} : Multiply elements of the second row of A with corresponding elements of the first column of B and add.

g_{22} : Multiply elements of the second column of A with corresponding elements of the second column and add.

Notes: 1. In general, if A and B are two matrices then AB may not be equal to BA. For example, if

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{then } AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and $BA = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. So, $AB \neq BA$

2. If product AB is defined, then it is not necessary that BA must also be defined. For example, if A is of order 2×3 and B is of order 3×1, then AB can be defined but BA cannot be defined (as the number of columns of $B \neq$ the number of rows of A).

It can be easily verified that,

(*i*) $A(BC) = (AB)C.$

(*ii*) $A(B + C) = AB + AC.$

(*iii*) $(A + B)C = AC + BC.$

Example 1.6: If $A = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 7 & 0 \\ -2 & -3 \end{pmatrix}$ write down AB.

Solution: $$AB = \begin{pmatrix} 2 \times 7 + (-1) \times (-2) & 2 \times 0 + (-1) \times (-3) \\ 0 \times 7 + 3 \times (-2) & 0 \times 0 + 3 \times (-3) \end{pmatrix}$$

$$= \begin{pmatrix} 16 & 3 \\ -6 & -9 \end{pmatrix}$$

Example 1.7: Verify the associative law $A(BC) = (AB)C$ for the following matrices.

$$A = \begin{pmatrix} -1 & 0 \\ 7 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 5 \\ 7 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & -1 \\ 2 & 0 \end{pmatrix}$$

NOTES

Solution:

$$AB = \begin{pmatrix} -1 & 0 \\ 7 & -2 \end{pmatrix}\begin{pmatrix} -1 & 5 \\ 7 & 0 \end{pmatrix} = \begin{pmatrix} 1+0 & -5+0 \\ -7-14 & 35+0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -5 \\ -21 & 35 \end{pmatrix}$$

$$BC = \begin{pmatrix} -1 & 5 \\ 7 & 0 \end{pmatrix}\begin{pmatrix} -1 & -1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 1+10 & 1+0 \\ -7+0 & -7+0 \end{pmatrix} = \begin{pmatrix} 11 & 1 \\ -7 & -7 \end{pmatrix}$$

$$A(BC) = \begin{pmatrix} -1 & 0 \\ 7 & -2 \end{pmatrix}\begin{pmatrix} 11 & 1 \\ -7 & -7 \end{pmatrix} = \begin{pmatrix} -11+0 & -1+0 \\ 77+14 & 7+14 \end{pmatrix}$$

$$= \begin{pmatrix} -11 & -1 \\ 91 & 21 \end{pmatrix}$$

$$(AB)C = \begin{pmatrix} 1 & -5 \\ -21 & 35 \end{pmatrix}\begin{pmatrix} -1 & -1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} -1-10 & -1+0 \\ 21+70 & 21+0 \end{pmatrix}$$

$$= \begin{pmatrix} -11 & -1 \\ 91 & 21 \end{pmatrix}$$

Thus, $A(BC) = (AB)C$

Example 1.8: If A is a square matrix, then A can be multiplied by itself. Define $A^2 = A.\ A$ which is called power of a matrix. Compute A^2 for the following matrix:

$$A = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}$$

Solution:

$$A^2 = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 15 & 16 \end{pmatrix}$$

Similarly, we can define $A^3, A^4, A^5, \ldots$ for any square matrix A.

Scalar Multiplication of Matrix

If k is any complex number and A, a given matrix, then kA is the matrix obtained from A by multiplying each element of A by k. The number k is called *Scalar*.

For example, if

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \text{ and } k = 2$$

Then,

$$kA = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}$$

It can be easily shown that,

(*i*) $k(A + B) = kA + kB$ (*ii*) $(k_1 + k_2)A = k_1A + k_2A$

(*iii*) $1A = A$ (*iv*) $(k_1k_2)A = k_1(k_2A)$

NOTES

Example 1.9: (*i*) If $A = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 3 & 4 \\ 4 & 5 & 6 \end{pmatrix}$ and $k_1 = i$, $k_2 = 2$, verify,

$$(k_1 + k_2)\, A = k_1 A + k_2 A$$

(*ii*) If $A = \begin{pmatrix} 0 & 2 & 3 \\ 2 & 1 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 7 & 6 & 3 \\ 1 & 4 & 5 \end{pmatrix}$, find the value of $2A + 3B$.

Solution: (*i*) Now $k_1 A = \begin{pmatrix} 0 & i & 2i \\ 2i & 3i & 4i \\ 4i & 5i & 6i \end{pmatrix}$ and $k_2 A = \begin{pmatrix} 0 & 2 & 4 \\ 4 & 6 & 8 \\ 8 & 10 & 12 \end{pmatrix}$

So, $k_1 A + k_2 A = \begin{pmatrix} 0 & 2+i & 4+2i \\ 4+2i & 6+3i & 8+4i \\ 8+4i & 10+5i & 12+6i \end{pmatrix}$

Also, $(k_1 + k_2)\, A = \begin{pmatrix} 0 & 2+i & 4+2i \\ 4+2i & 6+3i & 8+4i \\ 8+4i & 10+5i & 12+6i \end{pmatrix}$

Therefore, $(k_1 + k_2)A = k_1 A + k_2 A$

(*ii*) $2A = \begin{pmatrix} 0 & 4 & 6 \\ 4 & 2 & 8 \end{pmatrix}$

$3B = \begin{pmatrix} 21 & 18 & 9 \\ 3 & 12 & 15 \end{pmatrix}$

So, $2A + 3B = \begin{pmatrix} 21 & 22 & 15 \\ 7 & 14 & 23 \end{pmatrix}$

Example 1.10: If $A = \begin{pmatrix} 1 & 2 \\ -3 & 0 \end{pmatrix}$ find $A^2 + 3A + 5I$ where I is unit matrix of order 2.

Solution: $A^2 = \begin{pmatrix} 1 & 2 \\ -3 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -3 & 0 \end{pmatrix} = \begin{pmatrix} -5 & 2 \\ -3 & -6 \end{pmatrix}$

$3A = \begin{pmatrix} 3 & 6 \\ -9 & 0 \end{pmatrix}$

$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $5I = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$

So, $A^2 + 3A + 5I = \begin{pmatrix} -5 & 2 \\ -3 & -6 \end{pmatrix} + \begin{pmatrix} 3 & 6 \\ -9 & 0 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$

$= \begin{pmatrix} 3 & 8 \\ -12 & -1 \end{pmatrix}$

Example 1.11: If $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ show that, $AB = -BA$ and $A^2 = B^2 = I$.

Solution: Now, $AB = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$

$$BA = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

So, $AB = -BA$

Also, $$A^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

$$B^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

This proves the result.

NOTES

Example 1.12: In an examination of Mathematics, 20 students from college A, 30 students from college B and 40 students from college C appeared. Only 15 students from each college could get through the examination. Out of them 10 students from college A and 5 students from college B and 10 students from college C secured full marks. Write down the above data in matrix form.

Solution: Consider the matrix,

$$\begin{pmatrix} 20 & 30 & 40 \\ 15 & 15 & 15 \\ 10 & 5 & 10 \end{pmatrix}$$

First row represents the number of students in college A, college B, college C respectively.

Second row represents the number of students who got through the examination in three colleges respectively.

Third row represents the number of students who got full marks in the three colleges respectively.

Example 1.13: A publishing house has two branches. In each branch, there are three offices. In each office, there are 3 peons, 4 clerks and 5 typists. In one office of a branch, 6 salesmen are also working. In each office of other branch 2 head-clerks are also working. Using matrix notation find (*i*) the total number of posts of each kind in all the offices taken together in each branch, (*ii*) the total number of posts of each kind in all the offices taken together from both the branches.

Solution: (*i*) Consider the following row matrices,

$$A_1 = (3\ \ 4\ \ 5\ \ 6\ \ 0),\quad A_2 = (3\ \ 4\ \ 5\ \ 0\ \ 0),\quad A_3 = (3\ \ 4\ \ 5\ \ 0\ \ 0)$$

These matrices represent the three offices of the branch (say A) where elements appearing in the row represent the number of peons, clerks, typists, salesmen and head-clerks taken in that order working in the three offices.

Then, $A_1 + A_2 + A_3 = (3+3+3\ \ 4+4+4\ \ 5+5+5\ \ 6+0+0\ \ 0+0+0)$
$= (9\ \ 12\ \ 15\ \ 6\ \ 0)$

Thus, total number of posts of each kind in all the offices of branch A are the elements of matrix $A_1 + A_2 + A_3 = (9\ \ 12\ \ 15\ \ 6\ \ 0)$

Now consider the following row matrices,

$$B_1 = (3\ \ 4\ \ 5\ \ 0\ \ 2),\quad B_2 = (3\ \ 4\ \ 5\ \ 0\ \ 2),\quad B_3 = (3\ \ 4\ \ 5\ \ 0\ \ 2)$$

Then B_1, B_2, B_3 represent three offices of other branch (say B) where the elements in the row represents number of peons, clerks, typists, salesmen and head-clerks respectively.

NOTES

Thus, total number of posts of each kind in all the offices of branch B are the elements of the matrix $B_1 + B_2 + B_3 = (9 \ \ 12 \ \ 15 \ \ 0 \ \ 6)$

(*ii*) The total number of posts of each kind in all the offices taken together from both branches are the elements of matrix,

$$(A_1 + A_2 + A_3) + (B_1 + B_2 + B_3) = (18 \ \ 24 \ \ 30 \ \ 6 \ \ 6)$$

Example 1.14: Let $A = \begin{pmatrix} 10 & 20 \\ 30 & 40 \end{pmatrix}$ where first row represents the number of table fans and second row represents the number of ceiling fans which two manufacturing units A and B make in one day. The first and second column represent the manufacturing units A and B. Compute $5A$ and state what it represents.

Solution: $5A = \begin{pmatrix} 50 & 100 \\ 150 & 200 \end{pmatrix}$

It represents the number of table fans and ceiling fans that the manufacturing units A and B produce in five days.

Example 1.15: Let $A = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}$ where rows represent the number of items of type I, II, III, respectively. The four columns represents the four shops A_1, A_2, A_3, A_4 respectively.

Let, $B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 2 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 4 \end{pmatrix}$

Where elements in B represent the number of items of different types delivered at the beginning of a week and matrix C represent the sales during that week. Find,

(*i*) The number of items immediately after delivery of items.

(*ii*) The number of items at the end of the week.

(*iii*) The number of items needed to bring stocks of all items in all shops to 6.

Solution: (*i*) $A + B = \begin{pmatrix} 3 & 5 & 7 & 9 \\ 5 & 5 & 7 & 9 \\ 7 & 7 & 7 & 9 \end{pmatrix}$

Represent the number of items immediately after delivery of items.

(*ii*) $(A + B) - C = \begin{pmatrix} 2 & 3 & 5 & 6 \\ 4 & 3 & 4 & 5 \\ 5 & 4 & 3 & 5 \end{pmatrix}$

Represent the number of items at the end of the week.

(*iii*) We want that all elements in $(A + B) - C$ should be 6.

Let $D = \begin{pmatrix} 4 & 3 & 1 & 0 \\ 2 & 3 & 2 & 1 \\ 1 & 2 & 3 & 1 \end{pmatrix}$

Then $(A + B) - C + D$ is a matrix in which all elements are 6. So, D represents the number of items needed to bring stocks of all items of all shops to 6.

Example 1.16: The following matrix represents the results of the examination of B. Com. class:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$$

NOTES

The rows represent the three sections of the class. The first three columns represent the number of students securing 1st, 2nd, 3rd divisions respectively in that order and fourth column represents the number of students who failed in the examination.

(*i*) How many students passed in three sections respectively?

(*ii*) How many students failed in three sections respectively?

(*iii*) Write down the matrix in which number of successful students is shown.

(*iv*) Write down the column matrix where only failed students are shown.

(*v*) Write down the column matrix showing students in 1st division from three sections.

Solution: (*i*) The number of students who passed in three sections respectively are $1 + 2 + 3 = 6$, $5 + 6 + 7 = 18$, $9 + 10 + 11 = 30$.

(*ii*) The number of students who failed from three sections respectively are 4, 8, 12.

(*iii*) $\begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{pmatrix}$

(*iv*) $\begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix}$ represents column matrix where only failed students are shown.

(*v*) $\begin{pmatrix} 1 \\ 5 \\ 9 \end{pmatrix}$ represents column matrix of students securing 1st division.

Transpose of a Matrix

Let A be a matrix. The matrix obtained from A by interchange of its rows and columns, is called the *transpose* of A. For example,

If, $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix}$ then transpose of $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 2 & 0 \end{pmatrix}$

Transpose of A is denoted by A'.

It can be easily verified that,

(*i*) $(A')' = A$

(*ii*) $(A + B)' = A' + B'$

(*iii*) $(AB)' = B'A'$

Example 1.17: For the following matrices A and B verify $(A + B)' = A' + B'$.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 8 & 6 \end{pmatrix}$$

Solution: $A' = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$ $B' = \begin{pmatrix} 2 & 1 \\ 3 & 8 \\ 4 & 6 \end{pmatrix}$

NOTES

So, $A' + B' = \begin{pmatrix} 3 & 5 \\ 5 & 13 \\ 7 & 12 \end{pmatrix}$

Again, $A + B = \begin{pmatrix} 3 & 5 & 7 \\ 5 & 13 & 12 \end{pmatrix}$

So, $(A+B)' = \begin{pmatrix} 3 & 5 \\ 5 & 13 \\ 7 & 12 \end{pmatrix}$

Therefore, $(A+B)' = A' + B'$

Check Your Progress

5. What is a matrix?
6. When a matrix is termed as square matrix?
7. What is a diagonal matrix?
8. Define identity matrix.
9. What is matrix equality?
10. How is matrix Addition commutative?

1.8 KRONECKER SUM AND PRODUCT OF MATRICES

Let (x_1, x_2, x_3, x_4) and $(x_1 + dx_1, x_2 + dx_2, x_3 + dx_3, x_4 + dx_4)$ be the coordinates of two neighbouring events considered in Minkowski's four dimensional space. Then the interval *ds* between these two neighbouring events in any coordinate system, is given by

$$ds^2 = g_{11}\, dx_1^{\,2} + g_{22}dx_2^{\,2} + g_{33}dx_3^{\,2} + g_{44}dx_4^{\,2} + 2g_{12}dx_1dx_2 + 2g_{13}dx_1dx_3 + 2g_{14}dx_1dx_4 + 2g_{23}dx_2dx_3 + 2g_{24}dx_2dx_4 + 2g_{34}dx_3dx_4 \quad \text{... (1.59)}$$

where the coefficients $g_{\mu\nu}$ (μ, ν = 1, 2, 3, 4) are functions of x_1, x_2, x_3, x_4. This follows that ds^2 is some quadratic function of the difference of coordinates.

Adopting the convention that whenever a literal suffix appears twice in a term that term is to be summed for values of the suffix 1, 2, 3, 4; equation (1.59) can be written as

$$ds^2 = g_{\mu\nu}\, dx_\mu\, dx_\nu \ (\mu, \nu = 1, 2, 3, 4 \text{ and } g_{\mu\nu} = g_{\nu\mu}) \quad \text{... (1.60)}$$

Since μ and ν each appear twice, the right hand side of (2) indicates the summation

$$\sum_{\mu=1}^{4} \sum_{\nu=1}^{4}$$

Any literal suffix appearing twice in a term is said to be a *dummy suffix* and it may be changed freely to any other letter not already used in that term. Also two or more dummy suffixes can be interchanged, *e.g.*,

$$g_{\alpha\beta}\frac{\partial^2 x_\alpha}{\partial x'_\mu \partial x'_\nu}\frac{\partial x_\beta}{\partial x'_\lambda} = g_{\alpha\beta}\frac{\partial^2 x_\beta}{\partial x'_\mu\, \partial x'_\nu}\cdot\frac{\partial^2 x_\alpha}{\partial x'_\lambda}$$

(by intercharging the dummy suffixes α and β and using $g_{\beta\alpha} = g_{\alpha\beta}$)

NOTES

Illustration. *To prove that*

$$\frac{\partial x_\mu}{\partial x'_\alpha}\cdot\frac{\partial x'_\alpha}{\partial x_\nu} = \frac{\partial x_\mu}{\partial x_\nu} = 0 \quad \text{if} \quad \mu \neq \nu$$

$$= 1 \text{ if } \mu = \nu \text{ where } \alpha = 1, 2, 3, 4.$$

Here, R.H.S. $= \frac{\partial x_\mu}{\partial x'_1}\frac{\partial x'_1}{\partial x_\nu} + \frac{\partial x_\mu}{\partial x'_2}\frac{\partial x'_2}{\partial x_\nu} + \frac{\partial x_\mu}{\partial x'_3}\frac{\partial x'_3}{\partial x_\nu} + \frac{\partial x_\mu}{\partial x'_4}\frac{\partial x'_4}{\partial x'_\nu}$

$$= \frac{\partial x_\mu}{\partial x_\nu}$$

x_μ and x_ν being the coordinates of the same system, their variations are independent and so

$$dx_\mu = 0 \text{ when } \mu \neq \nu$$

and $dx_\mu = dx_\nu$ when $\mu = \nu$

$$\therefore \quad \frac{\partial x_\mu}{\partial x'_\nu} = \frac{\partial x_\mu}{\partial x'_\alpha}\cdot\frac{\partial x'_\alpha}{\partial x_\nu} = 0 \text{ when } \mu \neq \nu$$

$$= 1 \text{ when } \mu = \nu$$

Here the multiplier $\frac{\partial x_\mu}{\partial x'_\alpha}\frac{\partial x'_\alpha}{\partial x_\mu}$ acts as a *substitution operator*.

It is rather convenient to write

$\frac{\partial x_\mu}{\partial x_\nu} = \delta_{\mu\nu}$ or δ^μ_ν which is known as *Kronecker delta*.

As such the above results can be expressed as

$$\left.\begin{aligned}\delta^\mu_\nu &= 0 \text{ if } \mu \neq \nu\\ &= 1 \text{ if } \mu = \nu\end{aligned}\right\} \quad \text{... (1.61)}$$

Thus if $A(\mu)$ be an expression involving the suffix μ, then

$$\frac{\partial x_\mu}{\partial x'_\alpha}\frac{\partial x'_\alpha}{\partial x_\nu} = A(\mu) = A(\nu) \quad \text{... (1.62)}$$

for; the summation on the left, with respect to μ gives four terms corresponding to $\mu = 1, 2, 3, 4$; one of which will agree with ν. Denoting the other three values by σ, τ, ρ, the left hand side of (1.62) is

$$= 1 \,.\, A(\nu) + 0 \,.\, A(\sigma) + 0 \,.\, A(\tau) + 0 \,.\, A(\rho) \text{ by (16.1)}$$

$$= A(\nu)$$

i.e. $\delta^\mu_\nu A(\mu) = A(\nu)$ ... (1.63)

Evidently $\delta^\mu_\rho \delta^\rho_\nu = \delta^\mu_\nu$ ... (1.64)

and $\delta^\mu_\mu = 4$ ... (1.65)

for, in the latter case, $\delta^{\mu}_{\mu} = \delta^1_1 + \delta^2_2 + \delta^3_3 + \delta^4_4$

$= 1 + 1 + 1 + 1 = 4$ by (3)

NOTES

1.9 DEFINITION OF TENSOR

In this section, we will study about the concept of tensor.

Cartesian Tensor

In geometry and linear algebra, a Cartesian tensor uses an orthonormal basis to represent a tensor in a Euclidean space in the form of components. Converting a tensor's components from one such basis to another is through an orthogonal transformation.

The most familiar coordinate systems are the two-dimensional and three-dimensional Cartesian coordinate systems. Cartesian tensors may be used with any Euclidean space, or more technically, any finite-dimensional vector space over the field of real numbers that has an inner product.

Use of Cartesian tensors occurs in physics and engineering, such as with the Cauchy stress tensor and the moment of inertia tensor in rigid body dynamics. Sometimes general curvilinear coordinates are convenient, as in high-deformation continuum mechanics, or even necessary, as in general relativity. While orthonormal bases may be found for some such coordinate systems (for example, tangent to spherical coordinates), Cartesian tensors may provide considerable simplification for applications in which rotations of rectilinear coordinate axes suffice. The transformation is a passive transformation, since the coordinates are changed and not the physical system.

Tensor analysis is the generalization of vector analysis as is evident by considering a vector function $f(\mathbf{r})$ of a vector $\mathbf{r}$. This function is continuous at $\mathbf{r} = \mathbf{r}_0$ if,

$$\lim_{\mathbf{r} \to \mathbf{r}_0} f(\mathbf{r}) = f(\mathbf{r}_0)$$

and it is linear, if

$$f(\mathbf{r} + \mathbf{s}) = f(\mathbf{r}) + f(\mathbf{s}) \quad \text{... (1.66)}$$

and

$$f(\lambda \mathbf{r}) = \lambda f(\mathbf{r}) \quad \text{... (1.67)}$$

for arbitrary values of $\mathbf{r}$, $\mathbf{s}$, λ.

Now, we know that a linear vector function $f(\mathbf{r})$ is completely defined only if $f(\mathbf{a}_1)$, $f(\mathbf{a}_2)$ and $f(\mathbf{a}_3)$ are given for any three non-coplanar vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$. In terms of $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ as basis if we assume that

$$\mathbf{r} = x_1\, \mathbf{a}_1 + x_2\, \mathbf{a}_2 + x_3\, \mathbf{a}_3 \quad \text{... (1.68)}$$

then we have from Equations (6.1) and (6.2),

$$f(r) = x_1 f(\mathbf{a}_1) + x_2 f(\mathbf{a}_2) + x_3 f(\mathbf{a}_3)$$

As such Equation (6.3) yields

$$x_\alpha = \mathbf{r} \,.\, \mathbf{a}_\alpha, \; \alpha = 1, 2, 3$$

Let us put $\quad f(\mathbf{a}_\alpha) = \mathbf{b}_\alpha$.

So that $\quad f(\mathbf{r}) = (\mathbf{b}_1\mathbf{a}_1 + \mathbf{b}_2\mathbf{a}_2 + \mathbf{b}_3\mathbf{a}_3) \,.\, \mathbf{r}$

$$= \phi \,.\, \mathbf{r} \text{ (say)}$$

where the operator $\phi = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3$ consists of nine components in three dimensional coordinate geometry and hence it is neither a scalar nor a vector quantity but is a new mathematical symbol called as the *dyadic*.

NOTES

Suppose that there are two vectors **u** and **v** such that components of vector **v** are linear functions of the components of vector **u** defined as

$$\left.\begin{aligned} v_x &= a_{xx}\, u_x + a_{xy}\, u_y + a_{xz}\, u_z \\ v_y &= a_{yx}\, u_x + a_{yy}\, u_y + a_{yz}\, u_z \\ v_z &= a_{zx}\, u_x + a_{zy}\, u_y + a_{zz}\, u_z \end{aligned}\right\} \quad \text{... (1.69)}$$

In this way the vector **v** is placed in one-to-one correspondence with the vector **u**. The scheme of coefficients $a_{\alpha\beta}$ has thus an independent meaning if the correspondence is such that the passage from **u** to **v** is independent of the particular coordinate system in which the vectors are resolved into components. We call the coefficients $a_{\alpha\beta}$ in this case as the coefficients of a tensor.

It is observed that the nine components as mentioned above characterise the transformation of the components of one vector into those of other. The coeffcients $a_{\alpha\beta}$ in general transform u_β into one of three parts of v_α.

The Equations (6.4) are equivalent to a single vector equation,

$$\mathbf{v} = \phi\mathbf{u}$$

where the operator ϕ turns **u** into **v**. It is rather graphically known as **Tensor**.

The essential part of a tensor operation is the array of coefficients like $a_{\alpha\beta}$, written in the form of a matrix such as

$$\phi = \begin{bmatrix} a_{xx} & a_{xy} & a_{xz} \\ a_{yx} & a_{yy} & a_{yz} \\ a_{zx} & a_{zy} & a_{zz} \end{bmatrix}$$

As such the *dyadic* operator turns a vector **r** into the vector function $f(\mathbf{r})$ and is expressed as the sum of dyads **ab**, *i.e.*,

$$\phi = \sum_\alpha \mathbf{a}_\alpha\, \mathbf{b}_\alpha$$

In a similar way a *triadic* is expressed as the sum of the triads $\sum_\alpha \mathbf{a}_\alpha\, \mathbf{b}_\alpha\, \mathbf{c}_\alpha$.

Considering it as an opearor that converts vector **r** into the dyadic ϕ, we may write

$$\phi \,.\, \mathbf{r} = \sum(\mathbf{a}_\alpha\, \mathbf{b}_\alpha\, \mathbf{c}_\alpha).\mathbf{r}$$

Similarly a *tetradic* is the sum of tetrads, $\sum \mathbf{a}_\alpha\, \mathbf{b}_\alpha\, \mathbf{c}_\alpha\, \mathbf{d}_\alpha$ and etc.

All such physical quantities as scalars, vectors, dyadics, triadics, tetradics etc. are collectively known as tensors of rank 1, 2, 3, 4 etc. and as such the tensors can be regarded as generalized extended form of vectors.

The examples of dyadic, i.e., tensor of rank two are: an operator relating dielectric displacement vector with the electric vector of an electro-magnetic wave in an isotropic medium; a stress tensor relating stress and strain in an isotropic medium in which case a component of stress **T** is a function of every component of strain **S**.

i.e., $$T_\alpha = \sum_{\beta=1}^{3} \mathbf{a}_{\alpha\beta} S_\beta \quad \text{or} \quad \mathbf{T} = \varphi \mathbf{S}$$

NOTES

where φ is a nine coefficient operator in three dimensional space.

Note. The dyadic or tensor of rank two is also known as *Stress Tensor*.

An Explanatory Note on Dyads and Dyadics

We know that the gradient of a vector **f** such as

$$\nabla \mathbf{f} = \mathbf{i} f_x + \mathbf{j} f_y + \mathbf{k} f_z \qquad \text{... (1.70)}$$

is meaningless as it consists of sum of three ordered pairs of vectors, but we sometimes take it to define as *dyadic* and the ordered vector pairs as *dyads*. In Equation (1.70) we regard ∇**f** as an operator setting up a one-one correspondence between directions **e** at a point and its directional derivative $\frac{d\mathbf{f}}{ds}$, i.e., $\frac{d\mathbf{f}}{ds} = \mathbf{e} \cdot \nabla \mathbf{f}$

Actually the dyadic ∇**f** replaces an infinite number of vectors $\frac{d\mathbf{f}}{ds}$, so that any sum of dyads is called as dyadic, for example, the dyadic

$$\mathbf{D} = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \ldots + \mathbf{a}_n\mathbf{b}_n$$

is a general dyadic in which the vectors $\mathbf{a}_\alpha$ are known as *antecedents* and $\mathbf{b}_\alpha$ as *consequents*, while the dyadic

$$\mathbf{D}_c = \mathbf{b}_1\mathbf{a}_1 + \mathbf{b}_2\mathbf{a}_2 + \ldots\ldots\ldots\ldots \mathbf{b}_n\mathbf{a}_n$$

is said to be the conjugate of **D**, such that **D** is *symmetric* if $\mathbf{D} = \mathbf{D}_c$ and *skew* if $\mathbf{D} = -\mathbf{D}_c$.

In Equation (1.70) if **f** is replaced by $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$

such that

$$\mathbf{r}_x = \mathbf{i},\ \mathbf{r}_y = \mathbf{j},\ r_z = \mathbf{k} \text{ then}$$

$$\nabla \mathbf{r} = \mathbf{ii} + \mathbf{jj} + \mathbf{kk} = I$$

where the dyadic **I** is termed as *Idemfactor* as it transforms any vector **V** into itself, i.e.

$$\mathbf{V} \cdot \mathbf{I} = \mathbf{I} \cdot \mathbf{V} = \mathbf{V} \text{ for every } \mathbf{V}$$

Transformation of Coordinates

If we focus our attention on some point of Minkowski's four dimensional world and consider the transformation from one system of coordinates (x_1, x_2, x_3, x_4) to another system (x_1', x_2', x_3', x_4'), such that

$$x_1' = f_1 (x_1, x_2, x_3, x_4) \text{ etc.}$$

then we can solve x_1, x_2, x_3, x_4 in terms of x_1', x_2', x_3', x_4' such that

$$x_1 = \phi_1 (x_1', x_2', x_3', x_4') \text{ etc.}$$

and the differentials dx_1, dx_2, dx_3, dx_4 are then transformed as

$$dx_1 = \frac{\partial x_1'}{\partial x_1} dx_1 + \frac{\partial x_1'}{\partial x_2} dx_2 + \frac{\partial x_1'}{\partial x_3} dx_3 + \frac{\partial x_1'}{\partial x_4} dx_4 \text{ etc.}$$

or symbolically

$$dx_\mu' = \sum_{\alpha=1}^{4} \frac{\partial x'_\mu}{\partial x_\alpha}\, dx_\alpha; \ (\mu = 1, 2, 3, 4) \text{ etc.}$$

NOTES

The Summation Convention and Kronecker Delta Symbol

Let (x_1, x_2, x_3, x_4) and $(x_1 + dx_1, x_2 + dx_2, x_3 + dx_3, x_4 + dx_4)$ be the coordinates of two neighbouring events considered in Minkowski's four dimensional space. Then the interval *ds* between these two neighbouring events in any coordinate system, is given by

$$ds^2 = g_{11}\, dx_1^2 + g_{22}dx_2^2 + g_{33}dx_3^2 + g_{44}dx_4^2 + 2g_{12}dx_1dx_2 + 2g_{13}dx_1dx_3 + 2g_{14}dx_1dx_4 + 2g_{23}dx_2dx_3 + 2g_{24}dx_2dx_4 + 2g_{34}dx_3dx_4 \quad \text{... (1.71)}$$

where the coefficients $g_{\mu\nu}$ $(\mu, \nu = 1, 2, 3, 4)$ are functions of x_1, x_2, x_3, x_4. This follows that ds^2 is some quadratic function of the difference of coordinates.

Adopting the convention that whenever a literal suffix appears twice in a term that term is to be summed for values of the suffix 1, 2, 3, 4; Equation (1.71) can be written as

$$ds^2 = g_{\mu\nu}\, dx_\mu\, dx_\nu \ (\mu, \nu = 1, 2, 3, 4 \text{ and } g_{\mu\nu} = g_{\nu\mu}) \quad \text{... (1.72)}$$

Since μ and ν each appear twice, the right hand side of Equation (1.72) indicates the summation

$$\sum_{\mu=1}^{4} \sum_{\nu=1}^{4}$$

Any literal suffix appearing twice in a term is said to be a *dummy suffix* and it may be changed freely to any other letter not already used in that term. Also two or more dummy suffixes can be interchanged, for example,

$$g_{\alpha\beta} \frac{\partial^2 x_\alpha}{\partial x'_\mu \partial x'_\nu} \frac{\partial x_\beta}{\partial x'_\lambda} = g_{\alpha\beta} \frac{\partial^2 x_\beta}{\partial x'_\mu\, \partial x'_\nu} \cdot \frac{\partial^2 x_\alpha}{\partial x'_\lambda}$$

(by intercharging the dummy suffixes α and β and using $g_{\beta\alpha} = g_{\alpha\beta}$)

Illustration. *To prove that*

$$\frac{\partial x_\mu}{\partial x'_\alpha} \cdot \frac{\partial x'_\alpha}{\partial x_\nu} = \frac{\partial x_\mu}{\partial x_\nu} = 0 \quad \text{if} \quad \mu \neq \nu$$

$$= 1 \text{ if } \mu = \nu \text{ where } \alpha = 1, 2, 3, 4.$$

Here,

$$\text{R.H.S.} = \frac{\partial x_\mu}{\partial x'_1} \frac{\partial x'_1}{\partial x_\nu} + \frac{\partial x_\mu}{\partial x'_2} \frac{\partial x'_2}{\partial x_\nu} + \frac{\partial x_\mu}{\partial x'_3} \frac{\partial x'_3}{\partial x_\nu} + \frac{\partial x_\mu}{\partial x'_4} \frac{\partial x'_4}{\partial x'_\nu}$$

$$= \frac{\partial x_\mu}{\partial x_\nu}$$

x_μ and x_ν being the coordinates of the same system, their variations are independent and so

$$dx_\mu = 0 \text{ when } \mu \neq \nu$$

and

$$dx_\mu = dx_\nu \text{ when } \mu = \nu$$

NOTES

$$\therefore \qquad \frac{\partial x_\mu}{\partial x'_\nu} = \frac{\partial x_\mu}{\partial x'_\alpha}\cdot\frac{\partial x'_\alpha}{\partial x_\nu} = 0 \text{ when } \mu \neq \nu$$

$$= 1 \text{ when } \mu = \nu$$

Here the multiplier $\frac{\partial x_\mu}{\partial x'_\alpha}\frac{\partial x'_\alpha}{\partial x_\mu}$ acts as a *substitution operator*.

It is rather convenient to write

$$\frac{\partial x_\mu}{\partial x_\nu} = \delta_{\mu\nu} \text{ or } \delta^\mu_\nu$$ which is known as *Kronecker delta*.

As such the above results can be expressed as

$$\left.\begin{aligned}\delta^\mu_\nu &= 0 \text{ if } \mu \neq \nu\\ &= 1 \text{ if } \mu = \nu\end{aligned}\right\} \qquad \text{... (1.73)}$$

Thus if $A(\mu)$ be an expression involving the suffix μ, then

$$\frac{\partial x_\mu}{\partial x'_\alpha}\frac{\partial x'_\alpha}{\partial x_\nu} = A(\mu) = A(\nu) \qquad \text{... (1.74)}$$

for; the summation on the left, with respect to μ gives four terms corresponding to $\mu = 1, 2, 3, 4$; one of which will agree with ν. Denoting the other three values by σ, τ, ρ, the left hand side of Equation (1.74) is

$$= 1 \,.\, A(\nu) + 0 \,.\, A(\sigma) + 0 \,.\, A(\tau) + 0 \,.\, A(\rho) \text{ by Equation (1.73)}$$

$$= A(\nu)$$

i.e. $\qquad \delta^\mu_\nu A(\mu) = A(\nu) \qquad$ ... (1.75)

Evidently $\delta^\mu_\rho \delta^\rho_\nu = \delta^\mu_\nu$

and $\qquad \delta^\mu_\mu = 4$

for, in the latter case, $\delta^\mu_\mu = \delta^1_1 + \delta^2_2 + \delta^3_3 + \delta^4_4$

$$= 1 + 1 + 1 + 1 = 4 \text{ by Equation (6.8)}$$

Tensors as Classification of Transformation Laws

We have already mentioned that if we consider the transformation from one system of coordinates (x_1, x_2, x_3, x_4) to another system (x_1', x_2', x_3', x_4'), then the differentials dx_1, dx_2, dx_3, dx_4 are transformed as

$$dx'_1 = \frac{\partial x'_1}{\partial x_1}dx_1 + \frac{\partial x'_1}{\partial x_2}dx_2 + \frac{\partial x'_1}{\partial x_3}dx_3 + \frac{\partial x'_1}{\partial x_4}dx_4 \text{ etc.}$$

or in short as $\quad dx'_\mu = \sum_{\alpha=1}^{4}\frac{\partial x'_\mu}{\partial x_\alpha}dx_\alpha \quad \mu = 1, 2, 3, 4.$

Any set of four quantities transformed in accordance with this law is said to be a **Contravariant Vector**. Thus if a coordinate system (A^1, A^2, A^3, A^4) transforms to the new coordinate system (A'^1, A'^2, A'^3, A'^4) where

$$A'^\mu = \sum_{\alpha=1}^{4}\frac{\partial x'_\mu}{\partial x_\alpha}A^\alpha \qquad \text{... (1.76)}$$

Then (A^1, A^2, A^3, A^4) or briefly A^μ is a contravariant vector. Hence the upper position of the suffix (which is definitely, not an exponent) is reserved to indicate contravariant vectors.

Now, if we consider an operator ϕ such that it is an invariant function of position, i.e., it has a fixed value at each point independent of the coordinate system used, then the four quantities

$\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}, \frac{\partial\phi}{\partial x_3}, \frac{\partial\phi}{\partial x_4}$ are transformed as

$$\frac{\partial\phi}{\partial x_1'} = \frac{\partial x_1}{\partial x_1'}\frac{\partial\phi}{\partial x_1} + \frac{\partial x_2}{\partial x_1'}\frac{\partial\phi}{\partial x_2} + \frac{\partial x_3}{\partial x_1'}\frac{\partial\phi}{\partial x_3} + \frac{\partial x_4}{\partial x_1'}\frac{\partial\phi}{\partial x_4}, \text{ etc.}$$

or in short
$$\frac{\partial\phi}{\partial x_\mu'} = \sum_{\alpha=1}^{4}\frac{\partial x_\alpha}{\partial x_\mu'}\frac{\partial\phi}{\partial x_\alpha} \quad (\mu = 1, 2, 3, 4).$$

Any set of four quantities transformed in accordance with this law is said to be a **Covariant Vector**.

Thus if A_μ be a covariant vector, its transformation law is

$$A_\mu' = \sum_{\alpha=1}^{4}\frac{\partial x_\alpha}{\partial x_\mu'}A_\alpha \qquad \text{... (1.77)}$$

where the lower position of the suffix indicates covariance.

Hence the Relations (1.76) and (1.77) give the laws of transformation of vectors. If we denote by $A_{\mu\nu}$ a quantity with 16 components by assigning μ and ν the values 1, 2, 3, 4 independently, then a generalization of these laws yields quantities classified as

Contravariant Tensors $A'^{\mu\nu} = \frac{\partial x_\mu'}{\partial x_\alpha}\frac{\partial x_\nu'}{\partial x_\beta}A^{\alpha\beta}$

Covariant Tensors $A_{\mu\nu}' = \frac{\partial x_\alpha}{\partial x_\mu'}\frac{\partial x_\beta'}{\partial x_\nu}A_{\alpha\beta}$

Mixed Tensors $A_\mu'^\nu = \frac{\partial x_\alpha}{\partial x_\mu'}\frac{\partial x_\nu'}{\partial x_\beta}A_\alpha^\beta$

These are the tensors of *second rank*.

Similarly $A_{\mu\nu\sigma}$ has 64 components and $A_{\mu\nu\sigma\tau}$ has 256 components. Thus a tensor of higher rank is of the type.

$$A_{\mu\nu\sigma}'^{\tau} = \frac{\partial x_\alpha}{\partial x_\mu'}\frac{\partial x_\beta}{\partial x_\nu'}\frac{\partial x_\lambda}{\partial x_\sigma'}\frac{\partial x_\tau'}{\partial x_\delta'}A_{\alpha\beta\gamma}^{\delta}$$

Note. A vector is a tensor of *first rank* and an invariant or scalar is a tensor of *zero rank*.

Rank of a Tensor: The rank of a tensor is determined by the number of suffixes or indices attached to it. As a matter of fact the rank of a tensor when raised as power to the number of dimensions, yields the number of components of the tensor and hence the components of the matrix that represents the tensor. As such a tensor of rank n in four dimensional space has 4^n components. Consequently, the rank of a tensor gives the number of the mode of changes of a physical quantity when passing from one system to another system which is in rotation relative to the first. It is clear from this discussion that a quantity that does not change when the axes are rotated is a tensor of *zero rank*, since the number of mode of changes is then zero. These quantities named as tensor of zero rank are scalars while the tensors of rank one are vectors.

NOTES

Contravariant Tensor: If n quantities A^α ($\alpha = 1, 3, ...n$) in a coordinate system ($x_1, x_2, ... x_n$) are related to n other quantities A'^α ($\alpha = 1, 2, ...n$) in another coordinate system ($x_1', x_2', ..., x_n'$) by the transformation laws

$$A'^\mu = \frac{\partial x'_\mu}{\partial x_\alpha} A^\alpha \quad \text{(Contravariant law)}$$

on change of the coordinate x_β to x'_β according to summation convention, then A^α are termed as the components of a contravariant vector or a contravariant tensor of the first rank.

Covariant Tensor: If n quantities A_α ($\alpha = 1, 2,......n$) in a coordinate system ($x_1, x_2, ...x_n$) are related to n other quantities A'_α ($\alpha = 1, 2, ...n$) in another coordinate system ($x'_1, x'_2, ... x'_n$) by the transformation laws

$$A'_\mu = \frac{\partial x_\alpha}{\partial x'_\mu} A_\alpha \quad \text{(Covariant law)}$$

according to summation convention, then A_α are termed as the components of a covariant vector or a covariant tensor of the first rank.

Problem 1.3. *Show that the velocity of a fluid at any point is a contravariant vector of rank one.*

Assuming that $x_\alpha(t)$ is the coordinate of a moving particle with the time t, we have

$$v^\alpha = \frac{dx_\alpha}{dt}$$

as the velocity of the particle.

In transformed coordinates the components of velocity are

$$v'^\alpha = \frac{dx'_\alpha}{dt}$$

But $$v^\alpha = \frac{d}{dt}x_\alpha = \frac{\partial x'_\alpha}{\partial x_\beta}\frac{dx_\beta}{dt} = \frac{\partial x'_\alpha}{\partial x_\beta} v_\beta$$

which follows that velocity is a contravariant vector of rank one.

Problem 1.4. *Show that the law of transformation for a contravariant vector is transitive.*

We have $$A'^\mu = \frac{\partial x'_\mu}{\partial x_\alpha} A^\alpha$$

Let $$A''^\mu = \frac{\partial x''_\mu}{\partial x'_\alpha} A'^\alpha$$

$$\therefore \quad A''^\mu = \frac{\partial x''_\mu}{\partial x'_\beta} A^\beta = \frac{\partial x''_\mu}{\partial x'_\beta}\frac{\partial x'_\beta}{\partial x_\alpha} A^\alpha = \frac{\partial x''_\mu}{\partial x_\alpha} A^\alpha$$

which shows that contravariant law is transitive.

Problem 1.5. *Find the components of a vector in polar coordinates whose components in cartesian coordinates are* $\dot{x}, \dot{y}$ *and* $\ddot{x}, \ddot{y}$.

NOTES

As given, suppose that

$$x_1 = x,\ x_2 = y,$$
$$x_1' = r,\ x_2' = \theta$$

and $\quad (i) A^1 = \dot{x},\ A^2 = \dot{y}$

$\quad (ii) A^1 = \ddot{x},\ A^2 = \ddot{y}.$

Then, we have to find C'^1, A'^2.

We have the transformations

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

giving $\quad r^2 = x^2 + y^2 \quad \text{and} \quad \theta = \tan^{-1}\dfrac{y}{x}$

so that $\quad \dfrac{\partial r}{\partial x} = \dfrac{x}{r},\ \dfrac{\partial r}{\partial y} = \dfrac{y}{r},\ \dfrac{\partial \theta}{\partial x} = -\dfrac{y}{r^2} \text{ and } \dfrac{\partial \theta}{\partial y} = \dfrac{x}{r^2}$

Also $\quad r\dot{r} = x\dot{x} + y\dot{y} \text{ and } \dot{\theta} = \dfrac{x\dot{y} - y\dot{x}}{r^2}$

(i) Transformation law as already defined gives

$$A'^{\mu} = \frac{\partial x'_\mu}{\partial x_\alpha} A^\alpha \ (\alpha = 1, 2)$$
$$= \frac{\partial x'_\mu}{\partial x_1} A^1 + \frac{\partial x'_\mu}{\partial x_2} A^2$$
$$= \frac{\partial x'_\mu}{\partial x_1} \dot{x} + \frac{\partial x'_\mu}{\partial x_2} \dot{y} \quad \text{as } A^1 = \dot{x} \text{ and } A^2 = \dot{y}$$

$\therefore$
$$A'^1 = \frac{\partial x'_1}{\partial x_1} \dot{x} + \frac{\partial x'_1}{\partial x_2} \dot{y}$$
$$= \frac{\partial r}{\partial x} \dot{x} + \frac{\partial r}{\partial y} \dot{y} \quad \text{as} \quad x_1 = x,\ x_2 = y \quad \text{and} \quad x_1' = r$$
$$= \frac{x}{r}\dot{x} + \frac{y}{r}\dot{y}$$
$$= \frac{x\dot{x} + y\dot{y}}{r} = \frac{r\dot{r}}{r} = \dot{r}$$

and
$$A'^2 = \frac{\partial x'_2}{\partial x_1} \dot{x} + \frac{\partial x'_2}{\partial x_2} \dot{y}$$
$$= \frac{\partial \theta}{\partial x} \dot{x} + \frac{\partial \theta}{\partial y} \dot{y} \quad \text{as} \quad x'_2 = \theta,\ x_1 = x \quad \text{and} \quad x_2 = y$$
$$= -\frac{y}{r^2}\dot{x} + \frac{x}{r^2}\dot{y}$$
$$= -\frac{y\dot{x} - y\dot{x}}{r^2} = \dot{\theta}$$

(ii)
$$A'^1 = \frac{\partial r}{\partial x}\ddot{x} + \frac{\partial r}{\partial y}\ddot{y} \quad \text{as in part } (i)$$
$$= \frac{x\ddot{x} + y\ddot{y}}{r}$$

But $x\dot{x} + y\dot{y} = r\dot{r}$ gives on differentiation,

NOTES

$$x\ddot{x} + \dot{x}^2 + y\ddot{y} + \dot{y}^2 = r\ddot{r} + r\dot{r}^2,$$

i.e.,
$$x\ddot{x} + y\ddot{y} = r\ddot{r} + \dot{r}^2 - \dot{x}^2 - \dot{y}^2$$
$$= r\ddot{r} + \dot{r}^2 - (\dot{r}\cos\theta - r\sin\theta\,\dot{\theta})^2 - (\dot{r}\sin\theta + r\cos\theta\,\dot{\theta})^2$$
$$= r\ddot{r} + \dot{r}^2 - \dot{r}^2 - r^2\dot{\theta}^2$$
$$= r\ddot{r} - r^2\dot{\theta}^2$$

Thus
$$A'^1 = \frac{r\ddot{r} - r^2\dot{\theta}^2}{r}$$
$$= \ddot{r} - r\dot{\theta}^2$$

and
$$A'^2 = \frac{\partial\theta}{\partial x}\ddot{x} + \frac{\partial\theta}{\theta y}\ddot{y}$$
$$= \frac{x\ddot{y} + y\ddot{x}}{r^2}.$$

But $x\dot{y} - y\dot{x} = r^2\dot{\theta}$ gives on differentiation

$$\dot{x}\dot{y} + x\ddot{y} - \dot{y}\dot{x} - y\ddot{x} = r^2\ddot{\theta} + 2r\dot{r}\,\dot{\theta}$$

or
$$x\ddot{y} - y\ddot{x} = r^2\ddot{\theta} + 2r\dot{r}\,\dot{\theta}$$

$\therefore$
$$A'^2 = \frac{r^2\ddot{\theta} + 2r\dot{r}\,\dot{\theta}}{r^2}$$
$$= \ddot{\theta} + \frac{2\dot{r}\dot{\theta}}{r}$$

1.9.1 Symmetric and Anti-Symmetric Tensors

Let a tensor be such that contravariant or covariant indices of it can be interchanged without altering the value of the tensor, then the tensor is termed as symmetrical or symmetric in these indices.

If $A^{\mu\nu}$ and $A^{\nu\mu}$ be two contravariant tensors in a certain system of coordinates such that

$$A^{\mu\nu} = A^{\nu\mu}$$

then if $A^{\mu\nu}$ and $A^{\nu\mu}$ become $A'^{\mu\nu}$ and $A'^{\nu\mu}$ in another system of coordinates, the symmetry will be maintained in this system also if $A'^{\mu\nu} = A'^{\nu\mu}$.

To show it, let us consider

$$A'^{\mu\nu} = \frac{\partial x'_\mu}{\partial x_\alpha}\frac{\partial x'_\nu}{\partial x_\beta}A^{\alpha\beta}$$
$$= \frac{\partial x'_\nu}{\partial x_\beta}\frac{\partial x'_\mu}{\partial x_\alpha}A^{\alpha\beta} \text{ (on interchanging the indices)}$$
$$= A'^{\nu\mu} \text{ as } A^{\beta\alpha} = A^{\alpha\beta}$$

which shows the symmetry in the other system also.

Similarly if we consider two covariant tensors, $A_{\mu\nu}$ and $A_{\nu\mu}$ such that $A_{\mu\nu} = A_{\nu\mu}$ in one system, then if they become $A'_{\mu\nu}$ and $A'_{\nu\mu}$ in another system, we have

NOTES

$$A'_{\mu\nu} = \frac{\partial x_\alpha}{\partial x'_\mu}\cdot\frac{\partial x_\beta}{\beta x'_\nu}A_{\alpha\beta}$$

$$= \frac{\partial x_\beta}{\partial x'_\nu}\cdot\frac{\partial x_\alpha}{\partial x'_\mu}A_{\beta\alpha} \text{ (on interchanging the indices)}$$

$$= A'_{\nu\mu} \text{ as } A_{\alpha\beta} = A_{\beta\alpha}.$$

In case one index is contravariant and other covariant, the symmetry cannot be easily defined. But it is notable that Kronecker delta which is a mixed tensor is symmetrical with respect to its indices.

When $A^{\mu\nu}$ is symmetrical, we have

$$A^{11} = A^{11}, A^{22} = A^{22} \text{ etc.}$$

In all, there are ${}^4C_2 + 4 = {}^5C_2$ components.

As an example, the components of the angular momentum of a rigid body B_μ are connected with the components of its angular velocity A_α by the relations $B_\mu = \sum_{\alpha=1}^{3} T_{\mu\alpha}A_\alpha$, where $T_{\mu\alpha}$ is the inertia tensor. This tensor is symmetric because

$$T_{\mu\alpha} = T_{\alpha\mu} \qquad \text{... (1.78)}$$

Now it has been already mentioned that a tensor can be expressed as a matrix and the columns and rows of a matrix when interchanged, the resulting tensor is the conjugate tensor. As such the conjugate of the tensor may be written as

$$\phi_c = \begin{bmatrix} a_{xx} & a_{yx} & a_{zx} \\ a_{xy} & a_{yy} & a_{zy} \\ a_{xz} & a_{yz} & a_{zz} \end{bmatrix}$$

Thus the tensor ϕ_c is the conjugate of ϕ, when

$$a_{xy} = a_{yx}, \quad a_{xz} = a_{zx} \quad \text{and} \quad a_{yz} = a_{zy} \qquad \text{... (1.79)}$$

If a single tensor satisfies this condition, it is called a **symmetric tensor**. It means that Condition (1.78) is essential for a tensor to be symmetric. It has only six independent elements in three dimensional space and may be written as

$$\phi_{sym} = \begin{bmatrix} a_{xx} & a_{xy} & a_{xz} \\ a_{yx} & a_{yy} & a_{yz} \\ a_{zx} & a_{zy} & a_{zz} \end{bmatrix}$$

The Relations (1.79) between the components of tensors follow that any symmetrical tensor corresponds with a transformation from the principal axes to another rectangular system of axes. To express the symmetrical linear functions by graphical function, let us suppose that we have two vectors **u** and **v** such that

$$\mathbf{v} = \phi_{sym} \,.\, \mathbf{u} \qquad \text{... (1.80)}$$

Project **u** upon **v** so that we calculate

$$\mathbf{u} \,.\, \mathbf{v} = u_x v_x + u_y v_y + u_z v_z, \qquad \text{... (1.81)}$$

where u_x, u_y, u_z and v_x, v_y, v_z are the components of **u** and **v** respectively along the principal axes.

NOTES

Multiplying both sides of Equation (1.80) scalarly by **u**, we get

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \phi_{sym} \cdot \mathbf{u}$$
$$= u_x v_x + u_y v_y + u_z v_z \text{ from Equation (1.81)}$$
$$= u_x(a_{xx}u_x + a_{xy}u_y + a_{xz}u_z) + u_y(a_{yx}u_x + a_{yy}u_y + a_{yz}u_z) + u_z(a_{zx}u_x + a_{zy}u_y + a_{zz}u_z)$$

Say $S = a_{xx}\, u_x^2 + a_{yy}u_y^2 + a_{zz}\, u_z^2 + 2(a_{xy}u_x\, u_y + a_{yz}\, u_y\, u_z + a_{zx}\, u_z\, u_x)$

where S is a *scalar point function.*

We get

$$\frac{\partial S}{\partial u_x} = 2\,[a_{xx}u_x + a_{xy}u_y + a_{xz}u_z] = 2v_x \text{ by Equation (1.69)}$$

or $$v_x = \frac{1}{2}\frac{\partial S}{\partial u_x}$$

Similarly, $$v_y = \frac{1}{2}\frac{\partial S}{\partial u_y} \text{ and } v_z = \frac{1}{2}\frac{\partial S}{\partial u_z}$$

So that $\mathbf{v} = v_x\, \mathbf{i} + v_y\, \mathbf{j} + v_z\, \mathbf{k}$; **i**, **j**, **k** being unit vectors along principal axes

$$= \frac{1}{2}\left[\frac{\partial S}{\partial u_x}\mathbf{i} + \frac{\partial S}{\partial u_y}\mathbf{j} + \frac{\partial S}{\partial u_z}\mathbf{k}\right]$$

$$= \frac{1}{2} \text{ grad } S.$$

which shows that **v** is a vector perpendicular to the surface S = const. in the direction of the outward normal. But S = const. is an equation of the second degree in the rectangular components of **u** regarding these as coordinates defining the extremity P of the vector **u** the locus of P is a conicoid.

As a particular case if $S = 1$, the surface defined under certain conditions is the *tensor ellipsoid* as shown in Figure 1.4.

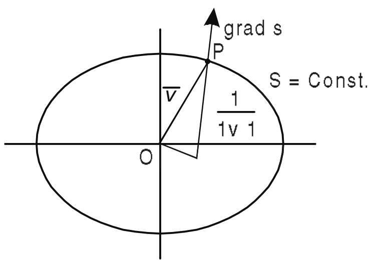

Fig. 1.4

Also $\mathbf{u} \cdot \mathbf{v} = S = 1$ = resolute of **u** in the direction of **v**.

In the direction of grade S, this resolute becomes $\frac{1}{|\mathbf{v}|}$.

Few other examples of symmetric tensor may be given as below

$$A_{\mu\alpha\beta} = A_{\alpha\mu\beta}$$

and $$A_{\mu\alpha\beta\gamma} = A_{\mu\beta\alpha\gamma} = A_{\alpha\mu\beta\gamma} = A_{\beta\alpha\mu\gamma} = A_{\alpha\beta\mu\gamma} = A_{\beta\mu\alpha\gamma}$$

Here the first tensor is symmetric in its first two indices and the second one is symmetric in first three indices.

If a tensor is such that two contravariant or covariant indices of it when interchanged, the components of the tensor alter in sign but not in magnitude, the tensor is said to be ***anti-symmetric*** *or* ***skew-symmetric****.*

NOTES

Hence if $A^{\mu\nu} = -A^{\nu\mu}$

then,
$$A'^{\mu\nu} = \frac{\partial x'_\mu}{\partial x_\alpha}\frac{\partial x'_\nu}{\partial x_\beta}A^{\alpha\beta}$$
$$= -\frac{\partial x'_\nu}{\partial x_\beta}\frac{\partial x'_\mu}{\partial x_\alpha}A^{\beta\alpha}$$
$$= -A'^{\nu\mu} \quad \text{as} \quad A^{\alpha\beta} = -A^{\beta\alpha}$$

Similarly if $A_{\mu\nu} = -A_{\nu\mu}$, then $A'_{\mu\nu} = -A'_{\nu\mu}$

Here $A^{12} = -A^{21}$ etc. and $A^{11} = 0 = A^{22} = A^{33}$

As such number of numerical components is $^4C_2$ only.

Evidently, the components of an antisymmetric tensor satisfy the relations

$$A_{\mu\nu} + A_{\nu\mu} = 0 \quad \text{or} \quad A_{\mu\nu} = -A_{\nu\mu} \qquad \text{... (1.82)}$$

which follows that the tensor changes its sign when indices are interchanged. If $\mu = \nu$, then Equation (1.82) yields

$$A_{\mu\mu} + A_{\mu\mu} = 0 \quad \text{or} \quad 2A_{\mu\mu} = 0 \quad \text{or} \quad A_{\mu\mu} = 0$$

So in terms of coefficients, $a_{xx} = a_{yy} = a_{zz} = 0$

and $a_{xy} = -a_{yx}$, $a_{xz} = -a_{zx}$, $a_{yz} = -a_{zy}$

give the conditions for a tensor to be anti- or skew-symmetric.

Thus ϕ will be antisymmetric if

$$\phi_{sk} = \begin{bmatrix} 0 & a_{xy} & a_{xz} \\ -a_{xy} & 0 & a_{yz} \\ -a_{xz} & -a_{yz} & 0 \end{bmatrix}$$

The matrix has only three components. The property of having only three components is possessed by vectors. This leads to the conclusion that an operation of ϕ_{sk} on the vector **u**, is exactly equivalent to the vector product of two vectors, since the final result is itself a vector **v**. For example, consider the product of the coordinates of two points

$$A_{\mu\nu} = x_\mu \xi_\nu - \xi_\mu x_\nu \qquad \text{... (1.83)}$$

In 3-dimensional space, suffixes can have only 3 values (1, 2, 3) and so any pair of suffixes can be replaced by a single one which is not present in the pair i.e.,

$$\left.\begin{aligned} A_{23} &= -A_{32} = A_1 \\ A_{12} &= -A_{21} = A_3 \\ A_{31} &= -A_{13} = A_2 \end{aligned}\right\} \qquad \text{... (1.84)}$$

Thus the anti-symmetric system may be replaced by

$$\xi_{\mu\nu\sigma} = +1 \text{ if } (\mu, \nu, \sigma) \text{ is an even permutation of } (1, 2, 3)$$
$$= -1 \text{ if } (\mu, \nu, \sigma) \text{ is an odd permutation of } (1, 2, 3)$$
$$= 0 \text{ if any of the suffixes are equal.}$$

NOTES

Now Equations (1.84) yield,

$$A_1 = \frac{1}{2}[2A_{23}],\ A_2 = \frac{1}{2}[2A_{13}],\ A_3 = \frac{1}{2}[2A_{12}]$$

i.e.,
$$A_1 = \frac{1}{2}\sum_{\mu,\nu=1}^{3} \xi_{\mu\nu\sigma} A_{\mu\nu} \qquad \text{... (1.85)}$$

There being only two terms in Sum (1.85).

Few other examples of antisymmetric tensors are

$$A_{\mu\nu\sigma} = -A_{\nu\mu\sigma}$$

and
$$A_{\mu\nu\sigma\rho} = -A_{\nu\sigma\mu\rho} = A_{\sigma\mu\nu\rho} = -A_{\mu\sigma\nu\rho} = -A_{\nu\mu\sigma\rho} = -A_{\sigma\nu\mu\rho}.$$

As an illustration, *if $A_{\mu\nu}$ is antisymmetric tensor of second order and B^{μ} is a tensor of rank one, then $A_{\mu\nu} B_{\mu} B_{\nu} = 0$, summation being taken over repeated indices.*

Intrerchange of dummy suffixes gives

$$A_{\mu\nu} B^{\mu} B^{\nu} = A_{\nu\mu} B^{\nu} B^{\mu} \qquad \text{... (1.86)}$$

where $A_{\mu\nu}$ being antisymmetric i.e.,

$$A_{\mu\nu} = -A_{\nu\mu}$$

renders

$$\begin{aligned} A_{\mu\nu} B^{\mu} B^{\nu} &= -A_{\nu\mu} B^{\mu} B^{\nu} \\ &= -A_{\nu\mu} B^{\nu} B^{\mu} \qquad \text{... (1.87)} \end{aligned}$$

The addition of Equation (1.86) and (1.88) yields

$$A_{\mu\nu} B^{\mu} B^{\nu} = 0 \qquad \text{... (1.88)}$$

1.9.2 Invariant Tensors

It is not known about any vector which has the same components in different systems of coordinates, but there exist tensors of higher ranks which have the same components in all the frames of reference. These tensors are called to have the *invariant components* or *invariant tensors* in general. One of the examples of such a tensor is Kronecker symbol defined as follows:

With respect to the old frame of the reference (i.e., before rotation)

$$\frac{\partial x_{\mu}}{\partial x_{\nu}} = \begin{cases} 0 \text{ if } \mu \neq \nu \\ 1 \text{ if } \mu = \nu \end{cases} \text{ since } x_{\mu} \text{ is independent of } x_{\nu}$$

But
$$\frac{\partial x_{\mu}}{\partial x'_{\nu}} = \frac{\partial x_{\mu}}{\partial x'_{\alpha}} \frac{\partial x'_{\alpha}}{\partial x_{\nu}}$$

Hence
$$\frac{\partial x_{\mu}}{\partial x_{\nu}} = \frac{\partial x_{\mu}}{\partial x'_{\nu}} \frac{\partial x'_{\alpha}}{\partial x_{\nu}} = \delta^{\mu}_{\nu}$$

where
$$\delta^{\mu}_{\nu} = \begin{cases} 1 \text{ if } \mu = \nu \\ 0 \text{ if } \mu \neq \nu \end{cases}$$

The symbol $\delta^{\mu}_{\nu} = \delta_{\mu\nu} = \delta^{\mu\nu} = \begin{cases} 1 \text{ if } \mu = \nu \\ 0 \text{ if } \mu \neq \nu \end{cases}$ is called as **Kronecker delta symbol**. In terms of new frame of reference (i.e., after the rotation,) we may write

$$\delta'^{\mu}_{\nu} = \frac{\partial x'_{\mu}}{\partial x'_{\nu}} = \frac{\partial x'_{\mu}}{\partial x_{\alpha}} \frac{\partial x'_{\alpha}}{\partial x'_{\nu}} = \begin{cases} 1 \text{ if } \mu = \nu \\ 0 \text{ if } \mu \neq \nu \end{cases}$$

NOTES

$$= \frac{\partial x'_\mu}{\partial x'_\beta} \cdot \frac{\partial x_\alpha}{\partial x'_\nu} \cdot \frac{\partial x_\beta}{\partial x_\alpha}$$

$$= \frac{\partial x'_\mu}{\partial x_\beta} \frac{\partial x_\alpha}{\partial x'_\nu} \delta^\beta_\alpha \text{ since } \delta^\beta_\alpha = \frac{\partial x_\beta}{\partial x_\alpha}$$

(in summation convention)

Hence δ^μ_ν is invariant and transforms as mixed tensor of rank two. Similarly $\delta_{\mu\nu}$ transforms as the components of covariant tensor of rank two while $\delta^{\mu\nu}$ transforms as a contravariant tensor of rank two.

Kronecker symbol can be used as *substitution multiplier*.

Since $$A_\mu = \frac{\partial x_\nu}{\partial x'_\alpha} \frac{\partial x'_\alpha}{\partial x_\mu} A^\nu \quad \text{(in summation convention)}$$

$$= \delta^\mu_\nu A^\nu \text{ (by substitution of index)}$$

Similarly $$A_\nu = \frac{\partial x'_\alpha}{\partial x_\nu} \frac{\partial x_\mu}{\partial x'_\alpha} A^\mu_\alpha$$

$$= \delta^\mu_\nu A_\mu$$

and $$\delta_{\mu\nu} \delta^{\nu\beta} = \frac{\partial x_\mu}{\partial x_\nu} \frac{\partial x_\nu}{\partial x_\beta} = \frac{\partial x_\mu}{\partial x_\beta} = \delta^\mu_\beta$$

while $$\delta^\mu_\mu = \delta^1_1 + \delta^2_2 + \delta^3_3 \text{ (summation convention)}$$

$$= 3 \text{ (in three-dimensional geometry)}$$

$$= 4 \text{ (in four-dimensional geometry).}$$

Secondly, we define the **generalized Kronecker delta symbol**, $\delta^{\mu\nu}_{\beta\gamma}$ by

$$\delta^{\mu\nu}_{\beta\gamma} = \begin{cases} 1 \text{ if } \mu = \beta,\ \nu = \gamma \text{ and } \mu \neq \nu \ \left(\text{as } \delta^{12}_{12} = 1\right) \\ -1 \text{ if } \mu = \gamma,\ \nu = \beta \text{ and } \mu \neq \nu \ \left(\text{as } \delta^{12}_{21} = -1\right) \\ 0 \text{ for all other combinations of indices.} \end{cases}$$

Similarly we can define $\delta^{\mu\nu\sigma}_{\beta\gamma\xi}$ as an absolute (invariant) tensor of rank six.

Conclusively, if both upper and lower indices of a gernalized delta consist of the same distinct numbers chosen from 1, 2, 3, the delta is 1 of –1 according as the upper indices form an even or odd permutation of the lower; in all other pemutations the delta is zero. We have, for example,

$$\delta^{12}_{12} = 1,\ \delta^{32}_{32} = -1,\ \delta^{23}_{11} = \delta^{13}_{21} = 0$$

$$\delta^{123}_{123} = \delta^{231}_{123} = 1,\ \delta^{213}_{123} = \delta^{321}_{123} = -1,\ \delta^{312}_{123} = 0.$$

Evaluation of the various possible combinations of indices shows that

$$\left.\begin{aligned} \delta^{\mu\beta}_{\beta\nu} &= -\delta^{\mu\nu}_{\nu\beta} = 2\,\delta^\mu_\beta \\ \delta^{\mu\nu}_{\beta\xi} &= \delta^\mu_\beta\,\delta^\nu_\xi - \delta^\mu_\xi\,\delta^\nu_\beta \end{aligned}\right\}$$

Alternating or Permutation Epsilon Tensor: This tensor is also an invariant component tensor of third rank and anti-symmetric in every pair of indices. Let $\varepsilon_{\mu\nu\sigma}$ be such a tensor, then

$$\varepsilon_{\mu\nu\sigma} = -\varepsilon_{\nu\mu\sigma} = \varepsilon_{\sigma\mu\nu} = -\varepsilon_{\mu\sigma\nu} = \varepsilon_{\nu\sigma\mu} = -\varepsilon_{\sigma\nu\mu}$$

But if $\mu = \nu$ then $\varepsilon_{\mu\nu\sigma} = -\varepsilon_{\nu\mu\sigma}$ gives $\varepsilon_{\mu\mu\sigma} = -\varepsilon_{\mu\mu\sigma}$

NOTES

or $\varepsilon_{\mu\mu\sigma} = 0$

It is clear that whenever two indices are equal, the component is zero. Moreover a tensor of third rank in three-dimensional geometry has 27 components. But in case of the tensor $\varepsilon_{\mu\nu\sigma}$ only 6 components are non-vanishing. All of them have the same absolute value, 3 being positive and the rest three are negative.

So

and
$$\left.\begin{array}{l}\varepsilon_{123} = \varepsilon_{xyz} = 1 = \varepsilon_{312} = \varepsilon_{231} \\ \varepsilon_{213} = \varepsilon_{321} = \varepsilon_{132} = -1\end{array}\right\}$$

while all other components are zero.

Thus all the even permutations of 1, 2 and 3 correspond to the components with value +1, while all the odd permutations correspond to –1.

The transformation law for this type of tensor in three dimensional space is given by

$$\varepsilon'_{\mu\nu\sigma} = \sum_{\alpha,\beta,\gamma=1}^{3} a_{\mu\alpha} a_{\nu\beta} a_{\sigma\gamma}\, \varepsilon_{\alpha\beta\gamma} = \varepsilon_{\mu\nu\sigma}$$

i.e., $\varepsilon_{\mu\nu\sigma}$ is invariant.

Consequently,

$$\varepsilon'_{\mu\nu\sigma} = \varepsilon_{\mu\nu\sigma}\begin{cases} 1 \text{ when } \mu,\nu,\sigma \text{ are an even permutation,} \\ -1 \text{ when } \mu,\nu,\sigma \text{ are an odd permutation,} \\ 0 \text{ when } \mu,\nu,\sigma \text{ contain two or more repeated indices.}\end{cases}$$

Similarly, the contravariant components can also be discussed,

$$\left.\begin{array}{l}\varepsilon^{123} = \varepsilon^{312} = \varepsilon^{231} = 1,\ \varepsilon^{132} = \varepsilon^{213} = \varepsilon^{321} = -1, \\ \varepsilon^{112} = \varepsilon^{233} = \varepsilon^{111} = \ldots = 0\end{array}\right\}$$

Pseudo Tensor: Let there be a tensor $\varepsilon_{\mu\sigma\tau\rho}$ of rank 4, defined such that

$$\varepsilon_{\mu\sigma\tau\rho} = \begin{cases} +1 \text{ if } \mu\sigma\tau\rho \text{ is an even permutation of } 0, 1, 2, 3 \\ -1 \text{ if } \mu\sigma\tau\rho \text{ is an odd permutation of } 0, 1, 2, 3 \\ 0 \text{ if two or more indices are equal}\end{cases}$$

These are termed as components of **pseudo tensor** of rank four.

In case ϕ is a scalar, the quantities $\phi\varepsilon_{\mu\sigma\tau\rho}$ are called as pseudo scalars since they have only one component.

NOTES

From every antisymmetric tensor $A_{\mu\sigma}$ of the second rank a pseudo tensor $A^{*}_{\mu\sigma}$ of the same rank can be obtained by multiplying the former with a pseudo-tensor of rank 4.

i.e., $$A^{*}_{\mu\sigma} = \frac{1}{2}\sum_{\alpha,\beta=0}^{3} \varepsilon^{\mu\sigma\tau\rho} A_{\alpha\beta}$$

Thus the product of a tensor with a pseudo-tensor is a pseudo-tensor. It is called dual of a given tensor.

A useful Property of ε Tensor

ε tensor can be used to write the cross-product of two vectors **A** and **B**.

Let $$\mathbf{D} = \mathbf{A} \times \mathbf{B},$$

then

$$D_1 = A_2B_3 - A_3B_2 = \varepsilon_{123}\, A_2B_3 + \varepsilon_{132}\, A_3B_2;$$

$A_1, A_2, A_3, B_1, B_2, B_3$ being components of **A** and **B**.

$$= \varepsilon_{1\nu\sigma}\, A_\nu B_\sigma$$

Similarly $$D_2 = \varepsilon_{2\nu\sigma}\, A_\nu B_\sigma \text{ (in summation convention)}$$

and $$D_3 = \varepsilon_{3\nu\sigma}\, A_\nu B_\sigma$$

or $$D_\mu = \varepsilon_{\mu\nu\sigma}\, A_\nu B_\sigma$$

Evaluating the various possible combinations we may have

$$\varepsilon^{\mu\nu\sigma}\varepsilon_{\sigma\alpha\beta} = \delta^{\mu\nu}_{\alpha\beta} = \delta^{\mu}_{\nu}\,\delta^{\nu}_{\beta} - \delta^{\mu}_{\beta}\,\delta^{\nu}_{\alpha}$$

Thus, if r = **C** . (**A** × **B**) = **C . D** for **D** = **A** × **B**.

then, $$r = C_\mu D^\mu \text{ (in summation convention)}$$

or $$r = C_\mu{}^{\mu\nu\sigma} A_\nu\, A_\sigma$$

$$= \varepsilon^{\mu\nu\sigma}\, C_\mu\, A_\nu\, B_\sigma$$

Similarly, vector triple product of three vectors can be given as

$$\mathbf{E} = \mathbf{C} \times (\mathbf{A} \times \mathbf{B}) = \mathbf{C} \times \mathbf{D}$$

i.e., $$E^\mu = \varepsilon^{\mu\nu\sigma}\, C_\nu\, D_\sigma$$

$$= \varepsilon^{\mu\nu\sigma}\, C_\nu\, (\varepsilon_{\alpha\alpha\beta}\, A^\alpha B^\beta)$$

$$= \varepsilon^{\mu\nu\sigma}\, \varepsilon_{\sigma\alpha\beta}\, C_\nu\, A^\alpha B^\beta$$

$$= \delta^{\mu\nu}_{\alpha\beta}\, C_\nu\, A^\alpha B^\beta = \left(\delta^{\mu}_{\alpha}\delta^{\nu}_{\beta} - \delta^{\mu}_{\beta}\delta^{\nu}_{\alpha}\right) C_\nu\, A^\alpha B^\beta$$

Since $\delta^{\mu}_{\alpha}\, A^\alpha = A^\mu$ etc., we get

$$E_\mu = \delta^{\nu}_{\beta}\, A^\mu C_\nu\, B^\beta - \delta^{\nu}_{\alpha}\, B^\mu\, C_\nu\, A^\alpha$$

$$= A^\mu\, (C_\beta B^\beta) - B^\mu\, (C_\alpha\, A^\alpha)$$

As such **E = A (C . B) – B (C . A).**

Evaluation of ∇ × (V × W) Using ε Tensor

Suppose, ∇ × (**V** × **W**)= ∇ × **Z**,

then $$(\nabla \times \mathbf{Z})_\mu = \varepsilon^{\mu\nu\sigma}\, \nabla_\nu\, Z_\sigma = \varepsilon^{\mu\nu\sigma}\, \nabla_\mu\, (\varepsilon_{\sigma\alpha\beta}\, V^\alpha\, W^\beta)$$

NOTES

$$\varepsilon^{\mu\nu\sigma}\,\varepsilon_{\sigma\alpha\beta}\,\nabla_\nu\,V^\alpha\,W^\beta = \delta^{\mu\nu}_{\alpha\beta}\,(V^\alpha\,\nabla_\nu\,W^\beta + W^\beta\,\nabla_\nu V^\alpha)$$

$$= \left(\delta^\mu_\nu\delta^\nu_\beta - \delta^\mu_\beta\delta^\nu_\alpha\right)(V^\alpha\,\nabla_\nu\,W^\beta + W^\beta\,\nabla_\nu\,V^\alpha)$$

$$= V^\mu\,\nabla_\beta\,W^\beta - V^\nu\,\nabla_\nu\,W^\mu + W^\nu\,\nabla_\nu\,V^\mu - W^\mu\,\nabla_\alpha\,V^\alpha$$

So $$\nabla \times \mathbf{Z} = \nabla \times (\mathbf{V} \times \mathbf{W})$$

$$= \mathbf{V}\,(\nabla\,.\,\mathbf{W}) - (\mathbf{V}\,.\,\nabla)\,\mathbf{W} + (\mathbf{W}\,.\,\nabla)\mathbf{V} - \mathbf{W}\,(\nabla\,.\,\mathbf{V})$$

Similarly all the vector relationships can be derived by using ε tensor.

Krutkov's Tensor: Let us consider a tensor $A^{\mu\gamma\beta\sigma}$ of fourth rank having following properties:

(1) Antisymmetric with respect to the first pair of indices

$$A^{\mu\gamma,\,\beta\alpha} = -A^{\gamma\mu,\,\beta\sigma}$$

(2) Antisymmetric in second pair of indices

$$A^{\mu\gamma,\,\beta\sigma} = -A^{\mu\gamma,\,\sigma\beta}$$

(3) Symmetric in cyclic order

$$A^{\mu\gamma,\,\beta\sigma} + A^{\mu\beta,\,\sigma\gamma} + A^{\mu\sigma,\,\gamma\beta} = 0$$

Then in terms of second derivatives of $A^{\mu\gamma,\beta\sigma}$ we can form a new tensor given by

$$B^{\mu\sigma} = \sum_{\gamma,\beta=0}^{3} \frac{\partial^2 A^{\mu\gamma,\beta\sigma}}{\partial x_\gamma\,\partial x_\beta} = 0 \qquad \text{... (1.89)}$$

This tensor is called as *Krutkov's tensor*. If we differentiate Equation (1.89) with respect to x, we have

$$\sum_{\sigma=0}^{3} \frac{\partial B^{\mu\sigma}}{\partial x_\sigma} = 0$$

This is an important property of Krutkov's tensor.

Problem 1.6. *Define a tensor. Prove that the Kronecker symbol* δ^k_i *is a tensor where components are the same in every coordinate system.*

We know that the tensors are quantities obeying certain transformation laws so that tensor may be regarded as an indispensable part of study which is rather suitable for the mathematical formulation of natural laws which remain invariant when one coordinate system is changed to another. The rank of a tensor measures the number of the mode of changes of a physical quantity when passing from one system to another which is in rotation relative to the first. As such tensor of zero rank is a scalar quantity and the tensor of rank one is a vector quantity.

The laws of transformation of vector being defined by

$$A'^\mu = \sum_{\alpha=1}^{4} \frac{\partial x'_\mu}{\partial x_\alpha} A^\alpha \quad \text{(contravariant vector)}$$

and $$A'_\mu = \sum_{\alpha=1}^{4} \frac{\partial x_\alpha}{\partial x'_\mu} A_\alpha \text{ (covariant vector)}$$

In Minkowski's four dimensional space, we define the tensors of rank two as follows:

NOTES

Contravariant tensor: $$A'^{\mu\nu} = \frac{\partial x'_\mu}{\partial x_\alpha}\frac{\partial x'_\nu}{\partial x_\beta} A^{\alpha\beta} \quad \ldots (1)$$

Covariant tensor: $$A'_{\mu\nu} = \frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial x_\beta}{\partial x'_\nu} A_{\alpha\beta} \quad \ldots (2)$$

Mixed tensor: $$A'^{\nu}_{\mu} = \frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial x'_\nu}{\partial x_\beta} A_{\alpha\beta} \quad \ldots (3)$$

Each one having 4^2, i.e., 16 components.

Similarly, we can define the tensors of higher ranks.

Now the Kronecker delta symbol δ_i^k is defined as

$$\delta_i^k = \frac{\partial x_k}{\partial x_i} = \frac{\partial x_k}{\partial x'_j}\frac{\partial x'_j}{\partial x_i}$$

which is easily deduced from Equation (3) by choosing A_α^β to be the Kronecker delta δ_α^β so that

$$A'^{\nu}_{\mu} = \frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial x'_\nu}{\partial x_\beta}\delta_\alpha^\beta = \frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial x'_\nu}{\partial x_\alpha}\frac{\partial x_\beta}{\partial x_\alpha} = \frac{\partial x'_\nu}{\partial x'_\mu} = \delta_\mu^\nu$$

and now replacing μ by i, ν by j this gives

$$\delta_i^k = A_i'^k \frac{\partial x_\alpha}{\partial x'_i}\frac{\partial x'_k}{\partial x_\beta} A_\alpha^\beta = \frac{\partial x_\alpha}{\partial x'_i}\frac{\partial x'_k}{\partial x_\beta}\delta_\alpha^\beta$$

From the definition of mixed tensor, it follows that δ_i^k is a mixed tensor of order two with 16 components in 4-dimensional space.

In order to show that the components of the tensor δ_i^k are the same in every coordinate system, let us define the symbol δ_i^k as

$$\delta_i^k = 1 \text{ if } i = k$$
$$= 0 \text{ if } i \neq k$$

which is evident from $$\delta_i^k = \frac{\partial x_k}{\partial x_i} = 0 \text{ when } i \neq k$$
$$= 1 \text{ when } i = k$$

In terms of new frame of reference (or new coordinate system), we may have

$$\delta_i'^k = \frac{\partial x'_k}{\partial x_i'} = \frac{\partial x'_k}{\partial x_j}\frac{\partial x_j}{\partial x'_i} = 1 \text{ if } i = k$$
$$= 0 \text{ if } i \neq k$$

or $$\delta_i'^k = \frac{\partial x'_k}{\partial x_l}\frac{\partial x_j}{\partial x'_i}\frac{\partial x_l}{\partial x_j} = \frac{\partial x'_k}{\partial x_l}\frac{\partial x_j}{\partial x'_i}\delta_j^l \text{ since } \delta_j^l = \frac{\partial x_l}{\partial x_j}$$

From which it is clear that δ_i^k is invariant and transforms as mixed tensor of rank two.

Problem 1.7. *Prove that Kronecker delta is a mixed tensor of rank two.*

NOTES

Its solution has been given in Problem 1.6.

Problem 1.8. *Show that symmetry properties of a tensor are invariant.*

If $A_{\lambda\mu\nu} = A_{\mu\lambda\nu}$ then we have to show that $A'_{\lambda\mu\nu} = A'_{\mu\lambda\nu}$

The definition follows:

$$A'_{\lambda\mu\nu} = \sum_{\alpha,\beta,\gamma=1}^{3} \frac{\partial x_\alpha}{\partial x'_\lambda}\frac{\partial x_\beta}{\partial x'_\mu}\frac{\partial x_\gamma}{\partial x'_\nu} A_{\alpha\beta\gamma}$$

and
$$A'_{\mu\lambda\nu} = \sum_{\beta,a,\gamma=1}^{3} \frac{\partial x_\beta}{\partial x'_\mu}\frac{\partial x_\alpha}{\partial x'_\lambda}\frac{\partial x_\gamma}{\partial x'_\nu} A_{\beta\alpha\gamma}$$

The given tensor being symmetrical in first two indices, we have

$$A_{\lambda\mu\nu} = A_{\mu\lambda\nu} \text{ and } A_{\alpha\beta\gamma} = A_{\beta\alpha\gamma}$$

Using this relation and comparing the two equations for $A'_{\lambda\mu\nu}$ and $A'_{\mu\lambda\nu}$ we find that both the equations are identical, i.e., $A'_{\lambda\mu\nu} = A'_{\mu\lambda\nu}$.

Which follows that the tensor in the other system is also symmetrical in first two indices. Hence the properties of symmetric tensors are invariant.

Rules Which Govern Tensor Analysis

Rule I. *The sum and difference of two tensors of the same rank result in a third tensor of the same rank. Moreover, if $F_{\lambda\mu}$... and $G_{\lambda\mu}$... are the tensors of the same rank, then $pF_{\lambda\mu}$ ... + $qG_{\lambda\mu}$ ... is also a tensor of the same rank* (*p*, *q being numbers*).

Suppose there are two tensors $A_{\lambda\mu}$ and $B_{\lambda\mu}$, then it will be shown that

$$A_{\lambda\mu} + B_{\lambda\mu} = C_{\lambda\mu}$$

is another tensor of the same rank.

Expressing the tensor $A_{\lambda\mu}$ in the form of a matrix, we have

$$A_{\lambda\mu} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad B_{\lambda\mu} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

and
$$C_{\lambda\mu} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

so that
$$A_{\lambda\mu} + B_{\lambda\mu} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \\ a_{31}+b_{31} & a_{32}+b_{32} & a_{33}+b_{33} \end{bmatrix}$$

If the relations between the coefficient *a*'s and *b*'s be such that

$$a_{\lambda\mu} + b_{\lambda\mu} = c_{\lambda\mu}$$

NOTES

then $$A_{\lambda\mu} + B_{\lambda\mu} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = C_{\lambda\mu}$$

which is a tensor of the same rank.

Similarly $$A_{\lambda\mu} - B_{\lambda\mu} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = D_{\lambda\mu}$$

where $a_{\lambda\mu} - b_{\lambda\mu} = d_{\lambda\mu}$.

Here $D_{\lambda\mu}$ is again a tensor of the same rank.

Further,

$$pA_{\lambda\mu} + qB_{\lambda\mu} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} = E_{\lambda\mu}$$

where $pa_{\lambda\mu} + qb_{\lambda\mu} = e_{\lambda\mu}$.
showing that $E_{\lambda\mu}$ is also a tensor of the same rank.

The rule of addition may be generalized for any number of tensors of any rank.

Suppose there are two mixed tensors T and S of rank N, having their r indices (from λ_1 to λ_r) contravariant and s indices (from μ_1 to μ_s) covariant, then laws of their transformation may be written as

$$T'^{\alpha_1\ldots\alpha_r}_{\beta_1\ldots\beta_s} = \left|\frac{\partial x}{\partial x'}\right|^N \left[\frac{\partial x_{\mu_1}}{\partial x'_{\beta_1}}\cdots\frac{\partial x_{\mu_s}}{\partial x'_{\beta_s}}\right]\left[\frac{\partial x'_{\alpha_1}}{\partial x_{\lambda_1}}\cdots\frac{\partial x'_{\alpha_r}}{\partial x_{\lambda_r}}\right] T^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}$$

$$S'^{\alpha_1\ldots\alpha_r}_{\beta_1\ldots\beta_s} = \left|\frac{\partial x}{\partial x'}\right|^N \left[\frac{\partial x_{\mu_1}}{\partial x'_{\beta_1}}\cdots\frac{\partial x_{\mu_s}}{\partial x'_{\beta_s}}\right]\left[\frac{\partial x'_{\alpha_1}}{\partial x_{\lambda_1}}\cdots\frac{\partial x'_{\alpha_r}}{\partial x_{\lambda_r}}\right] S^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}$$

If the sum of two tensors $T^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}$ and $S^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}$ be a third tensor

$U'^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}$, i.e., $U^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}$, $= T^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s} + S^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}$

Then, $$U'^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s} = T'^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s} + S'^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}$$

$$= \left|\frac{\partial x}{\partial x'}\right|^N \left[\frac{\partial x_{\mu_1}}{\partial x'_{\beta_1}}\cdots\frac{\partial x_{\mu_s}}{\partial x'_{\beta_s}}\right]\left[\frac{\partial x'_{\alpha_1}}{\partial x_{\lambda_1}}\cdots\frac{\partial x'_{\alpha_r}}{\partial x'_{\lambda_r}}\right]\left[T^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s} + S^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}\right]$$

$$= \left|\frac{\partial x}{\partial x'}\right|^N \frac{\partial x_{\mu_1}}{\partial x'_{\beta_1}}\cdots\frac{\partial x_{\mu_s}}{\partial x'_{\beta_s}}\,\frac{\partial x'_{\alpha_1}}{\partial x_{\lambda_1}}\cdots\frac{\partial x'_{\alpha_r}}{\partial x'_{\lambda_r}}\; U^{\lambda_1\ldots\lambda_r}_{\mu_1\ldots\mu_s}$$

Which is transformation equation for a tensor of rank N having r contravariant and s covariant indices and follows that the sum of two tensors of the same rank is a new tensor of the same rank.

Note. Here $\left|\frac{\partial x}{\partial x'}\right|$ is the *Jacobian of transformation* and the tensor $T^{\lambda_1 \ldots \lambda_r}_{\mu_1 \ldots \mu_s}$ is known as *Relative tensor of weight W*. For $W = 0$, the relative tensor becomes *Absolute tensor*, whereas for $W = 1$, the relative tensor is known as *Tensor density*.

NOTES

Rule II. *The direct product of two tensors gives a new tensor of rank equal to the sum of ranks of these tensors.*

Consider two tensors $T^{\lambda_1 \ldots \lambda_r}_{\mu_1 \ldots \mu_s}$ of rank N, weight W and $S^{\sigma_1 \ldots \sigma_\rho}_{\rho_1 \ldots \mu_q}$ of rank N' weight W'.

Their transformation may be given as,

$$T'^{\alpha_1 \ldots \alpha_r}_{\beta_1 \ldots \beta_s} = \left|\frac{\partial x}{\partial x'}\right|^N \frac{\partial x_{\mu_1}}{\partial x'_{\beta_1}} \cdots \frac{\partial x_{\mu_s}}{\partial x'_{\beta_s}} \frac{\partial x'_{\alpha_1}}{\partial x_{\lambda_1}} \cdots \frac{\partial x'_{\alpha_r}}{\partial x'_{\lambda_r}} T^{\lambda_1 \ldots \lambda_r}_{\mu_1 \ldots \mu_s}$$

$$S'^{\eta_1 \ldots \eta_r}_{\xi_1 \ldots \xi_s} = \left|\frac{\partial x}{\partial x'}\right|^{N'} \frac{\partial x_{\rho_1}}{\partial x'_{\xi_1}} \cdots \frac{\partial x_{\rho_q}}{\partial x'_{\xi_q}} \frac{\partial x'_{\eta_1}}{\partial x_{\sigma_1}} \cdots \frac{\partial x'_{\eta_p}}{\partial x'_{\sigma_p}} S^{\sigma_1 \ldots \sigma_\rho}_{\rho_1 \ldots \mu_q}$$

Then $$T'^{\alpha_1 \ldots \alpha_r}_{\beta_1 \ldots \beta_s} + S'^{\eta_1 \ldots \eta_\rho}_{\xi_1 \ldots \xi_q} = \left|\frac{\partial x}{\partial x'}\right|^{N+N'} \frac{\partial x_{\mu_1}}{\partial x'_{\beta_1}} \cdots \frac{\partial x_{\mu_s}}{\partial x'_{\beta_s}} \frac{\partial x'_{\alpha_1}}{\partial x_{\lambda_1}} \cdots \frac{\partial x'_{\alpha_r}}{\partial x_{\lambda_r}} \frac{\partial x_{\rho_1}}{\partial x'_{\xi_1}} \cdots \frac{\partial x'_{\rho_q}}{\partial x'\xi_q}$$

$$\frac{\partial x'_{\eta_1}}{\partial x'_{\sigma_1}} \cdots \frac{\partial x'_{\eta_p}}{\partial x_{\sigma_p}} \times \left[T^{\lambda_1 \ldots \lambda_r}_{\mu_1 \ldots \mu_s}\right]\left[S^{\sigma_1 \ldots \sigma_\rho}_{\rho_1 \ldots \rho_q}\right]$$

or $$U'^{\alpha_1 \ldots \alpha_s, \ldots \eta_1 \ldots \eta_p}_{\beta_1 \ldots \beta_s, \xi_1 \ldots \xi_q} = \left|\frac{\partial x}{\partial x'}\right|^{N+N'} \frac{\partial x_{\mu_1}}{\partial x'_{\beta_1}} \cdots \frac{\partial x_{\mu_s}}{\partial x'_{\beta_s}} \cdots \frac{\partial x'_{\alpha_1}}{\partial x_{\lambda_1}} \cdots\cdots \frac{\partial x_{\alpha_r}}{\partial x_{\lambda_r}}$$

$$\times \frac{\partial x_{\rho_1}}{\partial x'_{\xi_1}} \cdots \frac{\partial x_{\rho_q}}{\partial x'_{\xi_q}} \frac{\partial x'_{\eta_1}}{\partial x_{\sigma_1}} \cdots \frac{\partial x'_{\eta_p}}{\partial x_{\sigma_p}} \times U^{\lambda_1 \ldots \lambda_r, \sigma_1 \ldots \sigma_p}_{\mu_1 \ldots \mu_s, \rho_1 \ldots \rho_q} \quad \ldots (1.90)$$

where $$U^{\lambda_1 \ldots \lambda_r, \sigma_1 \ldots \sigma_p}_{\mu_1 \ldots \mu_s, \rho_1 \ldots \rho_q} = T^{\lambda_1 \ldots \lambda_r}_{\mu_1 \ldots \mu_s} S^{\sigma_1 \ldots \sigma_p}_{\rho_1 \ldots \rho_q}$$

The Equation (1.90) transforms a tensor of rank $N + N'$ and weight $W + W'$.

Note. This rule may also be stated as:

The outer product of two relative tensors is itself a relative tensor of rank and weight equal to the sum of the ranks and the sum of weights of the given relative tensors respectively.

Rule III. Contraction: *The algebraic operation by which the rank of a tensor may be lowered by 2 (or by any even number) is known as contraction.*

The contraction of a tensor may be affected by adding up all the components which have equal indices in a given pair. Any two indices are converted into a pair of dummy indices.

Consider a tensor of rank 3 with one contravariant index α and two covariant indices β and γ. Then we have

$$A'^{\lambda}_{\mu\nu} = \sum_{\alpha, \beta, \gamma = 0}^{3} A^{\alpha}_{\beta\lambda} \frac{\partial x_\beta}{\partial x'_\mu} \cdot \frac{\partial x_\gamma}{\partial x'_\nu} \cdot \frac{\partial x'_\lambda}{\partial x_\alpha}$$

NOTES

Replacing ν by λ, we have

$$A'^{\lambda}_{\mu\nu} = \sum_{\lambda,\mu,\nu=0}^{3} A^{\alpha}_{\beta\gamma} \frac{\partial x_\beta}{\partial x'_\mu}\cdot\frac{\partial x_\gamma}{\partial x'_\nu}\cdot\frac{\partial x'_\lambda}{\partial x_\alpha}$$

$$= \sum_{\alpha,\beta,\gamma=0}^{3} A^{\alpha}_{\beta\gamma} \frac{\partial x_\beta}{\partial x'_\mu}\cdot\frac{\partial x_\gamma}{\partial x_\alpha}$$

But $$\frac{\partial x_\gamma}{\partial x_\alpha} = \begin{Bmatrix} 0 \text{ if } \gamma \neq \alpha \\ 1 \text{ if } \gamma = \alpha \end{Bmatrix}$$

Choosing the second condition, i.e., if $\gamma = \alpha$, $\frac{\partial x_\gamma}{\partial x_\alpha} = 1$, above relation becomes

$$A'^{\lambda}_{\mu\lambda} = \sum_{\alpha,\beta=0}^{3} A^{\alpha}_{\beta\alpha} \frac{\partial x_\beta}{\partial x'_\mu}$$

i.e., $$A'_\mu = \sum_{\beta=0}^{3} \frac{\partial x_\beta}{\partial x'_\mu} A_\beta$$

Which denotes the law of transformation of tensors of rank one, i.e., vectors. As a general rule we equate a certain covariant index to a contravariant index, sum repeated indices, and obtain a new tensor of lower rank. This process is termed as **contraction**. The contraction of a tensor of rank 2 yields a scalar, i.e., tensor of rank zero.

Illustration: We know that the scalar product of two vectors is a scalar quantity. It follows that the scalar product of two tensors of rank one is a tensor of rank zero. As such the rank is lowered by two.

Rule IV. Extension of the Rank: *The differentiation of each component of a tensor of rank n with respect to x, y, z, gives a new tensor of rank* (*n* + *1*), *for example,*

$$\frac{\partial A_{\lambda\mu}}{\partial x_\nu} = B_{\lambda\mu\nu}$$

This rule may be proved for a simple case, where the original tensor is of rank zero, i.e., a scalar say $S(x_1, x_2, x_3, t)$ where derivatives relative to the axes K are

$$\frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, \frac{\partial S}{\partial z}.$$

In system K', the scalar is $S'(x'_1, x'_2, x'_3, t')$

$$S(x_1, x_2, x_3, t) = S'(x'_1, x'_2, x'_3, t')$$

So, $$\left.\begin{aligned} \frac{\partial S'}{\partial x'_1} &= \frac{\partial S}{\partial x_1}\frac{\partial x_1}{\partial x'_1} + \frac{\partial S}{\partial x_2}\frac{\partial x_2}{\partial x'_1} + \frac{\partial S}{\partial x_3}\frac{\partial x_{3.}}{\partial x'_1} \\ \frac{\partial S'}{\partial x'_2} &= \frac{\partial S}{\partial x_1}\frac{\partial x_1}{\partial x'_2} + \frac{\partial S}{\partial x_2}\frac{\partial x_2}{\partial x'_2} + \frac{\partial S}{\partial x_3}\frac{\partial x_3}{\partial x'_2} \\ \frac{\partial S'}{\partial x'_3} &= \frac{\partial S}{\partial x_1}\frac{\partial x_1}{\partial x'_3} + \frac{\partial S}{\partial x_2}\frac{\partial x_2}{\partial x'_3} + \frac{\partial S}{\partial x_3}\frac{\partial x_3}{\partial x'_3} \end{aligned}\right\}$$

which may be written as a single equation

$$\frac{\partial S'}{\partial x'_l} = \frac{\partial x_m}{\partial x'_l}\frac{\partial S}{\partial x_m} \quad (l, m = 1, 2, 3)$$

NOTES

This shows that $\frac{\partial S}{\partial x_m}$ transforms like the components of a tensor of rank one, i.e., vector. Thus the differentiation of the tensor of rank zero, yields a tensor of rank one. The rank of a tensor can also be extended when a tensor depends on another tensor and a differentiation is performed. For example, consider a scalar, i.e., tensor of rank zero, say S depending on tensor $A_{\lambda\mu}$, so that

$$\frac{\partial S}{\partial A_{\lambda\mu}} = B_{\mu\nu}$$

is also a tensor of rank two. Thus the rank of the tensor is extended by two.

Rule V. The Quotient Law: *If $A^{\lambda}B_{\mu\nu}$ is a tensor for all contravariant tensors A_{λ} then $B_{\mu\nu}$ is also a tensor.*

We have
$$A'^{\lambda} B'_{\mu\nu} = A^{\alpha} B_{\beta\gamma} \frac{\partial \alpha_{\beta}}{\partial x'_{\mu}} \frac{\partial x_{\gamma}}{\partial x'_{\nu}} \frac{\partial x'_{\lambda}}{\partial x_{\alpha}}$$
$$= A'^{\lambda} B_{\beta\gamma} \frac{\partial x_{\beta}}{\partial x'_{\mu}} \frac{\partial x_{\gamma}}{\partial x'_{\nu}}$$

as
$$A'^{\lambda} = A^{\alpha} \frac{\partial x'_{\lambda}}{\partial x_{\alpha}}$$

or $A'^{\lambda}\left[B'_{\mu\nu} - B_{\mu\nu}\frac{\partial x_{\alpha}}{\partial x'_{\mu}}\frac{\partial x_{\gamma}}{\partial x'_{\nu}}\right] = 0$

But A'^{λ} being arbitrary, $A'^{\lambda} \neq 0$ so that

$$B'_{\mu\nu} = B_{\mu\nu} \frac{\partial x_{\beta}}{\partial x'_{\mu}} \frac{\partial x_{\gamma}}{\partial x'_{\nu}}$$

which shows that $B_{\mu\nu}$ is a covariant tensor.

Note. If A be a symmetric convariant tensor of second order s.t. $|A_{\mu\nu}| = A \neq 0$, and we set,

$$A^{\mu\nu} = \frac{\text{Cofactor of } A_{\mu\nu} \text{ in } A}{A} = \frac{a_{\mu\nu}}{A} \qquad \text{... (1.91)}$$

$a_{\mu\nu}$ being cofactor of $A_{\mu\nu}$ in A, and $A_{\mu\nu}$ also being symmetric, then A and so $a_{\mu\nu}$ is symmetric. Consequently $A^{\mu\nu}$ is symmetric.

Also if B^{μ} be an arbitrary vector, then quotient law gives

$$B_{\nu} = A_{\mu\nu} B^{\mu}$$

as a convariant vector

$$\therefore \quad B_{\nu} A^{\nu\sigma} = A_{\mu\nu} B^{\mu} A^{\mu\sigma} = A_{\mu\nu} B^{\mu} \frac{a_{\nu\sigma}}{A}$$

$$= \frac{A_{\mu\nu}\, a_{\nu\sigma}}{A} B\mu = \delta^{\sigma}_{\mu} B^{\mu} \text{ by determinant theory.}$$

or $$B_\nu A^{\nu\sigma} = B\sigma$$

Here Relation (1.91) ⇒ *symmetric contravariant tensor of rank 2*, known as **conjugate** or **reciprocal tensor** of $A_{\mu\nu}$.

Problem 1.9. *Show that there exists no distinction between contravariant and covariant vectors if we restrict ourselves to trasformation of the type*

$$x'_\alpha = a^\alpha_\lambda \; x_\lambda + b^\alpha$$

where b^α are n constant which do not necessarily form the components of a contravariant vector and a^α_λ are constant (not necessarily forming a tensor) such that

$$a^\alpha_\mu \; a^\alpha_\lambda = \delta^\mu_\lambda$$

Given $$x'_\alpha = a^\alpha_\lambda \; x_\lambda + b^\alpha \qquad \text{... (1)}$$

i.e., $$a^\alpha_\lambda \; x_\lambda = x'_\alpha - b^\alpha \qquad \text{... (2)}$$

Multiplying Equation (2) throughout by a^α_μ and summing over the index α from 1 to n, we find

$$x_\mu = a^\alpha_\mu \; x'_\alpha - a^\alpha_\mu \, b^\alpha \qquad \text{... (3)}$$

Now Equations (1) and (3) yield,

$$\frac{\partial x'_\alpha}{\partial x_\beta} = a^\alpha_\beta \text{ and } \frac{\partial x_\beta}{\partial x'_\alpha} = a^\alpha_\beta$$

So that $$\frac{\partial x'_\alpha}{\partial x_\beta} = \frac{\partial x_\beta}{\partial x'_a} = a^\alpha_\beta$$

This follows that the tranformation laws

$$A'^\mu = \frac{\partial x'_\mu}{\partial x_\beta} A^\beta \text{ and } A'_\mu = \frac{\partial x_\alpha}{\partial x'_\mu} \; A_\alpha$$

define the same type of entity without any distinction between contravariant and covariant vectors.

1.10 CONTRAVARIANT, COVARIANT AND MIXED TENSORS

The covariant derivatives of a tensor of rank two are formed as:

$$A_{\mu\nu,\sigma} = \frac{\partial A_{\mu\nu}}{\partial x_\sigma} - \Gamma^\alpha_{\mu\sigma} \, A_{\alpha\nu} - \Gamma^\alpha_{\nu\sigma} \, A_{\mu\alpha} \qquad \text{... (1.92)}$$

$$A^\nu_{\mu\sigma} \text{ (or } A^\nu_{\mu,\sigma}) = \frac{\partial A^\nu_\mu}{\partial x_\sigma} - \Gamma^\alpha_{\mu\sigma} \, A^\nu_\alpha + \Gamma^\nu_{\alpha\sigma} \, A^\alpha_\mu \qquad \text{... (1.93)}$$

$$\text{A } A^{\mu\nu}_\sigma \text{ (or } A^{\mu}_{\sigma},{}^\nu) = \frac{\partial A^{\mu\nu}}{\partial x_\sigma} + \Gamma^\mu_{\alpha\sigma} \, A^{\alpha\nu} + \Gamma^\nu_{\alpha\sigma} \, A^{\mu\alpha} \qquad \text{... (1.94)}$$

and a general rule for giving covariant differentiation with respect to x_σ may be illustrated as follows:

$$A^{\rho}_{\lambda\mu\nu\sigma} = \frac{\partial}{\partial x_\sigma} A^{\rho}_{\lambda\mu\nu} - \Gamma^{\alpha}_{\lambda\sigma} A^{\rho}_{\sigma\mu\nu} - \Gamma^{\alpha}_{\mu\sigma} A^{\rho}_{\lambda\alpha\nu} - \Gamma^{\alpha}_{\nu\sigma} A^{\rho}_{\lambda\mu\alpha} + \Gamma^{\rho}_{\alpha\sigma} A^{\alpha}_{\lambda\mu\nu} \quad \text{... (1.95)}$$

NOTES

In order to show that the quantities on the right are actually tensors, we proceed as follows:

$$\frac{d}{ds}\left(A_{\mu\nu}\frac{dx_\mu}{ds}\frac{dx_\nu}{ds}\right) \text{ is invariant}$$

i.e., $$\frac{dA_{\mu\nu}}{dx_\sigma}\frac{dx_\sigma}{ds}\frac{dx_\mu}{ds}\frac{dx_\nu}{ds} + A_{\mu\nu}\frac{dx_\nu}{ds}\frac{d^2x_\mu}{ds^2} + A_{\mu\nu}\frac{dx_\mu}{ds}\frac{d^2x_\nu}{ds^2}$$

is invariant along a geodesic.

But along a geodesic, we have

$$A_{\mu\nu}\frac{d^2x_\mu}{ds^2} = A_{\alpha\nu}\frac{d^2x_\alpha}{ds^2} = -A_{\alpha\nu}\Gamma^{\alpha}_{\mu\sigma}\frac{dx_\mu}{ds}\frac{dx_\sigma}{ds}$$

and $$A_{\mu\nu}\frac{d^2x_\nu}{ds^2} = A_{\mu\alpha}\frac{d^2x_\alpha}{ds^2} = -A_{\mu\alpha}\Gamma^{\alpha}_{\nu\sigma}\frac{dx_\nu}{ds}\frac{dx_\sigma}{ds}.$$

$$\therefore \quad \left(\frac{\partial A_{\mu\nu}}{\partial x_\sigma} - \Gamma^{\alpha}_{\mu\sigma}A_{\alpha\nu} - \Gamma^{\alpha}_{\nu\sigma}A_{\mu\alpha}\right)\frac{dx_\mu}{ds}\frac{dx_\nu}{ds}\frac{dx_\sigma}{ds} \text{ is invariant,}$$

i.e., $$A_{\mu\nu,\sigma}\frac{dx_\mu}{ds}\frac{dx_\nu}{ds}\frac{dx_\sigma}{ds} \text{ is invariant}$$

which shows that $A_{\mu\nu\sigma}$ is a covariant tensor of rank three.

The results (1.93) and (1.94) may be obtained by raising the suffixes ν and μ as follows:

Since $A_{\mu\gamma} = g_{\gamma\varepsilon} A^{\varepsilon}_{\mu}$, therefore,

$$\begin{aligned}
A_{\mu\gamma,\sigma} &= \frac{\partial A_{\mu\nu}}{\partial x_\sigma} - \Gamma^{\alpha}_{\mu\sigma} A_{\alpha\gamma} - \Gamma^{\alpha}_{\gamma\sigma} A_{\mu\alpha} \\
&= \frac{\partial}{\partial x_\sigma}(g_{\gamma\varepsilon} A^{\varepsilon}_{\mu}) - \Gamma^{\varepsilon}_{\mu\sigma} A_{\varepsilon\gamma} - \Gamma^{\alpha}_{\gamma\sigma} g_{\alpha\varepsilon} A^{\varepsilon}_{\mu} \\
&= g_{\gamma\varepsilon}\frac{\partial A^{\varepsilon}_{\mu}}{\partial x_\sigma} + A^{\varepsilon}_{\mu}\frac{\partial g_{\gamma\varepsilon}}{\partial x_\sigma} - \Gamma^{\varepsilon}_{\mu\sigma} A_{\varepsilon\gamma} - \Gamma_{\gamma\sigma,\varepsilon} A^{\varepsilon}_{\mu} \\
&= g_{\gamma\varepsilon}\frac{\partial A^{\varepsilon}_{\mu}}{\partial x_\sigma} - \Gamma^{\varepsilon}_{\mu\sigma} A_{\varepsilon\gamma} + \left(\frac{\partial g_{\gamma\varepsilon}}{\partial x_\sigma} - \Gamma_{\gamma\sigma,\varepsilon}\right) A^{\varepsilon}_{\mu} \\
&= g_{\gamma\varepsilon}\frac{\partial A^{\varepsilon}_{\mu}}{\partial x_\sigma} - \Gamma^{\varepsilon}_{\mu\sigma} A_{\varepsilon\gamma} + \Gamma_{\varepsilon\sigma,\gamma} A^{\varepsilon}_{\mu}
\end{aligned}$$

Multiplying throughout by $g^{\nu\gamma}$, this becomes

$$g^{\nu\gamma} A_{\mu\gamma,\sigma} = g^{\nu\gamma} g_{\gamma\varepsilon}\frac{\partial A^{\varepsilon}_{\mu}}{\partial x_\sigma} - g^{\nu\gamma}\Gamma^{\varepsilon}_{\mu\sigma} A_{\varepsilon\gamma} + g^{\nu\gamma}\Gamma_{\varepsilon\sigma,\gamma} A^{\varepsilon}_{\mu}$$

i.e., $$A^{\nu}_{\mu,\sigma} = \frac{\partial A^{\nu}_{\mu}}{\partial x_\sigma} - \Gamma^{\varepsilon}_{\mu\sigma} g^{\nu\gamma} A_{\varepsilon\gamma} + \Gamma^{\nu}_{\varepsilon\sigma} A^{\varepsilon}_{\mu}$$

$$= \frac{\partial A_\mu^\nu}{\partial A_\sigma} - \Gamma_{\mu\sigma}^{\varepsilon} A_\varepsilon^\nu + \Gamma_{\varepsilon\sigma}^{\nu} A_\mu^\varepsilon$$

which is the result (1.93).

NOTES

Again, since $A_\gamma^\nu = g_{\gamma\varepsilon} A^{\varepsilon\nu}$, we have

$$A_{\gamma,\sigma}^\nu = \frac{\partial A_\gamma^\nu}{\partial x_\sigma} - \Gamma_{\gamma\sigma}^{\alpha} A_\alpha^\nu + \Gamma_{\alpha\sigma}^{\nu} A_\gamma^\alpha$$

$$= \frac{\partial}{\partial x_\sigma}(g_{\gamma\varepsilon} A^{\varepsilon\nu}) - \Gamma_{\gamma\sigma}^{\alpha} g_{\alpha\varepsilon} A^{\varepsilon\nu} + \Gamma_{\alpha\sigma}^{\nu} A_\gamma^\alpha$$

$$= g_{\gamma\varepsilon} \frac{\partial A^{\varepsilon\nu}}{\partial x_\sigma} + A^{\varepsilon\nu} \frac{\partial g_{\gamma\varepsilon}}{\partial x_\sigma} - \Gamma_{\gamma\sigma,\varepsilon} A^{\varepsilon\nu} + \Gamma_{\alpha\sigma}^{\nu} A_\gamma^\alpha$$

or $$A_{\gamma,\sigma}^\nu = g_{\gamma\varepsilon} \frac{\partial A^{\varepsilon\nu}}{\partial x_\sigma} + A^{\varepsilon\nu} \left(\frac{\partial g_{\gamma\varepsilon}}{\partial x_\sigma} - \Gamma_{\gamma\sigma,\varepsilon}\right) + \Gamma_{\alpha\sigma}^{\nu} A_\gamma^\alpha$$

$$= g_{\gamma\varepsilon} \frac{\partial A^{\varepsilon\nu}}{\partial x_\sigma} + A^{\varepsilon\nu} + \Gamma_{\varepsilon\sigma,\gamma} + \Gamma_{\alpha\sigma}^{\nu} A_\gamma^\alpha$$

Multiplying throughout by $g^{\mu\nu}$, this becomes

$$g^{\mu\gamma} A_{\gamma,\sigma}^\nu = g^{\mu\gamma} g_{\gamma\varepsilon} \frac{\partial A^{\varepsilon\nu}}{\partial x_\sigma} + g^{\mu\gamma} A^{\varepsilon\nu} \Gamma_{\varepsilon\sigma,\gamma} + g^{\mu\gamma} \Gamma_{\alpha\sigma}^{\nu} A_\gamma^\alpha$$

i.e., $$A_{\sigma}^{\mu,\nu} = \frac{\partial A^{\mu\nu}}{\partial x_\sigma} + \Gamma_{\varepsilon\sigma}^{\mu} A^{\varepsilon\nu} + \Gamma_{\alpha\sigma}^{\nu} A^{\mu\alpha}$$

$$= \frac{\partial A^{\mu\nu}}{\partial x_\sigma} + \Gamma_{\varepsilon\alpha}^{\mu} A^{\alpha\nu} + \Gamma_{\alpha\sigma}^{\nu} A^{\mu\alpha}$$

which is the result (1.94).

Problem 1.10. *Prove that $g_{\alpha\beta,\gamma} = 0$, i.e., covariant derivative of fundamental tensor vanishes identically.* (*Agra, 1963, 65, 68*)

$$g_{\alpha\beta,\gamma} = \frac{\partial g_{\alpha\beta}}{\partial x_\gamma} - \Gamma_{\alpha\gamma}^{\mu} g_{\mu\beta} - \Gamma_{\beta\gamma}^{\nu} g_{\mu\alpha} \qquad \text{... (1)}$$

But $$\Gamma_{\alpha\gamma}^{\mu} = \frac{1}{2} g^{\mu\beta} \left(\frac{\partial g_{\alpha\beta}}{\partial x_\gamma} + \frac{\partial g_{\beta\gamma}}{\partial x_\alpha} - \frac{\partial g_{\alpha\gamma}}{\partial x_\beta}\right)$$

so that $$g_{\mu\beta} \Gamma_{\alpha\nu}^{\mu} = \frac{1}{2} g_{\mu\beta} g^{\mu\beta} \left(\frac{\partial g_{\alpha\beta}}{\partial x_\gamma} + \frac{\partial g_{\beta\gamma}}{\partial x_\alpha} - \frac{\partial g_{\alpha\gamma}}{\partial x_\beta}\right)$$

$$= \frac{1}{2} \left(\frac{\partial g_{\alpha\beta}}{\partial x_\gamma} + \frac{\partial g_{\beta\gamma}}{\partial x_\gamma} - \frac{\partial g_{\alpha\gamma}}{\partial x_\beta}\right) \text{ since } g_{\mu\beta} g^{\mu\beta} = 1$$

Similarly $$g_{\mu\alpha} \Gamma_{\beta\gamma}^{\mu} = \frac{1}{2} \left(\frac{\partial g_{\beta\alpha}}{\partial x_\gamma} + \frac{\partial g_{\alpha\gamma}}{\partial x_\beta} - \frac{\partial g_{\beta\gamma}}{\partial x_\alpha}\right).$$

NOTES

Adding the last two relations, we get

$$g_{\mu\beta}\ \Gamma^{\mu}_{\alpha\gamma} + g_{\mu\alpha}\ \Gamma^{\mu}_{\beta\gamma} = \frac{\partial g_{\alpha\beta}}{\partial x_{\gamma}}.$$

Substituting this value in (1), we get

$$g_{\alpha\beta,\,\gamma} = g_{\mu\beta}\ \Gamma^{\mu}_{\alpha\gamma} + g_{\mu\alpha}\ \Gamma^{\mu}_{\beta\gamma} - \Gamma^{\mu}_{\alpha\gamma}\ g_{\mu\beta} - \Gamma^{\mu}_{\beta\gamma}\ g_{\mu\alpha} = 0.$$

Problem 1.11. *Show that the distributive law of ordinary differentation holds good in case of covariant differentiation of a product, i.e., to show that*

(*i*) $(B_{\mu}\, C_{\nu})_{,\sigma} = B_{\mu,\sigma} C_{\nu} + B_{\mu}\, C_{\nu,\sigma}$

(*ii*) $(B^{\mu}\, C^{\nu})_{,\sigma} = B^{\mu}{}_{,\sigma} C^{\nu} + B^{\mu}\, C^{\nu}{}_{,\sigma}$. (*Rohilkhand, 1990; Agra, 1959, 63*)

(*i*) R.H.S. $= B_{\mu,\sigma}\, C_{\nu} + B_{\mu}\, C_{\nu,\sigma}$

$$= \left(\frac{\partial B_{\mu}}{\partial x_{\sigma}} - \Gamma^{\alpha}_{\mu\sigma}\ B_{\alpha}\right) C_{\nu} + B_{\mu}\left(\frac{\partial C_{\nu}}{\partial x_{\sigma}} - \Gamma^{\alpha}_{\mu\sigma}\ C_{\alpha}\right)$$

$$= \frac{\partial}{\partial x_{\sigma}}(B_{\mu} C_{\nu}) - \Gamma^{\alpha}_{\mu\sigma}(B_{\alpha}\ C_{\nu}) - \Gamma^{\alpha}_{\nu\sigma}\ (B_{\mu}\ C_{\alpha})$$

$$= (B_{\mu}\ C_{\nu}),\ \sigma.$$

(*ii*) R.H.S $= B^{\mu}{}_{,\sigma}\ C^{\nu} + B^{\mu}\ C^{\nu}{}_{,\sigma}$

$$= \left(\frac{\partial B^{\mu}}{\partial x^{\sigma}} + \Gamma^{\mu}_{\alpha\sigma}\ B^{\alpha}\right) C^{\nu} + B_{\mu}\left(\frac{\partial C^{\nu}}{\partial x_{\sigma}} - \Gamma^{\nu}_{\nu\sigma}\ C^{\alpha}\right)$$

$$= \frac{\partial}{\partial x_{\sigma}}\ (B^{\mu}\ C^{\nu}) + \Gamma^{\mu}_{\alpha\sigma}\ (B^{\alpha}\ C^{\nu}) + \Gamma^{\nu}_{\alpha\sigma}\ (B^{\mu}\ C^{\alpha})$$

$$= (B^{\mu}\ C^{\nu}),\ \sigma.$$

Problem 1.12. *Prove that*

(*a*) $g,^{\mu\nu}_{\sigma} = 0.$ (*Agra, 1959, 65*)

(*b*) $\delta^{\nu}_{\mu,\,\sigma} = 0.$ (*Agra, 1965*)

(*a*) We have

$$g_{\xi\eta}\, g^{\xi\xi} = \delta^{\xi}_{\eta}$$

$$= 0 \text{ or } 1.$$

Differentiating with respect to x_{λ}, this gives

$$g^{\xi\xi}\ \frac{\partial\ g_{\xi\eta}}{\partial x_{\lambda}} + g_{\xi\eta}\ \frac{\partial g^{\xi\xi}}{\partial x_{\lambda}} = 0$$

Multiplying throughout by $g^{\eta\rho}$, we get

$$g^{\eta\rho}\, g^{\xi\xi}\ \frac{\partial g_{\xi\eta}}{\partial x_{\lambda}} + g^{\eta\rho}\ g_{\xi\eta}\ \frac{\partial g^{\xi\xi}}{\partial x_{\lambda}} = 0$$

or $$g^{\eta\rho}\ g^{\xi\xi}\left(\Gamma_{\xi\lambda,\,\eta} - \Gamma_{\lambda\eta,\,\xi}\right) + g^{\xi\rho}\ \frac{\partial g^{\xi\xi}}{\partial x_{\lambda}} = 0$$

or $$g^{\xi\xi}\,\Gamma^{\eta}_{\xi\lambda} + g^{\eta\rho}\,\Gamma^{\xi}_{\lambda\eta} + \frac{\partial g^{\rho\xi}}{\partial x_\lambda} = 0 \qquad \ldots (1)$$

Now $$g^{\mu\nu}_{,\sigma} = \frac{\partial g^{\mu\nu}}{\partial x_\sigma} + \Gamma^{\mu}_{\alpha\sigma}\, g^{\alpha\nu} + \Gamma^{\mu}_{\alpha\sigma}\, g^{\alpha\mu}$$

$$= 0 \text{ by } (1).$$

(*b*) $$\delta^{\nu}_{\mu,\sigma} = \frac{\partial \delta^{\nu}_{\mu}}{\partial x_\sigma} - \Gamma^{\alpha}_{\mu\sigma}\, g^{\nu}_{\alpha} + \Gamma^{\nu}_{\alpha\sigma}\,\delta^{\alpha}_{\mu}$$

$$= 0 - \Gamma^{\nu}_{\mu\sigma} + \Gamma^{\nu}_{\mu\sigma}$$

$$= 0.$$

Problem 1.13. *Show that the covariant derivative of an invariant is the same as its ordinary derivative.* (*Rohilkhand, 1990*)

Let I be an invariant, so that IA_μ is a covariant vector.

Now the covariant derivative of IA_μ is

$$(IA_\mu) = \frac{\partial (IA_\mu)}{\partial x_\nu} - \Gamma^{\alpha}_{\mu\nu}\,(IA_\alpha)$$

$$= A_\mu \frac{\partial I}{\partial x_\nu} - I\frac{\partial A_\mu}{\partial x_\nu} - \Gamma^{\alpha}_{\mu\nu}\, IA_\alpha$$

$$= A_\mu \frac{\partial I}{\partial x_\nu} + I\left(\frac{\partial A_\mu}{\partial x_\nu} - \Gamma^{\alpha}_{\mu\nu}\, A_\alpha\right)$$

$$= A_\mu \frac{\partial I}{\partial x_\nu} + IA_{\mu,\nu}$$

But from Problem 29,

$$(IA_\mu)_{,\nu} = I_{,\nu}\, A_\mu + IA_{\mu,\nu}$$

$$\therefore \quad I_{,\nu}\, A_\mu + IA_{\mu,\nu} = A_\mu \frac{\partial I}{\partial x_\nu} + IA_{\mu,\nu}$$

i.e. $$I_{,\nu}\, A_\mu = A_\mu \frac{\partial I}{\partial x_\nu}$$

or $$I_{,\,\nu} = \frac{\partial I}{\partial x_\nu}$$

which shows that the covariant derivative of an invariant is the same as its ordinary derivative.

1.11 CHRISTOFFEL SYMBOLS

Here, we introduce two expressions (not tensors) known as Christoffel's symbols of the first and second kinds. These will be found of great utility throughout our subsequent work.

Chirstoffel symbol of the first kind

$$\Gamma_{\mu\nu,\,\sigma} \equiv [\mu\nu,\,\sigma] = \frac{1}{2}\left(\frac{\partial g_{\mu\sigma}}{\partial x_\nu} + \frac{\partial g_{\nu\sigma}}{\partial x_\mu} - \frac{\partial g_{\mu\nu}}{\partial x_\sigma}\right) \qquad \ldots (1.96)$$

NOTES

Christoffel symbol of the second kind

$$\Gamma^{\sigma}_{\mu\nu} \equiv [\mu\nu, \sigma] = \frac{1}{2} g^{\sigma\lambda}\left(\frac{\partial g_{\mu\lambda}}{\partial x_{\nu}} + \frac{\partial g_{\nu\lambda}}{\partial x_{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x_{\lambda}}\right) \qquad ...(1.97)$$

It is obvious from these expressions that

$$[\mu\nu, \sigma] = [\nu\mu, \sigma], \ i.e., \ \Gamma_{\mu\nu,\ \sigma} = \Gamma_{\nu\mu,\ \sigma} \qquad ...(1.98)$$

and $\{\mu\nu, \sigma\} = \{\nu\mu, \sigma\}$, *i.e.*, $\Gamma^{\sigma}_{\mu\nu} = \Gamma^{\sigma}_{\nu\mu}$ ...(1.99)

showing that they are symmetrical with respect to μ and ν.

We also observe *the relations between two kinds*,

$$\{\mu\nu, \sigma\} = g^{\sigma\lambda} [\mu\nu, \lambda], \ i.e., \ \Gamma^{\sigma}_{\mu\nu} = g^{\sigma\lambda}\ \Gamma_{\mu\nu, \lambda} \qquad ... (1.100)$$

$$[\mu\nu, \sigma] = g_{\sigma\lambda}\ \{\mu\nu, \lambda\}, \ i.e., \ \Gamma_{\mu\nu,\sigma}\ \sigma = g_{\alpha\lambda}\ \Gamma^{\lambda}_{\mu\nu} \qquad ... (1.101)$$

To prove the result (1.100), we have from (1.96) on replacing σ by λ,

$$[\mu\nu, \lambda] = \frac{1}{2}\left(\frac{\partial g_{\mu\lambda}}{\partial x_{\nu}} + \frac{\partial g_{\nu\lambda}}{\partial x_{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x_{\lambda}}\right)$$

Multiplying both sides by $g^{\sigma\lambda}$, this becomes

$$g^{\sigma\lambda} [\mu\nu, \lambda] = \frac{1}{2} g^{\sigma\lambda}\left(\frac{\partial g_{\mu\lambda}}{\partial x_{\nu}} + \frac{\partial g_{\nu\lambda}}{\partial x_{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x_{\lambda}}\right)$$

$$= \{\mu\nu, \sigma\}, \text{ from (2).}$$

This proves the result (1.100).

Again to prove the result (1.101), interchanging λ and σ in (1.97), we have

$$\{\mu\nu, \lambda\} = \frac{1}{2} g^{\lambda\sigma}\left(\frac{\partial g_{\mu\sigma}}{\partial x_{\nu}} + \frac{\partial g_{\nu\sigma}}{\partial x_{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x_{\sigma}}\right)$$

$$= g^{\lambda\sigma} [\lambda\nu, \sigma], \text{ from (1.96).}$$

Multiplying both sides by $g_{\sigma\lambda}$, this becomes

$$g_{\sigma\lambda}\ \{\mu\nu, \lambda\} = g_{\sigma\lambda}\, g^{\sigma\lambda} [\mu\nu, \sigma]$$

$$= [\mu\nu, \sigma]$$

which proves the result (1.101).

Now we have from (1.96),

$$[\mu\nu, \sigma] + [\sigma\nu, \mu] = \frac{1}{2}\left(\frac{\partial g_{\mu\sigma}}{\partial x_{\nu}} + \frac{\partial g_{\nu\sigma}}{\partial x_{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x_{\sigma}}\right) + \frac{1}{2}\left(\frac{\partial g_{\mu\sigma}}{\partial x_{\nu}} + \frac{\partial g_{\mu\nu}}{\partial x_{\sigma}} - \frac{\partial g_{\nu\sigma}}{\partial x_{\mu}}\right)$$

$$= \frac{1}{2}\left(\frac{\partial g_{\mu\sigma}}{\partial x_{\nu}} + \frac{\partial g_{\nu\sigma}}{\partial x_{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x_{\sigma}} + \frac{\partial g_{\mu\sigma}}{\partial x_{\nu}} + \frac{\partial g_{\mu\nu}}{\partial x_{\sigma}} - \frac{\partial g_{\nu\sigma}}{\partial x_{\mu}}\right)$$

$$= \frac{\partial g_{\mu\sigma}}{\partial x_{\nu}}, \qquad ... (1.102)$$

NOTES

i.e. $\Gamma_{\mu\nu,\sigma} + \Gamma_{\sigma\nu,\mu} = \dfrac{\partial g_{\mu\sigma}}{\partial x_\nu}$. (*Agra, 1971, 77*)

Problem 1.14. *Find the Christoffel's symbols corresponding to*

(*a*) $ds^2 = a^2\, d\theta^2 + a^2 \sin^2\theta\, d\phi^2$. (*Rohilkhand, 1987, 90; Agra, 1974*)

(*b*) $ds^2 = dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\phi^2$. (*Rohilkhand, 1979, 90*)

(*a*) We have $ds^2 = a^2\, d\theta^2 + a^2 \sin^2\theta\, d\phi^2$,

so that $g_{11} = a^2,\ g_{22} = a^2\sin^2\theta,\ g_{12} = 0 = g_{21}$

$$\therefore \qquad g = \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} = \begin{vmatrix} a^2 & 0 \\ 0 & a^2\sin^2\theta \end{vmatrix} = a^4\sin^2\theta,$$

giving $$g^{11} = \frac{\text{cofactor of } g_{11}}{g} = \frac{a^2\sin^2\theta}{a^4\sin^2\theta} = \frac{1}{a^2}$$

Similarly, $$g^{22} = \frac{1}{a^2\sin^2\theta} \text{ and } g^{12} = 0 = g^{21}.$$

Thus, Christoffel symbols of first kind are

$$[22, 1] = \frac{1}{2}\left(\frac{\partial g_{21}}{\partial x_2} + \frac{\partial g_{21}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_1}\right), \text{ where } x_1 = \theta,\ x_2 = \phi \text{ etc.}$$

$$= \frac{1}{2}\left\{0 + 0 - \frac{\partial}{\partial\theta}(a^2\sin^2\theta)\right\}$$

$$= -a^2\sin\theta\cos\theta.$$

Similarly, $$[12, 2] = \frac{1}{2}\frac{\partial g_{22}}{\partial x_1} = \frac{1}{2}\frac{\partial}{\partial\theta}(a^2\sin^2\theta) = a^2\sin\theta\cos\theta.$$

The rest of all are zero, and the Christoffel symbols of second kind are

$$\{22, 1\} = g^{1\lambda}[22, \lambda]$$

$$= g^{11}[22, 1] + g^{12}[22, 1]$$

$$= \frac{1}{a^2}(-a^2\sin\theta\cos\theta) + 0 = -\sin\theta\cos\theta.$$

Similarly, $$\{12, 2\} = g^{2\lambda}[12, \lambda]$$

$$= g^{21}[12, 2] + g^{22}[12, 2]$$

$$= 0 + \frac{1}{a^2\sin^2\theta}(a^2\sin\theta\cos\theta) = \cot\theta$$

and the rest of all zero.

(*b*) We have $ds^2 = dr^2 + r^2\, d\theta^2 + r^2\sin^2\theta\, d\phi^2$,

where $x_1 = r_1,\ x_2 = \theta,\ x_3 = \phi$

and $g_{11} = 1,\ g_{22} = r^2,\ g_{33} = r^2\sin\theta,$

$g_{12} = 0 = g_{13} = \ldots$etc.

NOTES

$$\therefore \qquad g = \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{vmatrix} = r^4 \sin^2\theta, \qquad \ldots (1)$$

giving $$g^{11} = \frac{\text{cofactor of } g_{11}}{g} = \frac{\begin{vmatrix} r^2 & 0 \\ 3 & r^2 \sin^2\theta \end{vmatrix}}{r^4 \sin^2\theta} = \frac{r^4 \sin^2\theta}{r^4 \sin^2\theta} = 1.$$

Similarly $$g^{22} = \frac{\text{cofactor of } g_{22}}{g} = \frac{r^2 \sin^2\theta}{r^4 \sin^2\theta} = \frac{1}{r^2},$$

$$g^{33} = \frac{\text{cofactor of } g_{33}}{g} = \frac{r^2}{r^4 \sin^2\theta} = \frac{1}{r^2 \sin^2\theta}$$

and the rest all are zero.

Now the Christoffel's symbols of the first kind are

$$[11, 1] = \frac{1}{2}\left(\frac{\partial g_{11}}{\partial x_1} + \frac{\partial g_{11}}{\partial x_1} - \frac{\partial g_{11}}{\partial x_1}\right) = 0.$$

Similarly $[11, 2] = 0 = [11, 3]$.

In a similar manner, it is easy to show that

$[22, 1] = -r$, $[22, 2] = 0 = [22, 3]$

$[33, 1] = -r \sin^2\theta$, $[33, 2] = -r^2 \sin\theta \cos\theta$, $[33, 3] = 0$

$[12, 1] = 0 = [21, 1]$, $[13, 1] = 0 = [31, 1]$

$[12, 2] = 0 = [21, 2]$, $[12, 3] = 0 = [21, 3]$

$[13, 1] = 0 = [31, 1]$, $[13, 2] = 0 = [31, 2]$

$[13, 3] = r \sin^2\theta = [31, 3]$, $[23, 1] = 0 = [32, 1]$

$[23, 2] = 0 = [32, 2]$, $[23, 3] = r^2 \sin\theta \cos\theta = [32, 3]$

and the Christoffel symbols of the second kind are

$$\{22, 1\} = g^{1\lambda} [22, \lambda]$$

$$g^{11} [22, 1] + g^{12} [22, 2] + g^{13} [22, 3]$$

$$= -r + 0 + 0 = -r.$$

Similarly $\{3, 3, 1\} = -r \sin^2\theta$, $\{11, 1\} = 0$

$$\{33, 2\} = -\sin\theta \cos\theta,\ \{13, 3\} = \frac{1}{2},\ \{23, 3\} = \cot\theta$$

and the rest all are zero.

Note. We get the **Metric tensor** in *spherical polar coordinates* as

$$g_{\mu\nu} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{bmatrix} \text{ just as in (1)}$$

The metric tensor in *cylindrical coordinates* is given by

$$g_{\mu\nu} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Since in this case

$$ds^2 = dr^2 + r^2\, d\theta^2 + dz^2$$

with $x_1 = r,\ x_2 = \theta,\ x_3 = z$

so that $g_{11} = 1,\ g_{22} = r^2,\ g_{33} = 1$

and $g_{12} = 0 = g_{13} = \ldots$ etc.

Check Your Progress

11. Give examples of dyadic tensors.
12. Define a tensor of zero rank.
13. What is stress tensor?
14. What rank does direct product of two tensors give?
15. What is anti-symmetric tensor?
16. What is contraction?

1.12 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Suppose that the three surfaces u = const., v = const., w = const., intersect in a point P of the region R. The values of u, v, w for the three surfaces intersecting at P are called the curvilinear co-ordinates of the point P.
2. The three surfaces intersect pairwise in three curves known as *co-ordinate curves.*
3. A scalar field is the region in which the scalar point function specifies the scalar physical quantity. It is represented by a continuous scalar function giving the value of the quantity at each point.
4. In scalar field all the points having same value of f can be connected by means of surfaces, which are called *equal* or *level* surfaces.
5. A matrix is simply an arrangement of elements and has no numerical value.
6. A matrix in which the number of rows is equal to the number of columns is called a square matrix.

   For example, $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ is a 2×2 square matrix.

NOTES

NOTES

7. A square matrix whose every element other than diagonal elements is zero, is called a diagonal matrix. For example,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$ is a diagonal matrix.

8. A diagonal matrix whose diagonal elements are all equal to 1 (unity) is called identity matrix or (unit matrix). For example,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$ is an identity matrix.

9. Two matrices A and B are said to be equal if,
   (*i*) A and B are of same order.
   (*ii*) Corresponding elements in A and B are same. For example, the following two matrices are equal,

$$\begin{pmatrix} 3 & 4 & 9 \\ 16 & 25 & 64 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 9 \\ 16 & 25 & 64 \end{pmatrix}$$

10. Matrix addition is commutative.
    i.e., $A + B = B + A$
    For, (i, j)th element of $A + B$ is $(a_{ij} + b_{ij})$ and of $B + A$ is $(b_{ij} + a_{ij})$, and they are same as a_{ij} and b_{ij} are real numbers.
11. The examples of dyadic, i.e., tensor of rank two are: an operator relating dielectric displacement vector with the electric vector of an electro-magnetic wave in an isotropic medium; a stress tensor relating stress and strain in an isotropic medium in which case a component of stress T is a function of every component of strain S.
12. A vector is a tensor of first rank and an invariant or scalar is a tensor of zero rank.
13. The dyadic or tensor of rank two is also known as stress tensor.
14. The direct product of two tensors gives a new tensor of rank equal to the sum of ranks of these tensors.
15. If a tensor is such that two contravariant or covariant indices of it when interchanged, the components of the tensor alter in sign but not in magnitude, the tensor is said to be anti-symmetric or skew-symmetric.
16. The algebraic operation by which the rank of a tensor may be lowered by 2 (or by any even number) is known as contraction.

1.13 SUMMARY

- Suppose that the three surfaces u = const., v = const., w = const., intersect in a point P of the region R. The values of u, v, w for the three surfaces intersecting at P are called the *curvilinear co-ordinates* of the point P. The three surfaces are then known as *co-ordinate surfaces.* The three surfaces intersect pairwise in three curves known as *co-ordinate curves.*

- A system of orthogonal curvilinear co-ordinates is one which corresponds to the points of intersection of a triply orthogonal system of three families of surfaces $u(x, y, z)$ = const., $v(x, y, z)$ = const., $w(x, y, z)$ = const.
- Gradient of any scalar-point function is a vector.
- Gradient of sum of two scalar-point functions is equal to the sum of their gradients.
- The divergence of the sum of two vector functions is equal to the sum of their divergences.
- If a rigid body is in motion, the curl of its linear velocity at any point gives twice its angular velocity.
- Curl of sum of two vector point functions is equal to the sum of their curls.
- A matrix polynomial of degree n is an expression such as P(l) = A0 + A1l+ A2l2 +........+ A*m*l*m*; A*m* ¹ 0, where all the square matrices A0, A1, A2,....., A*m* are of the same order, say, n, *i.e.,* the matrix polynomial is n-rowed. Thus if A be an n-rowed square matrix and I be an r-rowed unit matrix, then A–lI is the matrix polynomial of the first degree, and known as the characteristic or proper or latent or eigen or varient matrix of the matrix A and its determinant A – lI is known as the characteristic polynomial of the matrix A whereas A – lI = O is the characteristic equation of the matrix A.
- Every square matrix satisfies its own characteristic equation.
- A matrix which has exactly one row is called a row matrix.
- A matrix which has exactly one column is called a column matrix.
- A matrix in which the number of rows is equal to the number of columns is called a square matrix.
- A matrix each of whose elements is zero is called a null matrix or zero matrix.
- A diagonal matrix whose diagonal elements are equal, is called a scalar matrix.
- If A and B are two matrices of the same order then addition of A and B is defined to be the matrix obtained by adding the corresponding elements of A and B.
- The product AB of two matrices A and B is defined only when the number of columns of A is same as the number of rows in B and by definition the product AB is a matrix G of order $m \times p$ if A and B were of order $m \times n$ and $n \times p$, respectively.
- If k is any complex number and A, a given matrix, then kA is the matrix obtained from A by multiplying each element of A by k. The number k is called scalar.
- Let A be a matrix. The matrix obtained from A by interchange of its rows and columns, is called the transpose of A.
- A square matrix A = [*aij*] is called a symmetric matrix if *aij* = *aji* for all i and j. If *A*¢ denotes the transpose of a square matrix A, then A is symmetric when A = *A*¢ and A is skew symmetric when *A*¢ = –*A*.
- A square matrix A is said to be orthogonal if *A*¢*A* = *I*.
- A square matrix A is said to be unitary if $A * A = I$.

NOTES

- The adjoint matrix of A is obtained by replacing the elements of A by their respective cofactors and then transposing.
- If A is a square matrix of order n, then a square matrix B of the same order n is said to be inverse of A if $AB = BA = I$ (unit matrix).
- Let A be an $m \times n$ matrix. We say rank of A is r if (i) at least one minor of order r is non zero and (ii) every minor of order (r + 1) is zero.
- Matrix obtained from identity matrix by a single elementary operation is called an elementary matrix.
- Any literal suffix appearing twice in a term is said to be a *dummy suffix* and it may be changed freely to any other letter not already used in that term. Also two or more dummy suffixes can be interchanged
- Tensor analysis is the generalization of vector analysis as is evident by considering a vector function $f(\mathbf{r})$ of a vector $\mathbf{r}$. This function is continuous at $\mathbf{r} = \mathbf{r}_0$ if,

  $f(\mathbf{r}) = \quad f(\mathbf{r}_0)$
- $f(\mathbf{r}) = (\mathbf{b}_1\mathbf{a}_1 + \mathbf{b}_2\mathbf{a}_2 + \mathbf{b}_3\mathbf{a}_3)\,.\,\mathbf{r}$

  $= \quad \mathrm{f}\,.\,\mathbf{r}$
- $\mathrm{f} = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3$ consists of nine components in three dimensional coordinate geometry and hence it is neither a scalar nor a vector quantity but is a new mathematical symbol called as the *dyadic*.
- A tetradic is the sum of tetrads, $\sum \mathbf{a}_\alpha\, \mathbf{b}_\alpha\, \mathbf{c}_\alpha\, \mathbf{d}_\alpha$ and etc.
- $dx_m¢ = \sum_{\alpha=1}^{4} \frac{\partial x'_\mu}{\partial x_\alpha}\, dx_a$; (m = 1, 2, 3, 4)
- $\frac{\partial x_\mu}{\partial x'_\alpha} \cdot \frac{\partial x'_\alpha}{\partial x_\nu} = \quad \frac{\partial x_\mu}{\partial x_\nu} = 0 \quad \text{if} \quad \mu \neq \nu$
- if A_m be a covariant vector, its transformation law is

  $$A'_\mu = \sum_{\alpha=1}^{4} \frac{\partial x_\alpha}{\partial x'_\mu} A_\alpha$$

  where the lower position of the suffix indicates covariance.
- Contravariant Tensors $A¢\mathrm{mn} = \frac{\partial x'_\mu}{\partial x_\alpha} \frac{\partial x'_\nu}{\partial x_\beta} A^{\alpha\beta}$
- Mixed Tensors $A'^{\nu}_{\mu} = \frac{\partial x_\alpha}{\partial x'_\mu} \frac{\partial x'_\nu}{\partial x_\beta} A^{\beta}_{\alpha}$
- The rank of a tensor is determined by the number of suffixes or indices attached to it.
- Let a tensor be such that contravariant or covariant indices of it can be interchanged without altering the value of the tensor, then the tensor is termed as symmetrical or symmetric in these indices.

NOTES

- If a tensor is such that two contravariant or covariant indices of it when interchanged, the components of the tensor alter in sign but not in magnitude, the tensor is said to be anti-symmetric or skew-symmetric.
- It is not known about any vector which has the same components in different systems of coordinates, but there exist tensors of higher ranks which have the same components in all the frames of reference. These tensors are called to have the invariant components or invariant tensors in general.
- The outer product of two relative tensors is itself a relative tensor of rank and weight equal to the sum of the ranks and the sum of weights of the given relative tensors respectively.

1.14 KEY TERMS

- **Vector**: Vector is a quantity having direction as well as magnitude, especially as determining the position of one point in space relative to another**.**
- **Curl:** Curl is a vector operator that describes the infinitesimal rotation of a vector field in three-dimensional Euclidean space
- **Divergence:** Divergence is a vector operator that produces a scalar field, giving the quantity of a vector field's source at each point.
- **Gradient**: The gradient is a vector operation which operates on a scalar function to produce a vector whose magnitude is the maximum rate of change of the function at the point of the gradient and which is pointed in the direction of that maximum rate of change.
- **Row matrix:** A matrix which has exactly one row is called a row matrix.
- **Column matrix:** A matrix which has exactly one column is called a column matrix.
- **Square matrix:** In this matrix, the number of rows is equal to the number of columns.
- **Equivalent matrices:** Two matrices A and B of the same order are said to be equivalent if one of them can be obtained from the other by performing a sequence of elementary transformations. Equivalent matrices have the same order or rank.
- **Rank of a matrix:** It is the largest of the orders of all the non-vanishing minors of that matrix. Rank of a matrix is denoted as $R(A)$ or $P(A)$. If A is an $m \times n$ matrix then $R(A)$ or $P(A)$ £ minimum of (m, n).
- **Tensor:** A tensor is a geometric object that maps in a multi-linear manner geometric vectors, scalars, and other tensors to a resulting tensor. Vectors and scalars which are often used in elementary physics and engineering applications, are considered as the simplest tensors.
- **Rank of a tensor:** The rank of a tensor is determined by the number of suffixes or indices attached to it. As a matter of fact the rank of a tensor when raised as power to the number of dimensions, yields the number of components of the tensor and hence the components of the matrix that represents the tensor.

1.15 SELF-ASSESSMENT QUESTIONS AND EXERCISES

NOTES

Short Answer Questions

1. What do you mean by orthogonal curvilinear co-ordinates?
2. What do you mean by eigen vector and eigen value?
3. Why is Cayley-Hamilton theorem used?
4. What is a matrix? What are its types?
5. Find matrices A and B such that A and B are both non-zero matrices but the product AB is a zero matrix.
6. Is it necessary that if two matrices A and B are such that AB and BA are both defined and $AB = 0$, then BA must also be a zero matrix?
7. Define orthogonal and unitary matrices.
8. What do you mean by the inverse of a matrix?
9. How will you determine the rank of a matrix?
10. What is meant by elementary transformation?
11. How will you determine that the given square matrix is of order one, order two or order three?
12. Differentiate between consistent and inconsistent equations.
13. Define any three properties of eigenvalues and eigenvectors.
14. What is an inner product?
15. What do you understand by a tensor and its rank?
16. Define contravariant and covariant tensors.
17. What are antecedents and consequents?
18. Write in brief about the transformation of coordinates in Minkowski's four dimensional world.
19. What are symmetric and anti-symmetric tensors?
20. Define generalized Kronecker delta symbol.
21. What do you mean by alternating or permutation epsilon tensor?
22. State and prove the quotient law.

Long Answer Questions

1. Express gradient of scalar product in terms of curl.
2. Show that curl of sum of two vector point functions is equal to the sum of their curls.
3. Show that the divergences of the sum of the two vector functions is equal to the sum of their divergences.
4. Show that the gradient of a scalar-point functions is a vector.
5. Explain the iterative method for dominant eigen roots.

NOTES

6. Demonstrate the Jacobi's methods.
7. Discuss various types of matrices.
8. Define eigen values and eigen vectors of a matrix. Prove that the eigen values of an orthogonal matrix are unimodular.

   Obtain the eigen values and eigen vectors of the matrix $\mathbf{A} = \begin{bmatrix} 8 & -12 & 5 \\ 15 & -25 & 11 \\ 24 & -24 & 19 \end{bmatrix}$
9. Prove that $\frac{\partial x_\mu}{\partial x'_\alpha} \cdot \frac{\partial x'_\alpha}{\partial x_\nu} = \frac{\partial x_\mu}{\partial x_\nu} = 0 \; if \; \mu \neq \nu$
10. Show that the law of transformation for a contravariant vector is transitive.
11. Show that $\frac{\partial A_\lambda}{\partial x_\mu}$ is not a tensor although A_λ is a covariant tensor of rank one.
12. If xy, $2y - z^2$ and xz are the components of a covariant tensor in rectangular coordinates, then find its covariant components in spherical coordinates.
13. Show that every tensor can be expressed as the sum of two tensors, one of which is symmetric and other skew-symmetric in a pair of covariant or contravariant indices.
14. If the components of a tensor are zero in one coordinate system, then prove that the components are zero in all coordinate systems.
15. Evaluate $\nabla \times (V \times W)$ using e tensor.
16. Show that symmetry properties of a tensor are invariant.
17. Explain the covariant derivative of a tensor
18. Define the Christoffel three-index symbol of the second kind. Prove that it is not a tensor. Verify that it behaves as a constant in covariant differentiation.
19. Explain what is meant by covariant, contravariant and mixed tensors.

1.16 FURTHER READING

Dass, HK. 2008. *Mathematical Physics*. New Delhi: S. Chand

Chattopadhyay, PK. 1990. *Mathematical Physics*. New Delhi: New age International

Hassani, Sadri. 2013. *Mathematical Physics: A Modern Introduction to Its Foundations* Berlin: Springer Science & Business Media

Khanna, V.K. and S.K. Bhambri. 2008. *A Course in Abstract Algebra*, 3rd edition. New Delhi: Vikas Publishing House

Grewal, B.S. 1998. *Higher Engineering Mathematics,* 34th edition. New Delhi: Khanna Publishers

Narayan, Shanti. 1996. *Differential Calculus,* 14th edition. New Delhi: S. Chand and Company

Kreyszig, Erwin. *Advanced Engineering Mathematics,* 7th edition. Mumbai: Wiley Eastern Ltd

NOTES

Bali, N.P. 2007. *A Textbook of Engineering Mathematics*. New Delhi: Laxmi Publications (P) Ltd

Narayan, Shanti. 2010. *A Text Book of Vector Analysis.* New Delhi: S. Chand and Sons

Spiegel. 1959. *Vector Analysis (Schaum'S Outline).* New York: McGraw Hill Professional

Gupta, B. D. 2009. *Mathematical Physics*, 4th Edition. Noida: Vikas Publishing House Pvt. Ltd.

UNIT 2 PARTIAL DIFFERENTIAL EQUATIONS AND GROUP THEORY

NOTES

Structure

2.0 INTRODUCTION

In this unit, you will learn about partial differential equations, which are used to formulate, and thus aid, the solution of problems involving functions of several variables. You will learn how to differentiate a given equation with respect to x and y and then find the solution for the same by forming partial differential equations. You will understand the importance of arbitrary functions and arbitrary constants in the formation of these equations. You will learn to form partial differential equations by eliminating arbitrary functions and constants and solve different types of partial differential equations by the direct integration method. Some well-known partial differential equations like Laplace equations, Poisson equations and wave equations are also explained in detail. In mathematics and abstract algebra, group theory studies the algebraic structures known as groups. The concept of a group is central to abstract algebra: other well-known algebraic structures, such as rings, fields, and vector spaces, can all be seen as groups endowed with additional operations and axioms. Groups recur throughout mathematics, and the methods of group theory have influenced many parts of algebra. Linear algebraic groups and Lie groups are two branches of group theory that have experienced advances and have become subject areas in their own right. Various physical systems, such as crystals and the hydrogen atom, may be modelled by symmetry groups. Thus group theory and the closely related representation theory have many important applications in physics, chemistry, and materials science. Group theory is also central to public key cryptography. One of the most important mathematical achievements of the 20th century was the collaborative effort, taking up more than 10,000 journal pages and mostly published between 1960 and 1980, that culminated in a complete classification of finite simple groups. In this unit, you will study about the group theory, homomorphism and isomorphism and various problems and theorems related to group theory in detail.

2.1 OBJECTIVES

After going through this unit, you will be able to:

- Analyse various types of partial differential equations
- Discuss the method of forming partial differential equations by eliminating arbitrary constants and functions
- Explain Laplace equations, Poisson equations and wave equations
- Describe equations of heat flow
- Discuss the group theory
- Explain homomorphism and isomorphism
- Discuss various problems and theorems related to group theory

2.2 PARTIAL DIFFERENTIAL EQUATIONS

Let $z = f(x, y)$ be a function of two independent variables x and y. Then $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$ are the first order partial derivatives; $\frac{\partial^2 z}{\partial x^2}, \frac{\partial^2 z}{\partial y^2}, \frac{\partial^2 z}{\partial x \partial y}$ are the second order partial derivatives.

Any equation which contains one or more partial derivatives is called a partial differential equation. $x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = z; \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} + \frac{\partial^2 z}{\partial x \partial y} = 0$ are examples for partial differential equation (PDE) of first order and second order respectively.

We use the following notations for partial derivatives,

$$p = \frac{\partial z}{\partial x},\ q = \frac{\partial z}{\partial y},\ r = \frac{\partial^2 z}{\partial x^2},\ s = \frac{\partial^2 z}{\partial x \partial y},\ t = \frac{\partial^2 z}{\partial y^2}$$

Partial differential equation may be formed by eliminating *(i)* arbitrary constants *(ii)* arbitrary functions.

Example 2.1: Form the parital differential equation by eliminating the arbitrary constants from $z = ax + by + a^2 + b^2$.

Solution: Given, $z = ax + by + a^2 + b^2$ (1)

Here we have two arbitrary constants a and b. Therefore, we need two more equations to eliminate a and b. Differentiating equation (1) partially with respect to x and y respectively we get,

$$\frac{\partial z}{\partial x} = p = a \quad (2)$$

$$\frac{\partial z}{\partial y} = q = b \quad (3)$$

From equations (2) and (3), we get,

$$a = p,\ b = q$$

Substituting values of a and b in (1) we get,

$$z = px + qy + p^2 + q^2$$

This is the required partial differential equation.

Example 2.2: Eliminate a and b from $z = (x + a)(y + b)$.

Solution: Differentiating partially with respect to x and y,

$$p = y + b,\ q = x + a$$

Eliminating a and b, we get $z = pq$.

Example 2.3: Form the partial differential equation by eliminating the arbitrary constants in $z = (x - a)^2 + (y - b)^2$.

Solution: Given, $z = (x - a)^2 + (y - b)^2$ (1)

Here we have two arbitrary cosntants a and b. To eliminate these two arbitrary constants we need two more equations connecting a and b. Therefore, differentiating equation (1) partially with respect to x and y, we get,

$$\frac{\partial z}{\partial x} = p = 2(x - a) \quad (2)$$

$$\frac{\partial z}{\partial y} = q = 2(y - b) \quad (3)$$

From equation (2), we get,

$$x - a = \frac{p}{2} \quad (4)$$

From equation (3), we get,

$$y - b = \frac{q}{2} \quad (5)$$

NOTES

Substituting equations (4) and (5) in (1) we get,

$$z = \left(\frac{p}{2}\right)^2 + \left(\frac{q}{2}\right)^2$$

Simplifying we get, $4z = p^2 + q^2$

This gives the partial differential equation after elimination of a and b.

Example 2.4: Form the partial differential equation by eliminating the arbitrary constants from $z = (x^2 + a)(y^2 + b)$.

Solution: Given, $z = (x^2 + a)(y^2 + b)$ (1)

Here we have two arbitrary constants a and b.

Differentiating equation (1) partially with respect to x and y we get,

$$\frac{\partial z}{\partial x} = p = 2x(y^2 + b) \quad (2)$$

$$\frac{\partial z}{\partial y} = q = 2y(x^2 + a) \quad (3)$$

From equation (2) we get, $\frac{p}{2x} = y^2 + b$ (4)

From equation (3) we get, $\frac{q}{2y} = x^2 + a$ (5)

Substituting equations (4) and (5) in (1), we get,

$$z = \frac{p}{2x} . \frac{q}{2y}$$

$$pq = 4xyz$$

This gives the required partial differential equation.

Example 2.5: Form the partial differential equation by eliminating a, b, c from $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

Solution: Given, $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ (1)

Differential partially with respect to x and y we get,

$$\frac{2x}{a^2} + \frac{2z}{c^2} \cdot p = 0 \quad (2)$$

$$\frac{2y}{b^2} + \frac{2z}{c^2} \cdot q = 0 \quad (3)$$

NOTES

Differentiating equation (2) partially with respect to y,

$$0 + \frac{2}{c^2}(zs + qp) = 0$$

$$zs + qp = 0$$

Note: More than one partial differential equation is possible in this problem. These partial differential equations are,

$$xzr + xp^2 - zp = 0, \quad yzt + yq^2 - zq = 0$$

Formation of Partial Differential Equation by Eliminating Arbitrary Functions

The partial differential equations can be formed by eliminating arbitrary junctions. The following examples will make the concept clear.

Example 2.6: Eliminate arbitrary function from,

$$z = f(x^2 + y^2) \qquad (1)$$

Solution: Differentiating partially with respect to x and y, we get,

$$p = f'(x^2 + y^2).2x \qquad (2)$$

$$q = f'(x^2 + y^2).2y \qquad (3)$$

Eliminating $f'(x^2 + y^2)$ from equation (2) and (3), we get, $py = qx$

Example 2.7: Form the partial differential equation by eliminating the arbitrary function ϕ from $xyz = \phi(x^2 + y^2 - z^2)$.

Solution: Given, $xyz = \phi(x^2 + y^2 - z^2)$ (1)

This equation contains only one arbitrary function ϕ and we have to eliminate it.

Differentiating equation (1) partially with respect to x and y we get,

$$yz + xyp = \phi'(x^2 + y^2 - z^2)(2x - 2zp) \qquad (2)$$

$$xz + xyq = \phi'(x^2 + y^2 - z^2)(2y - 2zq) \qquad (3)$$

From equation (2), we get,

$$\phi'(x^2 + y^2 - z^2) = \frac{yz + xyp}{2x - 2zp} \qquad (4)$$

From equation (3), we get,

$$\phi'(x^2 + y^2 - z^2) = \frac{xz + xyq}{2y - 2zq} \qquad (5)$$

Since, LHS of equations (4) and (5) are equal, we have,

$$\frac{yz + xyp}{2x - 2zp} = \frac{xz + xyq}{2y - 2zq}$$

$$(yz + xyp)(y - zq) = (xz + xyq)(x - zp)$$

i.e., $$y(z + xp)(y - zq) = x(z + yq)(x - zp) \quad (6)$$

NOTES

On simplifying equation (6) we get,

$$px(y^2 + z^2) - qy(z^2 + x^2) = z(x^2 - y^2)$$

Which gives the required partial differential equation.

Example 2.8: Eliminate the arbitrary function from $z = (x + y)\, f(x^2 - y^2)$

Solution: Given, $z = (x + y)\, f(x^2 - y^2)$ (1)

Differentiating partially with respect to x and y we get,

$$p = (x + y)f'(x^2 - y^2)2x + f(x^2 - y^2) \cdot 1 \quad (2)$$

$$q = (x + y)f'(x^2 - y^2)(-2y) + f(x^2 - y^2) \cdot 1 \quad (3)$$

Eliminating $f'(x^2 - y^2)$ from equations (2) and (3) we get,

$$\frac{2x(x+y)}{-2y(x+y)} = \frac{p - f(x^2 - y^2)}{q - f(x^2 - y^2)}$$

$$2x[q - f(x^2 - y^2)] = -2y[p - f(x^2 - y^2)]$$

$$xq - xf(x^2 - y^2) = -yp + yf(x^2 - y^2)$$

$$xq + yp = (x + y)\, f(x^2 - y^2)$$

$$= (x + y)\frac{z}{(x+y)}$$

$$\therefore \quad z = xq + yp$$

This is a required equation.

Example 2.9: Eliminate the arbitrary function from $z = xy + f(x^2 + y^2)$

Solution: Given, $z = xy + f(x^2 + y^2)$ (1)

Differentiating partially equation (1) with respect to x and y we get,

$$p = y + f'(x^2 + y^2) \cdot 2x \quad (2)$$

$$q = x + f'(x^2 + y^2) \cdot 2y \quad (3)$$

Eliminating $f'(x^2 + y^2)$ from equations (2) and (3) we get,

$$(p - y)y = (q - x)x$$

$$py - y^2 = qx - x^2$$

$$py - qx = y^2 - x^2$$

Which is a required equation.

Example 2.10: Eliminate the arbitrary functions f and ϕ from the relation $z = f(x + ay) + \phi(x - ay)$

Solution: Differentiating partially with respect to x and y we get,

$$p = f'(x + ay) + \phi'(x - ay) \quad (1)$$

$$q = af'(x + ay) - a\phi'(x - ay) \quad (2)$$

Differentiating these again, with respect to x and y we get,

$$\frac{\partial^2 z}{\partial x^2} = r = f''(x + ay) + \phi''(x - ay) \quad (3)$$

$$\frac{\partial^2 z}{\partial y^2} = t = a^2 f''(x + ay) + a^2\phi''(x - ay) \quad (4)$$

From equations (3) and (4) we get,

$$t = a^2 r$$

Equations Solvable by Direct Integration

A solution in which the number of arbitrary constants is equal to the number of independent variables is called complete integral or complete solution.

In complete integral, if we give particular values to the arbitrary constants, we get particular integral. If $\varphi(x, y, z, a, b) = 0$, is the complete integral of a partial differential equation, then the eliminant of a and b from the equations $\frac{\partial \phi}{\partial a} = 0$, $\frac{\partial \phi}{\partial b}$ = 0, is called singular integral.

Let us consider four standard types of nonlinear partial differential equations and the procedure for obtaining their complete solution.

Type I Equations of the form $F(p, q) = 0$. In this type of equations we have only p and q and there is no x, y and z. To solve this type of problems, let us assume that $z = ax + by + c$ be the solution and then proceed as in the following examples.

Example 2.11: .Solve $p^2 + q^2 = 4$

Solution: Given, $p^2 + q^2 = 4$ (1)

Let us assume that $z = ax + by + c$ be a solution of equation (1). (2)

Partially differentiating equation (1) with respect to x and y, we get,

$$\frac{\partial z}{\partial x} = p = a \text{ and } \frac{\partial z}{\partial y} = q = b \quad (3)$$

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

$$a^2 + b^2 = 4 \quad (4)$$

To get the complete integral we have to eliminate any one of the arbitrary constants from equation (2).

From equation (4) we get,

$$b = \pm\sqrt{4 - a^2} \quad (5)$$

NOTES

Substituting equation (5) in (2) we get,

$$z = ax \pm y\sqrt{4-a^2} + C \qquad (6)$$

Which contains only two constants (equal to number of independent variables). Therefore, it gives the complete integral.

To check for Singular Integral:

Differentiating equation (6) partially with respect to a and c and equating to zero, we get,

$$\frac{\partial z}{\partial a} = x \pm \frac{1}{2\sqrt{4-a^2}}(-2a) = 0 \qquad (7)$$

and, $$\frac{\partial z}{\partial c} = 1 = 0$$

Here, 1 = 0 is not possible.

Hence, there is no singular integral.

Example 2.12: Solve $p^2 + q^2 = npq$

Solution. The solution is, $z = ax + by + c$, where $a^2 + b^2 = nab$

Solving, $b = \dfrac{a\left(n \pm \sqrt{n^2-4}\right)}{2}$

The compete integral is,

$$z = ax + \frac{ay}{2}\left(n \pm \sqrt{n^2-4}\right) + c$$

Differentiating partially with respect to c, we see that there is no singular integral, as we get an absurd result.

Example 2.13: Solve $p + q = pq$

Solution: This equation is of the type, $F(p, q) = 0$.

∴ The complete solution is of the form, $z = ax + by + c$ (1)

Differentiating equation (1) partially with respect to x and y we get,

$$p = a,\ q = b$$

Therefore, the given equation becomes,

$$a + b = ab$$

$$a = b(a-1); \qquad b = \frac{a}{a-1}$$

Therefore, the complete solution is,

$$z = ax + \left(\frac{a}{a-1}\right)y + c$$

This type of equation has no singular solution.

Let, $c = \varphi(a)$

$$z = ax + \left(\frac{a}{a-1}\right)y + \phi(a) \tag{2}$$

NOTES

Differentiating partially with respect to *a,*

$$0 = x + \left[\frac{(a-1)1-a}{(a-1)^2}\right]y + \phi'(a)$$

$$0 = -\frac{1}{(a-1)^2}y + \phi'(a) \tag{3}$$

The elimination of *a* between equations (2) and (3) gives the general solution.

Type II Equation of the form $z = px + qy + F(p, q)$ (Clairaut's form). In this type of problems assume that, $z = ax + by + F(a, b)$ be the solution.

Example 2.14: Solve $z = px + qy + ab$

Solution: This equation is of Clairaut's type. Therefore, the complete solution is obtained by replacing *p* by *a* and *q* by *b,* where *a* and *b* are arbitrary constants.

i.e., the complete solution is, $z = ax + by + ab$ (1)

Differentiating equation (1) partially with respect to *a* and *b,* and equating these to zero we get,

$$0 = x + b \tag{2}$$

$$0 = y + a \tag{3}$$

Eliminating *a* and *b* from equations (1), (2) and (3) we get,

$$z = -xy - xy + xy$$

i.e., $z + xy = 0$

This gives the singular solution of the given partial differential equation and to get the general solution.

Put, $b = \phi(a)$ in equation (1)

$$\therefore \quad z = ax + \phi(a)y + a\phi(a) \tag{4}$$

Differentiating partially with respect to *a* we get,

$$0 = x + \phi'(a)y + a\phi'(a) + \phi(a) \tag{5}$$

Eliminating *a* from equations (4) and (5) we get the general solution.

Example 2.15: Obtain the complete solution and singular solution of,

$$z = px + qy + p^2 + pq + q^2$$

Solution: This equation is of Clairaut's form. Therefore, the complete solution is,

$$z = ax + by + a^2 + ab + b^2 \tag{1}$$

NOTES

Where, a and b are arbitrary constants.

Differentiating equation (1) partially with respect to a and b we get,

$$0 = x + 2a + b \qquad (2)$$

$$0 = y + 2b + a \qquad (3)$$

$$2x - y = 3a, \text{ and } 2y - x = 3b$$

$$a = \frac{2x-y}{3}, \; b = \frac{2y-x}{3}$$

Substituting this in equation (1) we get,

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

$$+\frac{(2x-y)(2y-x)}{9} + \left(\frac{2y-x}{3}\right)^2$$

Simplifying we get, $3z = xy - x^2 - y^2$. This is the singular solution.

To find singular integral:

Differentiating equation (2) partially with respect to a and b, and then equating to zero, we get,

$$\frac{\partial z}{\partial a} = x + \frac{a}{\sqrt{1+a^2+b^2}} = 0 \qquad (3)$$

$$\frac{\partial z}{\partial b} = y + \frac{b}{\sqrt{1+a^2+b^2}} = 0 \qquad (4)$$

From equation (3), we get,

$$x^2 = \frac{a^2}{1+a^2+b^2} \qquad (5)$$

From equation (4), we get,

$$y^2 = \frac{b^2}{1+a^2+b^2} \qquad (6)$$

From equations (5) and (6) we get,

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

$$\therefore \qquad 1 - (x^2 + y^2) = 1 - \frac{a^2+b^2}{1+a^2+b^2}$$

$$= \frac{1}{1+a^2+b^2}$$

NOTES

i.e., $$1 - x^2 - y^2 = \frac{1}{1+a^2+b^2}$$

$$\therefore \quad \sqrt{1+a^2+b^2} = \frac{1}{1-x^2-y^2} \qquad (7)$$

Substituting equation (7) in (3) and (4) we get,

$$a = \frac{-x}{\sqrt{1-x^2-y^2}},\ b = \frac{-y}{\sqrt{1-x^2-y^2}} \qquad (8)$$

Substituting equations (7) and (8) in (2) we get,

$$z = \frac{-x^2}{\sqrt{1-x^2-y^2}} - \frac{y^2}{\sqrt{1-x^2-y^2}} + \frac{1}{\sqrt{1-x^2-y^2}}$$

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

$$\therefore \quad z = \sqrt{1-x^2-y^2} \text{ or, } z^2 = 1 - x^2 - y^2$$

$$\therefore \quad x^2 + y^2 + z^2 = 1$$

This is the singular integral.

Type III Equation of the form, $F(z, p, q) = 0$

Example 2.16: Solve $z = p^2 + q^2$

Solution: Given, $z = p^2 + q^2$ (1)

Assume that, $z = f(x + ay)$ is a solution of equation (1). (2)

Put, $x + ay = u$ in equation (2)

Then, $z = f(u)$ (3)

Partially differentiating equation (3) with respect to x and y we get,

$$p = \frac{dz}{du},\ q = a\frac{dz}{du} \qquad (4)$$

$$\left(\because \frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} \text{ and } \frac{\partial z}{\partial y} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial y}\right)$$

Substituting equation (4) in (1) we get,

$$z = \left(\frac{dz}{du}\right)^2 + a^2\left(\frac{dz}{du}\right)^2$$

i.e., $$\left(\frac{dz}{du}\right)^2 (1+a^2) = z$$

NOTES

i.e., $$\frac{dz}{du} = \frac{\sqrt{z}}{\sqrt{1+a^2}}$$

i.e., $$\frac{dz}{\sqrt{z}} = \frac{du}{\sqrt{1+a^2}} \qquad (5)$$

Integrating equation (5) we get,

$$\int \frac{dz}{\sqrt{z}} = \frac{1}{\sqrt{1+a^2}}\int du$$

$$2\sqrt{z} = \frac{u}{\sqrt{1+a^2}} + b$$

i.e., $$2\sqrt{z} = \frac{x+ay}{\sqrt{1+a^2}} + b$$

This gives the complete integral.

Example 2.17: Solve $ap + bq + cz = 0$

Solution: Given, $ap + bq + cz = 0$ (1)

Let us assume that, $z = f(x + ky)$ (2)

By the solution of equation (2).

Put $x + ky = u$ in equation (2)

∴ $$z = f(u) \qquad (3)$$

$$p = \frac{dz}{du};\ q = k\frac{dz}{du} \qquad (4)$$

Substituting equation (4) in (5) we get,

$$a \cdot \frac{dz}{du} + b \cdot k\frac{dz}{du} + c \cdot z = 0$$

i.e., $$\frac{dz}{du}(a+bk) = -cz$$

∴ $$\frac{dz}{du} = -\frac{cz}{a+bk}$$

i.e., $$\frac{dz}{z} = -\frac{c}{a+bk}du \qquad (5)$$

Integrating equation (5) we get,

$$\int \frac{dz}{z} = -\frac{c}{a+bk}\int du$$

$$\log z = -\frac{c}{a+bk}(u) + \log b$$

NOTES

i.e., $\log z = A[x + ky] + \log b, \quad \text{where } A = -\dfrac{c}{a + bk}$

i.e., $\log z - \log b = A(x + ky)$

$$\log\left(\frac{z}{b}\right) = A(x + ky)$$

$$\frac{z}{b} = e^{A(x+ky)}$$

$$\therefore \quad z = be^{A(x+ky)}$$

This gives the complete integral.

Type IV Equation of the form, $F_1(x, p) = F_2(y, q)$

Example 2.18: Solve the equation, $p + q = x + y$

Solution: We can write the equation in the form, $p - x = y - q$

Let, $p - x = a$, then $y - q = a$

Hence, $p = x + a$, $q = y - a$

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy = pdx + qdy$$

$$= (x + a)dx + (y - a)dy$$

On Integrating,

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

There is no singular integral and the general integral is found as usual.

Example 2.19: Solve $p^2 + q^2 = x + y$

Solution: Given, $p^2 + q^2 = x + y$

$$p^2 - x = y - q^2 = k$$

$$\therefore \quad p^2 - x = k; \; y - q^2 = k$$

$$p = \pm\sqrt{x+k}, \; q = \pm\sqrt{y-k}$$

$$dz = pdx + qdy$$

$$= \pm(\sqrt{x+k})dx \pm (\sqrt{y-k})dy$$

Integrating we get the complete solution.

$$z = \pm\frac{2}{3}(x+k)^{3/2} \pm \frac{2}{3}(y-k)^{3/2} + C$$

$$= \pm\frac{2}{3}\left[(x+k)^{3/2} + (y-k)^{3/2}\right] + C$$

NOTES

Example 2.20: Solve $p + q = \sin x + \sin y$

Solution:

$$p - \sin x = \sin y - q = k$$

$$\therefore \quad p = k + \sin x;\ q = \sin y - k$$

$$dz = pdx + qdy$$

$$= (k + \sin x)dx + (\sin y - k)dy$$

On integrating, we get,

$$z = (kx - \cos x) - (ky + \cos y) + C$$

$$z = k(x - y) - (\cos x + \cos y) + C$$

This is the complete solution.

Check Your Progress

1. What are partial differential equations?
2. How is a partial differential equation formed?
3. Define the terms 'complete integral' and 'particular integral'.

2.3 LAPLACE EQUATION

In mathematics, Laplace equation is a second order partial differential equation. It is named after Pierre-Simon Laplace and is written as,

$$\nabla^2 \varphi = 0$$

Here ∇^2 is the Laplace operator and φ is a scalar function of 3 variables. Laplace equation and Poisson equation are examples of elliptic partial differential equations. The universal theory of solutions to Laplace equation is termed as potential theory. The solutions of Laplace equation are harmonic functions and have great important in many fields of science.

Twice differentiable real-valued functions f of real variables x, y and z are found using the following notations.

In **Cartesian coordinates**:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0$$

In **Cylindrical coordinates**:

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial f}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2} = 0$$

In **Spherical coordinates**:

$$\frac{1}{r^2}\frac{\partial}{\partial r}+\left(r^2\frac{\partial f}{\partial r}\right)+\frac{1}{r^2\sin\varphi}\frac{\partial}{\partial\varphi}\left(\sin\varphi\frac{\partial f}{\partial\varphi}\right)+\frac{1}{r^2\sin 2\varphi}\frac{\partial^2 f}{\partial\theta^2}=0$$

NOTES

The Laplace equation $\nabla^2\varphi = 0$ can also be written as $\nabla\cdot\nabla\varphi = 0$.

It is also sometimes written using the notation $\nabla\varphi = 0$, where Δ is also the Laplace operator.

Solutions of Laplace equation are harmonic functions. If the right-hand side is specified as a given function, $f(x, y, z)$ then the whole equation can be written as,

$$\nabla\varphi = f$$

This is the Poisson equation. The Laplace equation is also considered as a special type of the Helmholtz equation.

Laplace Equation in Two Dimensions

The Laplace equation in two independent variables has the form,

$$\varphi_{xx}+\varphi_{yy}=0$$

Analytic functions: Both the real and imaginary parts of a complex analytic function satisfy the Laplace equation. If $z = x + iy$ and also if $f(z) = u(x,y) + iv(x,y)$ then the necessary condition that $f(z)$ be analytic is that it must satisfy the Cauchy-Riemann equations $u_x = v_y$ and $v_x = -u_y$, where u_x is the first partial derivative of u with respect to x. It follows the notation,

$$u_{yy}=(-\mathrm{v}_x)_y=-(\mathrm{v}_y)_x=-(u_x)_x$$

Thus u satisfies the Laplace equation. Similarly it can be proved that v also satisfies the Laplace equation. Conversely, for a harmonic function it is the real part of an analytic function $f(z)$.

For a trial form, $f(z) = \varphi$ (x, y) + iψ(x, y), the Cauchy-Riemann equations is satisfied if,

$$\Psi_x=-\varphi_y,\quad \Psi_y=\varphi_x$$

This relation does not determine ψ, but only its increments as $d\psi = \varphi_y\, dx + \varphi_x\, dy$.

The Laplace equation for φ implies that the integrability condition for ψ satisfied as ψ_{xy} is ψ_{yx}; and thus ψ can be defined using a line integral. Both the integrability condition and Stokes' theorem implies that the value of the line integral connecting two points is independent of the path. The resulting pair of solutions of the Laplace equation is termed as **conjugate harmonic functions**. This construction is valid only locally or provided that the path does not loop around a singularity. There is a close connection between the Laplace equation and analytic functions which implies that any solution of the Laplace equation has derivatives of all orders and can be expanded in a power series within a circle that does not enclose a singularity. Also there is a close association between power series and Fourier series. If a function f is expanded in a power series inside a circle of radius R then it means that,

$$f(z)\sum_{n=0}^{\infty}C_n Z^n$$

NOTES

These are correctly defined coefficients whose real and imaginary parts are given as,

$$c_n = a_n + ib_n.$$

Therefore,

$$f(z) = \sum_{n=0}^{\infty}[a_n r^n \cos n\theta - b_n r^n \sin n\theta] + i\sum_{n=1}^{\infty}[a_n r^n \sin n\theta + b_n r^n \cos n\theta],$$

This is a Fourier series for *f*.

Laplace Equation in Three Dimensions

A fundamental solution of Laplace's equation satisfies the equation,

$$\Delta u = u_{xx} + u_{yy} + u_{zz} = -\delta(x-x', y-y', z-z'),$$

Here the Dirac delta function δ denotes a unit source concentrated at the points (x', y', z'). No other function has this specific property. It can be taken as a limit of functions whose integrals over space are unity and which can shrink to a point in the region where the function is non-zero. Basically, a different sign convention is taken for this equation while defining fundamental solutions. This sign is very helpful because $-\Delta$ is a positive operator. Thus, the definition of the fundamental solution implies that if the Laplacian of *u* is integrated over any volume that encloses the source point, then it is denoted as,

$$\iiint_V \nabla\cdot\nabla u\, dV = -1$$

The Laplace equation remains unchanged during rotation of coordinates and hence a fundamental solution can be obtained that only depends upon the distance *r* from the source point. For example, if we consider the volume of a ball of radius *a* around the source point, then Gauss' divergence theorem implies that,

$$-1 = \iiint_V \nabla\cdot\nabla u\, dV = \iint_S u_r\, dS = 4\pi a^2 u_r(a)$$

It can be denoted as,

$$u_r(r) = -\frac{1}{4\pi r^2},$$

It is on a sphere of radius *r* which is centered around the source point.

Thus,

$$u = \frac{1}{4\pi r}$$

2.4 POISSON EQUATION

In mathematics, Poisson equation is a partial differential equation. It is named after the French mathematician, geometer and physicist Siméon-Denis Poisson. The Poisson equation is,

$$\Delta\varphi = f$$

Here Δ is the Laplace operator and *f* and φ are real or complex-valued functions on a manifold. If the manifold is Euclidean space, then the Laplace operator is denoted as Δ^2 and hence Poisson equation can be written as,

$$\nabla^2\varphi = f$$

In three dimensional Cartesian coordinates, it takes the form:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\varphi(x, y, z) = f(x, y, z).$$

For disappearing *f*, this equation becomes Laplace equation and is denoted as,

$$\Delta\varphi = 0.$$

The Poisson equation may be solved using a Green's function; a general exposition of the Green's function for the Poisson equation is given in the article on the screened Poisson equation. There are various methods for numerical solution. The relaxation method, an iterative algorithm, is one example.

A second order partial differential equation is of the form, $\nabla^2\Psi = -4\pi\rho$. If $\rho = 0$, then it reduces to Laplace equation. It can also be considered as Helmholtz differential equation of the form,

$$\nabla^2\Psi + k^2\Psi = 0$$

2.5 WAVE EQUATION

The wave equation is an important second-order linear partial differential equation of waves. It is analysed on the basis of sound waves, light waves and water waves. The wave equation is considered as a hyperbolic partial differential equation. In its simplest form, the wave equation refers to a scalar function $u=(x_1, x_2,...,x_n,t)$ that satisfies,

$$\frac{\partial^2 u}{\partial t^2} = c^2\nabla^2 u$$

Here ∇^2 is the spatial Laplacian and *c* is a fixed constant equal to the propagation speed of the wave and is also known as the non-dispersive wave equation. For a sound wave in air at 20°C this constant is about 343 m/s (speed of sound). For a spiral spring, it can be as slow as a meter per second. The differential equations for waves are based on the speed of wave propagation that varies with the frequency of the wave. This specific phenomenon is known as dispersion. In such a case, *c* must be replaced by the phase velocity as shown below:

$$v_{\mathrm{p}} = \frac{\omega}{k}$$

The speed can also depend on the amplitude of the wave which will lead to a

NOTES

nonlinear wave equation of the form:

$$\frac{\partial^2 u}{\partial t^2} = c(u)^2 \nabla^2 u$$

A wave can be superimposed onto another movement. In that case the scalar u will contain a Mach factor which is positive for the wave moving along the flow and negative for the reflected wave.

The elastic wave equation in three dimensions describes the propagation of waves in an isotropic homogeneous elastic medium. Most of the solid materials are elastic, hence this equation is used to analyse the phenomena such as seismic waves in the Earth and ultrasonic waves which detect flaws in materials. In its linear form, this equation has a more complex form compared to the equations discussed above because it accounts for both longitudinal and transverse motion using the notation:

$$\rho \ddot{\mathbf{u}} = \mathbf{f} + (\lambda + 2\mu)\nabla(\nabla \cdot \mathbf{u}) - \mu \nabla \times (\nabla \times \mathbf{u})$$

Where:

- λ and μ are termed as Lamé parameters which describe the elastic properties of the medium.
- ρ is the density.
- **f** is the source function or driving force.
- **ü** is the displacement vector.

In this equation, both the force and the displacement are vector quantities. Hence, this equation is also termed as the vector wave equation.

General Solution of One Dimensional Wave Equation

The one dimensional wave equation for a partial differential equation has a general solution of the form that defines new variables as,

$$\xi = x - ct \quad ; \quad \eta = x + ct$$

It changes the wave equation into,

$$\frac{\partial^2 u}{\partial \xi \partial \eta} = 0$$

This leads to the general solution of the form,

$$u(\xi, \eta) = F(\xi) + G(\eta) \quad \Rightarrow \quad u(x,t) = F(x - ct) + G(x + ct)$$

Basically, solutions of the one dimensional wave equation are sums of a right traveling function F and a left traveling function G. Here the term 'Traveling' refers the shape of the individual arbitrary functions with respect to x which stays constant, though the functions are transformed left and right with time at the speed c.

As per the Helmholtz equation, named for Hermann von Helmholtz, is the elliptic partial differential equation of the form $\nabla^2 A + k^2 A = 0$, where ∇^2 is the Laplace operator, k is the wavenumber and A is the amplitude.

Check Your Progress

4. State the Laplace equation.
5. What is the other way to write a Laplace equation?
6. What is analytical function?
7. What is Poisson function?
8. How can a Poisson equation be correlated to a second order partial differential equation and a Helmholtz equation?
9. Define wave equation.

2.6 HEAT CONDUCTION EQUATION

Assuming that the temperature at any point (x, y, z) *of a solid at time t is u* (x, y, z, t), *the thermal conductivity of the solid is K, the density of the solid is* ρ *and specific heat is* σ, *the heat equation*

$$\frac{\partial u}{\partial t} = h^2 \nabla^2 u \qquad \text{... (2.1)}$$

where $h^2 = \dfrac{K}{\rho\sigma} = k$ *(say), k being known as* diffusivity, *is said to be the equation of diffusion or the Fourier equation of heat flow.*

We know that heat flows from points at higher temperature to the points at lower temperature and the rate of decrease of temperature at any point varies with the direction. In other words the amount of heat say ΔH crossing an element of surface ΔS in Δt seconds is proportional to the greatest rate of decrease of the temperature *u,* i.e.

$$\Delta H = K\, \Delta S\, \Delta t \left|\frac{du}{dt}\right| \qquad \text{... (2.2)}$$

and **v** the velocity of heat flow is given by

$$\mathbf{v} = -K \text{ grad } u = -K \cdot \nabla u \qquad \text{... (2.3)}$$

Here $u(x, y, z, t)$ is the temperature of the solid at (x, y, z) at an instant of time t and K the thermal conductivity of the solid is a positive constant in cal./cm-sec °C units.

Let S be the surface of an arbitrary volume V of the solid. Then the total flux of heat flow across S per unit time is given by,

$$H = \iint_S (-K\, \nabla u) \cdot \hat{\mathbf{n}}\, dS \qquad \text{... (2.4)}$$

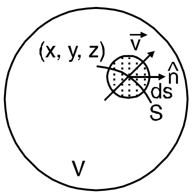

Fig. 2.1

NOTES

where $\hat{\mathbf{n}}$ is the positive outward drawn normal vector to the element dS and the negative sign shows the increase of temperature with the increase of x so that $\frac{\partial u}{\partial x}$ is positive and heat flows towards negative x from points of higher temperature to those of lower temperature, thereby rendering the flux to be negative.

Now applying Gauss's divergence theorem according to which if V be the volume bounded by a closed surface S and A be a vector function of position with continuous derivative, we have the quantity of heat entering S per unit time as

$$\iint_S (K\,\nabla u)\,\hat{\mathbf{n}}\,dS = \iiint_V \nabla\cdot(K\,\nabla u)\,dV \qquad \text{... (2.5)}$$

i.e., $$\iiint_V \nabla\cdot\mathbf{A}\,dV = \iint_S \mathbf{A}\cdot\hat{\mathbf{n}}\,dS = \oint\!\!\oint_S \mathbf{A}\cdot d\mathbf{S}. \qquad \text{... (2.6)}$$

Taking volume element $dv = dx\,dy\,dz$, the heat contained in

$$V = \iiint_V \sigma\rho u\,dV.$$

$\therefore$ The time rate of increase of heat is given by

$$\frac{\partial}{\partial t}\iiint_V \sigma\rho u\,dV = \iiint_V \sigma\rho\frac{\partial u}{\partial t}\,dV \qquad \text{... (2.7)}$$

Equating R.H.S.'s of (2.5) and (2.7), we find

$$\iiint_V \left[\sigma\rho\frac{\partial u}{\partial t} - \nabla\cdot(K\,\nabla u)\right] dV = 0. \qquad \text{... (2.8)}$$

But V being arbitrary and the integrand being assumed to be continuous the relation (2.8) will be identically zero for every point if

$$\sigma\rho\frac{\partial u}{\partial t} = \nabla\cdot(K\,\nabla u)$$

or $$\frac{\partial u}{\partial t} = \frac{K}{\sigma\rho}\nabla\cdot\nabla u = h^2\nabla^2 u = k\nabla^2 u \text{ where } h^2 = k = \frac{K}{\sigma\rho}.$$

or $$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{1}{k}\frac{\partial u}{\partial t} = \frac{1}{h^2}\frac{\partial u}{\partial t} \qquad \text{... (2.9)}$$

This is three dimensional diffusion equation.

Corollary 1. If the temperature within a substance be assumed to be independent of z, i.e., there being no heat flow in direction of z, then (2.9) reduces to

$$\frac{\partial u}{\partial t} = h^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) \qquad \text{... (2.10)}$$

which is known as two dimensional diffusion equation or the equation for two dimensional flow parallel to x-y plane.

Corollary 2. Putting $\frac{\partial u}{\partial y} = \frac{\partial u}{\partial z} = 0$ in (9) we get $\frac{\partial u}{\partial t} = h^2 \frac{\partial^2 u}{\partial x^2}$ ... (2.11)

which is the equation for the one dimensional flow of heat along a bar.

NOTES

Corollary 3. For steady-state heat flow, u is independent of time, i.e., $\frac{\partial u}{\partial t} = 0$ and hence (2.9) reduces to

$$\nabla^2 u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 \quad \text{... (2.12)}$$

which is known as three dimensional *Laplace's equation.*

One Dimensional Diffusion Equation

$$\frac{\partial \mathbf{u}}{\partial \mathbf{t}} = h^2 \frac{\partial^2 \mathbf{u}}{\partial \mathbf{x}^2} = k \frac{\partial^2 \mathbf{u}}{\partial \mathbf{x}^2}$$

[A] Independent derivation of $\frac{\partial \mathbf{u}}{\partial \mathbf{t}} = k \frac{\partial^2 \mathbf{u}}{\partial \mathbf{x}^2}$

Consider one dimensional flow of electricity in a long insulated cable and specify the current i and voltage E at any time in the cable by x-coordinate and time-variable t.

The potential drop E in a line-element δx of length at any point x is given by

$$-\delta E = iR\,\delta x + L\,\delta x \frac{\partial i}{\partial t} \quad \text{... (2.13)}$$

where R and L are respectively resistance and induction per unit length.

If C and G are respectively capacitance to earth and conductance per unit length, then we have

$$-\delta i = GE\,\delta x + C\,\delta x \frac{\partial E}{\partial t} \quad \text{... (2.14)}$$

Rewriting (2.13) and (2.14), $\frac{\partial E}{\partial x} + Ri + L\frac{\partial i}{\partial t} = 0$ ... (2.15)

and $\frac{\partial i}{\partial x} + GE + C\frac{\partial E}{\partial t} = 0$ ... (2.16)

Differentiating (2.15) with respect to. x and (2.16) with respect to. t, we have

$$\frac{\partial^2 E}{\partial x^2} + R\frac{\partial i}{\partial x} + L\frac{\partial^2 i}{\partial x \partial t} = 0 \quad \text{... (2.17)}$$

and $\frac{\partial^2 i}{\partial x \partial t} + G\frac{\partial E}{\partial t} + C\frac{\partial^2 E}{\partial t^2} = 0$ ... (2.18)

Eliminating $\frac{\partial^2 i}{\partial x \partial t}$ from (2.17) and (2.18), we get

$$\frac{\partial^2 E}{\partial x^2} = CL\frac{\partial^2 E}{\partial t^2} + LG\frac{\partial E}{\partial t} - R\frac{\partial i}{\partial x} \quad \text{... (2.19)}$$

NOTES

Again eliminating $\frac{\partial i}{\partial x}$ from (2.16) and (2.19), we find

$$\frac{\partial^2 E}{\partial x^2} = CL\frac{\partial^2 E}{\partial t^2} + (CR + GL)\frac{\partial E}{\partial t} + RGE \quad \text{... (2.20)}$$

Differentiation of (2.15) with respect to 't' and (2.16) with respect to. 'x' yields

$$\frac{\partial^2 E}{\partial x\,\partial t} + R\frac{\partial i}{\partial t} + L\frac{\partial^2 i}{\partial t^2} = 0 \quad \text{... (2.21)}$$

and
$$\frac{\partial^2 i}{\partial x^2} + G\frac{\partial E}{\partial x} + C\frac{\partial^2 E}{\partial x\,\partial t} = 0 \quad \text{... (2.22)}$$

Elimination of $\frac{\partial E}{\partial x}$ and $\frac{\partial^2 E}{\partial x\,\partial t}$ from (2.15), (2.21) and (2.22) gives

$$\frac{\partial^2 i}{\partial x^2} = CL\frac{\partial^2 i}{\partial t^2} + (CR + GL)\frac{\partial i}{\partial t} + RGi \quad \text{... (2.23)}$$

(2.19) and (2.23) follow that E and i satisfy a second order partial differential equation

$$\frac{\partial^2 u}{\partial x^2} = CL\frac{\partial^2 u}{\partial t^2} + (CR + GL)\frac{\partial u}{\partial t} + RGu \quad \text{... (2.24)}$$

which is known as *telegraphy equation.*

If the leakage to the ground is small then $G = 0 = L$ and hence (2.24) reduces to $\frac{\partial^2 u}{\partial x^2} = CR\frac{\partial u}{\partial t} = \frac{1}{k}\frac{\partial u}{\partial t}$ where $k = \frac{1}{CR}$.

which is one dimensional diffusion equation.

[B] Solution of $\frac{\partial \mathbf{u}}{\partial \mathbf{t}} = \mathbf{h}^2 \frac{\partial^2 \mathbf{u}}{\partial \mathbf{x}^2}$ or $\mathbf{u}_t = \mathbf{h}^2\, \mathbf{u}_{xx}$

Here we will discuss the solution in different conditions.

[b_1] (Both the ends of a bar at temperature zero)

If both the ends of a bar of length l are at temperature zero and the initial temperature is to be prescribed function F(x) in the bar, then find the temperature at a subsequent time t.

One dimensional heat equation is $\frac{\partial u}{\partial t} = h^2\frac{\partial^2 u}{\partial x^2}$ ... (2.25)

we have to find a function $u(x, t)$ satisfying (2.25) with the boundary conditions u

$(0, t) = u\,(l, t) = 0,\ t \geq 0$, l being the length of bar ... (2.26)

and $u\,(x, 0) \ = \ F\,(x),\ 0 < x < l$ ... (2.27)

In order to apply the method of seperation of variables, let us assume that

NOTES

$u\,(x, t) = X\,(x)\ T(t)$, X and T being respectively the function of x and t alone.

So that $\dfrac{\partial u}{\partial t} = X\dfrac{dT}{dt}$ and $\dfrac{\partial^2 u}{\partial x^2} = T\dfrac{d^2X}{dx^2}$.

Their substitution in (2.25) gives

$$\frac{1}{X}\frac{d^2X}{dx^2} = \frac{1}{h^2T}\frac{dT}{dt} \qquad \text{... (2.28)}$$

The L.H.S. and R.H.S. of (2.28) are constants because of variables being separated and hence we can write $\dfrac{1}{X}\dfrac{d^2X}{dx^2} = \dfrac{1}{h^2T}\dfrac{dT}{dt} = -\lambda^2$ (constant of separation).

Here $\dfrac{1}{X}\dfrac{d^2X}{dx^2} = -\lambda^2$ *i.e.,* $\dfrac{d^2X}{dx^2} + \lambda^2 X = 0$ gives $X = A\cos\lambda x + B\sin\lambda x$... (2.29)

and $\dfrac{1}{h^2T}\dfrac{dT}{dt} = -\lambda^2$ *i.e.,* $\dfrac{dT}{dt} + h^2\lambda^2 T = 0$ gives $T = Ce^{-\lambda^2 h^2 t}$ ... (2.30)

In view of condition (2.26), i.e., $u = 0$ at $x = 0$ or (2.29) gives $A = 0$ and λ be chosen such that $\sin\lambda l = 0$, i.e., $\lambda = \dfrac{n\pi}{l}$, n being an integer.

Hence the solution (2.25), i.e., $u = XT$ takes the form

$$u = B\sin\frac{n\pi}{l}x\,e^{-\frac{n^2\pi^2h^2}{l}t} \qquad \text{... (2.31)}$$

Summing over for all values of n, this becomes

$$u\,(x, t) = \sum_{n=1}^{\infty} B_n \sin\frac{n\pi}{l}x\,e^{-\frac{n^2\pi^2h^2}{l}t} \qquad \text{... (2.32)}$$

Applying condition (2.27) i.e., $u\,(x, 0) = F\,(x)$ at $t = 0$, we have

$$F\,(x) = \sum_{n=1}^{\infty} B_n \sin\frac{n\pi}{l}x \ \text{ for } 0 < x < l \qquad \text{... (2.33)}$$

So that
$$B_n = \frac{2}{l}\int_0^t F(x)\sin\frac{n\pi x}{l}\,dx \qquad \text{... (2.34)}$$

which is obtained by multiplying (2.33) by $\sin\dfrac{n\pi x}{l}$ and then integrating from $x = 0$ to $x = l$.

Hence the required solution is

$$u(x, t) = \frac{2}{l}\sum_{n=1}^{\infty} e^{-\frac{n^2\pi^2h^2}{l}t} \sin\frac{n\pi}{l}x \int_0^t F(u)\sin\frac{n\pi}{l}u\,du \qquad \text{... (2.35)}$$

NOTES

Deduction: (Insulated faces)

If instead of the ends of a bar of length l having kept at temperature zero, they are impervious to heat and the initial temperature is the prescribed function F(x) in the bar, then to find the temperature at a subsequent time t, we have the boundary conditions

$$\frac{\partial u}{\partial x} = 0 \text{ at } x = 0 \text{ or } l \text{ for all } t \qquad \text{... (2.36)}$$

$$u(x, 0) = F(x),\ 0 < x < l \qquad \text{... (2.37)}$$

Then the solution follows from (2.29) as

$$u = A\cos\lambda x + B\sin\lambda x$$

which in view of (2.36) requires $B = 0$ and $\sin\lambda x = 0$ i.e., $\lambda = \frac{n\pi}{l}$, $n = 0, 1, 2, 3...$

So that the general solution of the one dimensional diffusion equation will be of the form

$$u(x, t) = B_0 + \sum_{n=1}^{\infty} A_n e^{-\frac{n^2\pi^2h^2}{l}t} \cos\frac{n\pi x}{l} \qquad \text{... (2.38)}$$

where B_0 corresponds to $n = 0$.

By (2.37), this yields, $F(x) = u(x, 0) = B_0 + \sum_{n=1}^{\infty} A_n \cos\frac{n\pi x}{l}$ ... (2.39)

from which we can easily find the coefficients

$$A_n = \frac{2}{l}\int_0^l F(x)\cos\frac{n\pi x}{l}\,dx \qquad \text{... (2.40)}$$

and $B_0 = \frac{1}{2}A_0.$ ... (2.41)

Note 1. The temperature in a slab having initial temperature $F(x)$ and the faces $x = 0$, $x = \pi$ thermally insulated is given by

$$u(x, t) = B_0 + \sum_{n=1}^{\infty} B_n e^{-n^2h^2t}\cos nx. \qquad \text{... (2.42)}$$

where $A_n = \frac{2}{\pi}\int_0^\pi f(x)\cos nx\,dx$ ... (2.43)

and $B_0 = \frac{1}{2}A_0 = \frac{1}{\pi}\int_0^\pi F(x)\,dx.$ ... (2.44)

Note 2. The temperature in a slab having initial temperature $F(x)$ and the faces $x = 0$, $x = l$ thermally insulated is given by

$$u(x, t) = \frac{1}{l}\int_0^l F(x)\,dx + \frac{2}{l}\sum_{n=1}^{\infty} e^{\frac{-n^2\pi^2 h^2}{l}t} \cos\frac{n\pi x}{l}\int_0^l F(x)\frac{n\pi x}{l}\,dx \qquad \text{... (2.45)}$$

NOTES

[b_2] (One end of a bar at temperature u_0 and other at zero temperature).

If a bar of length l is at a temperature v_0 such that one of its ends x = 0 is kept at zero temperature and the other end x = l is kept at temperature u_0, then find the temperature at any point x of the bar at an instant of time t > 0.

or

A rod of length l and thermal conductivity h^2 is maintained at a uniform temperature v_0. At t = 0 the end x = 0 is suddenly cooled to 0°C by application of ice and the end x = l is heated to the temperature u_0 by applying steam, the rod being insulated along its length so that no heat can transfer from the sides. Find the temperature of the rod at any point at any time.

The equation is $\frac{\partial u}{\partial t} = h^2 \frac{\partial^2 u}{\partial x^2}$, $0 < x < l, t > 0$ ... (2.46)

With boundary conditions $u(0, t) = 0$, $u(l, t) = u_0$ for all t ... (2.47)

and $u(x, 0) = v_0$ ... (2.48)

Let the solution of (2.46) be $u(x, t) = X(x)\, T(t)$ ... (2.49)

where X is a function of x alone and T is a function of t alone.

Substituting from (2.49) $\frac{\partial u}{\partial t} = X\frac{dT}{dt}$ and $\frac{\partial^2 u}{\partial x^2} = T\frac{\partial^2 X}{\partial x^2}$, in (2.46) we get $\frac{1}{X}\frac{d^2X}{dx^2} = \frac{1}{h^2 T}\frac{dT}{dt}$ where variables are separated and hence terms on either side are constants.

Now there arise these possibilities:

[1] $\frac{d^2X}{dx^2} = 0, \frac{dT}{dt} = 0$ whence the solution is $X = Ax + B$, $T = C$ ... (2.50)

[2] $\frac{d^2X}{dx^2} = \lambda^2 x, \frac{dT}{dt} = h^2\lambda^2 T$, the solution being $X = Ae^{\lambda x} + Be^{-\lambda x}$, $T = Ce^{h^2\lambda^2 t}$

... (2.51)

[3] $\frac{d^2X}{dx^2} = -\lambda^2 x, \frac{dT}{dt} = -h^2\lambda^2 T$, the solution being $X = A\cos\lambda x + B\sin\lambda x$, $T = Ce^{-h^2\lambda^2 t}$. ... (2.52)

The combined solution in any of the three cases is $u = XT$. But $u = XT$ increases indefinitely with time t so possibility [2] is ruled out since then $u \to 0$ as $t \to \infty$.

NOTES

Conclusively the possibilities [1] and [3] determine the solution of (2.46) in the form

$$u(x, t) = u_S(x) + u_T(x, t) \qquad \text{... (2.53)}$$

where $u_S(x)$ is the temperature distribution after a long interval of time when there exists steady state of temperature and $u_T(x, t)$ is the transient effects which die down when the time passes. Consequently there exists uniform temperature after one and $x = 0$ being kept at zero temperature and the end $x = l$ at $u = u_0$ so that

$$u_S(x) = \frac{u_0}{l}x, \text{ whence (2.53) yields } u(x, t) = \frac{u_0}{l}x + u_T(x, t) \qquad \text{... (2.54)}$$

with boundary conditions $u_T(0, t) = 0 = u_T(l, t)$ by (2.47) ... (2.55)

and $$u_T(x, 0) = v_0 - \frac{u_0}{l}x \qquad \text{[by (2.48)]} \qquad \text{... (2.56)}$$

Hence the possibility [3] i.e., the solution (2.52) reduces to

$$u_T(x, t) = (A\cos\lambda x + B\sin\lambda x)\, e^{-\lambda^2 h^2 t} \qquad \text{... (2.57)}$$

whence in view of (2.47), this requires $A = 0$ and $\sin\lambda l = 0$ i.e., $\lambda = \frac{n\pi}{l}$, being an integer.

We thus obtain a solution

$$u_T(x, t) = \sum_{n=1}^{\infty} B_n\, e^{-\frac{n^2\pi^2 h^2}{l^2}t} \sin\frac{n\pi}{l}x \qquad \text{... (2.58)}$$

In view of (2.56), this gives $u_T(x, 0) = v_0 - \frac{u_0}{l}x = \sum_{n=1}^{\infty} B_n \sin\frac{n\pi}{l}x.$

$$\therefore \quad B_n = \frac{2}{l}\int_0^l \left(v_0 - \frac{u_0}{l}x\right)\sin\frac{n\pi}{l}x\,dx$$

$$= \frac{2}{n\pi}[v_0 - (-1)^n(v_0 - u_0)]$$

(on integrating by parts)

Hence the general solution of (2.46) with the help of (2.54) and (2.58) is

$$u(x, t) = \frac{u_0}{l}x + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}[v_0 - (-1)^n(v_0 - u_0)]\, e^{-\frac{n^2\pi^2 h^2}{l^2}t}\sin\frac{n\pi}{l}x \qquad \text{... (2.59)}$$

which gives temperature at any point x of the bar at any time $t > 0$.

Note. If we set $v_0 = 0$, then (2.59) takes the form

$$u(x, t) = \frac{u_0}{l}\left[x + \frac{2l}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}(-1)^n e^{-\frac{n^2\pi^2 h^2}{l^2}t}\sin\frac{n\pi}{l}x\right] \qquad \text{... (2.60)}$$

[b_3] (Temperature in an infinite bar)

If an infinite bar of small cross-section is insulated such that there is no transfer of heat at the surface and the temperature of the bar at t = 0 is given by an arbitrary function F(x) of x (taking the bar along x-axis), then find the temperature of the rod at any point of the bar at any time t.

The boundary value example is $\frac{\partial u}{\partial t} = h^2 \frac{\partial^2 u}{\partial x^2}$ ... (2.61)

With initial condition, $u(x, 0) = F(x), -\infty < x < \infty$ ... (2.62)

Let the solution be $u(x, t) = X(x)\, T(t)$ ... (2.63)

whence (2.61) gives $\frac{1}{X}\frac{d^2X}{dx^2} = \frac{1}{h^2T}\frac{dT}{dt} = -\lambda^2$ (say) ... (2.64)

Then the solution of (2.61) is

$$u(x, t) = XT = (A \cos \lambda x + B \sin \lambda x)\; e^{-h^2\lambda^2 t} \quad ... (2.65)$$

Here the arbitrary constants A and B being periodic may be taken as $A = A(\lambda)$, $B = B(\lambda)$ and due to the linearity and homogeneity of the heat equation we may write

$$u(x, t) = \int_0^\infty u(x, t, \lambda)\, d\lambda = \int_0^\infty e^{-h^2\lambda^2 t}[A(\lambda) \cos \lambda x + B(\lambda) \sin(\lambda x)]\, d\lambda \quad ... (2.66)$$

The condition (2.62) claims that

$$u(x, 0) = F(x) = \int_0^\infty [A(\lambda) \cos \lambda x + B(\lambda) \sin \lambda x]\, d\lambda$$

In view of Fourier's integrals we have

$$A(\lambda) = \frac{1}{\pi}\int_{-\infty}^{\infty} F(\mu) \cos(\mu\lambda)\, d\mu \text{ and } B(\lambda) = \frac{1}{\pi}\int_0^\infty F(\mu) \sin(\lambda\mu)\, d\mu$$

so that $u(x, 0) = \frac{1}{\pi}\int_0^\infty \left[\int_{-\infty}^{\infty} F(\mu) \cos \lambda(x-\mu)\, d\mu\right] d\lambda$

As such (2.66) takes the form

$$u(x, t) = \frac{1}{\pi}\int_0^\infty \left[\int_{-\infty}^{\infty} F(\mu) \cos \lambda(x-\mu)\, e^{-h^2\lambda^2 t}\, d\mu\right] d\lambda$$

$$= \frac{1}{\pi}\int_{-\infty}^{\infty} F(\mu)\left[\int_0^\infty e^{-h^2\lambda^2 t} \cos \lambda(x-\mu)\, d\lambda\right] d\mu \quad ... (2.67)$$

But we know that $\int_0^\infty e^{-x^2} \cos 2bx\, dx = \sqrt{\pi}\, e^{-b^2/2}$

So that $\int_0^\infty e^{-h^2\lambda^2 t} \cos \lambda(x-\mu)\, d\lambda = \frac{\sqrt{\pi}}{2h\sqrt{t}}\, e^{-\left(\frac{x-\mu}{2h\sqrt{t}}\right)^2}$

$$= \frac{\sqrt{\pi}}{2h\sqrt{t}}\, e^{-\frac{(x-\mu)^2}{4h^2 t}}$$

NOTES

Hence (2.67) gives

$$u(x, t) = \frac{1}{2h\sqrt{\pi t}} \int_{-\infty}^{\infty} F(\mu)\, e^{-\frac{(x-\mu)^2}{4h^2 t}}\, d\mu \qquad \text{... (2.68)}$$

which gives the required temperature at any point at any time.

Example 2.21. *Solve* $\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}$ *where u = 0 for t = ∞ and x = 0 or l.*

Taking $u(x, l) = X(x)\, T(t)$, the solution of the given equation is

$$X = A \cos \lambda x + B \sin \lambda x,\; T = Ce^{-\lambda^2 t}$$

with boundary conditions, $u(0, t) = 0$ and $u(x, \infty) = 0$.

Hence putting $h = 1$ the required solution is

$$u(x, t) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{l}\, e^{-\frac{n^2\pi^2}{l^2}t}$$

Example 2.22. *Solve* $\frac{\partial \theta}{\partial t} = h^2 \frac{\partial^2 \theta}{\partial x^2}$ *under the boundary conditions*

$$\theta(0, t) = \theta(l, t) = 0,\; t > 0 \qquad \text{... (1)}$$

and $$\theta(x, 0) = x,\; 0 < x < l, \qquad \text{... (2)}$$

l being the length of the bar.

$$\theta(x, t) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{l}\, e^{-n^2\pi^2 h^2/l^2\, t}$$

where, $$B_n = \frac{2}{l}\int_0^l F(x) \sin \frac{n\pi x}{l}\, dx = \frac{2}{l}\int_0^l x \sin \frac{n\pi}{l} x\, dx$$

$$= -\frac{2l}{n\pi} \cos n\pi$$

$$= \begin{cases} \dfrac{2l}{n\pi} & \text{when } n \text{ is odd} \\ \dfrac{-2l}{n\pi} & \text{when } n \text{ is even} \end{cases}$$

Hence,

$$\theta(x, t) = \frac{2l}{\pi}\left[e^{-\pi^2 h^2 t/l^2} \sin \frac{\pi}{l} - \frac{1}{2} e^{-2^2\pi^2 h^2 t/l^2} \sin \frac{2\pi}{l} x + \frac{1}{3} e^{-3^2\pi^2 h^2 t/l^2} \sin \frac{3\pi}{l} x ... \right]$$

Example 2.23. *Find the temperature u (x, t) in a bar of length l, perfectly insulated, and whose ends are kept at temperature zero while the initial temperature is given by*

NOTES

$$F(x) = \begin{cases} x, 0 < x < l/2 \\ l - x, \dfrac{l}{2} < x < l. \end{cases}$$

The boundary value example is $\dfrac{\partial u}{\partial t} = h^2 \dfrac{\partial^2 u}{\partial x^2}$.

With conditions $u(0, t) = u(l, t) = 0$ and $u(x, 0) =$

$$F(x) = \begin{cases} x, 0 < x < l/2 \\ l - x, \dfrac{l}{2} < x < l. \end{cases}$$

The required solution is

$$u(x, t) = \sum_{n=1}^{\infty} B_n\, F(x) \sin \frac{n\pi x}{l}\, dx$$

where $B_n = \dfrac{2}{l}\displaystyle\int_0^l F(x) \sin \frac{n\pi x}{l}\, dx = \frac{2}{l}\int_0^{l/2} x \sin \frac{n\pi}{l}\, x\, dx + \int_{l/2}^{l} (l - x) \sin \frac{n\pi x}{l}\, dx$

$$= \frac{4l}{n^2\pi^2} \sin \frac{n\pi}{2} = \begin{cases} \dfrac{4l}{n^2\pi^2} & \text{for } n = 1, 5, 9, \ldots \\ 0 & \text{for } n = 2, 4, 6, \ldots \\ -\dfrac{4l}{n^2\pi^2} & \text{for } n = 3, 7, 11, \ldots \end{cases}$$

Hence the solution is

$$u(x, t) = \frac{4l}{\pi^2}\left[-\frac{1}{1^2} \sin \frac{\pi x}{l} e^{-h^2\pi^2 t/l^2} - \frac{1}{3^2} \sin \frac{3\pi x}{l} e^{-3^2 h^2\pi^2 t/l^2} + \ldots\right]$$

Note. Had we considered the case of slab with its ends $x = 0$ and $n = l$ maintained at temperature zero and initial temperature being

$$F(x) = \begin{cases} T_0, & 0 < x < l/2 \\ 0, & l/2 < x < l \end{cases}$$

Then we should have

$$B_n = \frac{2}{l}\int_0^l F(x) \sin \frac{n\pi x}{l}\, dx = \frac{2T_0}{l}\int_0^l \sin \frac{n\pi x}{l}\, dx$$

$$= \frac{4T_0}{n\pi} \sin^2 \frac{n\pi}{4}$$

and the solution would be

$$u(x, t) = \frac{4T_0}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin^2 \frac{n\pi}{4}\, e^{-n^2\pi^2 t/l^2} \sin \frac{n\pi x}{l}.$$

Example 2.24. *Solve* $\dfrac{\partial u}{\partial t} = h^2 \dfrac{\partial^2 u}{\partial x^2}$, $0 < x < \pi$, $t > 0$, *under the boundary conditions* $u_x(0, t) = 0 = u_x(\pi, t)$ *and* $u_x(x, 0) = \sin x$.

$$u(x,\ t) = B_0 + \sum_{n=1}^{\infty} A_n \cos \frac{n\pi x}{l}\, e^{-n^2\pi^2h^2t/l^2} \text{ where } B_0 = \frac{A_0}{2} \text{ and } l = \pi$$

NOTES

$$= \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos nx\, e^{-n^2h^2\cdot t}$$

where $$A_n = \frac{2}{\pi}\int_0^{\pi} \sin x \cos nx\, dx = \begin{cases} 0, \text{ when } n \text{ is odd} \\ -\dfrac{4}{\pi\left(4m^2-1\right)}, \text{ when } n = 2m \end{cases}$$

and $$A_n = \frac{2}{\pi}\int_0^{\pi} \sin x\, dx = \frac{4}{\pi}.$$

Hence the required solution is

$$u(x,\ t) = \frac{2}{\pi} - \frac{4}{\pi}\sum_{n=1}^{\infty} (4m^2-1)^{-1}\, e^{-4m^2h^2t} \cos 2mx.$$

Example 2.25. *The face $x = 0$ of a slab is maintained at temperature zero and heat is supplied at constant rate at the face $x = \pi$, so that $\frac{\partial u}{\partial x} = \mu$ when $x = \pi$. If the initial temperature is zero, show that*

$$u(x,\ t) = \mu x + \sum_{j=1}^{\infty} \frac{(-1)^j}{\left(j-\frac{1}{2}\right)^2} \sin\left(j - \tfrac{1}{2}\right) x\, e^{-(j-\frac{1}{2})^2 t}$$

where the unit of time is so chosen that $k = 1$.

Taking $u(x, t)$ as the temperature of the slab, the boundary value example is

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2},\ 0 < x < \pi,\ t > 0. \quad \text{... (1)}$$

with condition $u(0, t) = 0$ ... (2)

$u(x, 0) = 0$ ... (3)

and $\frac{\partial}{\partial x} u(\pi, t) = \mu$ ... (4)

Applying the method of seperation of variables, the solutions of the given equation are

(*i*) $u = (Ae^{\lambda x} + Be^{-\lambda x})\ e^{\lambda^2 t}$

(*ii*) $u = A_1 + B_1 x$

(*iii*) $u = (A_2 \cos \pi x + B_2 \sin \pi x)\ e^{-\lambda^2 t}$

according as the constant of variation is λ^2 or $-\lambda^2$.

Here (*i*) is inadmissible as $u \to \infty$ when $t \to \infty$.

(ii) alone is inadequate to give the complete solution and hence the complete solution is given by (ii) and (iii) jointly

i.e. $$u(x, t) = u_S(x) + u_T(x, t) \quad \text{... (5)}$$

where $u_S(x)$ is the temperature distribution after a long period of time when the slab has reached the steady state of the temperature distribution and $u_T(x, t)$ denotes the transient effects which die down with the passage of time.

From (ii) $u_S(x) = A_1 + B_1 x$ ... (6)

and from (iii) $u_T(x, t) = (A_2 \cos \lambda x + B_2 \sin \lambda x)\ e^{-\lambda^2 t}$ ... (7)

Applying (2), (6) gives $A_1 = 0$ and by (4), (6) gives $\mu = B_1$

so that $u_S = \mu x$ ... (8)

Thus with the help of (7) and (8), (5) reduces to

$$u(x, t) = \mu x + (A_2 \cos \lambda x + B_2 \sin \lambda x)\ e^{-\lambda^2 t} \quad \text{... (9)}$$

Applying (2), i.e., $u(0, t) = 0$, we get $A_2 = 0$ ... (10)

Applying (4) i.e., $\mu = \dfrac{\partial}{\partial x} u(\pi, t)$,

we have $(\mu + \lambda B_2 \cos \lambda\pi)\ e^{-\lambda^2 t} = \mu$

i.e. $$\cos \lambda\pi = 0 \text{ giving } \lambda\pi = (2j - 1)\frac{\pi}{2} \text{ i.e., } \lambda = j - \frac{1}{2} \quad \text{... (11)}$$

As such $u_T(x, t) = B_j \sin\left(j - \frac{1}{2}\right) x \cdot e^{\left(j - \frac{1}{2}\right)^2 t}$ where we have set $B_j = B_2$.

Summing over all j, the general solution is

$$u_T(x, t) = \sum_{j=1}^{\infty} B_j \sin\left(j - \frac{1}{2}\right) x\, e^{-(j - \frac{1}{2})^2 t} \quad \text{... (12)}$$

Hence from (5)

$$u(x, t) = \mu x + \sum_{j=1}^{\infty} B_j \sin\left(j - \frac{1}{2}\right) x\, e^{-(j - \frac{1}{2})^2 t} \quad \text{... (13)}$$

Applying the condition (3), $0 = \mu x + \sum_{j=1}^{\infty} B_j \sin\left(j - \frac{1}{2}\right) x$

i.e. $$-\mu x = \sum_{j=1}^{\infty} B_j \sin\left(j - \frac{1}{2}\right) x \text{ so that } B_j = \frac{1}{\pi}\int_j^{\pi} (-\mu x) \sin\left(j - \frac{1}{2}\right) x\, dx$$

$$= \frac{2\mu}{\pi} \frac{(-1)^j}{\left(j - \frac{1}{2}\right)^2}$$

NOTES

Hence (13) reduces to

$$u(x, t) = \mu x + \frac{2\mu}{\pi}\sum_{j=1}^{\infty}\frac{(-1)^j}{\left(j-\frac{1}{2}\right)^2}\sin\left(j-\frac{1}{2}\right)x\, e^{-(j-\frac{1}{2})^2 t}$$

which is the required relation.

Two Dimensional Diffusion Equation

Given is the following equation:

i.e. $$\frac{\partial \mathbf{u}}{\partial \mathbf{t}} = \mathbf{h}^2\left(\frac{\partial^2 \mathbf{u}}{\partial \mathbf{x}^2}+\frac{\partial^2 \mathbf{u}}{\partial \mathbf{y}^2}\right)$$

Consider a thin rectangular plate whose surface is impervious to heat flow and which has an arbitrary function of temperature F(x, y) at t = 0. Its four edges say x = 0, x = a, y = 0, y = b are kept at zero temperature. We have to determine the subsequent temperature at a point of the plate as t increases.

The boundary value example is $\frac{\partial u}{\partial t} = h^2\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)$ ... (2.69)

Subject to the conditions for all t, (*i*) $u(0, y, t) = 0$, (*ii*) $u(a, y, t) = 0$, (*iii*) $u(x, 0, t) = 0$, (*iv*) $u(x, b, t) = 0$ and the initial condition (*v*) $u(x, y, 0) = F(x, y)$

In order to apply the method of separation of variables, let us assume that

$$u(x, y, t) = X(x)\, Y(y)\, T(t) \qquad \text{... (2.70)}$$

where X is a function of x alone, Y is a function of y alone and T is a function of t alone.

From (2.70) we have

$$\frac{\partial u}{\partial t} = XY\frac{dT}{dt}, \frac{\partial^2 u}{\partial x^2} = YT\frac{d^2X}{dx^2} \text{ and } \frac{\partial^2 u}{\partial y^2} = XT\frac{d^2Y}{dy^2}$$

Substituting them in (2.69), we find

$$\frac{1}{h^2T}\frac{dT}{dt} = \frac{1}{X}\frac{d^2X}{dx^2}+\frac{1}{Y}\frac{d^2Y}{dy^2}, \text{ after dividing by } XYT \qquad \text{... (2.71)}$$

In (2.71), the variables being separated, we can assume

$$\frac{1}{X}\frac{d^2X}{dx^2} = -\lambda_1^2; \frac{1}{Y}\frac{d^2Y}{dy^2} = -\lambda_2^2 \text{ and } \frac{1}{h^2T}\frac{dT}{dt} = -\lambda^2$$

so that $$\lambda^2 = \lambda_1^2 + \lambda_2^2 \qquad \text{... (2.72)}$$

The general solutions of (2.72) are

$X = A\cos\lambda_1 x + B\sin\lambda_1 x$; $Y = C\cos\lambda_2 y + D\sin\lambda_2 y$:

$$T = Ee^{-\lambda^2 h^2 t} \qquad \text{... (2.73)}$$

NOTES

So that with the help of (2.73), (2.70) gives the solution of (2.69) in the form

$u(x, y, t) = (A \cos \lambda_1 x + B \sin \lambda_1 x)\,(C \cos \lambda_2 y + D \sin \lambda_2 y)\; e^{-\lambda^2 h^2 t}$ ... (2.74)

In view of condition (*i*), $0 = u(0, y, t) = A\,(C \cos \lambda_2 y + D \sin \lambda_2 y)\; e^{-\lambda^2 h^2 t}$, giving $A = 0$.

In view of condition (*ii*), we claim $\sin \lambda_1 x = 0$ i.e. $\lambda_1 = \frac{m\pi}{a}$, m being an integer.

Similarly applying conditions (*iii*) and (*iv*) to (2.74), we get $C = 0$ and $\lambda^2 = \frac{n\pi}{b}$, n being an integer. As such (2.73) takes the form

$$u_{mn}(x, y, t) = B_{mn}\, e^{-\lambda_{mn}^2 h^2 t} \sin \frac{m\pi}{a} x \sin \frac{n\pi}{b} y$$

where $$\lambda^2 = \lambda^2_{mn} = \pi^2 \left(\frac{m^2}{a^2} + \frac{n^2}{b^2} \right).$$

Summing over all the possible values of m and n, the general solution is

$$u(x, y, t) = \sum_{m, n=1}^{\infty} B_{mn}\, e^{-\lambda_{mn}^2 h^2 t} \sin \frac{m\pi}{a} x \sin \frac{n\pi}{b} y \qquad \text{... (2.75)}$$

where B_{mn} are arbitrary constants to be determined by the condition (*v*)

i.e., $$F(x, y) = u(x, y, 0) = \sum_{m,n=1}^{\infty} B_{mn} \sin \frac{m\pi}{a} x \sin \frac{n\pi}{b} y \qquad \text{... (2.76)}$$

Multiplying both sides of (2.76) by $\sin \frac{m\pi}{a} x \sin \frac{n\pi}{b} y\, dx\, dy$ and integrating with regard to x from 0 to a and with regard to y from 0 to b we get on using orthogonality properties of the sines,

$$B_{mn} = \frac{4}{ab} \int_0^a \int_0^b F(x, y) \sin \frac{m\pi}{a} x \sin \frac{n\pi}{b} y\, dx\, dy \qquad \text{... (2.77)}$$

which gives the arbitrary constants of (2.75).

Example 2.26. *A rectangular plate bounded by the lines $x = 0$, $y = 0$, $x = a$, $y = b$ has an initial distribution of temperature given by $F(x, y) = B \sin \frac{\pi x}{a} \sin \frac{\pi y}{b}$. The edges are maintained at zero temperature and the plane faces are impervious to heat. Find the temperature at any point at any time.*

The general solution is

$$u(x, y, t) = \sum_{m, n=1}^{\infty} B_{mn}\, e^{-\lambda_{mn}^2 h^2 t} \left\{ \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} \right\}$$

where $B_{mn} = \frac{4}{ab} \int_0^a \int_0^b F(x, y) \sin \frac{m\pi}{a} x \sin \frac{n\pi}{b} y\, dx\, dy$

$$= \frac{4B}{ab}\int_0^a\int_0^b \sin\frac{\pi x}{a}\sin\frac{\pi y}{b}\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}\,dx\,dy$$

NOTES

$\because\ F(x, y) = B\sin\frac{\pi x}{a}\sin\frac{\pi y}{b}$

$$\therefore\quad B_{m1} = \frac{4B}{ab}\int_0^a \frac{b}{2}\sin\frac{\pi x}{a}\sin\frac{m\pi x}{a}\,dx$$

$$\because\quad \int_0^b \sin\frac{\pi y}{b}\sin\frac{n\pi y}{b}\,dy = \begin{cases} 0, & \text{for } n = 2, 3, 4, \ldots \\ b/2, & \text{for } n = 1. \end{cases}$$

So that $B_{11} = B$.

Also $\lambda_{11}^2 = \pi^2\left(\frac{1}{a^2}+\frac{1}{b^2}\right)$

Hence the solution is,

$$u(x, y, t) = B_{11}\,e^{-\lambda_{mn}^2 h^2 t}\sin\frac{\pi x}{a}\sin\frac{\pi y}{b}$$

$$= B\,e^{-h^2\pi^2\left(\frac{1}{a^2}+\frac{1}{b^2}\right)}\sin\frac{\pi x}{a}\sin\frac{\pi y}{b}.$$

Example 2.27. *A semi-infinite plate having width π has its faces insulated. The semi-infinite edges are kept at 0°C while the infinite edge is maintained at 100°C. Assuming that the initial temperature is 0°C, find the temperature at any point at any time.*

Taking the diffusivity, i.e., $h^2 = 1$, the boundary value example is

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \qquad \ldots (1)$$

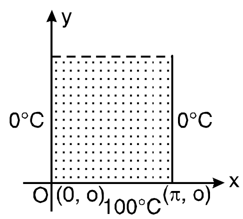

Fig. 2.2

with conditions (*i*) $u(0, y, t) = 0$

(*ii*) $u(\pi, y, t) = 0$, (*iii*) $u(x, y, o) = 0$;

(*iv*) $u(x, o, t) = 100$ and

(*v*) $|u(x, y, t)| < M$

where $0 < x < \pi$, $y > 0$, $t > 0$.

Taking Laplace transform of (1) and assuming $L\{u(x, y, t)\} = U(x, y, t)$,

NOTES

We have $\dfrac{\partial^2 U}{\partial x^2} + \dfrac{\partial^2 U}{\partial y^2} = sU$ [by using condition (*iii*)] ... (2)

But the finite Fourier sine transform of a function $F(x)$, $0 < x < l$ is defined as

$$f_s(x) = \int_0^l F(x) \sin \frac{n\pi x}{l}\, dx,\ n \text{ being an integer} \quad \text{... (3)}$$

Multiplying (2) by sin nx and integrating from 0 to π, we get

$$\int_0^\pi \frac{\partial^2 U}{\partial x^2} \sin nx\, dx + \int_0^\pi \frac{\partial^2 U}{\partial y^2} \sin nx\, dx = \int_0^\pi sU \sin nx\, dx$$

Setting $\tilde{U} = \int_0^\pi U \sin nx\, dx$, this beocmes

$$-n^2 \tilde{U} + nU(\pi, y, s) \cos n\pi + nU(o, y, s) + \frac{d^2\tilde{U}}{dy^2} = s\tilde{U} \quad \text{... (4)}$$

But from the Laplace transforms of conditions (*i*) and (*ii*), we have

$$U(o, y, s) = 0,\ U(\pi, y, s) = 0$$

$\therefore$ (4) reduces to $\dfrac{d^2\tilde{U}}{dy^2} - (n^2 + s)\tilde{U} = 0$ ... (5)

Its solution is $\tilde{U} = Ae^{y\sqrt{n^2+s}} + Be^{-y\sqrt{n^2+s}}$ ... (6)

By condition (*v*), $\tilde{U}$ being bounded, as $y \to \infty$, we have $A = 0$ so that (6) yields, $\tilde{U} = Be^{-y\sqrt{n^2+s}}$ .. (7)

Applying condition (*iv*),

$$\tilde{U}(n, o, s) = \int_0^\pi \frac{100}{s} \sin nx\, dx = \frac{100}{s}\left(\frac{1 - \cos n\pi}{n}\right) \quad \text{... (8)}$$

In (7) if we put $y = 0$, we get with the help of (8),

$$B = \tilde{U} = \frac{100}{s}\left(1 - \frac{\cos n\pi}{n}\right)$$

Hence $\tilde{U} = \dfrac{100}{s}\dfrac{(1 - \cos n\pi)}{n} e^{-y\sqrt{n^2+s}}$.

Applying Fourier sine inversion formula, we find

$$U = \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{100}{s}\left(\frac{1 - \cos n\pi}{n}\right) e^{-y\sqrt{n^2+s}} \sin nx \quad \text{... (9)}$$

Now we have $L^{-1}\left\{e^{-y\sqrt{s}}\right\} = \dfrac{y}{2\sqrt{\pi t^3}} e^{-\frac{y^2}{4t}}$

NOTES

so that $$L^{-1}\left\{e^{-y\sqrt{s+n^2}}\right\} = \frac{y}{2\sqrt{\pi t^3}}\, e^{-\frac{y^2}{4t}}\, e^{-n^2 t}$$

Thus $$L^{-1}\left\{\frac{e^{-y\sqrt{s+n^2}}}{s}\right\} = \int_0^t \frac{y}{2\sqrt{\pi t^3}}\, e^{-y^2/4v}\, e^{-n^2 v}\, dv$$

$$= \frac{2}{\sqrt{\pi}} \int_{y/2\sqrt{t}}^{\infty} e^{-(p^2 + n^2 y^2/4p^2)}\, dp \text{ where } p^2 = \frac{y^2}{4v}.$$

Hence taking the inverse Laplace transform of (9) term by term, we get

$$u(x, y, t) = \frac{400}{\pi^{3/2}} \sum_{n=1}^{\infty} \left(\frac{1 - \cos n\pi}{n}\right) \sin nx \int_{y/2\sqrt{t}}^{\infty} e^{-(p^2 + n^2 y^2/4p^2)}\, dp.$$

2.7 DEFINITION OF GROUP

A group is the simplest algebraic structure found in nature wherever symmetry exists.

A group (G, o) is a system consisting of a non-empty set G such as $G = \{a, b, c, ...\}$ and a binary operation 'o' satisfying the following axioms:

G_1—(*Closure*). If $a \in G$, $b \in G$ then $aob \in G$ or in other words if $a \in G$, $b \in G$ then $aob = c$ (closure) where $c \in G$.

G_2—(*Associativity*). If $a, b, c \in G$ then $ao\,(boc) = (aob)oc$.

G_3—(*Existence of an identity*). If $a \in G$, then $\exists$ an identity element $e \in G$ s.t. $eoa = a \;\forall\; a \in G$

G_4—(*Existence of an inverse*). If $a \in G$, $\exists$ an inverse $a^{-1} \in G$ s.t. $a^{-1}\, oa = e$

where $e \in G$, being an identity element.

In addition to these four axioms if a fifth axiom of commutativity namely G_5 is also satisfied, i.e.,

G_5—(*Commutation*). If $a, b \in G$ the $aob = boa$ so that G_3 and G_4 take the forms

$$G_3: eoa = aoe = a$$

$$G_4: aoa^{-1} = a^{-1}oa = e$$

then the group is said to be an *Abelian group*.

for example, the set **I** (of all integers) with the binary operation 'o' taken as additive (+) is a group, for it satisfies all the four axioms

G_1—if $a, b \in$ **I**, then $a + b \in$ **I**

G_2— if $a, b, c, \in$ **I**, then $a + (b + c) = (a + b) + c$

NOTES

G_3—if $a \in \mathbf{I}$ then $\exists$ an integer zero (0) such that $0 + a = a$

G_4— if $a \in \mathbf{I}$, then $\exists$ an inverse $(-a)$ s.t. $-a + a = 0$ (identity element).

This group is also abelian or commutative as G_5 is also satisfied, i.e., $G_5 - a, b \in \mathbf{I}$, $a + b = b + a$.

Finite and Infinite Groups: A group (G, o) consisting of a finite number of elements is said to be a *finite group*, for example, the set $S = \{1, \omega, \omega^2\}$ where $\omega^3 = 1$, is a finite group under multiplication composition.

A group (G, o) consisting of an infinite number of elements is said to be an *infinite group*, for example, the set **I** (of all integers) is an infinite group under the addition composition.

Order of a Group: The number of elements in a finite group is known as the order of the group. The infinite set is said to be of infinite order. As an example the set $\{1, -1\}$ under multiplication composition is a group of order 2.

Elementary Properties of a Group

I. **Uniqueness of Identity**, i.e., *the identity element of a group* (G, o) *is unique*.

If possible let us assume that e and e' are two identity elements of the group (G, o), then

$$eoa = aoe = a \ \forall \ a \in G \quad \text{... (2.78)}$$

and

$$e'oa = aoe' = a \ \forall \ a \in G \quad \text{... (2.79)}$$

Putting $a = e'$ in Equation (2.78), we have

$$eoe' = e'oe = e' \quad \text{... (2.80)}$$

Also putting $a = e$ in Equation (2.78), we have

$$e'oe = eoe' = e \quad \text{... (2.81)}$$

From Equations (2.80) and (2.81) it follows that $e' = e$, i.e., there cannot be two identity elements for (G, o) and hence the identity element of a group is unique.

II. Uniqueness of Inverse, i.e., *in a group* (G, o) *every element possesses a unique inverse.*

If possible let us assume that a' and a^{-1} are two inverses of a. Also let $a \in G$ and e be the identity element of G. Then we have

$$a^{-1}oa = aoa^{-1} = e \quad \text{...(2.82)}$$

$$a'oa = aoa' = e \quad \text{...(2.83)}$$

Post-multiplying Equation (2.82) by a' we get

$$(a^{-1}oa)oa' = (aoa^{-1})oa' = eoa' = a' \quad \text{... (2.84)}$$

and premultiplying Equation (2.83) by a^{-1} we get

$$a^{-1}o(a'oa) = a^{-1}o(aoa') = a^{-1}oe = a^{-1} \quad \text{... (2.85)}$$

But group-postulate G_2 gives

$$(a^{-1}oa)oa' = a^{-1}o(aoa')$$

NOTES

$\therefore$ Equation (2.84) and (2.85) follow that $a' = a^{-1}$, i.e., the inverse of an element in a group is unique.

Aliter. Taking a^{-1} and a' two inverses of $a \in G$, we have

$$aoa' = a'oa = e \text{ and } a^{-1}oa = aoa^{-1} = e$$

$$\therefore \quad a^{-1} = a^{-1}oe = a^{-1}o(aoa') = (a^{-1}oa)oa' = eoa' = a'$$

which follows the uniqueness of inverse of $a \in G$.

III. Cancellation Laws, i.e., *for any group* (G, o), *and* $a, b, c \in G$, *the following laws hold*

(*i*) $aob = aoc \Rightarrow b = c$ (*left cancellation law*)

(*ii*) $boa = coa \Rightarrow b = c$ (*right cancellation law*)

(*i*) Taking $a^{-1} \in G$ as the inverse of $a \in G$, we have

$aob = aoc \Rightarrow a^{-1}o(aob) = a^{-1}o(aoc)$ on premultiplying by a^{-1}

$\Rightarrow (a^{-1}oa)ob = (a^{-1}oa)oc$, the composition being associative by G_2

$\Rightarrow eob = eoc$, since $\exists$ an identity element $e \in G$ for 'o'

$\Rightarrow b = c.$

(*ii*) Again taking $a^{-1} \in G$ as the inverse of $a \in G$, we have

$boa = coa \Rightarrow (boa)oa^{-1} = (coa)oa^{-1}$ on postmultiplying by a^{-1}

$\Rightarrow bo(aoa^{-1}) = co(aoa^{-1})$ by G_2

$\Rightarrow boe = coe \quad \because \exists$ an identity element $e \in G$ for 'o'

$\Rightarrow b = c.$

Note. $aoc = cob \Rightarrow a = b$ unless the group is abelian.

IV. Uniqueness of Solutions, i.e., *if* $a, b \in G$, *then the equations* $aox = b$ *and* $yoa = b$ *have unique solutions in* G.

If a^{-1} be the inverse of $a \in G$, then $a^{-1} \in G$ and $aoa^{-1} = e$ (identity element).

Now $a^{-1} \in G$ and $b \in G \Rightarrow a^{-1}\, ob \in G$

Putting $x = a^{-1}ob$ in the equaiton $aox = b$, we get

$$ao(a^{-1}ob) = b$$

or $(aoa^{-1})ob = b$ by G_2

or $eob = b$, i.e., $b = b$

which follows that $x = a^{-1}ob$ is a solution of $aox = b$.

To show that this solution is unique, let us assume that y is an element different from $a^{-1}ob$ in G s.t. it satisfies the equation $aox = b$. Then,

$$aoy = b = eob = (aoa^{-1})ob = ao(a^{-1}ob) \text{ by } G_2.$$

so that left cancellation law yields $y = a^{-1}ob$.

NOTES

As such $x = y$, i.e., the solution is unique.

Again, $b \in G$ and $a^{-1} \in G \Rightarrow boa^{-1} \in G$

Putting $y = boa^{-1}$ in $yoa = b$, we get

$(boa^{-1})oa = b$ or $bo(a^{-1}\ oa) = b$ by G_2

or $\qquad boe = b$, i.e., $b = b$

which follows that $y = boa^{-1}$ is a solution of $yoa = b$.

To show that this solution is unique, let us assume that z is an element different from boa^{-1} in G s.t. it satisfies the equation $yoa = b$.

Then $\quad zoa = b = boe = bo\,(a^{-1}oa) = (boa^{-1})\ oa$ by G_2.

The right cancellation law gives $z = boa^{-1}$

so that $y = z$ and hence the solution is unique.

Note. The unique solution of $xox = x$ is $x = e$ in group (G, o)

V. Inverse of the Inverse is Itself, i.e., *if* $a \in G$ *then* $(a^{-1})^{-1} = a$.

Inverse law gives $(a^{-1})^{-1}oa^{-1} = e$, e being identity element in G.

Postmultiplying by a, we get

$$[(a^{-1})^{-1}oa^{-1}]oa = eoa$$

or $\qquad (a^{-1})^{-1}o(a^{-1}oa) = a$ by G_2

or $\qquad (a^{-1})^{-1}oe = a$ by G_4

or $\qquad (a^{-1})^{-1} = a$ by G_3

which proves the proposition.

VI. Reversal Law, i.e., if $a, b \in G$ *then* $(aob)^{-1} = b^{-1}\ oa^{-1}$.

Let e be the identity element in G.

Now $a \in G \Rightarrow a^{-1} \in G$ and $b \in G \Rightarrow b^{-1} \in G$

$\therefore \quad (b^{-1}oa^{-1})o(aob) = b^{-1}o[a^{-1}o(aob)]$ by G_2

$= b^{-1}o[(a^{-1}oa)ob]$ by G_2

$= b^{-1}o[eob]$ by G_4

$= b^{-1}ob$ by G_3

$= e.$

Hence by definition of inverse element of a group $b^{-1}oa^{-1}$ is the inverse of aob, i.e.,

$$(aob)^{-1} = b^{-1}oa^{-1}.$$

Note. The result may be generalized for any number of elements

$a_1, a_2, a_3, \ldots\ldots a_n \in G$, where we have

$(a_1oa_2o\ldots.oa_n)^{-1} = a_n^{-1}oa_{n-1}^{-1}\,o\ldots.oa_2^{-1}oa_1^{-1}$.

Some Definitions

Semi-Group: A set S with a binary operation 'o' is said to be a semi-group if it satisfies the following two axioms:

S_{G1}—(*Closure*). $a \in S,\ b \in S \Rightarrow aob \in S$.

S_{G2}—(*Associativity*). If $a, b, c \in S$ then $(aob)\ oc = ao\ (boc)$.

Theorem 2.1. *A semi-group* (G, o) *satisfying the following postulates is a group*:

(1) *G has a left identity e s.t.* $eoa = a\ \forall\ a \in G$.

(2) *Every element a in G has a left inverse* a^{-1} *in G s.t.* $a^{-1}oa = e$.

Since (G, o) is a semi-group, therefore by definition it follows that

(*i*) (G, o) satisfies the closure law.

(*ii*) (G, o) satisfies the associative law.

(*iii*) a^{-1} being the left inverse of a and e the left identity we have

$$\begin{aligned} a^{-1}o(aoe) &= (a^{-1}oa)oe && \text{by } (ii) \\ &= eoe && \text{by Postulate (2)} \\ &= e && \text{by Postulate (1)} \\ &= a^{-1}oa && \text{by Postulate (2)} \end{aligned}$$

So that by left cancellation law, $aoe = a$

which follows that e is also a right identity.

Hence the identity element exists for the composition.

(*iv*) We have $a^{-1}o(aoa^{-1}) = (a^{-1}oa)oa^{-1}$ by (*ii*)

$$\begin{aligned} \text{or} \qquad a^{-1}o(aoa^{-1}) &= eoa^{-1} \text{ by (2)} \\ &= a^{-1} \text{ by (1)} \\ &= a^{-1}oe, \text{ since identity element exists for } `o\text{'}, \\ \therefore \qquad aoa^{-1} &= e \text{ by left cancellation law.} \end{aligned}$$

which shows that a^{-1} is also a right inverse of a.

Hence every element of G has an inverse.

Since all the four group axioms are satisfied, therefore a semi-group with the given two postulates is a group.

Sub-Group: *A sub-group of a group* (G, o) *is any collection of elements of G satisfying the axioms of G. In other words, a non-empty subset say H of a group G is said to be the sub-group of G, if the binary operation 'o' in G induces a binary operation in H and the elements of H obey the group axioms.*

In other words, *a* non-empty subset H of a group G is said to be a sub-group of G if it satisfies the following two axioms:

(*i*) $a, b \in H \Rightarrow aob \in H$,

(*ii*) $a \in H \Rightarrow a^{-1} \in H$

for example, the set of even integers is a sub-group of the additive group of integers.

Proper Sub-Group: A sub-group of a group (G, o) other than G itself and the group consisting of the identity element alone is termed as a proper sub-group of

NOTES

G. for example, the additive group of integers is a proper sub-group of the additive group of rational numbers.

Improper or Trival Sub-Groups: The group (G, o) itself and the group consisting of identity alone, i.e., ($\{e\}$, o) are known as trivial or improper sub-groups of (G, o).

Order of an Element of a Group. Let a be an element of a group (G, o), i.e., $a \in G$. Then the order of a is the least positive integer n s.t. $a^n = e$.

In case $\nexists$ such integer, the order of a is said to be zero or infinite. for example, the order of the element –1 in the multiplicative group $\{1, -1, i, -i\}$ is 2 since $(-1)^2 = 1$, the identity element. The order of i is 4 since $i^4 = 1$.

Addition Modulo m, (m being an integer). If a and b are two integers and m a positive integer, then 'addition modulo m' is denoted by $a +_m b$ and defined as $a +_m b = r$, $o \le r < m$ where r is the least positive remainder obtained on dividing the sum of a and b by m.

for example $12 +_3 5 = 2$ since $12 + 5 = 3(5) + 2$

and $-5 +_4 10 = 1$ since $-5 + 10 = 4(1) + 1$

Multiplication Modulo p, (p being a prime). If a and b are two integers and p, a positive integer, then 'multiplication modulo p', is denoted by $a \times_p b$ and defined as

$a \times_p b = r$, $o \le r < p$ where r is the remainder obtained on dividing the ordinary product ab by p.

for example, $9 \times_6 7 = 3$ since $9 \times 7 = (6)\ 10 + 3$

and $-7 \times_5 8 = 4$ since $-7 \times 8 = (5)(-12) + 4$

Group Table or Composition Table: It is commonly observed that a 'table' is a convenient way of either defining a binary operation in a finite set S or tabulating the effect of a binary operation in a set S. In forming a table or say a group table we arrange the elements of a group in rows and columns of a square array such that each element of the group occurs once and only once in each row or column. The composition element aob occurs at the intersection of row and column of the elements a and b of the group after the binary operation has been performed. For example consider a set

$S = \{1, 2, 3\}$ and let '·' be the binary operation in S defined by

$$\therefore (1, 1) \to 1, (1, 2) \to 2, (1, 3) \to 3, (2, 1) \to 2, (2, 2) \to 1, (2, 3) \to 2,$$
$$(3, 1) \to 3, (3, 2) \to 2, (3, 3) \to 1$$

then these operations can be arranged in a table as follows:

It is clear that (i, j)th square $(i, j = 1, 2, 3)$ is the intersection of ith row (i.e., row labelled or faced by i) and jth column (i.e., column labelled or faced by j) and in this square we have put the element obtained by the binary operation '·' on (i, j) such as, $\therefore (1, 2) \to 2$

As another example if $a, b \in G$ and e be an identity element of the group (G, o), such that $aoa = b$ and $aob = e$ etc., then the group table is as shown here.

NOTES

Problem 2.1. *Show that three cube roots of unity form an abelian finite group under multiplication.*

We have the set $G = \{1, \omega, \omega^2\}$, where $\omega^3 = 1$.

The composition table under multiplication is as shown here.

The set G forms an abelian finite group, since it satisfies all the five axioms:

G_1—since all the elements in group table belong to G, hence closure axiom is true.

G_2—since multiplication of complex numbers is associative, therefore G_2 is satisfied.

G_3—since $\exists$ an identity element $1 \in G$, G_3 is satisfied.

G_4—since the inverses of 1, ω, ω^2 are respectively 1, ω^2, $\omega \in G$, G_4 is satisfied.

G_5—commutative property is apparently satisfied since $1 \cdot \omega = \omega \cdot 1 = \omega$ etc.

Moreover, the set consists of finite number of elements and hence G_1 is an abelian finite group.

Problem 2.2. *Show that the set of all nth roots of unity form a finite abelian group G of order n under ordinary multiplication as composition.*

By De Movre's theorem, nth roots of unity are given by

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

$$= \cos\frac{2r\pi}{n} + i\sin\frac{2r\pi}{n}, \text{ where } r = 0, 1, 2,......, n-1$$

So n, nth roots of unity are

$$1,\ \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}, \cos\frac{2.2\pi}{n} + i\sin\frac{2.2\pi}{n},....., \cos\frac{(n-1)\cdot 2\pi}{n} + i\sin\frac{(n-1)\cdot 2\pi}{n}$$

i.e. $1, e^{1.2\pi i/n}, e^{2.2\pi i/n}, e^{3.2\pi i/n},....., e^{(n-1)\ .\ 2\pi i/n}$

Now, G_1 is satisfied since the product of any two elements of the set is the element of the set such as if $a = e^{p\cdot 2\pi i/n}$, $b = e^{q\cdot 2\pi i/n} \in G$, where $0 \le \pi \le n - 1$, $0 \le q \le n-1$, then $a\ .\ b = e^{2\pi i/n\ (p+q)}$ will belong to G if $p + q \le n-1$. Let us assume the contrary, i.e., $p + q > n - 1$ so that $p + q = n + m$ where $m \le n - 2$ since the maximum value of $p + q$ can be $2(n-1)$, i.e., $2n - 2$.

$$\therefore\quad a \cdot b = e^{2\pi i/n(n+m)} = e^{2\pi i}\ e^{2\pi im/n} = e^{2\pi im/n}$$

$$\because\quad e^{2\pi i} = \cos 2\pi + i \sin 2\pi = 1$$

which follows that $a \cdot b \in G$ since $m \le n - 2$.

G_2 is satisfied since multiplication of complex numbers is associative.

G_3 is satisfied since there exists an identity element $e^{2\pi i\cdot o/n} = 1$

G_4 is satisfied since $\exists$ inverse of $e^{2\pi ir/n}$ as $e^{2\pi i(n-r)n}$ since

$$e^{2\pi ir/n}\ e^{2\pi i(n-r)/n} = e^{2\pi in/n} = e^{2\pi i} = 1$$

G_5 is also satisfied since the multiplication of complex numbers is commutative.

Moreover the set consists of finite number of elements. Hence (G, o) is a finite abelian group.

Problem 2.3. *If every element of a group (G, o) is its own inverse then show that (G, o) is abelian.*

Given that (G, o) is a group.

$\therefore$ if $a, b \in G$ then $a^{-1}, b^{-1} \in G$

also if $aob \in G$ then $(aob)^{-1} \in G$

But we have $a = a^{-1}$ and $b = b^{-1}$

As such $(aob) = (aob)^{-1} = b^{-1}\ oa^{-1} = (boa)$

i.e., (G, o) is commutative. Hence (G, o) is abelian.

Problem 2.4. *If (G, o) be a group and $a^2 = e$ (identity) $\forall\ a \in G$, then show that the group must be commutative.*

Given that (G, o) is a group and $a^2 = aoa = e$ also $aoa^{-1} = e$

$\therefore$ $aoa = e = aoa^{-1}$

So that left cancellation law gives $a = a^{-1}$

i.e., every element of the group is its own inverse and hence by Problem 9.3 it follows that the group (G, o) is commutative.

Problem 2.5. *Show that if a group has 3, 4 or 5 elements, then it is abelian.*

We prove the proposition for 4 elements, similar procedure can be adopted for other two.

Suppose that $G = \{e, a, b, c\}$ is a set forming the group $\{G, o\}$ where e is the identity element.

Let $a^{-1} = b$. Then the only alternative is that $c^{-1} = c$, so that

$aob = boa = e$ and $coc = e$ ... (1)

Now

$aoc \neq e$ as $c^{-1} \neq a$

$aoc \neq a$ as $c \neq e$

$aoc \neq c$ as $a \neq e$

So the only alternative is that $aoc = b$.

Similar argument will give that $coa = b$

$\therefore$ $aoc = coa$ ... (2)

Also $boc \neq e$ as $b^{-1} \neq c$

$boc \neq b$ as $c \neq e$

$boc \neq c$ as $b \neq e$

leading that $boc = a$ and similarly $cob = a$

$\therefore$ $boc = cob$ ... (3)

Equations (1), (2) and (3) follow that the group (G, o) is commutative and hence it is abelian. The group table is as shown here.

NOTES

Problem 2.6. *Show that any non-commutative group has at least six elements.*

Let (G, o) be a non-commutative group. It will be so if it has at least one pair of non-commuting elements a and b (say).

We shall first show that a set $\{e, a, b, aob, boa\}$ having a, b non-commuting elements, i.e., $aob \neq boa$, consists of distinct elements. Taking two at a time, there are ten possibilities leading to a contradiction of $aob \neq boa$:

(*i*) $e = a \Rightarrow aob = eob = b = boe = boa.$

(*ii*) $e = b \Rightarrow aob = aoe = a = eoa = boa$

(*iii*) $e = aob \Rightarrow aoe = eoa = (aob)\ oa = ao\ (boa)$

i.e., $e = boa$ or $aob = boa$

(*iv*) $e = boa \Rightarrow eoa = aoe = ao\ (boa) = (aob)\ oa$, i.e., $e = aob$ or $boa = aob$

(*v*) $a = b \Rightarrow aob = aoa = boa$

(*vi*) $a = aob \Rightarrow e = b$ thereby reducing to (*ii*)

(*vii*) $a = boa \Rightarrow e = b$ thereby reducing to (*ii*)

(*viii*) $b = aob \Rightarrow e = a$ thereby reducing to (*i*)

(*ix*) $b = boa \Rightarrow e = a$ thereby reducing to (*i*)

(*x*) $aob = boa$

Hence the elements of the set $\{e, a, b, aob, boa\}$ having (a, b) non-commuting, are all distinct.

We shall now show that at least one of the group elements aoa or $ao\ boa$ is distinct from these five namely, e, a, b, aob, boa.

To show that aoa is different from each element a, b, aob, boa, we see that

(*xi*) $aoa = a \Rightarrow a = e$ thereby reducing to (*i*)

(*xii*) $aoa = b \Rightarrow aob = ao\ (aoa) = (aoa)oa = boa$

(*xiii*) $aoa = aob \Rightarrow a = b$ thereby reducing to (*v*)

(*xiv*) $aoa = boa \Rightarrow a = b$ thereby reducing to (*v*)

These possibilities lead that either $aoa \neq e$ in which case aoa is the sixth distinct element of G or else $aoa = e$.

Again we shall show that $ao\ boa$ is different from each element e, a, b, aob, boa so that it will be the sixth distinct element of G.

Obviously $ao\ (ao\ boa) = (aoa)\ o\ (boa) = eo\ (boa) = boa.$

Now consider the case

(*xv*) $ao\ boa = e \Rightarrow boa = ao\ (ao\ boa) = aoe = o$ thereby reducing to (*vii*)

(*xvi*) $ao\ boa = a \Rightarrow aob = e$ thereby reducing to (*iii*)

(*xvii*) $ao\ boa = b \Rightarrow aob = ao\ (ao\ boa) = boa$ when $aoe = e$

(*xviii*) $ao\ boa = aob \Rightarrow a = e$ thereby reducing to (*i*)

(*xix*) $ao\ boa = boa \Rightarrow a = e$ thereby reducing to (*i*)

Conclusively a group with upto 5 elements is essentially abelian but for it to be non-abelian there should be at least six elements.

Problem 2.7. *Show that non-empty semi-group* (G, o) *forms a group if the equations* $ax = b$ *and* $ya = b$ *have unique solutions in* G $\forall$ *pair of elements* $a, b \in G$.

Since $ya = b$ is solvable for any $b \in G$, therefore by taking $b = a$, we find that $ya = a$ has a solution in G. Call this solution as e_1 so that $e_1\, a = a$ where a is a fixed element of G.

Let $c \in G$, then $ax = c$ has a solution in G.

Thus $e_1 c = e_1\,(ax) = (e_1 a)x = ax = c$

which follows that $e_1\, c = c\ \forall\ c \in G$, i.e., e_1 is the left identity in G.

As e_1 exists in G, so $ya = e_1$ has a solution in G. Call this solution as a^{-1}. This follows that every element in G has a left inverse relative to the left identity. It follows that (G, o) is a group.

Problem 2.8. *Show that a finite-non-empty semi-group* (G, o) *forms a group if* $ab = ac \Rightarrow b = c$ *and* $ba = ca \Rightarrow b = c\ \forall\ a, b, c \in G$.

Consider a set $G = \{a_1, a_2, ...a_r, ..., a_p\}$ consisting of p distinct elements. Take an element a_m and multiply it to all the elements of this group,

$$a_m a_1,\ a_m a_2, ...a_m a_r, ..., a_m a_p.$$

All these elements will be distinct save possibly arranged in different order. If possible let us assume that

$$a_m a_r = a_m a_p \Rightarrow a_r = a_p$$

which contradicts the hypothesis that a_r and a_p are distinct elements of G. Thus

$G = \{a_m a_1,\ a_m a_2, ...,\ a_m a_r,\ ....\ a_m a_p\}$ consists of p distinct elements and $a_m a_1$ will be some element say a_r of G, i.e.,

$$a_m a_1 = a_r \Rightarrow ax = b \text{ has a unique solution in } G$$

Similarly we can show that

$G = \{a_1\, a_m,\ a_2 a_m, ..., a_r a_m, ..., a_p a_m\} \Rightarrow ya = b$ has a unique solution in G.

The semi-group (G, o) under given conditions forms a group.

Problem 2.9. *Show that the set of subsets of a set with the union composition is a semi-group.*

If $S_1 = \{A, B, C, ...\}$ be the set of subsets of a set S, then

S_{G1} is satisfied since, $A, B \in S$, and $A \subset S, B \subset S \Rightarrow A \cup B \subset S$ and $A, B \in S_1, \Rightarrow A \cup B \in S_1$, i.e., the closure law is satisfied.

S_{G2} is satisfied since if $A, B, C \in S$, then associative property of union yields, $(A \cup B) \cup C = A \cup (B \cup C)$

Hence S_1 is a semi-group.

Problem 2.10. *Show that the identity of a subgroup of a group is the same as that of the group.*

Let (H, o) be a subgroup of the group (G, o) and let e, e' be the identities of (G, o) and (H, o) respectively. Then

$$aoe' = a \ \forall \ a \in H$$

This equality will also hold in (G, o) as $a \in H \Rightarrow a \in G$.

Now if b be the inverse of $a \in G$, then we have

$$aoe' = a \Rightarrow bo(aoe')$$

$$\Rightarrow (boa)oe' = boa \text{ by } G_2 \text{ for } G$$

$$\Rightarrow eoe' = e \qquad boa = e$$

$$\Rightarrow e' = e.$$

Problem 2.11. *Show that the inverse of an element of a subgroup of a group is the same as the inverse of the same element regarded as an element of the group.*

Let (H, o) be a subgroup of the group (G, o) and let b_1 and b_2 be the inverses of an element a as member of H and G respectively. Also let e and e' be the identities of G and H respectively. $e = e'$,

Now $aob_1 = e' = e \Rightarrow b_2o(aob_1) = b_2oe$

$$\Rightarrow (b_2oa)\ ob_1 = b_2 \text{ by } G_2, G_3 \text{ for } G.$$

$$\Rightarrow eob_1 = b_2 \qquad \because\ b_2oa = e$$

$$\Rightarrow b_1 = b_2.$$

Problem 2.12. *Show that the necessary and sufficient conditions for a complex H to be a subgroup* (H, o) *of a group* (G, o) *are*

(i) $a, b \in H \Rightarrow aob \in H \ \forall \ a, b$; *and* (ii) $a \in H \Rightarrow a^{-1} \in H \ \forall \ a$

The *conditions are necessary*, since (H, o) being a subgroup of (G, o) the composition in H (being also the composition in G) satisfies the closure law

i.e. $$a, b \in H \Rightarrow aob \in H \ \forall \ a, b$$

which proves the first condition.

Also by Problem 2.10, the identity of H being the same as that of G and by Problem 2.11, the inverse of any element of H being the same as its inverse in G, we have

$$a \in H \Rightarrow a^{-1} \in H \ \forall \ a$$

which proves the second condition.

The *conditions are also sufficient*, since if the conditions (i) and (ii) hold then

G_1 is satisfied, for $a, b \in H \Rightarrow aob \in H$ by condition (i)

G_2 is satisfied, for $a, b \in H \Rightarrow aob \in H$ by (i) leads to

$$aob, c \in H \text{ and } a, boc \in H \ \forall \ a, b, c \in H$$

$\Rightarrow$ the same element $aoboc \in H$, i.e., associative law is satisfied.

G_3 is satisfied since $a \in H \Rightarrow a^{-1} \in H$ by (ii) leads to

$$a \in H \quad \text{and} \quad a^{-1} \in H \Rightarrow aoa^{-1} \in H \text{ by } (i)$$

But $aoa^{-1} = e$, (identity of G)

$\therefore$ $e \in H$ is an identity in H, which is also identity in G, thereby showing the existence of an identity element in H.

NOTES

G_4 is satisfied since from G_3 and condition (*ii*), every element of H has an inverse.

Hence (H, o) which is a sub-group of the group (G, o) satisfies all the four exioms of group.

Problem 2.13. *Show that a necessary and sufficient condition for a complex H to be a subgroup (H, o) of a group (G, o) is that $a \in H, b \in H \Rightarrow aob^{-1} \in H$.*

The *condition is necessary*, since when (H, o) is a subgroup of (G, o) then by condition (*ii*) of Problem 2.12, we have $b \in H \Rightarrow b^{-1} \in H$.

$a, b^{-1} \in H \Rightarrow aob^{-1} \in H.$

Combining these two conditions we have $a \in H, b \in H \Rightarrow aob^{-1} \in H.$

The *condition is sufficient*, since if $a, b \in H \Rightarrow aob^{-1} \in H$, then we can show as below that (H, o) is a subgroup of (G, o).

The given condition yields,

$$a \in H, e \in \text{H} \Rightarrow aoa^{-1} = e \in H, e \text{ being identity of } G.$$

This follows that G_3 is satisfied, i.e., $\exists$ an identity $e \in H$.

Also $\quad e \in H, a \in H \Rightarrow eoa^{-1} = a^{-1} \in H$

i.e., G_4 is satisfied or in other words every element in H is invertible.

As such any $b \in H \Rightarrow b^{-1} \in H$

So that $a \in H, b^{-1} \in H \Rightarrow ao\ (b^{-1})^{-1} = aob \in H$

which follows that H satisfies closure law under 'o', i.e., G_1 is satisfied.

Now associativity of G w.r.t. 'o' immediately follows the associativity of H w.r.t. 'o', i.e., G_2 is satisfied.

Hence (H, o) is a group.

But $\quad (H, o)$ is a subset of (G, o).

Therefore (H, o) is a sub-group of (G, o).

Problem 2.14. *Show that the intersection of two subgroups of a group (G, o) is a subgroup of (G, o)*

Let (H_1, o) and (H_2, o) be the subgroups of (G, o), then

$$H_1 \cap H_2 \subset G.$$

Now $a, b \in H_1 \cap H_2 \Rightarrow a, b \in H_1, a, b \in H_2$

$\Rightarrow aob \in H_1, aob \in H_2$ since H_1, H_2 being subgroups,

satisfy group axioms.

$\Rightarrow aob \in H_1 \cap H_2 \ \forall a, b \in H_1 \cap H_2$

Also $\quad a \in H_1 \cap H_2 \Rightarrow a \in H_1$ and $a \in H_2$

$\Rightarrow a^{-1} \in H_1$ and $a^{-1} \in H_2$ since H_1, H_2 being subgroups satisfy group axioms.

$\Rightarrow a^{-1} \in H_1 \cap H_2.$

NOTES

Hence by Problem 2.12, $H_1 \cap H_2$ is a subgroup of G.

Problem 2.15. *Show that the union of two subgroups of a group* (G, o) *may not be subgroup of* G.

Let (H_1, o) and (H_2, o) be the two subgroups of (G, o) and let

$$a \in H_1,\ b \in H_2, \text{ so that } a, b \in H_1 \cap H_2.$$

Now $a, b \in H_1 \cup H_2 \Rightarrow a \in H_1, b \in H_2 \not\Rightarrow aob \in H_1 \cup H_2$ for b may not belong to H_1.

Hence the union of two subgroups of a group may not be subgroup of the group.

Problem 2.16. *Show that the order of any power of any element a of a group is utmost equal to the order of the element.*

Assuming that order $a = m$ and order of $(a^p) = n$, $p \in \mathbf{I}$ (set of integers), we have order of $a = m$

$\Rightarrow$ $a^m = e$, e being identity element.

$$\Rightarrow (a^m)^p = e^p$$

$$\Rightarrow a^{mp} = e$$

$$\Rightarrow (a^p)^m = e$$

$$\Rightarrow \text{order of } (a^p) \le m$$

which proves the proposition.

Problem 2.17. *Show that the order of any element of a group is always equal to the order of its inverse.*

Taking the orders of a and a^{-1} as m and n respectively, we have

$$a^m = e \text{ and } (a^{-1})^n = e$$

Now a^{-1} being an exponent power of a, the Problem 9.16 leads to order of $(a^{-1}) \le$ order of a, i.e., $n \le m$.

Also since $a = (a^{-1})^{-1}$, i.e., a is an exponent power of a^{-1}, so by Problem 2.16, we have order of $a \le$ order of (a^{-1}), i.e., $m \le n$.

Hence $m \le n$ and $n \le m \Rightarrow m = n$.

The Centre of a Group

If (G, o) *be a group and H be the set of those elements* $x \in G$, *which commute with each element in G,* i.e., *the set*

$$H = \{x : x \in G \text{ and } aox = xoa \ \forall \ a \in G\}$$

then the set H is known as the **centre** *of G.*

Theorem 2.2. *The centre of G is a subgroup of* (G, o).

If H be the centre of G, then we have by definition

$$H = \{x : x \in G \text{ and } aox = xoa \ \forall \ a \in G\}$$

$\therefore \quad x_1, x_2 \in H \Rightarrow aox_1 = x_1oa$ and $aox_2 = x_2oa \ \forall \ a \in G.$

But $aox_1 = x_1oa = x_1o(x_2^{-1}ox_2)oa$, since $x_2^{-1}ox_2 = e$, the identity in H and

$$x_1oeoa = x_1oa$$
$$= (x_1ox_2^{-1})\ o\ (x_2oa)$$
$$= (x_1ox_2^{-1})\ o\ (aox_2) \qquad \because \quad aox_2 = x_2oa.$$

NOTES

$$\therefore \qquad aox_1 = (x_1ox_2^{-1})\ o\ (aox_2) \Rightarrow (aox_1)ox_2^{-1} = (x_1ox^{-2})\ oa$$
$$\Rightarrow ao(x_1ox_2^{-1}) = (x_1ox_2^{-1})\ oa$$
$$\Rightarrow x_1ox_2^{-1} \text{ commutes with } a \in G$$
$$\Rightarrow x_1ox_2^{-1} \in H$$

Conclusively $x_1 \in H, x_2 \in H \Rightarrow x_1\ o\ x_2^{-1} \in H.$

Which follows by the definition of a subgroup that (H, o) is a subgroup of (G, o).

Cosets or Cosets of a Subgroup

Let (G, o) *be a group*, (H, o) *be a subgroup of* (G, o) *and* 'a' *be an element in G,* i.e., $a \in G$. *Then the set*

$$aH = \{ah : h \in H\} \text{ (not using the binary operation)}$$

i.e., *the collection,*

$$aoH = \{aoh_1, aoh_2..., aoh_i,...\}, h_i \in H$$
$$= \{aox : x \in H \text{ and } a \in G\}$$

is said to be the **Left Coset** *of H in G*;

and the set $Ha = \{ha : h \in H\}$

i.e., *the collection,* $Hoa = \{h_1oa, h_2oa,...., h_ioa,...\}, h_i \in H$

$$= \{xoa : x \in H \text{ and } a \in G\}$$

is said to be the **Right Coset** *of H in G*.

Since $eH = H = He$, therefore H is itself a coset.

If the cosets aH and bH are such that $aH \cap bH \neq \phi$, then $aH = bH$, hence the cosets have no elements in common with H, i.e., two cosets contain either the same elements or have no elements in common. Also the cosets do not form a group.

The number of left (or right) cosets of H in G is said to be the *Index* of H in G and denoted by $(G : H)$.

Theorem 2.3. *If* (H, o) *or simply H be a subgroup of* (G, o) *or simply G, then H is both a left coset and a right coset.*

If e be the identity in G, then $He = eH = H$, which follows that H is both a left coset as well as a right coset of H in G.

Theorem 2.4. *If H be a subgroup of G, then* $aH = H \Leftrightarrow a \in H$.

If e be the identity in G and so is in H, then

$$aH = H \Rightarrow ae \in H$$

i.e. $\qquad aH = H \Rightarrow a \in H \qquad$ ... (2.86)

Again, $\qquad$ if $a \in H$ and $h \in H$ then

$$a \in H \Rightarrow ah \in H\ \forall\ h \in H$$

NOTES

$\therefore \quad aH \subset H$

Also $a \in H \Rightarrow a^{-1} \in H$, H being a subgroup of the group G, satisfies group axioms.

$\Rightarrow a^{-1}\, h \in H \;\forall\; h \in H$ by closure law in H

$\Rightarrow a\,(a^{-1}h) \in H \;\forall\; h \in H$ by closure law in H

$\Rightarrow h \in aH \;\forall\; h \in H$

$\therefore \quad H \subset aH$

So $aH \subset H$ and $H \subset aH \Rightarrow aH = H$

Ultimately $\quad a \in H \Rightarrow aH = H$ ... (2.87)

Hence $\quad aH = H \Leftrightarrow a \in H$ by Equations (2.86) and (2.87).

Theorem 2.5. *If $a, b \in G$ and $a \neq b$ then $aH = bH \Leftrightarrow a^{-1}b \in H$ where H is a subgroup of the group G.*

We have,

$aH = bH \Rightarrow a^{-1}aH = a^{-1}bH$

$\Rightarrow (a^{-1}a)H = (a^{-1}b)H$

$\Rightarrow eH = (a^{-1}b)H$, e being the identity in G and so in H.

$\Rightarrow H = (a^{-1}b)H$

$\therefore \quad aH = bH \Rightarrow a^{-1}b \in H$ ... (2.88)

Also, if $a^{-1}\,b \in H$, then

$bH = e(bH) = (aa^{-1})(bH) = a(a^{-1}b)H = bH$ ... (2.89)

Equations (2.88) and (2.89) follow that $aH = bH \Leftrightarrow a^{-1}b \in H$.

Theorem 2.6. *The two left cosets aH and bH of a subgroup H of a group G are either identical or disjoint.*

There arise two cases:

Case I. If $aH \neq bH$, then we have to show that aH and bH are disjoint.

Let us assume if possible that $x \in aH$ and $x \in bH$.

Then $\quad x = ay, y \in H$ and $x = bz, z \in H$

$\therefore \quad ay = bz \Rightarrow ayz^{-1} = bzz^{-1} \Rightarrow \quad a(yz^{-1}) = b(zz^{-1}) = be = b$

$\Rightarrow (a^{-1}a)\,(yz^{-1}) = a^{-1}b$

$\Rightarrow e(yz^{-1}) = a^{-1}b$

$\Rightarrow yz^{-1} = a^{-1}b$

Thus, $yz^{-1} \in H \Rightarrow a^{-1}b \in H$.

So that by Theorem 2.5, it follows that $aH = bH$, which contradicts the hypothesis and hence two unequal cosets cannot have any element in common, i.e., aH and bH are disjoint.

Case II. If aH and bH are not disjoint, then we have to show that $aH = bH$.

aH and bH are not disjoint $\quad \Rightarrow \exists$ an element common to aH and bH

$\Rightarrow \exists\ h_i, h_j$ s.t. $ah_i = bh_j$
$\Rightarrow a(h_i h_i^{-1}) = bh_j h_i^{-1}$
$\Rightarrow a = bh_j h_i^{-1}$
$\Rightarrow ah = b(h_j h_i^{-1} h)\ \forall\ h \in H$
$\Rightarrow ah \in bh\ \forall\ h \in H$
$\Rightarrow aH \subset bH$ ... (2.90)

Similarly it can be shown that

$$ah_i = bh_j \Rightarrow bH \subset aH \qquad \text{... (2.91)}$$

Equations (2.90) and (2.91) follow that $aH = bH$, i.e., aH and bH are identical.

Theorem 2.7. *If H be a subgroup of the group G and $a \in G$ but $a \notin H$ then $\exists$ one-one mapping of H onto aH.*

Taking $f : H \to aH$ defined by $f(h) = ah$, $h \in H$, we have to show that the map f is onto.

Every element of the left coset aH being of the form ah, $h \in H$, and so being the f-image of h in H, the mapping f is onto.

Again to show that f is one-one, let $h_i, h_j \in H$ s.t. $ah_i = ah_j$.

Then $ah_i = ah_j \Rightarrow h_i = h_j$ by left cancellation law.

So f is one-one.

Conclusively f is a one-one mapping of H onto aH.

Note. This theorem follows that if H be a finite subgroup, the number of elements in each of its left cosets is the same as the number of elements in H, i.e., equal to the order of H.

Theorem 2.8. (*Lagrange's Theorem*). *The order of every subgroup of a finite group is a divisor of the order of the group.*

Let H be a subgroup of a finite group G. So being finite, H is also finite.

Let m and n be the order of H and G respectively.

Since the order of H is m, therefore H consists of exactly m elements or in other words every coset aH has exactly m elements, for if $h_1, h_2, \in H$, $ah_1 = ah_2$ iff $h_1 = h_2$, hence aH has the same number of elements as H.

Now if $m = n$, the theorem is self-evident.

But if $n > m$, then G being of finite order, there are only a finite number say k, of different cosets of H in G.

Taking $H = \{h_1, h_2 ...., h_m\}$, if $a \in G$ but $a \notin H$ and binary operation of G being denoted multiplicatively, the distinct m elements

$ah_1, ah_2, ..., ah_m \notin H$ but belong to G by closure axiom.

Denoting the set formed by these m elements by H, i.e.,

$H' = \{ah_1, ah_2 ..., ah_m\}$

we observe that if $H \cup H'$ is a proper subset of G then there is an element say $b \in G$ s.t. $b \notin H \cup H'$. We thus have again a set of m distinct elements

$bh_1, bh_2 .... bh_m$ which belong to G but not to $H \cup H'$

NOTES

Denoting the set of these m elements by H'', i.e.,

$$H'' = \{bh_1, bh_2, ..., bh_m\}$$

and continuing this process, we see that G can be divided into k subsets each consisting of m elements.

$$\therefore \quad \text{Order of } G = \text{number of elements in } G$$

i.e.,
$$n = k \times \text{order of } H = km$$

which follows that the order of H is a divisor of the order of G.

Corollary 1. *The order of an element of a group G of finite order is a divisor of the order of the group.*

Let m be the order of the group G and $a \in G$. Then by definition,

$a^m = e$, m being least positive integer and e being identity in G.

Evidently the elements $a, a^2, a^3, ... a^{m-1}, a^m \in G$, are all distinct and form a subgroup of order m. Also by definition, m is the order of a.

Therefore, the order m of a is a divisor of the order of the group.

Corollary 2. *A finite group of prime order has no proper subgroup.*

Let G be a finite group of order p, where p is a prime. Then by Lagrange's theorem, the order of any subgroup of G is divisor of p. But p being prime has no divisor and hence there is no proper sub-group of G.

Corollary 3. (*Fermats' Theorem*). *If p be a prime and 'a' a natural number not divisible by p, then*

$$a^{p-1} = 1 \ (mod\ p)$$

Taking the multiplicative group of non-zero residue classes modulo p and a not divisible by p, we have the equivalence class $a \neq 0$, i.e., $[a] \neq 0$.

But the order of the group being $p - 1$, it follows from Cor. 1, that

$$[a^{p-1}] = [1]$$

which yields $\quad a^{p-1} = 1 \ (\text{mod } p)$.

Problem 2.18. *If H be a subgroup of a group G and m, n are the orders of m and n respectively then prove that $a^n = e$, e being identity in G.*

Lagrange's theorem gives $n = km$, k being some positive integer.

$$\therefore \quad a^n = a^{km} = (a^m)^k = e^k = e.$$

Problem 2.19. *Find the cosets of the additive subgroup* (2**I**, +) *of the additive group* (**I**, +), **I** *being set of all integers.*

We have

$$\mathbf{I} = \{...-3, -2, -1, 0, 1, 2, 3...\}$$

and say $H = (2\mathbf{I}, +) = \{... -6, -4, -2, 0, 2, 4, 6...\}$

If $a \in \mathbf{I}$ then the coset of H in $\mathbf{I}$ corresponding to a is $2\mathbf{I} + a$ since the group being abelian, $\mathbf{I} + a = a + \mathbf{I}$

$$\therefore \quad 2\mathbf{I} + 0 = \{..., -6, -4, -2, 0, 2, 4, 6,....\}$$

NOTES

$$2\mathbf{I} + 1 = \{..., -5, -3, -1, 1, 3, 5, 7,....\}$$

$$2\mathbf{I} + 2 = \{..., -4, -2, 0, 2, 4, 6, 8,...\} = 2\mathbf{I}$$

$$2\mathbf{I} + 3 = \{..., -3, -1, 1, 3, 5, 7, 9,....\} = 2\mathbf{I} + 1$$

$$2\mathbf{I} + 4 = \{..., -2, 0, 2, 4, 6, 8, 10,....\} = 2\mathbf{I}$$

$$2\mathbf{I} + 5 = \{...., -1, 1, 3, 5, 7, 9, 11, ...\} = 2\mathbf{I} + 1 \text{ and so on.}$$

Thus the distinct cosets of H in $\mathbf{I}$ are $2\mathbf{I}$ and $2\mathbf{I} + 1$.

Clearly $2\mathbf{I} \cup (2\mathbf{I} + 1) = \mathbf{I}$.

Homomorphism and Isomorphism.

Homomorphism of Groups: *If* (G, o) *and* (G', o') *be two groups, then a mapping* $f : G \to G'$ *which retains the structure and is many one is called* **Homomorphism of the Group** G *with the group* G' *s.t.*

$$f(aob) = f(a)\, o'\, f(b),\ \forall\ a, b \in G.$$

We sometimes use to say that G is homomorphic to G' and denote it by $G \simeq G'$ if $\exists$ a mapping $f : G \to G'$ s.t. $f(aob) = f(a)\, o'\, f(b)\ \forall\ a, b \in G$.

Properties of Homomorphism

(1) *The group* (G', o') *is a homomorphic image of the group* (G, o).

(2) *The relation of homomorphism is not symmetric,* i.e.,

$$G \simeq G' \not\Rightarrow G' \simeq G$$

(3) *The homomorphic image of the identity of the group* (G, o) *is the identity of the group* (G', o') i.e. *if* e, e' *be the identities in* G, G' *respectively then* $f(e) = e'$.

We have $a \in G \Rightarrow f(a) \in G'$

and $f(aoe) = f(a)\, o'\, f(e)\ \forall\ a \in G$ by definition of homomorphism.

$\therefore\ f(a)\, o'\, e = f(a) = f(aoe) = f(a)\, o'\, f(e)$ since $aoe = a$

and $f(a)\, o'\, e' = f(a)$

So left cancellation law gives $e' = f(e)$.

(4) *The homomorphic image of the inverse of any element a of a group* (G, o) *is the inverse of the image of a,* i.e., $f(a^{-1}) = [f(a)]^{-1}\ \forall\ a \in G$

We have $a^{-1}, a \in G \Rightarrow f(a^{-1}), f(a) \in G'$

$\therefore\ f(a^{-1})\, o'\, f(a) = f(a^{-1}oa)$, by definition of homomorphism

$= f(e) = e'$ by Property (3)

But $f(a^{-1})\, o'\, f(a) = e' \Rightarrow f(a^{-1}) = [f(a)]^{-1}\quad \because\ f(a), f(a^{-1}) \in G'$

Isomorphism of Groups: *If* (G, o) *and* (G', o') *are two groups and* $\exists$ *a one-one onto mapping* $f : G \to G'$ *s.t.* $aob \xrightarrow{\text{mapped that}} a'\, ob'$ *where* $a \to a'$, $b \to b'$, $\forall\ a, b \in G$ *and* $a', b' \in G'$, *then the mapping f is called as* **Isomorphism** *and we say that G is* **isomorphic** *to* G' *and write* $G \cong G'$.

for example, if G is an additive group of all integers, i.e.,

$$G = \{...., -4, -3, -2, -1, 0, 1, 2, 3, 4,....\}$$

NOTES

and G' is a multiplicative group of all positive and negative powers of an integer 2 i.e.

$$G' = \{2^m : m = 0, \pm 1, \pm 2....\}$$

$$= \left\{....\frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 2, 4, 8, 16,.....\right\}$$

Then we have $f(m) = 2^m$, m being an integer

and $f(m + n) = 2^{m+n} = 2^m \cdot 2^n = f(m) \cdot f(n)$, m, n being integers.

This shows that f is one-one onto and retains the group structure and hence $G \cong G'$.

Properties of Isomorphism

(*i*) *The order of G = the order of G′*

(*ii*) *For isomorphic groups* (G, o) *and* (G', o') *the identity* e' *of* G' *is the image of identity* e *of* G, i.e., $f(e) = e'$.

If $a \in G$ and $a' \in G'$ then $a' = f(a)$

$\therefore \quad f : G \to G'$ is one-one onto $\Rightarrow f(a) \in G' \ \forall \ a \in G$.

$$\Rightarrow f(e) \in G' \quad \because e \in G$$

Now $aoe = a \Rightarrow f(aoe) = f(a)$

$\Rightarrow f(a)\ o'\ f(e) = f(a)\ o'\ e'$ by definition of isomorphism

$\Rightarrow a'\ o'\ f(e) = a'\ o'\ e'$

$\Rightarrow f(e) = e'$ by left cancellation law.

(*iii*) *For isomorphic groups* (G, o) *and* (G', o') *the image of inverse of any element a of G is the inverse of the image of a,* i.e.

$$f(a^{-1}) = [f(a)]^{-1}$$

If e, e' are identities of G, G' respectively then by property (*ii*) $f(e) = e'$

Also we have $a^{-1}oa = e = aoa^{-1} \ \forall \ a \in G$

But

$a^{-1}oa = e \Rightarrow f(a^{-1}oa) = f(e) \ \forall \ a \in G$

$\Rightarrow \quad f(a^{-1})o'\ f(a) = e'$ by definition of isomorphism

$\Rightarrow \quad f(a^{-1}) = [f(a)]^{-1}$ by definition of inverse of an element in G'

(*iv*) *For isomorphic groups* (G, o) *and* (G', o'), *the order of an element a* $\in$ *G is the same as the order of its image* $a' \in G'$.

$$f : G \to G' \text{ is one-one and onto.}$$

If e, e' be identities in G, G' respectively, then

$$f(e) = e' \text{ and } f(aob) = f(a)\ o'\ f(b) \ \forall \ a, b \in G.$$

If n be the order of an element $a \in G$ then $a^n = e$

Also if m be the order of $f(a)$ then $[f(a)]^m = e'$

But $\quad a^n = e \Rightarrow f(a^n) = f(e)$

$\Rightarrow f(aoaoa...n \text{ times}) = e'$

NOTES

$\Rightarrow f(a)\ o'\ f(a)\ o'\ ....n$ times $= e'$ by definition of isomorphism

$\Rightarrow [f(a)]^n = e'$

$\Rightarrow$ order of $f(a) \leq n$

$\Rightarrow m \leq n.$

Also $[f(a)]^m = e' \Rightarrow f(a)\ o', f(a)\ o', ...m$ times $= f(e)$

$\Rightarrow f(aoaoa...m$ times$) = f(e)$ by definition of isomorphism

$\Rightarrow f(a^m) = f(e)$

$\Rightarrow a^m = e \quad \because f$ is one-one

$\Rightarrow$ order of $a \leq m$

$\Rightarrow n \leq m$

So that $m \leq n$ and $n \leq m \quad \Rightarrow m = n$

$\Rightarrow$ order of a = order of a'.

(v) *If f is isomorphic mapping of $G \to G'$, then f^{-1} is also isomorphic.*

If f is one-one and onto then f^{-1} exists and is also one-one onto.

Also if $x = f(a),\ y = f(b)$ for $a, b \in G$ and $x, y \in G'$, then

$$a = f^{-1}(x),\ b = f^{-1}(y)$$

But $\quad f^{-1}(xo'y) = f^{-1}[f(a)\ o'\ f(b)]$

$= f^{-1}[f(aob)] \because f$ is isomorphic mapping

$\Rightarrow aob \quad \because f^{-1} f(p) = p.$

$\Rightarrow f^{-1}(x)\ o\ f^{-1}(y)$

which follows that f^{-1} retains the group structure and hence f^{-1} is isomorphic.

Automorphism of Groups: *An isomorphism of a group onto itself is said to be an* **automorphism** *of the group* for example $f : G \to G'$ *given by* $f(a) = a^{-1}$, $a \in G$ *is an automorphism iff G is an abelian group.*

In other words an automorphism f of G is a one-one transformation of G onto itself s.t. $(xy)f = (xf)(yf)\ \forall\ x, y \in G$

i.e., $\quad f(xy) = f(x)f(y)$

As another example the identity mapping $i : G \to G$ is an automorphism of group G.

Product of Automorphisms: If $x \leftrightarrow xf = x'$ be an automorphism of A where x' is the element of A in some order, then the mapping is automorphism and so $(xy)f = (xf)(yf) = x'y'$.

Take $x \leftrightarrow z'$ another automorphism and denote z' by $f\phi$, so that

$$(xy)f\phi = [(xy)f]\phi = [(xf)(yf)]\phi = [(xf)\phi][(yf)\phi]$$
$$= [(x)f\phi][(y)f\phi]\ \forall\ x, y \in A$$

which shows that $f\phi$ is an automorphism of A and the mapping $f\phi$ is termed as product of automorphisms of f and ϕ.

The automorphism of a mathematical system forms a group.

NOTES

The mapping $x \leftrightarrow x$ is said to be the *Identity automorphism* in the identity element of the automorphism group. So axiom G_3 is satisfied.

G_1 is satisfied since product of two automorphisms is an automorphism.

G_2 is satisfied since if we arrange the mappings

$$f : x \leftrightarrow x',\ \phi : x \leftrightarrow x'',\ \psi : x \leftrightarrow x'''$$

as

$$f : x \leftrightarrow x',\ \phi : x' \leftrightarrow x'',\ \psi : x'' \leftrightarrow x'''$$

then $x'' = (x) f \phi\psi$ corresponding to x under the automorphism $f \phi\psi$ is uniquely determined whether it is obtained as $[(x) f]\ \phi\psi$ from the automorphism $f(\phi\psi)$ or as $[(x)\phi]\psi$ from $(f\phi\psi)$. Ultimately $x \leftrightarrow xf$ is an automorphism and so

$$\begin{aligned}(xy) f^{-1} &= [(xf^{-1}f)\ (yf^{-1}f)] f^{-1} \\ &= [(xf^{-1})\ (yf^{-1}) f] f^{-1} \\ &= (xf^{-1})\ (yf^{-1})\end{aligned}$$

showing that f^{-1} is an automorphism and hence G_4 is satisfied.

Conclusively the automorphisms of a mathematical system form a group.

Endomorphism of Groups: *A homomorphism of a group onto itself is said to be an* **endomorphism** *of the group.*

Regular Permutation Group: *A permutation group to which a group G is isomorphic is said to be a* **regular permutation group**.

Theorem 2.9. (*Transference of Group Structures*)

If (G, o) is a group and G' is a set with the multiplicative composition 'o' and if $\exists$ a one-one onto mapping $f : G \to G'$ s.t. $f(aob) = f(a)\ o'\ f(b)\ \forall$ $a, b \in G$ then G' is also a group isomorphic to G for the given composition.

We have to show that G' is a group and $G' \cong G$.

Let $a' = f(a)$, $b' = f(b)$, $c' = f(c)$; $a, b, c \in G$ and $a', b', c' \in G'$ then $a'o'b' = f(a)\ o'\ f(b) = f(aob)$ is given.

G_1 is satisfied since $a', b' \in G' \Rightarrow f(a), f(b) \in G'$

$$\begin{aligned}&\Rightarrow a, b \in G \\ &\Rightarrow aob \in G \\ &\Rightarrow f(aob) \in G' \\ &\Rightarrow f(a)\ o'\ f(b) \in G' \\ &\Rightarrow a'o'b' \in G' \ \forall\ a', b' \in G'\end{aligned}$$

G_2 is satisfied, since $(a'o'b')o'c' = [f(a)\ o'\ f(b)]\ o'\ f(c)$

$$\begin{aligned}&= f[ao\ (boc)] \because \text{'}o\text{' is associative} \\ &= f(a)\ o'\ f(boc) \\ &= f(a)\ o'\ [f(b)\ o'\ f(c)] \\ &= a'\ o'\ (b'o'c')\end{aligned}$$

G_3 is satisfied, since if e be the identity in G then

$$f(e)\ o'\ a' = [f(e)\ o'\ f(a)]$$

$$= f(eoa) = f(a)$$
$$= a' \qquad \because eoa = a$$

and $a' \, o' f(e) = f(a) \, o' f(e) = f(aoe) = f(a) = a'$

$\therefore \quad f(e) \, o' \, a' = a' \, o' f(e) = a'$

NOTES

G_4 is satisfied, since if $a \in G$ then $a^{-1} \in G$ so that $aoa^{-1} = e = a^{-1}oa$ and

$$aoe^{-1} = e \Rightarrow f(aoa^{-1}) = f(e)$$
$$\Rightarrow f(a) \, o' f(a^{-1}) = f(e)$$
$$\Rightarrow a' \, o' f(a^{-1}) = f(e)$$

Also $f(a^{-1}oa) = f(e) \Rightarrow f(a^{-1}) \, o' f(a) = f(e)$

$$\Rightarrow f(a^{-1}) \, o' \, a' = f(e)$$

$\therefore \quad a'o' f(a^{-1}) = f(a^{-1}) \, o'a' = f(e).$

Thus $f(a^{-1})$ is the inverse of $a' \in G'$, i.e., $f(a^{-1}) = (a')^{-1} = [f(a)]^{-1}$

These axioms show that G' is a group.

Again $G \cong G'$ and the relation of isomorphism is symmetric.

$\therefore \quad G' \cong G$

Theorem 2.10. *The relation of isomorphism in the set of all groups is an equivalence relation.*

If G be a group belonging to the set of all groups and $x \in G$, then consider a one-one onto mapping $f : G \to G$ defined by $f(x) = x \;\forall\; x \in G$.

The relation $\cong$ is reflexive, since $f(x) = f(y) \Rightarrow x = y$, i.e., f is one-one and $f(xy) = xy = f(x) \,.\, f(y)$, operation being multiplicative.

Therefore, the group structure is retained and so $G \cong G \;\forall\; G \in S$, S being the set of all groups.

The relation $\cong$ is symmetric, since if f is isomorphism of G to G', then f is one-one onto and so f^{-1} exists s.t. $f^{-1} : G' \to G$

Therefore, by property (v) of isomorphism, f^{-1} is isomporphic

Thus, $G \cong G' \Rightarrow G' \cong G$.

The relation $\cong$ is transitive, since if $f : G \to G'$ and $g : G' \to G''$ be two isomorphic mappings, then composite mapping gof is also one-one onto when $gof : G \to G''$

Now $x, y \in G \Rightarrow f(x), f(y) \in G' \Rightarrow g[f(y)] \in G''$

So that $\quad (gof)(xy) = g[f(xy)]$

$$= g[f(x) f(y)], f \text{ being isomorphic.}$$
$$\Rightarrow g f(x) \, g f(y), g \text{ being isomorphic.}$$

i.e., gof retains the group compositions and also it is one-one onto, so gof is isomorphism and maps $G \to G''$, i.e.,

$$G \cong G', \; G' \cong G'' \; G \Rightarrow G \cong G''$$

Hence $\cong$ is an equivalence relation.

Theorem 2.11. (*Cayley's Theorem*)

Every finite group G of order n (say) is isomorphic with a subgroup of symmetric group S_n.

Or

Every finite group G of order n is isomorphic to a permutation group (or transformation group).

Let $G = \{a_1\ a_2, \ldots\ a_n\}$ be a finite group of order n, with multiplicative composition and $a \in G$. Then n products

$a\,a_1, a\,a_2, \ldots, a\,a_n$ are all distinct elements of G, for if possible let us assume that $aa_i = aa_j$.

Let cancellation law give, $a_i = a_j$.

But $a_i \neq a_j \therefore a\,a_i \neq a\,a_j$ so that $a\,a_1, a\,a_2, \ldots, a\,a_n$ are all distinct elements of G in some order.

Therefore, the mapping $f_a : G \to G$ s.t. $f_a(x) = ax$, $a \in G$, $x \in G$ is one-one and onto.

Thus $f_a = \begin{pmatrix} a_1 & a_2 \ldots\ldots a_n \\ aa_1 & aa_2 \ldots\ldots aa_n \end{pmatrix}$ is a permutation on n symbols.

Replacing a by $a_1, a_2 \ldots a_n$ in succession, we shall have n permutations $f_{a1}, f_{a2}, \ldots, f_{an}$ of which no two can be equal since if $a_1, a_2 \in G$, then

$$\begin{aligned} f_{a1} = f_{a2} &\Rightarrow f_{a1}(x) = f_{a2}(x) \ \forall\ x \in G \\ &\Rightarrow a_1 x = a_2 x \ \forall\ x \in G \\ &\Rightarrow a_1 = a_2 \end{aligned}$$

Denoting the n permutations by G', i.e.,

$$G' = \{f_a : a \in G\}$$

We have

$$\begin{aligned} f_a f_b &= \begin{pmatrix} a_1 & a_2 \ldots\ldots a_n \\ aa_1 & aa_2 \ldots\ldots aa_n \end{pmatrix} \begin{pmatrix} a_1 & a_2 \ldots\ldots a_n \\ ba_1 & ba_2 \ldots\ldots ba_n \end{pmatrix} \\ &= \begin{pmatrix} a_1 & a_2 \ldots\ldots a_n \\ aa_1 & aa_2 \ldots\ldots aa_n \end{pmatrix} \begin{pmatrix} aa_1 & aa_2 \ldots\ldots aa_n \\ aba_1 & aba_2 \ldots\ldots aba_n \end{pmatrix} \\ &= \begin{pmatrix} a_1 & a_2 \ldots\ldots a_n \\ aba_1 & aba_2 \ldots\ldots aba_n \end{pmatrix} = f_{ab} \end{aligned}$$

But $a, b \in G \Rightarrow ab \in G$ and so $f_a, f_b \in G' \Rightarrow f_{ab} \in G'$

$\therefore$ Closure axiom is satisfied.

Now to show that the set G' with the composite composition is a group isomorphic to the given group G, let us consider a mapping

$$g : G \to G' \text{ s.t. } g(a) = f_a \ \forall\ a \in G$$

Then, $\quad g(a) = g(b) \Rightarrow f_a = f_b$

$$\Rightarrow ax = bx \ \forall\ x \in G$$

$$\Rightarrow a = b \text{ by right cancellation law.}$$

So that g is one-one and therefore G, G' consist of the same number of elements. But g being one-one mapping of G to G', g is also onto. Moreover, $g(ab) = f_{ab} = f_a f_b = g(a)\, g(b)$

NOTES

i.e., the group-composition is retained (preserved) by g.

Hence $G \cong G'$.

Therorem 2.12. *Every cyclic group of infinite order is isomorphic to the additive group of integers.*

If G be an infinite cyclic group generated by a, then $G = \{a\}$ and all the powers of a are distinct.

Consider the mapping $f : G \to \mathbf{I}$ given by $f(a^i) = i$

This mapping is onto and also one-one since $i \neq j \Rightarrow a^i \neq a^j$

$\therefore \quad f(a^i \cdot a^j) = f(a^{i+j} \cdot) = i + j = f(a^i) + f(a^j)$

So that f preserves the operation and hence f is an isomorphism i.e.

$$(G, .) \cong (\mathbf{I}, +)$$

Problem 2.20. *Show that the multiplicative group $G = \{1, -1, i, -i\}$ is isomorphic to the permutation group G' $\{I, (abcd), (ac)\,(bd)\,(adcb)\}$ on four symbols.*

Isomorphism of G and G' will be established if we define mapping of $G \to G'$ s.t. identity element of G is mapped to identity element of G' and inverses are mapped to inverses since then the elements of same order are mapped to elements of the same order.

In G', the order of $(ac)\,(bd)$ is 2 and the order of each of $(abcd)$ and $(adcb)$ is 4.

Now

$$[(ac)\,(bd)]^2 = (ac)\,(bd)\,(ac)\,(bd) = (ac)\,(ac)\,(bd)\,(bd),$$

product of disjoint cycles being abelian

$$= (ac)^2\,(bd)^2 = II = I \text{ as } (ac)^2 = I,\ (bd)^2 = I.$$

$$\text{and}(abcd)^2 = \begin{pmatrix} a & b & c & d \\ b & c & d & a \end{pmatrix}\begin{pmatrix} a & b & c & d \\ b & c & d & a \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ c & d & a & b \end{pmatrix} = (ac)\,(bd)$$

so that $(abcd)^4 = [(ac)\,(bd)]^2 = I$ as above

Similarly $(adcb)^4 = I$

We can thus define a mapping $f : G \to G'$ given by

$$f(1) = I, f(-1) = (ac)\,(bd) = A \text{ (say)}, f(i)$$
$$= (abcd) = B \text{ (say)}, f(-i)$$
$$= (adcb) = C \text{ (say)}$$

The mapping is evidently one-one and onto. The composition tables for G and G' are as shown here.

Clearly in the Table of G if $1, -1, i, -i$ are replaced by I, A, B, C respectively then it transforms to the Table for G'.

Hence $G \cong G'$.

Problem 2.21. *Find the regular permutation group isomorphic to the group $G = \{a, b, c, d\}$ with the composition table.*

Let G' be the required regular permutation group. Then by Cayley's theorem G' will consist of four permutations p_1, p_2, p_3, p_4 given by

$$p_1 = \begin{pmatrix} a & b & c & d \\ aa & ab & ac & ad \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ a & b & c & d \end{pmatrix} = I \text{ (by given composition table)}$$

$$p_2 = \begin{pmatrix} a & b & c & d \\ ba & bb & bc & bd \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ b & a & d & c \end{pmatrix} = \begin{pmatrix} a & b \\ b & a \end{pmatrix}\begin{pmatrix} c & d \\ d & c \end{pmatrix} = (ab)\ (cd)$$

$$p_3 = \begin{pmatrix} a & b & c & d \\ ca & cb & cc & cd \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ c & d & a & b \end{pmatrix} = \begin{pmatrix} a & c \\ c & a \end{pmatrix}\begin{pmatrix} b & d \\ d & b \end{pmatrix} = (ac)\ (bd)$$

$$p_4 = \begin{pmatrix} a & b & c & d \\ da & db & dc & dd \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ d & c & b & a \end{pmatrix} = \begin{pmatrix} a & d \\ d & a \end{pmatrix}\begin{pmatrix} b & c \\ c & b \end{pmatrix} = (ad)\ (bc)$$

Hence $G' = [\{p_1, p_2, p_3, p_4\}, o]$

$= [\{I, (ab)\ (cd), (ac)\ (bd), (ad)\ (bc)\}, o]$

2.7.1 Classes and Invariant Subgroups

Let us discuss about class and invariant subgroup.

Class. The set of elements of a group that are conjugate are said to form a class. The importance of the class concept is that the character (trace) is identical for the matrix representation of all the members of the same class. Thus, in character tables of group representations, the group elements are gathered into the different classes.

Invariant subgroup. A subgroup that consists of complete classes is called an invariant subgroup (or sometimes a normal divisor or normal subgroup). By this, we mean that T is an invariant subgroup of G if it is a subgroup of G *and* if, for any g in G and t in T, $g^{-1}tg$ is an element of T. From this we see that $g^{-1}Tg = T$ or $Tg = gT$. T consists of complete classes belonging to G. Thus, for an invariant subgroup, the left and right cosets are the same. We may also say that an invariant subgroup is self-conjugate.

2.8 FACTOR GROUP COMPLEXES

Complex of a group. *A non-empty subset H of a group G is called as a* **complex** *of the group G.*

Properties of complexes

(*i*) *If Z be a complex containing the elements a, b, c of a group G then $Z = \{a, b, c\}$*

(*ii*) *If $Z = \{a, b, c\}$ be a complex then $aZ = \{a^2, ab, ac\}$ etc.*

(*iii*) *If Z_1 and Z_2 be two complexes of a group G, then the product of Z_1, Z_2 is defined as*

$$Z_1 Z_2 = \{x : x = z_1 z_2, z_1 \in Z_1, z_2 \in Z_2\}$$

NOTES

Now since $z_1 \in Z_1, z_2 \in Z_2$ and $Z_1, Z_2 \subset G$

$\therefore \quad z_1 z_2 = x \in G$ by closure axiom.

As such $Z_1 Z_2 \subset G$.

Which follows that Z_1Z_2 is also a complex of G, obtained by multiplying every element in Z_1 with every element in Z_2.

(*iv*) *The subgroup H of a group G also gives a complex s.t. HH =* H^2 *= H.*

(*v*) *A group can be expressed as a sum of complexes.*

If $x \in G$ and $x \notin H$, H being a subgroup of G, then the complex Hx is a right coset and xH is a left coset of H in G. But cosets are not groups and they are complexes, therefore if the group G as a whole is capable of forming a complex Z which consists of all the elements of the group, then we have

$$Z = H + Hx + Hy + \ldots\ldots$$

(*vi*) *The number of complexes in a group is equal to the index of a subgroup H in G and in fact it is the order of the group divided by the order of the subgroup H.*

(*vii*) *The product of complexes is associative.*

Let Z_1, Z_2 and Z_3 be three complexes of a group G and let

$z_1 \in Z_1, z_2 \in Z_2, z_3 \in Z_3$, then

$z_1 \in Z_1, z_2 \in Z_2 \Rightarrow z_1 z_2 \in Z_1 Z_2$

$\therefore \quad z_1 z_2 \in Z_1 Z_2, z_2 \in Z_3 \Rightarrow \qquad (z_1 z_2) z_3 \in (Z_1 Z_2) Z_3$

$\Rightarrow \qquad z_1 z_2 z_3 \in (Z_1 Z_2) Z_3$

But $\quad z_1 z_2 z_3 = z_1 (z_2 z_3) \qquad \because \quad z_1 z_2 z_3 \in Z_1 (Z_2 Z_3)$

Thus $z_1 z_2 z_3 \in (Z_1 Z_2) Z_3 \Rightarrow z_1 z_2 z_3 \in Z_1 (Z_2 Z_3)$

$\therefore \quad (Z_1 Z_2) Z_3 \subset Z_1 (Z_2 Z_3)$

Similarly $Z_1 (Z_2 Z_3) \subset (Z_1 Z_2) Z_3$

So that $(Z_1 Z_2) Z_3 = Z_1 (Z_2 Z_3)$

Inverse of complex. If Z be a complex of a group G, then its inverse is given by $Z^{-1} = \{z^{-1} : z \in Z\}$

In other words, the inverse of a complex Z is the set of inverses of all elements of Z.

Properties of inverse of a complex

(1) *If* Z_1, Z_2 *be two complexes of a group G, then* $(Z_1Z_2)^{-1} = Z_2^{-1}Z_1^{-1}$

And $x \in (Z_1Z_2)^{-1} \qquad \Rightarrow x = (z_1 z_2)^{-1}$ for $z_1 \in Z_1, z_2 \in Z_2$

$\Rightarrow x = z_2^{-1}z_1^{-1}$ by reversal law of inverses.

$\Rightarrow x \in Z_2^{-1}Z_1^{-1}$ by definition

$\therefore \quad (Z_1Z_2)^{-1} \subset Z_2^{-1}Z_1^{-1} \qquad$ .... (**A**)

NOTES

Similarly if $y \in Z_2^{-1}Z_1^{-1}$ $\Rightarrow y = z_2^{-1}z_1^{-1}$ when $z_2^{-1} \in Z_2^{-1}, z_1^{-1} \in Z_1^{-1}$

$\Rightarrow y = (z_1 z_2)^{-1}$ where $z_1 \in Z_1, z_2 \in Z_2$

$\Rightarrow y \in (Z_1Z_2)^{-1}$ by definition

$\therefore \quad Z_2^{-1}Z_1^{-1} \subset (Z_1Z_2)^{-1}$ .... (**B**)

(**A**) and (**B**) follow that $(Z_1Z_2)^{-1} = Z_2^{-1}Z_1^{-1}$

(2) *If H be a subgroup of a group G, then* $H^{-1} = H$.

An $h^{-1} \in H^{-1} \Rightarrow h \in H$

$\Rightarrow h^{-1} \in H$, H being a group

So $H^{-1} \subset H$

Similarly an $h \in H \Rightarrow h^{-1} \in H$, H being a group

$\Rightarrow h = (h^{-1}) \in H$ by definition of inverse of a complex.

So $H \subset H^{-1}$

$\therefore \quad H^{-1} \subset H, H \subset H^{-1} \Rightarrow H^{-1} = H.$

(3) *If H, K be two subgroups of a group G, then HK is also a subgroup of G iff HK = KH.*

Taking $HK = KH$, we have $(HK)^{-1} = (KH)^{-1}$

$= K^{-1}H^{-1}$ by Property (1)

$= KH$ by Property (2)

$= HK \quad \because HK = KH$

which shows that HK is a subgroup of G.

Again taking HK as subgroup of G, we have

$(HK)^{-1} = HK$ by Property (2)

$\therefore \quad K^{-1}H^{-1} = HK$ by Property (1)

i.e., $KH = HK$ by Property (2)

Hence the proposition.

(4) *A necessary and sufficient condition for a complex H of a group G to be a subgroup is that* $HH^{-1} = H$.

The condition is necessary since if H is a subgroup of G and $ab^{-1} \in HH^{-1}$ then

$a \in H, b \in H \Rightarrow a \in H, b^{-1} \in H$

$\Rightarrow ab^{-1} \in H$

So $ab^{-1} \in HH^{-1} \Rightarrow ab^{-1} \in H, b \in H, b^{-1} \in H^{-1}$

i.e. $HH^{-1} \subset H$

Also H is a subgroup of $G \Rightarrow$ identity $e \in H$

If $h \in H$, then $h = he = he^{-1} \in HH^{-1}, h \in H, e^{-1} \in H^{-1}$

$\therefore \quad H \subset HH^{-1}$

Thus $HH^{-1} \subset H, H \subset HH^{-1} \Rightarrow HH^{-1} = H$

The condition is sufficient since if $HH^{-1} = H$, then we have

$$HH^{-1} \subset H$$

Now suppose that $a, b \in H$ so that $ab^{-1} \in HH^{-1}$

$\therefore \quad HH^{-1} \subset H$ and $ab^{-1} \in HH^{-1} \Rightarrow ab^{-1} \in H$

Ultimately $a \in H, b \in H \Rightarrow ab^{-1} \in H$

Which follows that H is a subgroup as is evident from the following discussion: Taking H a subgroup of G with the same composition as in G, the identity in H and G is the same. Also $a \in H$ and $b \in H$ give $b^{-1} \in H$, H being a group.

$$\therefore \quad a \in H, b^{-1} \in H \Rightarrow ab^{-1} \in H \qquad .... (\alpha)$$

Further taking H to be a non-empty subset of G s.t. $a \in H, b \in H$, and assuming that $a \in H, b \in H \Rightarrow ab^{-1} \in H$, we observe that H is non-empty and $\exists$ an $a \in H$ so that by setting $b = a$ in (α), we find

$$a \in H, a \in H \Rightarrow aa^{-1} \in H$$

$$\Rightarrow e \in H, e \text{ also being identity in } G.$$

Now $\quad e \in H, b \in H \Rightarrow eb^{-1} \in H$ by (α) $\qquad .... (\beta)$

$$\Rightarrow b^{-1} \in H$$

$$\therefore \quad a \in H, b^{-1} \in H \Rightarrow a\,(b^{-1})^{-1} \in H \Rightarrow ab \in H$$

i.e. $\quad a \in H, b \in H \Rightarrow ab \in H$

Here (α) and (β) fulfil the requirements for H which is a complex of G, to be its subgroup.

Image of a group G under a mapping f. *If $f : G \to G'$ be a homomorphism of a group G into a group G', then $f(G) = \{f(x) \in G' : x \in G\}$ is a subset of G' and is termed as the* **Image of G under *f*** *and denoted by Im (f).*

Kernel of f. *If $f : G \to G'$ be a homomorphism of G into G', then the subset of those elements of G which are mapped onto the identity of G' under f' is said to be the* **Kernel of f** *and denoted by ker (f) or $f^{-1}(e')$.*

i.e., $\ker(f) = \{x \in G : f(x) = e'\}$

Propositions relating to Kernel

I. *A homomorphism $f : G \to G'$ is an isomorphism iff $\ker f = \{e\}$.*

Assuming that $f : G \to G'$ is an isomorphism, if $a \in \ker f$ then

$$f(a) = e' = f(e), \; e' \text{ being identity in } G'.$$

Now f being one-one and $a = e$, kernel of f consists of e only. Conversely i f
$\ker f = \{e\}$ for f to be homomorphism, and if $a, b \in G$ s.t. $f(a) = f(b)$, then

$$f(ab^{-1}) = f(a) f(b^{-1})$$

NOTES

NOTES

$$= f(a)\,[f(b)]^{-1}$$

$$= e' \qquad \because f(a) = f(b)$$

$$\therefore \quad ab^{-1} \in \ker f$$

$$\text{or} \quad ab^{-1} = e$$

$$\text{or} \quad a = b$$

So f is one-one and hence f is an isomorphism.

II. *If f be homomorphism of G' then ker (f) is an invariant subgroup of G.*

If $a, b \in \ker(f)$, then $f(a) = e' = f(b)$, e' being identity of G.

$\therefore \quad f(ab) = f(a)\,f(b) = e'\,e' = e'$

which implies that $ab \in \ker(f)$, *i.e.*, closure axiom is satisfied.

Now ker (f) being a subset of G, associativity axiom is self-evident.

Again $f(e) = e' \Rightarrow e \in \ker(f)$, e being identity in G.

Therefore, there exists an identity in G.

Further if $a \in \ker(f)$ then $f(a^{-1}) = [f(a)]^{-1} = (e')^{-1} = e'$

which shows that $a^{-1} \in \ker(f)$ when $a \in \ker(f)$.

This follows the existence of an inverse in G.

As such ker (f) is a subgroup of G, as ker (f) satisfies all the four group axioms.

Moreover ker (f) is an invariant subgroup of G as is shown below :

If $g \in G$ and $h \in \ker(f)$, then

$$f(g^{-1}hg) = f(g^{-1})\,f(h)\,f(g)$$

$$= [f(g)]^{-1}\,e'\,f(g) \quad \because \quad h \in \ker(f) \Rightarrow f(h) = e'$$

$$= [f(g)]^{-1} f(g)$$

$$= e'$$

$\therefore \quad g^{-1}hg \in \ker(f)$.

Hence ker (f) is an invariant subgroup of G.

Note 1. It is easy to show that Im (f) is a subgroup of G.

III. *If H be a normal subgroup of a group G, then there is a homomorphism of G onto G/H.*

Let $f : G \to G/H$ be given by $f(x) = Hx \;\forall\; x \in G$

$\because \;\forall\; x \in G$, $\exists$ a unique coset Hx, f is a mapping.

Also the binary operation in G/H being defined by

$$(Hx)(Hy) = H(xy)$$

We have

$$f(xy) = H(xy) = (Hx)(Hy) = f(x)\,f(y)$$

Which follows that f is a homomorphism and it is onto since every coset $Hz \in G/\text{H}$ has z as its preimage in G.

Note 2. **Natural Homomorphism**. The homorphism $f : G \to G/H$ given by $f(x) = Hx$ is known as **Natural Homomorphism** or **Canonical Homomorphism** of G onto G/H.

IV. *If f be a homomorphism of a group G onto a group G′ with kernel k, then*

$$G/K \cong G'$$

Consider the mapping $\phi : G/K \to G'$ defined by $\phi(Kx) = f(x)$

Taking $Kx = Ky$, we have $xy^{-1} \in K$ and $f(xy^{-1}) = e'$, e' being identity in G', *i.e.*, $f(x) f(y^{-1}) = e'$

or $\quad f(x) [f(y)]^{-1} = e'$

or $\quad f(x) = f(y)$.

This follows that ϕ is uniquely defined.

Now if $f(y) \in G'$ then Ky is the preimage of $f(y)$ in G/K under ϕ.

This follows that ϕ is onto.

Again f will be one-one if $Kx = Ky$ provided $f(x) = f(y)$.

Take an element $z = xy^{-1} \in G$ *i.e.* $zy = x$

$$\therefore \quad f(z) = f(xy^{-1}) = f(x) f(y^{-1})$$
$$= f(x) [f(y)]^{-1}$$
$$= e' \qquad \because f(x) = f(y)$$

So that $z \in K$ and $Kx = K(zy) = (Kz)y = Ky$

$\therefore \quad \phi$ is one-one.

Further to show that f preserves the structure, we have

$$\phi(Kx)\,\phi(Ky) = f(x) f(y) = f(xy) = \phi\,[K(xy)] = \phi\,[(Kx)(Ky)]$$

Hence ϕ is isomorphism and thus $G/K \cong G'$.

V. *If f is a homomorphism from the group* (G, o) *into the group* (G', o') *then the pair* $(\ker f, o)$ *is a normal subgroup of* (G, o).

Evidently $\ker f \neq \phi$ (non-empty) since $e \in \ker f$ and $\ker f \subset G$

Now $a, b \in \ker f \Rightarrow f(a) = e', f(b) = e'$

But $f(b^{-1}) = [f(b)]^{-1} = [e']^{-1} = e'$

$\therefore f(aob^{-1}) = f(a)\, o\, [f(b)]^{-1} = e' o e' = e'$

$\therefore a, b \in \ker f \Rightarrow aob^{-1} \in \ker f$

Hence $(\ker f, o)$ is a subgroup.

Again $\forall\ a \in G$ and $h \in \ker f$, we have

$$f(ao\,hoa^{-1}) = f(a)\, o\, f(h)\, o\, f(a^{-1})$$

NOTES

$$= f(a) \, o \, f(h) \, o \, [f(a)]^{-1}$$
$$= f(a) \, oe' \, o \, [f(a)]^{-1}$$
$$= f(a) \, o \, [f(a)]^{-1}$$
$$= e'$$

$\therefore \quad \forall \; a \in G$ and $h \in \ker f \Rightarrow aohoa^{-1} \in \ker f$

Hence $(\ker f, o)$ is a normal subgroup.

Note 3. Similarly it can be shown that image of $[Im(f), o]$ is a subgroup of $(G'o')$ when f is a homomorphism of $(G\,o)$ into (G', o').

Problem 2.22. *If GL (n, R) is the multiplication group of all $n \times n$ singular matrices with elements as real numbers and that G' is the multiplicative group of all non-zero real numbers, then show that the mapping $f: G \to G'$ s.t. $f(A) = |A| \; \forall \; A \in G$ is a homomorphism of G onto G' and also show that*

$$\ker f = \{A \in GL\,(n, R) : |A| = e', \text{the identity in } G'\}.$$

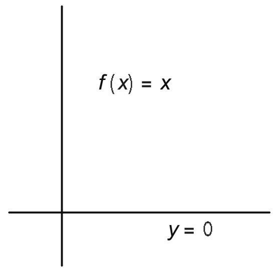

Fig. 2.3

Let $f: (\mathbf{C}, +) \to (\mathbf{R}^+, +)$ s.t. $f(x + iy) = x$.

We have to show that f is a homomorphism of $(\mathbf{C}, \cdot)$ onto $(\mathbf{R}, \cdot)$ and $\ker f = \{z \in \mathbf{C} : x = 0\}$, *i.e.*, $\ker f$ is the imaginary y-axis.

If $z_1 = x_1 + iy_1 \in \mathbf{C}$, $z_2 = x_2 + iy_2 \in \mathbf{C}$, then

$$f(x_1 + iy_1) + f(x_2 + iy_2)$$
$$= f(x_1 + x_2) + if(y_1 + y_2)$$
$$= x_1 + x_2 \text{ by hypothesis}$$
$$= f(x_1 + iy_1) + f(x_2 + iy_2)$$

This shows that f is onto since $f(x + io) = x$ if $x \in \mathbf{R}$.

As such f is a homomorphism of $(\mathbf{C}^*, +)$ onto $(\mathbf{R}^+, +)$ and 1 is the identity element in $(\mathbf{R}^+, \cdot)$.

Also $\ker f$ is given by $f(x + iy) = 0 = x \; \forall \; x \in \mathbf{R}$.

i.e., $\quad \ker f = \{z \in \mathbf{C} : x = 0 = e'$, the identity in $\mathbf{R}\}$

which follows that $\ker f$ is the imaginary axis.

Problem 2.23. *If $(\mathbf{R}, \cdot)$ be a multiplicative group and $x \in \mathbf{R}$, then find homomorphisms and their kernels in the following mappings:*

(i) $x \to |x|$

$(ii)\ x \to \frac{1}{x^2}$.

$(i)\ x \to |x| \Rightarrow x \to |x|$ and $-x \to |x|$

$\therefore\quad x \to |x|$ and $y \to |y| \Rightarrow xy \to |xy| \to |x||y|$

Thus $x \to |x|$ is homomorphism.

Now $f : x \to |x| \Rightarrow f(x) = |x| \Rightarrow f(xy) = |xy| \Rightarrow f(xy) = |xy|$

$$= |x||y| = f(x) f(y)$$

i.e., f is two-one mapping since $|f(x)| \Rightarrow f(x) = -1, 1$.

Its kernel is $|f(x)| = 1 \Rightarrow -1, +1$, *i.e.*, $\{-1, 1\}$.

(ii) Say $g : x \to \frac{1}{x^2} \Rightarrow g(x) = \frac{1}{x^2} \Rightarrow g(xy) = \frac{1}{(xy)^2} = \frac{1}{x^2 y^2} = \frac{1}{x^2}\cdot\frac{1}{y^2} = g(x)\cdot g(y)$

So g is a homomorphism and it is two-one mapping since

$$g(x) = \frac{1}{x^2} \Rightarrow g(-x) = \frac{1}{x^2}$$

Now $g(1) = e' \Rightarrow \frac{1}{x^2} = 1 \Rightarrow x = -1, 1$

$\therefore\quad \ker g = \{-1, 1\}$.

2.9 DIRECT SUM AND PRODUCT

Let us discuss direct sum and direct product.

Direct Sum

Direct sums are defined for a number of different sorts of mathematical objects, including subspaces, matrices, modules, and groups.

The matrix direct sum is defined by

$$\bigoplus_{i=1}^{n} A_i = \text{diag}(A_1, A_2, \ldots, A_n)$$

$$\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{bmatrix}$$

The direct sum of two subspaces and is the sum of subspaces in which and have only the zero vector in common.

The significant property of the direct sum is that it is the coproduct in the category of modules (i.e., a module direct sum). This general definition gives as a consequence the definition of the direct sum of Abelian groups and (since they are -modules, i.e., modules over the integers) and the direct sum of vector spaces (since they are modules over a field). Note that the direct sum of Abelian groups is the same as the group direct product, but that the term direct sum is not used for groups which are non-Abelian.

Note that direct products and direct sums differ for infinite indices. An element of the direct sum is zero for all but a finite number of entries, while an element of the direct product can have all nonzero entries.

NOTES

Direct Product

The direct product is defined for a number of classes of algebraic objects, including sets, groups, rings, and modules. In each case, the direct product of an algebraic object is given by the Cartesian product of its elements, considered as sets, and its algebraic operations are defined component wise. For instance, the direct product of two vector spaces of dimensions m and n is a vector space of dimension $m+n$.

Direct products satisfy the property that, given maps $\alpha : S \to A$ and $\beta : S \to B$, there exists a unique map $S \to A \times B$ given by $(\alpha(s), \beta(s))$. The notion of map is determined by the category, and this definition extends to other categories such as topological spaces. Note that no notion of commutativity is necessary, in contrast to the case for the coproduct. In fact, when A and B are Abelian, as in the cases of modules (e.g., vector spaces) or Abelian groups (which are modules over the integers), then the direct sum $A \oplus B$ is well-defined and is the same as the direct product. Although the terminology is slightly confusing because of the distinction between the elementary operations of addition and multiplication, the term "direct sum" is used in these cases instead of "direct product" because of the implicit connotation that addition is always commutative.

Note that direct products and direct sums differ for infinite indices. An element of the direct sum is zero for all but a finite number of entries, while an element of the direct product can have all nonzero entries.

Some other unrelated objects are sometimes also called a direct product. For example, the tensor direct product is the same as the tensor product, in which case the dimensions multiply instead of add. Here, "direct" may be used to distinguish it from the external tensor product.

2.10 REDUCIBLE AND IRREDUCIBLE REPRESENTATIONS

The irreducible representation as defined earlier. In sequence to the above special Unitary group, we have

$$Uf(x) = f(x') = f(\alpha x_1 + \beta x_2, -\bar{\beta} x_1 + \bar{\alpha} x_2)$$

As such if U operates on a set of $(n + 1)$ homogeneous products

$$f_m^{(n)} = x_1^{n-m} x_2^{n-m}, \quad m = 0, 1, 2, ...n$$

then we get

$$Uf_m^{(n)} = (\alpha x_1 + \beta x_2)^m (-\bar{\beta} x_1 + \bar{\alpha} x_2)^{n-m}$$

$$= \sum_{j=0}^{n} U_{mj}^{(n)} x_1^j x_2^{n-j}$$

where $$U_{mj}^{(n)} = \sum_k (-1)^k \frac{\underline{|m}\,\underline{|(n-m)}}{\underline{|(m-k)}\,\underline{|k}\,\underline{|(n-m-j+k)}\,\underline{|(j-k)}} \times \alpha^k \beta^{m-k} (\bar{\alpha})^{n-m-j+k} (\bar{\beta})^{j-k}$$

NOTES

Character of a Special Unitary Group SU(2): The character of a special unitary group is found out if a typical matrix by means of unitary transformation is transformed to diagonal form.

Take a unitary matrix V such that

$$V^{-1}UV = \begin{bmatrix} \gamma & 0 \\ 0 & \gamma \end{bmatrix} = U'$$

Now $|U'| = +1$ is apparantly satisfied if $U' = \begin{bmatrix} e^{i\phi/2} & o \\ o & e^{-i\phi/2} \end{bmatrix}$

All the other matrices of the group will belong to the same class as U and U' since the class constitutes elements obtained by unitary transformation while a unitary matrix remains unitary under such a transformation.

Thus the character of one element of the class is given by

$$\chi^{(1/2)} = e^{i\,\phi/2} + e^{-i\,\phi/2} \text{ by using } U'$$

In general the character of special unitary group for any value j is given by

$$\chi^{(j)} = \sum_{m=-j}^{j} e^{im\phi}$$

If $\chi = e^{i\phi}$, then $\chi^{(j)} = e^{-ij\phi}\,(1 + \chi + \chi^2 + ...+ \chi^{2j})$

$$= e^{-ij\phi}\left(\frac{1-\chi^{2j+1}}{1-\chi}\right)$$

$$= \frac{\sin\,(2j+1)\frac{\phi}{2}}{\sin\frac{\phi}{2}}$$

on multiplying numerator and denominator by $i\dfrac{\phi}{2}$.

***n*-Dimensional Rotation Group:** A continuous group formed from the set of all orthogonal n-dimensional matrices is said to be *n-dimensional rotational group*. In fact this is a subgroup of a full linear group provided all elements are real unitary matrices whose determinant is +1.

For example, if a point $P(x, y, z)$ is taken on the surface of a unit sphere and the sphere is rotated in any manner keeping its centre fixed; then the new coordinates of P say (x', y', z') related to (x, y, z) by some matrix $R(\alpha, \beta, \gamma)$ which is an element of a 3-dimensional rotation group $R^+(3)$ give a rotation factorised as product of three plane rotations described by the Eulerian angles (α, β, γ) (discussed in classical mechanics), i.e., $R(\alpha, \beta, \gamma) = R_z(\gamma)\, R_y(\beta)\, R_z(\alpha)$, where R_z, R_y, R_x are rotations about z, y, x axes respectively.

As an other example if we define a 2-*dimensional rotation group* $R^+(2)$ as a subgroup of $R^+(3)$, then its elements are obtained by proper rotation in a plane perpendicular to a fixed axis, say z-axis. Taking $R(\theta)$ as one element of this group and $T(\theta)$ an operator transforming a vector $\mathbf{x}$ with components x_1, x_2 to another vector $\mathbf{x}'$ with components x_1', x_2', i.e.,

$$\mathbf{x}' = T(\theta)\mathbf{x},\; 0 \le \theta \le 2\pi$$

NOTES

such that $$T(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \text{ we have}$$

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or equivalently, $$x_1' = x_1 \cos\theta + x_2 \sin\theta$$

$$x_2' = x_2 \cos\theta - x_1 \sin\theta$$

But if $R(\theta')$ is another element of the group with transformation $R(\theta')$ then

$$T(\theta)\, T(\theta') = T(\theta + \theta') = T(\theta')\, T(\theta)$$

which follows that the group is commutative, i.e., Abelian.

Point Group: The inversion operation in 3-dimensional space is given by the matrix

$$T = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

and an identity operation I is given by the unit matrix

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Evidently $$TI = T$$

and $$T^2 = I, \text{ i.e., } T = T^{-1}$$

Here T and I form a group with matrix multiplication. Such a group is said to be a *point group* since one point (say origin) remains fixed in all operations. The fixed point is sometimes known as *centre of inversion.*

Consider a point group $\{c_n\}$ with operations on a regular polygon of n sides such that there exist

(*i*) a rotation through an angle $2\pi/n$ about an n-fold axis of rotation properly.

(*ii*) a rotation through $-2\pi/n$ about an n-fold axis of rotation improperly.

(*iii*) a reflction in a plane given by σ_H, σ_V; H, V denoting Horizontal and Vertical planes.

(*iv*) an inversion.

Such operations form a point group $\{c_n\}$.

Quaternion Group: If we define a group G of order 8 such that

$$G = \{1, a, a^2, a^3, b, ab, a^2b, a^3b\}$$

with the properties, $$ab = ba^3$$

$$b^2 = a^2$$

and $$a^4 = 1$$

whence $ab = ba^3$ and $b^2 = a^2 \Rightarrow ba = a^3b$

NOTES

	1	a	a^2	a^3	b	ab	a^2b	a^3b
1	1	a	a^2	a^3	b	ab	a^2b	a^3b
a	a	a^2	a^3	1	ab	a^2b	a^3b	b
a^2	a^2	a^3	1	a	a^2b	a^3b	b	ab
a^3	a^3	1	a	a^2	a^3b	b	ab	a^2b
b	b	a^3b	a^2b	ab	a^2	a	1	a^3
ab	ab	b	a^3b	a^2b	a^3	a^2	a	1
a^2b	a^2b	ab	b	a^3b	1	a^3	a^2	a
a^3b	a^3b	a^2b	ab	b	a	1	a^3	a^2

Since $a^3b = a^2(ba^3) = b^3a^3 = b(a^2a^3) = ba$.

The composition table is as shown here. It is clear from this table that the group of order 8 under consideration does actually exist and defines a group. Such a group is known as *quaternion group.* All of its subgroups are normal, though it is not abelian. Clearly a quaternion group is also a Hamiltonian group.

2.11 SCHUR'S LEMMAS

Schur's lemma is a fundamental result in representation theory, an elementary observation about irreducible modules, which is nonetheless noteworthy because of its profound applications.

Lemma 1: Let G be a finite group and let V and W be irreducible G-modules. Then, every G-module homomorphism $f: V \to W$ is either invertible or the trivial zero map.

Proof: Both the kernel, ker f and the image, im f are G-submodules of V and W, respectively. Since V is irreducible, ker f is either trivial or all of V. In the former case, im f is all of W also because W is irreducible and hence f is invertible. In the latter case, f is the zero map.

Given below is one of the most important consequences of Schur's lemma:

Corollary: Let V be a finite-dimensional, irreducible G-module taken over an algebraically closed field. Then, every G-module homomorphism $f: V \to V$ is equal to a scalar multiplication.

Proof: Since the ground field is algebraically closed, the linear transformation $f: V \to V$ has an eigenvalue λ, say. By definition, $f - \lambda$ is not invertible, and hence equal to zero by Schur's lemma. In other words, $f = \lambda$, i.e., a scalar.

2.11.1 Orthogonality Theorem

The Orthogonality Theorem is a fundamental result in group representation theory. It can be described to anyone familiar with groups and matrices.

Let G be a finite group with n elements. Every finite group has an associated set of "irreducible representations." An irreducible representation is simply some function that assigns a (unitary) matrix to each element of the group. (Naturally there are further details: the functions must be group homomorphisms, etc.)

For example, the group D4 (i.e., the n=8 symmetries of a square) has 5 irreducible representations. Four of these representations assign 1-by-1 matrices (i.e., scalars) to each of the 8 elements of D4. This includes the trivial representation, which assigns 1 to every group element: we might think of it as the n-dimensional vector (1,1,1,1,1,1,1,1). Another assigns 1 to four elements and -1 to the other four (the "flipped" elements): this could be written as the vector (1,1,1,1,-1,-1,-1,-1). The fifth representation assigns 2-by-2 matrices to each element of D4D4. We can make four more nn-d vectors out of these. Taking the (1,1) elements of the 8 matrices yields the vector (1,0,-1,0,0,1,0,-1); all the (1,2) elements form another vector, etc. In total we get *n n-d* vectors this way.

The Orthogonality Theorem states that the *n* vectors so obtained are orthogonal, for any finite group of order *n*. That is, the (complex) dot product of any two distinct vectors is 0. It also says the dot product of any vector with itself is *n* over the order of the representation. E.g., 8/1 = 8 for (1,1,1,1,1,1,1,1), and 8/2 = 4 for (1,0,-1,0,0,1,0,-1).

2.12 CHARACTER OF A REPRESENTATION

In mathematics, more specifically in group theory, the **character** of a group representation is a function on the group that associates to each group element the trace of the corresponding matrix. The character carries the essential information about the representation in a more condensed form. Georg Frobenius initially developed representation theory of finite groups entirely based on the characters, and without any explicit matrix realization of representations themselves. This is possible because a complex representation of a finite group is determined (up to isomorphism) by its character. The situation with representations over a field of positive characteristic, so-called "modular representations", is more delicate, but Richard Brauer developed a powerful theory of characters in this case as well. Many deep theorems on the structure of finite groups use characters of modular representations.

2.12.1 Application of Group Theory in Physics

In physics, groups are important as they explain the symmetries which the laws of physics appear to obey. According to Noether's theorem, every continuous symmetry of a physical system corresponds to a conservation law of the system. Lie groups are often point the way to the "possible" physical theories. Examples of the use of groups in physics include the Standard Model, gauge theory, the Lorentz group, and the Poincaré group.

Group theory can be used to resolve the incompleteness of the statistical interpretations of mechanics developed by Willard Gibbs, relating to the summing of an infinite number of probabilities to yield a meaningful solution

2.12.2 Classification of States and Elementary Particle

Classification of States

States are classified into classes: **transient, ergodic and periodic**. If two vertices belong to the same communicating class they have the same classification.

Classification of Elementary Particle

In particle physics, an elementary particle or fundamental particle is a subatomic particle that is not composed of other particles. Particles currently thought to be elementary include the fundamental fermions (quarks, leptons, antiquarks, and antileptons), which generally are "matter particles" and "antimatter particles", as well as the fundamental bosons (gauge bosons and the Higgs boson), which generally are "force particles" that mediate interactions among fermions. A particle containing two or more elementary particles is a composite particle.

2.12.3 Splitting of Energy Levels

In quantum physics, energy level splitting or a split in an energy level of a quantum system occurs when a perturbation changes the system. The perturbation changes the corresponding Hamiltonian and the result is change in eigenvalues; several distinct energy levels emerge in place of the former degenerate (multi-state) level. This may take place because of external fields, quantum tunnelling between states, or other effects. The term is commonly used in reference to the electron configuration in atoms or molecules.

The simplest case of level splitting is a quantum system with two states whose unperturbed Hamiltonian is a diagonal operator: $\$_0 = E_0\ I$, where I is the 2 × 2 identity matrix. Eigenstates and eigenvalues (energy levels) of a perturbed Hamiltonian

$$\hat{H}_\varepsilon = \hat{H}_0 + \varepsilon\sigma_3 = \begin{pmatrix} E_0 + \varepsilon & 0 \\ 0 & E_0 - \varepsilon \end{pmatrix}$$

will be:

(i) the $E_0 + \varepsilon$ level, and

(ii) the E_0 " ε level,

so this degenerate E_0 eigenvalue splits in two whenever $\varepsilon \neq 0$. Though, if a perturbed Hamiltonian is not diagonal for this quantum states basis {|0>, |1>|, then Hamiltonian's eigenstates are linear combinations of these two states.

2.12.4 Matrix Elements and Selection Rules

The direct (outer) product of two irreducible representations A and B of a group G,

$$A \otimes B = C = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & \\ a_{21}b_{11} & & & \\ & & & \end{pmatrix}$$

gives us the chance to find out the representation for which the product of two functions forms a basis. This representation will in general be reducible.

Note that we are dealing with the representations within one group, since we deal with a system with well-defined symmetry. Thus "products" of functions such as , where each function forms the basis for an irreducible representation Γa and Γb of the group, form a basis for an (often reducible) representation Γa ⊗ Γb.

Using the reduction formula, we can decompose the reducible representation of the products into the irreducible ones of the group.

$$a_j = \frac{1}{h}\sum_R \chi_i\left[\Gamma^j(R)^*\right]\chi_i\left[\Gamma^{red}(R)\right]$$

An expression of the form, where O is some operator which connects the initial with the final state, i.e. a transition matrix element is an integral over all space. 0 ‘“ 0 = r If the product function does not contain a function that forms a basis for the totally symmetric irreducible representation, i.e. if that function is not of even parity under all symmetry operations of the group, then by necessity that integral will be zero.

Check Your Progress

10. What is diffusion equation?
11. When is a group said to be semi-group?
12. What is a sub group?
13. What is the centre of a group?
14. State Fermat's theorem.

2.13 ANSWERS TO ‘CHECK YOUR PROGRESS’

1. Any equation which contains one or more partial derivatives is called a partial differential equation. $x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = z; \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} + \frac{\partial^2 z}{\partial x \partial y} = 0$ are examples for partial differential equation of first order and second order respectively.
2. Partial differential equation may be formed by eliminating *(i)* arbitrary constants *(ii)* arbitrary functions.
3. A solution in which the number of arbitrary constants is equal to the number of independent variables is called complete integral or complete solution.

   In complete integral, if we give particular values to the arbitrary constants, we get particular integral. If $\varphi(x, y, z, a, b) = 0$, is the complete integral of a partial differential equation, then the eliminant of a and b from the equations $\frac{\partial \phi}{\partial a} = 0$, $\frac{\partial \phi}{\partial b} = 0$, is called singular integral.
4. In mathematics, Laplace equation is a second order partial differential equation. It is named after Pierre-Simon Laplace and is written as,

   $\nabla^2 \varphi = 0$

   Here is the Laplace operator and φ is a scalar function of 3 variables. Laplace equation and Poisson equation are examples of elliptic partial differential equations. The universal theory of solutions to Laplace equation is termed as potential theory. The solutions of Laplace equation are harmonic functions and have great important in many fields of science.

NOTES

5. The Laplace equation $\nabla^2\varphi = 0$ can also be written as $\nabla \cdot \nabla\varphi = 0$.
6. Both the real and imaginary parts of a complex analytic function satisfy the Laplace equation. If $z = x + iy$ and also if $f(z) = u(x,y) + iv(x,y)$ then the necessary condition that $f(z)$ be analytic is that it must satisfy the Cauchy-Riemann equations $u_x = v_y$ and $v_x = -u_y$, where u_x is the first partial derivative of u with respect to x. It follows the notation,

$u_{yy} = (-v_x)_y = -(-v_y)_x = -(u_x)_x$.

7. In mathematics, Poisson equation is a partial differential equation. It is named after the French mathematician, geometer and physicist Siméon-Denis Poisson. The Poisson equation is,

$$\Delta\varphi = f$$

Here Δ is the Laplace operator and f and φ are real or complex-valued functions on a manifold. If the manifold is Euclidean space, then the Laplace operator is denoted as Δ^2 and hence Poisson equation can be written as,

$$\nabla^2\varphi = f$$

8. A second order partial differential equation is of the form, $\nabla^2\Psi = -4\pi\rho$.

If $\rho = 0$, then it reduces to Laplace equation. It can also be considered as Helmholtz differential equation of the form,

$$\nabla^2\Psi + k^2\Psi = 0$$

9. The wave equation is an important second-order linear partial differential equation of waves. It is analysed on the basis of sound waves, light waves and water waves. The wave equation is considered as a hyperbolic partial differential equation. In its simplest form, the wave equation refers to a scalar function $u=(x_1, x_2,...,x_n, t)$ that satisfies,

$$\frac{\partial^2 u}{\partial t^2} = c^2\nabla^2 u$$

Here is the spatial Laplacian and c is a fixed constant equal to the propagation speed of the wave and is also known as the non-dispersive wave equation.

10. A Parabolic equation of the type

$$\nabla^2 V = \frac{1}{k}\frac{\partial V}{\partial t}$$

is said to be a *diffusion equation*. However, if p be the source density, then $\nabla^2 V = \rho$ is known as *Poisson's equation*.

11. A set S with a binary operation 'o' is said to be a semi-group if it satisfies the following two axioms:

S_{G1}—(*Closure*). $a \in S$, $b \in S \Rightarrow aob \in S$.

S_{G2}—(*Associativity*). If $a, b, c \in S$ then $(aob)\, oc = ao\, (boc)$.

12. A sub-group of a group (G, o) is any collection of elements of G satisfying the axioms of G. In other words, a non-empty subset say H of a group G is

said to be the sub-group of G, if the binary operation 'o' in G induces a binary operation in H and the elements of H obey the group axioms.

13. If (G, o) be a group and H be the set of those elements $x \in G$, which commute with each element in G, i.e., the set

    $$H = \{x : x \in G \text{ and } aox = xoa \;\; a \in G\}$$

    then the set H is known as the centre of G.

14. If p be a prime and 'a' a natural number not divisible by p, then

    $a^{p-1} = 1 \; (mod \, p)$

2.14 SUMMARY

- Partial differential equation may be formed by eliminating *(i)* arbitrary constants

  (ii) arbitrary functions.

- The partial differential equations can be formed by eliminating arbitrary junctions.
- A solution in which the number of arbitrary constants is equal to the number of independent variables is called complete integral or complete solution.
- In mathematics, Laplace equation is a second order partial differential equation. It is named after Pierre-Simon Laplace and is written as,

  $\nabla^2 \varphi = 0$

- Both the real and imaginary parts of a complex analytic function satisfy the Laplace equation. If $z = x + iy$ and also if $f(z) = u(x,y) + iv(x,y)$ then the necessary condition that $f(z)$ be analytic is that it must satisfy the Cauchy-Riemann equations $u_x = v_y$ and $v_x = -u_y$, where u_x is the first partial derivative of u with respect to x.
- The Laplace equation remains unchanged during rotation of coordinates and hence a fundamental solution can be obtained that only depends upon the distance r from the source point.
- In mathematics, Poisson equation is a partial differential equation. It is named after the French mathematician, geometer and physicist Siméon-Denis Poisson. The Poisson equation is,

  $$\Delta\varphi = f$$

- The wave equation is an important second-order linear partial differential equation of waves. It is analysed on the basis of sound waves, light waves and water waves. The wave equation is considered as a hyperbolic partial differential equation.
- A wave can be superimposed onto another movement. In that case the scalar u will contain a Mach factor which is positive for the wave moving along the flow and negative for the reflected wave.

- As per the Helmholtz equation, named for Hermann von Helmholtz, is the elliptic partial differential equation of the form $\nabla^2 A + k^2 A = 0$, where is the Laplace operator, k is the wavenumber and A is the amplitude.
- One dimensional heat equation is of the form $\frac{\partial u}{\partial t} = h^2 \frac{\partial^2 u}{\partial x^2}$.
- A group is the simplest algebraic structure found in nature wherever symmetry exists.
- Uniqueness of Inverse, i.e., in a group (G, o) every element possesses a unique inverse.
- Uniqueness of Solutions, i.e., *if* $a, b \in G$, *then the equations* $aox = b$ *and* $yoa = b$ *have unique solutions in* G.
- Reversal Law, i.e., if $a, b \in G$ *then* $(aob)^{-1} = b^{-1}\, oa^{-1}$.
- A set S with a binary operation 'o' is said to be a semi-group if it satisfies the following two axioms:

  S_{G1}—(*Closure*). $a \in S, b \in S \Rightarrow aob \in S$.

  S_{G2}—(*Associativity*). If $a, b, c \in S$ then $(aob)\, oc = ao\,(boc)$.
- Multiplication Modulo **p**, (p being a prime). If a and b are two integers and p, a positive integer, then 'multiplication modulo p', is denoted by $a \times_p b$ and defined as $a \times_p b = r, o \leq r < p$ where r is the remainder obtained on dividing the ordinary product ab by p.
- It is commonly observed that a 'table' is a convenient way of either defining a binary operation in a finite set S or tabulating the effect of a binary operation in a set S.
- In forming a table or say a group table we arrange the elements of a group in rows and columns of a square array such that each element of the group occurs once and only once in each row or column.
- The composition element aob occurs at the intersection of row and column of the elements a and b of the group after the binary operation has been performed.
- If (G, o) be a group and H be the set of those elements $x \in G$, which commute with each element in G, i.e., the set

  $H = \{x : x \in G \text{ and } aox = xoa \;\; a \in G\}$

  then the set H is known as the centre of G.
- The order of every subgroup of a finite group is a divisor of the order of the group.
- The order of an element of a group G of finite order is a divisor of the order of the group.
- A finite group of prime order has no proper subgroup.
- If p be a prime and 'a' a natural number not divisible by p, then

  $a^{p-1} \equiv 1 \; (mod\, p)$

- If $x \leftrightarrow xf = x'$ be an automorphism of A where x' is the element of A in some order, then the mapping is automorphism and so $(xy)f = (xf)(yf) = x'y'$.
- A homomorphism of a group onto itself is said to be an endomorphism of the group.
- Every finite group G of order n (say) is isomorphic with a subgroup of symmetric group S_n.
- Every cyclic group of infinite order is isomorphic to the additive group of integers.
- The character of a special unitary group is found out if a typical matrix by means of unitary transformation is transformed to diagonal form.
- A continuous group formed from the set of all orthogonal n-dimensional matrices is said to be *n-dimensional rotational group*.

2.15 KEY TERMS

- **First order partial derivatives:** $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$ are the first order partial derivatives.
- **Second order partial derivatives:** $\frac{\partial^2 z}{\partial x^2}, \frac{\partial^2 z}{\partial y^2}, \frac{\partial^2 z}{\partial x \partial y}$ are the second order partial derivatives.
- **Complete integral:** A solution in which the number of arbitrary constants is equal to the number of independent variables is called complete integral or complete solution.
- **Partial differential equation:** Any equation which contains one or more partial derivatives is called a partial differential equation. $x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = z$; $\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} + \frac{\partial^2 z}{\partial x \partial y} = 0$ are examples for partial differential equation of first order and second order respectively.
- **Laplace equation:** In mathematics, Laplace equation is a second order partial differential equation. It is named after Pierre-Simon Laplace and is written as,

  $$\nabla^2 \varphi = 0$$

  Here is the Laplace operator and φ is a scalar function of 3 variables.
- **Poisson equation:** In mathematics, Poisson equation is a partial differential equation. It is named after the French mathematician, geometer and physicist Siméon-Denis Poisson. The Poisson equation is,

  $$\Delta \varphi = f$$

  Here Δ is the Laplace operator and f and φ are real or complex-valued functions on a manifold. If the manifold is Euclidean space, then the Laplace operator is denoted as and hence Poisson equation can be written as,

$$\nabla^2\varphi = f$$

- **Wave equation:** In its simplest form, the wave equation refers to a scalar function $u=(x_1, x_2,...,x_n,t)$ that satisfies,

NOTES

$$\frac{\partial^2 u}{\partial t^2} = c^2\nabla^2 u$$

Here is the spatial Laplacian and c is a fixed constant equal to the propagation speed of the wave and is also known as the non-dispersive wave equation.

- **Finite and Infinite Groups:** A group (G, o) consisting of a finite number of elements is said to be a *finite group*, for example, the set $S = \{1, \omega, \omega^2\}$ where $\omega^3 = 1$, is a finite group under multiplication composition. A group (G, o) consisting of an infinite number of elements is said to be an *infinite group*, for example, the set **I** (of all integers) is an infinite group under the addition composition.
- **Order of a Group:** The number of elements in a finite group is known as the order of the group.
- **Sub-Group.** A sub-group of a group (G, o) is any collection of elements of G satisfying the axioms of G. In other words, a non-empty subset say H of a group G is said to be the sub-group of G, if the binary operation *'o'* in G induces a binary operation in H and the elements of H obey the group axioms.
- **Proper Sub-Group:** A sub-group of a group (G, o) other than G itself and the group consisting of the identity element alone is termed as a proper sub-group of G. for example, the additive group of integers is a proper sub-group of the additive group of rational numbers.
- **Improper or Trival Sub-Groups:** The group (G, o) itself and the group consisting of identity alone, i.e., ($\{e\}$, o) are known as trivial or improper sub-groups of (G, o).

2.16 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. Define partial differential equation.
2. How will you identify the order of a partial differential equation? Give an example.
3. Which equations are termed as singular integral?
4. Write the first order partial derivative of the function $z(x,y)$.
5. Name some partial differential equations.

NOTES

6. Name two methods of converting ordinary equations to partial differential equations.
7. Define complete integral.
8. What is a Laplace equation?
9. What is a wave equation?
10. What is the equation of heat flow?
11. Give the diffusion equation or Fourier equation of heat flow?
12. Write about the uniqueness of a solution.
13. What is group/composition table?
14. What is homomorphism of a group? State the properties of homomorphism.
15. What is isomorphism of groups? Write the properties of isomorphism.
16. Define the term automorphism. Give the product of automorphism.
17. What is Schur's Lemma?
18. What is the application of group theory in physics?

Long Answer Questions

1. Obtain a partial differential equation by eliminating the arbitrary constants of the following:

   (i) $z = ax + by + \sqrt{a^2 + b^2}$ *(ii)* $\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{b^2} = 1$

   (iii) $z = xy + y\sqrt{x^2 - a^2} + b$ *(iv)* $z = ax^3 + by^3$

   (v) $(x-a)^2 + (y-b)^2 + z^2 = a^2 + b^2$ *(vi)* $2z = (ax + y)^2 + b$

2. Eliminate the arbitrary function from the following:

   (i) $z = e^y f(x + y)$ *(ii)* $z = f(my - lx)$

   (iii) $z = f(x^2 + y^2 + z^2)$ *(iv)* $z = x + y + f(xy)$

   (v) $z = f(x) + e^y g(x)$ *(vi)* $z = f(x + 4y) + g(x - 4y)$

   (vii) $z = f(2x + 3y) + y\,g(2x + 3y)$ *(viii)* $z = f(x + y) \cdot \phi(x - y)$

3. Solve the following differential equations:

   (i) $(3z - 4y)p + (4x - 2z)q = 2y - 3x$ *(ii)* $y^2zp + x^2zq = y^2x$

   (iii) $x^2p - y^2q = (x - y)z$ *(iv)* $xp + yq = 2z$

   (v) $x(z^2 - y^2)p + y(x^2 - z^2)q = z(y^2 - x^2)$

4. Eliminate the arbitrary function(s) from the following and form the partial differential equations:

   (i) $xy + yz + zx = f\left(\frac{z}{x + y}\right)$ *(ii)* $z = f(x^2 + y^2 + z^2)$

   (iii) $u = e^y f(x - y)$ *(iv)* $z = f(\sin x + \cos y)$

   (v) $\phi(x + y + z, x^2 + y^2 - z^2) = 0$ *(vi)* $z = f(2x + 3y) + \phi(y + 2x)$

   (vii) $u = f(x^2 + y) + g(x^2 - y)$ *(viii)* $u = x\,f(ax + by) + g(ax + by)$

NOTES

5. Find the complete solution of the following partial differential equations:

   (*i*) $pq + p + q = 0$ (*ii*) $p^3 = q^3$

   (*iii*) $p = e^q$ (*iv*) $z = px + qy + p^2 + pq + q^2$

   (*v*) $z = px + qy + \log pq$ (*vi*) $z = px + qy + p^2 - q^2$

   (*vii*) $z^2 = 1 + p^2 + q^2$
6. Explain Laplace equations with the help of examples.
7. Explain wave equations with the help of examples.
8. Establish the equation for heat conduction $k\nabla^2 T = \frac{\partial T}{\partial t}$.
9. Show that the set of all nth roots of unity form a finite abelian group G of order n under ordinary multiplication as composition.
10. Show that any non-commutative group has at least six elements.
11. Show that non-empty semi-group (G, o) forms a group if the equations $ax = b$ and $ya = b$ have unique solutions in G pair of elements $a, b \in G$.
12. Show that the set of subsets of a set with the union composition is a semi-group.
13. Show that the necessary and sufficient conditions for a complex H to be a subgroup (H, o) of a group (G, o) are

    $(i)\ a, b \in H \Rightarrow aob \in H\ \ a, b;$ *and* $(ii)\ a \in H \Rightarrow a^{-1} \in H\ \ a$
14. Show that the order of any element of a group is always equal to the order of its inverse.
15. Prove that the two left cosets aH and bH of a subgroup H of a group G are either identical or disjoint.
16. Show that the multiplicative group $G = \{1, -1, i, -i\}$ is isomorphic to the permutation group $G'\ \{I, (abcd), (ac)(bd)(adcb)\}$ on four symbols.
17. Explain direct sum and direct product.

2.17 FURTHER READING

Dass, HK. 2008. *Mathematical Physics*. New Delhi: S. Chand

Chattopadhyay, PK. 1990. *Mathematical Physics*. New Delhi: New age International

Hassani, Sadri. 2013. *Mathematical Physics: A Modern Introduction to Its Foundations* Berlin: Springer Science & Business Media

Khanna, V.K. and S.K. Bhambri. 2008. *A Course in Abstract Algebra*, 3rd edition. New Delhi: Vikas Publishing House

Grewal, B.S. 1998. *Higher Engineering Mathematics,* 34th edition. New Delhi: Khanna Publishers

Narayan, Shanti. 1996. *Differential Calculus,* 14th edition. New Delhi: S. Chand and Company

NOTES

Kreyszig, Erwin. *Advanced Engineering Mathematics,* 7th edition. Mumbai: Wiley Eastern Ltd

Bali, N.P. 2007. *A Textbook of Engineering Mathematics*. New Delhi: Laxmi Publications (P) Ltd

Narayan, Shanti. 2010. *A Text Book of Vector Analysis.* New Delhi: S. Chand and Sons

Spiegel. 1959. *Vector Analysis (Schaum'S Outline).* New York: McGraw Hill Professional

Gupta, B. D. 2009. *Mathematical Physics*, 4th Edition. Noida: Vikas Publishing House Pvt.

UNIT 3 FUNCTIONS OF COMPLEX VARIABLE

NOTES

Structure

3.0 INTRODUCTION

Cantor, Dedekind and Weierstrass etc., extended the conception of rational numbers to a larger field known as real numbers which constitute rational as well as irrational numbers. Evidently the system of real numbers is not sufficient for all mathematical needs, for example, there is no real number (rational or irrational) which satisfies $x^2 + 1 = 0$. It was, therefore, felt necessary by Euler Gauss, Hamilton, Cauchy, Riemann and Weierstrass, etc., to extend the field of real numbers to the still large field of complex numbers. Euler for the first time introduced the symbol i with the property $i^2 = -1$ and then Gauss introduced a number of the form

a + ib which satisfies every algebraic equation with real coefficients. Such a number a + ib with $i = -1$ and a, b being real, is known as a complex number.

In the field of complex analysis in mathematics, the Cauchy–Riemann equations consist of a system of two partial differential equations which, together with certain continuity and differentiability criteria, form a necessary and sufficient condition for a complex function to be complex differentiable, that is, holomorphic. Cauchy's integral formula, named after Augustin-Louis Cauchy, is a central statement in complex analysis. It expresses the fact that a holomorphic function defined on a disk is completely determined by its values on the boundary of the disk. This unit discusses Cauchy–Riemann equations, which gives a necessary and sufficient condition for a complex function to be complex differentiable. The derivation of Cauchy integral formula is given to provide integral formulas for all derivatives of a holomorphic function. Cauchy's formula shows that, in complex analysis, 'differentiation is equivalent to integration': complex differentiation, like integration, behaves well under uniform limits – a result denied in real analysis. A conformal map is a function that preserves orientation and angles locally. In the most common case, the function has a domain and an image in the complex plane.

In complex analysis (a branch of mathematics), a pole is a certain type of singularity of a function, nearby which the function behaves relatively regularly, in contrast to essential singularities, such as 0 for the logarithm function, and branch points, such as 0 for the complex square root function. A zero of a meromorphic function f is a complex number z such that

$f(z) = 0$. A pole of f is a zero of $1/f$. A function f of a complex variable z is meromorphic in the neighbourhood of a point z_0 if either f or its reciprocal function $1/f$ is holomorphic in some neighbourhood of z_0 (that is, if f or $1/f$ is complex differentiable in a neighbourhood of z_0). In this unit, you will study about the zeros and poles, Cauchy's Residue theorem and contour integration.

3.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the concept of conformal mapping
- State the purpose of Argand diagram
- Explain the method of analytic continuation
- Evaluate the Cauchy-Riemann conditions
- Discuss the problems using Cauchy integral formula
- Describe the meaning of zeroes and poles
- State Cauchy's Residue theorem
- Discuss the method of contour integration

3.2 COMPLEX NUMBER: DEFINITION

An ordered pair of real numbers such as (x, y) is termed as a complex number. If we write

$$z = (x, y) \text{ or } x + iy, \text{ where } i = \sqrt{-1}, \text{ then}$$

x is called the *real part* and y the *imaginary part* of the complex number z and denoted by

$$x = R_z \text{ or } R(z) \text{ or } Re(z)$$

$$y = I_z \text{ or } I(z) \text{ or } Im(z)$$

Equality of complex numbers. Two complex numbers (x, y) and (x', y') are equal iff $x = x'$ and $y = y'$.

Modulus of a complex number. If $z = x + iy$ be a complex number then its modulus (or module) is denoted by $|z|$ and given by

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

Evidently $|z| = 0$ iff $x = 0, y = 0$.

3.3 FUNCTION OF A COMPLEX VARIABLE

NOTES

Exponential Functions: If $z = x + iy$ and y is used as radian measure of the angle to define cos y, sin y etc., then the exponential function in terms of real valued functions is defined by

$$e^x = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y) \qquad ...(3.1)$$

In case z is purely real, i.e., $y = 0$, we have $e^z = e^x$ ...(3.2)

and if z is purely imaginary, i.e., $x = 0$, we have

$$e^{iy} = \cos y + \sin y \qquad ...(3.3)$$

As such Maclaurin series representation of e^t on replacing t by iy, gives

$$\sum_{n=0}^{\infty} \frac{(iy)^n}{\lfloor n} = \sum_{n=0}^{\infty} \frac{i^{2n} y^{2n}}{\lfloor 2n} + \sum_{n=0}^{\infty} \frac{i^{2n+1} y^{2n+1}}{\lfloor 2n+1} \text{ where } \lfloor 0 = 1$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n}}{\lfloor 2n} + i \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n+1}}{\lfloor 2n+1} \qquad ...(3.4)$$

which are Maclaurin series for cos y and sin y respectively.

The Exponential function given by (3.1) is an *entire function* since

$$\frac{d}{dz} e^z = e^z \qquad ...(3.5)$$

Similarly $$\frac{d}{dz} e^w = e^w \frac{dw}{dz} \qquad ...(3.6)$$

w being an analytic function of z.

Polar form of Equation (3.1) is $e^z = r(\cos\theta + i \sin\theta) = re^{i\theta}$ ...(3.7)

where $r = e^x$, $\theta = y$.

$\therefore\ |e^z| = r = e^x$ and arg $e^z = \theta = y$ ...(3.8)

Also $|e^z| > 0$ i.e. $e^z \neq 0$ for every value of z ...(3.9)

So the range of the exponential function is the entire complex plane excluding the origin where r = 0.

Now $$e^z = -1 = \cos(\pi \pm 2m\pi) + i \sin(\pi \pm 2m\pi) \text{ gives}$$

$$x = 0 \text{ and } y = \pi \pm 2m\pi,\ m = 0, 1, 2,...$$

If $$e^{z_1} = r_1 (\cos\theta_1 + i \sin\theta_1) \text{ so that } r_1 = e^{x_1},\ \theta_1 = y_1$$

and $$e^{z_2} = r_2 (\cos\theta_2 + i \sin\theta_2) \text{ so that } r_2 = e^{x_2},\ \theta_2 = y_2$$

then, $$e^{z_1} \cdot e^{z_2} = r_1 r_2 \{\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)\}$$

$$= e^{x_1} e^{x_2} [\cos(y_1 + y_2) + i \sin(y_1 + y_2)]$$

$$= e^{x_1 + x_2} \cdot e^{i(y_1 + y_2)} = e^{x_1 + x_2 + i(y_1 + y_2)}$$

$$= e^{(x_1 + iy_1) + (x_2 + iy_2)} = e^{z_1 + z_2} \qquad ...(3.10)$$

Similarly $$e^{z_1} / e^{z_2} = e^{z_1 - z_2} \qquad ...(3.11)$$

$$\frac{1}{e^z} = e^{-z} \qquad ...(3.12)$$

$$(e^z)^n = e^{nz} \qquad ...(3.13)$$

n being a positive integer.

$$(e^z)^{m/n} = e^{m/n}(z + 2\pi pi) \qquad ...(3.14)$$

NOTES

$p = 0, 1, 2,...., n - 1$ and m, n are integers (+ ve).

Also $$e^{z + 2\pi i} = e^z e^{2\pi i} = e^z \text{ as } e^{2\pi i} = 1 \qquad ...(3.15)$$

From Equation (3.15) it follows that the exponential function is *periodic*.

Again $$e^{\bar{z}} = \overline{(e^z)} \qquad ...(3.16)$$

In polar form $z = re^{i\theta}$, $\bar{z} = re^{-i\theta}$ so that

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}, \; \frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1 - \theta_2)} \qquad ...(3.17)$$

Trigonometric Functions

$$e^{iy} = \cos y + i \sin y \quad \text{and} \quad e^{-iy} = \cos y - i \sin y \text{ yield}$$

$$\cos y = \frac{e^{iy} + e^{-iy}}{2}, \; \sin y = \frac{e^{iy} - e^{-iy}}{2i} \qquad ...(3.18)$$

$\sin z$ and $\cos z$ are *entire functions* as

$$\frac{d}{dz} \sin z = \cos z, \; \frac{d}{dz} \cos z = - \sin z \text{ etc.} \qquad ...(3.19)$$

Now,

$$\cos z = \cos (x + iy) = \tfrac{1}{2}[e^{i(x+iy)} - e^{i(z+iy)}]$$

$$= \frac{1}{2}\left[e^{ix-y} + e^{-ix+y}\right].$$

$$= \frac{1}{2} e^{-y} (\cos x + i \sin x) + \frac{1}{2} e^{y} (\cos x - i \sin x)$$

$$= \frac{e^y + e^{-y}}{2} \cos x - i \frac{e^y - e^{-y}}{2} \sin x \qquad ...(3.20)$$

Introducing the *hyperbolic functions* with the properties

$$\sin z = \frac{e^z - e^{-z}}{2}, \; \cosh z = \frac{e^z + e^{-z}}{2} \qquad ...(3.21)$$

$$\frac{d}{dz} \sinh z = \cosh z, \; \frac{d}{dz} \cosh z = \sinh z \text{ etc.} \qquad ...(3.22)$$

$\cosh^2 z - \sinh^2 z = 1$

$$\cos iz = \cosh z \quad \text{and} \quad \sin iz = i \sinh z \qquad ...(3.23)$$

the Relation (3.20) becomes

$$\cos z = \cos (x + iy) = \cos x \cosh y - i \sin x \sinh y \qquad ...(3.24)$$

Similarly $\sin z = \sin (x + iy) = \sin x \cosh y + i \cos x \sinh y$ ...(3.25)

$$\text{Also } \overline{\sin z} = \sin \bar{z} \text{ and } \overline{\cos z} = \cos \bar{z} \qquad ...(3.26)$$

$$\cos (z + \pi) = - \cos z, \quad \sin (z + \pi) = - \sin z, \text{ etc.} \qquad ...(3.27)$$

It is easy to show that

$$|\sin z|^2 = \sin^2 x + \sinh^2 y \qquad ...(3.28)$$

NOTES

$$|\cos z|^2 = \cos^2 x + \sinh^2 y \qquad ...(3.29)$$

A value of z for which $f(z) = 0$ is known as a *zero* of the function f.

The real zeros of sin z and cos z are their only zeros, since

$\sin z = 0$ from Equation (3.25) gives

$\sin x \cosh y = 0, \cos x \sinh y = 0$

x and y being real, $\cosh y \geq 1$ and $\sin x = 0$ only when $x = 0, \pm\pi, \pm 2\pi, ...$ and for these values of x, $\cos x \neq 0$ and thus $\sinh y = 0$, i.e., $y = 0$.

Also

$$\left.\begin{aligned} &\sin z = 0 \Rightarrow z = 0 \text{ or } \pm n\pi, \quad n = 1, 2, 3, ... \\ &\cos z = 0 \Rightarrow z = \pm\frac{(2n-1)\pi}{2}, n = 1, 2, 3, ... \end{aligned}\right\}$$

We may also show that

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

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

So that $\quad |\sinh z|^2 = \sinh^2 x + \sin^2 y$

$\qquad |\cosh z|^2 = \sinh^2 x + \cos^2 y$

$\sinh z$ and $\cosh z$ are periodic with *period* $2\pi i$.

Logarithmic Functions and Branch Points

$z = re^{i\theta}$, θ being measured in radian, gives

$$\left.\begin{aligned} \log z = \log re^{i\theta} &= \log r + i\theta \\ &= \log|z| + i \arg\theta \end{aligned}\right\} r > 0 \text{ and } -\pi < \theta < \pi \qquad ...(3.30)$$

If $-\pi < \theta \leq \pi$, then $z = re^{i(\theta \pm 2n\pi)}$, $n = 0, 1, 2, ...$

So that $\log z = \log r + i(\theta + i2n\pi)$, $n = 0, 1, 2, ...$

We write log z for *principal value* of log z and Log z for its *general value*.

In Equation (3.30) if we put $\log r = u$, $\theta = v$ so that

$$\frac{\partial u}{\partial r} = \frac{1}{r}, \quad \frac{\partial u}{\partial \theta} = 0, \quad \frac{\partial v}{\partial r} = 0, \quad \frac{\partial v}{\partial \theta} = 1,$$

are all continuous functions of z.

Also, $$\frac{d}{dz}\log z = e^{-i\theta}\left(\frac{\partial u}{\partial r} + i\frac{\partial v}{\partial r}\right) = \frac{1}{re^{i\theta}} = \frac{1}{z}$$

where $z \neq 0$ and $-\pi < \arg z < \pi$.

A *branch F* of a many-valued function f is any single-valued function which is analytic in some domain at each point of which the value $F(z)$ is one of the values $f(z)$. The Equation (3.30) gives the *principal branch* log z. Each point of the negative real axis $\theta = \pi$ along with the origin is a singular point of the principal branch log z. Then (say) $\theta = \pi$ is said to be a *branch cut* for the principal branch and the singular point $z = 0$ common to all branch cuts for the many-valued function log z is known as a *branch point*.

NOTES

Now if $w = \log z$, then $e^w = e^{\log z} = e^{(\log r + i\theta)}$

$$= e^{\log r} e^{i\theta} = re^{i\theta} = z$$

i.e. $e^{\log z} = z, z \neq 0$

and if $e^z = w$, $\log w = \log e^{x+iy} = \log e^x + \log e^{iy}$.

$$= x + i\,(y \pm 2p\pi),\ p = 0, 1, 2, 3,...$$
$$= x + iy \pm 2p\pi i$$
$$= z \pm 2p\pi i$$

So $\log w = z$ when $e^z = w$

and $\log e^x = z$ for appropriate choice of logarithm

Again if $z_1 = r_1 e^{i\theta_1}$, $z_2 = r_2 e^{i\theta_2}$, $r_1 > 0$, $r_2 > 0$, then

$$\log z_1 + \log z_2 = \log r_1 e^{i\theta_1} + \log r_2 e^{i\theta_2}$$
$$= \log r_1 + \log r_2 + i(\theta_1 + \theta_2)$$
$$= \log r_1 r_2 + i(\theta_1 + \theta_2)$$
$$= \log z_1 z_2$$

Similarly $\log z_1 - \log z_2 = \log \frac{z_1}{z_2}$

It is easy to verify that

$$\log z^m = m \log z$$
$$\log z^{1/n} = \frac{1}{n} \log z$$
$$z^{m/n} = e^{m/n \log z}$$

In case of *complex exponents*, we define

$$z^c = e^{c \log z},\ z, c \text{ being complex and } z \neq 0$$
$$\frac{d}{dz} z^c = e^{c \log z} \cdot \frac{c}{z} = c \cdot \frac{e^{c \log z}}{e^{\log z}} = ce^{(c-1)\log z} = cz^{c-1}$$

Also $c^z = e^{z \log c},\ c \neq 0,$

$$\frac{d}{dz} c^z = c^z \log c,\ c \neq 0$$

Inverse Trigonometric Functions

Defining the inverse of sine function as $w = \sin^{-1} z$, we have

$$z = \sin w = \frac{e^{iw} - e^{-iw}}{2i},\ \text{i.e., } e^{2iw} - 2ize^{iw} - 1 = 0.$$

Being qudratic in e^{iw}, this gives, $e^{iw} = iz + \sqrt{(1-z^2)}$.

So that $w = \sin^{-1} z = -i \log \left\{iz + \sqrt{1-z^2}\right\}$

Similarly $\cos^{-1} z = -i \log \left\{z + \sqrt{z^2 - 1}\right\}$

$$\tan^{-1} z = \frac{i}{2} \log \frac{1-iz}{1+iz} = \frac{i}{2} \log \frac{1+z}{1-z}$$

So that $\frac{d}{dz} \sin^{-1} z = \frac{1}{\sqrt{1-z^2}},\ \frac{d}{dz} \tan^{-1} z = \frac{1}{1+z^2}$

Also

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

$$\cosh^{-1} z = \log\{z+\sqrt{z^2-1}\}$$

$$\tan^{-1} z = \frac{1}{2}\log\frac{1+z}{1-z}$$

NOTES

Mapping

If $w = f(z)$ and corresponding to each point (x, y) in z-plane in a domain of function f, there is a point (u, v) in w-plane where $z = x + iy$ and $w = u + iv$, then this correspondence between the points of two planes is said to be a *mapping* or a *transformation* of points in the z-plane into points of the w-plane by the function f. Corresponding points or set of points are known as *images* of each other. The use of graphic terms as *translation*, *rotation* or *reflection* is rather convenient in mapping, for example, the mapping $w = z + c$, c being a complex constant gives the translation of every point z through the vector c, i.e., if $z = x + iy$, $w = u + iv$, $c = c_1 + ic_2$, then the image of any point (x, y) in z-plane is the point $(x + c_1, y + c_2)$.

The mapping $w = Bz$ where $B = be^{i\beta}$ and $z = re^{i\theta}$

i.e., $$w = br\, e^{i(\theta+\beta)}$$

maps the point (r, θ) in z-plane into a point $(br, \theta + \beta)$ into w-plane, i.e., the mapping consists of a rotation of the radius vector of z about the origin through an angle $\beta = \arg B$ and an extension or contraction of radius vector r by $b = |B|$.

As an illustration the function $w = z^2$ maps the entire first quadrant of the z-plane, $0 \le \theta \le \pi/2$, $r \ge 0$, into the entire upper half of the w-plane.

The transformation

$$T : w = \frac{az+b}{cz+d}$$

a, b, c, d being complex constants is termed as the *linear fractional transformation* or *bilinear transformation* or *Mobius transformation*.

Here T^{-1} i.e. inverse of T is $z = \dfrac{-dw+b}{cw-a}$

Any set of elements which satisfies all the following conditions is called a *group*:

(*i*) There is a rule of combination such that product TT' for each distinct pair T, T' of elements is an element of the set.

(*ii*) The product is associative, i.e., $T(T'T'') = (TT')T''$

(*iii*) The set contains an identity T_0 such that $TT_0 = T_0T = T$ for each element T.

(*iv*) Each element T has an inverse T^{-1} s.t. $TT^{-1} = T^{-1}T = T_0$.

It may be shown that the set of all linear fractional transformations is a group.

If besides the above four properties a group also satisfies the commutative property then it is called as *Commutative group* or an *Abelian group*.

Problem 3.1. *Show that the set of complex numbers form an abelian group under addition.*

NOTES

Take three complex numbers, z_1 (x_1, y_1), $z_2 = (x_2, y_2)$, $z_3 = (x_3, y_3)$ belonging to the set **C** of complex numbers. Then the addition is commutative since

$$z_1 + z_2 = (x_1 + x_2, y_1 + y_2) = (x_2 + x_1, y_2 + y_1)$$
$$= z_2 + z_1 \ \forall \ z_1, z_2 \in \mathbf{C}$$

Also $\quad z_1, z_2 \in \mathbf{C} \Rightarrow z_1 + z_2 \in \mathbf{C}$

The addition is associative since

$$\begin{aligned} z_1 + (z_2 + z_3) &= (x_1, y_1) + (x_2 + x_3, y_2 + y_3) \\ &= (x_1 + x_2 + x_3, y_1 + y_2 + y_3) \\ &= (x_1 + x_2, y_1 + y_2) + (x_3, y_3) \\ &= (z_1 + z_2) + z_3. \end{aligned}$$

There exists an additive identity **o** = (o, o) such that

$$z + \mathbf{o} = (x, y) + (o, o) = (x, y) = z$$

There exists an additive inverse $(-z)$ $\forall$ z such that

$z + (-z) = (x, y) + (-x, -y) = (o, o)$ the identity element.

Hence the set of complex numbers form an abelian group.

Problem 3.2. *Show that the set of complex numbers form an abelian group under multiplication.*

It is easy to show that

$(x_1, y_1), (x_2, y_2) \in \mathbf{C} \Rightarrow (x_1, y_1)\,(x_2, y_2)$ i.e. $(x_1x_2 - y_1y_2, x_1y_2 + x_2y_1) \in \mathbf{C}$

$$(x_1, y_1)\ \{(x_2, y_2)\,(x_3, y_3)\} = \{(x_1, y_1)\,(x_2, y_2)\}\ (x_3, y_3),$$
$$(x_1, y_1), (x_2, y_2), (x_3, y_3) \in \mathbf{C}$$

$\exists$ multiplicative identity $(1, 0)$ s.t. $(x, y)\,(1, 0) = (x, y)\ \forall\ (x, y) \in \mathbf{C}$

$\exists$ multiplicative inverse s.t. $(x, y)\left(\dfrac{x}{x^2+y^2}, \dfrac{-y}{x^2+y^2}\right) = (1, 0)$

The commutative law holds, i.e.,

$$(x_1, y_1) + (x_2, y_2) = (x_2, y_2) + (x_1, y_1).$$

Hence the given set is an abelian group.

Complex Differential Operators

If $z = x + iy$ so that $\bar{z} = x - iy$ and F be a continuous differential function, then

$$\frac{\partial F}{\partial x} = \frac{\partial F}{\partial z}\cdot\frac{\partial z}{\partial x} + \frac{\partial F}{\partial \bar{z}}\frac{\partial \bar{z}}{\partial x} = \frac{\partial F}{\partial z} + \frac{\partial F}{\partial \bar{z}} \text{ giving } \frac{\partial}{\partial x} \equiv \frac{\partial}{\partial z} + \frac{\partial}{\partial \bar{z}}$$

and $$\frac{\partial F}{\partial y} = \frac{\partial F}{\partial z}\frac{\partial z}{\partial y} + \frac{\partial F}{\partial \bar{z}}\frac{\partial \bar{z}}{\partial y} = i\frac{\partial F}{\partial z} - i\frac{\partial F}{\partial \bar{z}} \text{ giving } \frac{\partial}{\partial y} \equiv i\left(\frac{\partial}{\partial z} - \frac{\partial}{\partial \bar{z}}\right)$$

We then define ∇ (Del operator) $\equiv \dfrac{\partial}{\partial x} + i\dfrac{\partial}{\partial y} = \dfrac{\partial}{\partial z} + \dfrac{\partial}{\partial \bar{z}} + i^2\left(\dfrac{\partial}{\partial z} - \dfrac{\partial}{\partial \bar{z}}\right) = 2\dfrac{\partial}{\partial \bar{z}}$

and $\overline{\nabla}$ (Del bar) $\equiv \dfrac{\partial}{\partial x} - i\dfrac{\partial}{\partial y} = \dfrac{\partial}{\partial z} + \dfrac{\partial}{\partial \bar{z}} - i^2\left(\dfrac{\partial}{\partial z} - \dfrac{\partial}{\partial \bar{z}}\right) = 2\dfrac{\partial}{\partial z}$

NOTES

Now taking $F(x, y)$ as real continuously differentiable function of scalars x, y and $A(x, y) = P(x, y) + i\,Q(x, y)$ as complex continuously differentiable function of vectors x and y, then

$$F(x, y) = F\left(\frac{z+\bar{z}}{2}, \frac{z+\bar{z}}{2i}\right) = G(z,\bar{z}) \text{ say and } A(x, y) = B(z,\bar{z}) \text{ say}$$

The *Gradient* of a real scalar function F is defined by

$$\text{grand } F = \nabla F = \frac{\partial F}{\partial x} + i\frac{\partial F}{\partial y} = 2\frac{\partial F}{\partial \bar{z}}$$

and the gradient of a complex vector function A is defined as

$$\text{grand } A = \nabla A = \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)(P+iQ) = \frac{\partial P}{\partial x} - \frac{\partial Q}{\partial y} + i\left(\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right) = 2\frac{\partial B}{\partial \bar{z}}$$

Geometrically interpreted, ∇F is a vector normal to the curve $F(x, y) =$ constant, and if B is analytic function of $\bar{z}$ so that $\frac{\partial B}{\partial \bar{z}} = 0$, then gradient is also zero i.e. $\frac{\partial P}{\partial x} = \frac{\partial Q}{\partial y}, \frac{\partial P}{\partial y} = -\frac{\partial Q}{\partial x}$ showing that Cauchy-Riemann equations are satisfied.

The *Divergence* of a complex vector function A is defined as

$$\text{div } A = \nabla \cdot A = Re\,\{\bar{\nabla} \cdot A\} = Re\left\{\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(P+iQ)\right\} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}$$

$$= 2\,Re\left\{\frac{\partial B}{\partial z}\right\} \text{ where } Re \text{ denotes real part.}$$

It is notable that the divergence of a real or complex function is always a real function.

The *Curl* of a complex vector function A is defined as

$$\text{curl } A = \nabla \times A = Im\,(\bar{\nabla} \times A) = Im\left\{\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\times(P+iQ)\right\}$$

$$= i\left\{\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right\} = 2Im\left\{\frac{\partial B}{\partial z}\right\}$$

The *Laplacian operator* is defined as the scalar or dot product of ∇ with itself

i.e. $$\nabla \cdot \nabla = \nabla^2 = Re\,(\bar{\nabla}\,\nabla) = R\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)$$

$$= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = 4\frac{\partial^2}{\partial z \partial \bar{z}}$$

Here below we summarize few identities involving grad, div and curl of two complex diffentiable functions A_1 and A_2.

$$\text{grad } (A_1 + A_2) = \text{grad } A_1 + \text{grad } A_2$$
$$\text{div } (A_1 + A_2) = \text{div } A_1 + \text{div } A_2$$
$$\text{curl } (A_1 + A_2) = \text{curl } A_1 + \text{curl } A_2$$
$$\text{grad } (A_1 A_2) = A_1 (\text{grad } A_2) + (\text{grad } A_1) A_2$$

$$\text{curl (grad } A) = 0, \text{ when } A \text{ is real or } Im\ (A) \text{ is harmonic}$$

$$\text{div (grad } A) = 0, \text{ when } A \text{ is imaginary or } Re\ (A) \text{ is harmonic}$$

NOTES

3.3.1 Conformal Mapping

We can set a correspondence between a domain D of x-y plane and a domain D' of u-v plane by the transformation or mapping $w = f(z)$, i.e., $u = u(x, y)$ and $v = v(x, y)$, for example, if $u = x^2$ and $v = y^2$, then the circular domain $x^2 + y^2 \leq 1$ in z-plane corresponds to the triangle in w-plane bounded by the lines $u = 0$, $v = 0$ and $u + v = 1$.

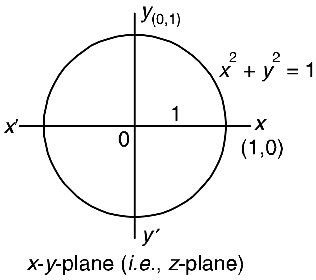

Fig. 3.1

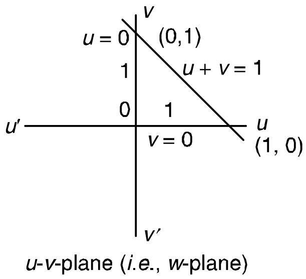

Fig. 3.2

Conformal and Isogonal Transformations. If the curves in z-plane intersect at a point z_0 (x_0, z_0) at an angle θ, then if the two corresponding curves in w-plane intersect at w_0 (u_0, v_0) at the same angle θ, where w_0 corresponds z_0, the transformation or mapping is known as **Isogonal**. In other words if only the magnitude of the angle is preserved, the mapping is Isogonal, but if the sense of rotation as well as the magnitude of the angle is preserved, the mapping is **conformal**.

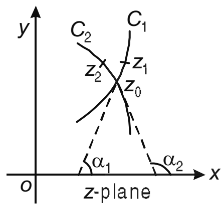

Fig. 3.3

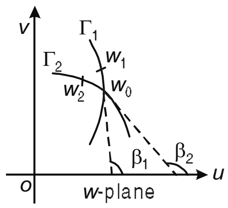

Fig. 3.4

Necessary and Sufficient Conditions for Conformality

I. If $f(z)$ is analytic, then the mapping is conformal, i.e., the necessary condition for conformality is that $f(z)$ must be analytic.

In other words, an analytic function is necessarily conformal.

Take points z_1, z_2 on curves C_1, C_2 in z-plane near to z_0 at distance r such that tangents at z_1, z_2 make angles, α_1, α_2 with the real axes. Let the corresponding points in w-plane be w_1, w_2 on curves Γ_1, Γ_2, near to w. Then $z_1 - z_0 = r\, e^{i\theta_1}$, $z_2 - z_0 = re^{i\theta_2}$ and if as $r \to 0$, $\theta_1 \to \alpha_1$, $\theta_2 \to \alpha_2$.

Also $w_1 - w_0 = \rho_1 e^{i\phi_1}$, $w_2 - w_0 = \rho_2 e^{i\phi_2}$ and if $\rho_1 \to 0$, $\phi \to \beta_1$, $\phi_2 \to \beta_2$

then $f'(z_0) = \lim_{z \to z_0} \frac{w_1 - w_0}{z_1 - z_0}$

or say, $$Re^{i\lambda} = \text{Lim} \frac{\rho_1 e^{i\phi_1}}{re^{i\theta_1}}$$ or $f'(z_0) \neq 0$ and it may be written as

$$Re^{i\lambda} = \text{Lim} \frac{\rho_1}{r} e^{i(\phi_1 - \theta_1)}$$

Equating moduli and arguments on either side, we get

$$R = \text{Lim} \frac{\rho_1}{r} = |f'(z_0)| \text{ and } \lambda = \text{Lim } (\phi_1 - \theta_1) = \text{Lim } \phi_1 - \text{Lim } \theta_1 = \beta_1 - \alpha_1$$

giving $\beta_1 = \alpha_1 + \lambda$ and similarly $\beta_2 = \alpha_2 + \lambda$

$\therefore$ $\beta_1 - \beta_2 = \alpha_1 - \alpha_2 \Rightarrow$ The angle between the curves C_1 and C_2 in z-plane is the same as the angle between the curves Γ_1 and Γ_2 in w-plane, i.e., the magnitude as well as the sense of rotation of the two angles in z and w planes is the same. This means the transformation is conformal.

II. If the mapping is conformal, then the function w = f (z) is analytic, i.e., **the sufficient condition for conformality is that f (z) is analytic.** Consider a pair of differentiable relations

$$u = u(x, y),\ v = v(x, y)$$

defining a transformation from z-plane (i.e., x-y plane) to w-plane (i.e., u-v plane). The transformation being of the form $w = f(z)$, where $f(z)$ is regular, is conformal.

Take the elements of lengths $d\sigma$ and ds in (u, v) and (x, y) planes respectively, so that

$$ds^2 = dx^2 + dy^2 \quad ...(3.31)$$

and $$d\sigma^2 = du^2 + dv^2 \quad ...(3.32)$$

But $\partial u = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy$ and $\partial v = \frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy$

$$\therefore \quad d\sigma^2 = \left\{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2\right\}(dx)^2 + 2\left\{\frac{\partial u}{\partial x}\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\frac{\partial v}{\partial y}\right\} dx\, dy + \left\{\left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2\right\}(dy)^2$$

Mapping being conformal, the ratio $d\sigma : ds$ is independent of direction if

$$\frac{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2}{1} = \frac{\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} + \frac{\partial v}{\partial x}\frac{\partial v}{\partial y}}{0} = \frac{\left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2}{1}$$ by Equation (3.31) and (3.32)

giving $$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 = \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 = h^2 \text{ (say)} \quad ...(3.33)$$

and $$\frac{\partial u}{\partial x}\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\frac{\partial v}{\partial y} = 0 \quad ...(3.34)$$

Equation (3.33) is satisfied if

$$u_x\left(i.e., \frac{\partial u}{\partial x}\right) = h\cos\alpha,\ v_x = h\sin\alpha,\ u_y = h\cos\beta,\ v_y = h\sin\beta$$

NOTES

and Equation (3.34) is satisfied if $\alpha - \beta = \pm\frac{\pi}{2}$.

Thus the correspondence is Isogonal if either

$(a)\ u_x = v_y,\ u_x = -u_y$ or $(b)\ u_x = -v_y,\ v_x = u_y$

Here (a) *are Cauchy-Riemann equations* and expressed as $w = f(z)$ i.e. $u + iv = f(x + iy)$, where $f(z)$ is regular function of z and equations (b) reduce to (a) by writing $-v$ for v, *i.e.*, by taking image formed by reflection in the real axis of the w-plane so that (b) correspond to an Isogonal but not conformal transformation.

Hence the only conformal transformation of z-plane into w-plane is of the form $w = f(z)$, where $f(z)$ is a regular function of z.

III. The case f ′(z) = 0. Assuming that $f'(z_0)$ has a zero of order $(n-1)$ at the point z_0, we have in the neighbourhood of z_0 by Taylor's theorem,

$$f(z) = f(z_0) + a_n(z - z_0)^n + \cdots \text{ where } a_n = \frac{f_n(z_0)}{\lfloor n} \neq 0$$

$$\therefore \quad f(z) - f(z_0) = a_n(z - z_0)^n + \cdots$$

i.e., $$w_1 - w_0 = a_n(z_1 - z_0)^n + \cdots$$

or $$\rho_1 e^{i\phi 1} = |a_n| r^n e^{i(n\theta_1 + \lambda)} + \cdots \text{ where } \lambda = \arg a_n$$

$$\therefore \quad \text{Lim}\ \phi_1 = \text{Lim}\ (n\theta_1 + \lambda) = n\alpha_1 + \lambda$$

Similarly, $$\phi_2 = n\alpha_2 + \lambda.$$

As such the curve Γ_1, Γ_2 still have definite tangents at w_0, but the angle between the tangents is

$\text{Lim}\ (\phi_2 - \phi_1) = n(\alpha_2 - \alpha_1) \Rightarrow$ the magnitude of the angle is not preserved, but magnified.

Linear magnification $R = \text{Lim}\frac{1}{r} = 0 \Rightarrow$ the conformal property does not hold good at such a point where $f'(z) = 0$.

Note. The points at which $\frac{dw}{dz} = 0$ *or* ∞ *are called* **critical points** *of the transformation defined by* $w = f(z)$.

IV. Transformations which are Isogonal but not Conformal. In this case, the magnitude of the angles is conserved but their sign is changed such as

$$w = x - iy = \bar{z}$$

which replaces every point by its reflection in the real axis, so that angles are conserved but signs are changed. In general it is true for the transformation of the type

$$w = f(\bar{z}),\ f(z) \text{ being regular.}$$

It is a combination of two transformations

(*i*) $\zeta = \bar{z}$, (*ii*) $w = f(\zeta)$

In (*i*), angles are conserved but their signs are changed and in (*ii*) angles as well as signs are conserved. Thus $w = f(\bar{z})$ gives a transformation which is isogonal but not conformal.

Linear or Bilinear or Möbius' Transformation

Its form is $$w = \frac{az+b}{cz+d} \qquad ...(3.35)$$

where z, w are complex variables and a, b, c, d are complex constants.

If we write Equation (3.55) in the form

$$cwz + dw - az - b = 0 \qquad ...(3.36)$$

then it is linear in z as well as w and hence it is called bilinear. It was studied by A.F. Möbius (1790–1868) as mentioned by Caratheodory and hence bears the name of Mobius.

Writing Equation (3.35) as $w = \dfrac{a\,z + b/a}{c\,z + d/c}$

it follows that for every value of z, there is the same value of w, if $\frac{b}{a} = \frac{d}{c}$, *i.e.*, $ad - bc = 0$ and there correspond different values of w to different values of z if $ad - bc \neq 0$.

Here the expression ($ad - bc$) is known as the **Determinant of the Transformation**.

Equation (3.36) $\Rightarrow$ For every $z \neq -\dfrac{d}{c}$, there exists a value of w ...(3.37)

Equation (3.35) can also be written is

$$z = \frac{dw-b}{-cw+a} = -\frac{d}{c}\frac{w-b/d}{w-a/c} \qquad ...(3.38)$$

This implies, that for every $w \neq \dfrac{a}{c}$, there exists a value of z

Now, Equations (3.37) and (3.38) $\Rightarrow$ In (3.35), the correspondence between w and z is one-one, except that when $c \neq 0$ for $z = -\dfrac{d}{c}$, $|w| \to \infty$ so that we may regard the point at infinity in w-plane (or extended plane) as corresponding to the point $z = -\dfrac{d}{c}$ in z-plane. Similarly the point at infinity in z-plane corresponds to $w = \dfrac{a}{c}$ in w-plane.

When $c = 0$, then Equation (3.35) gives $w = \dfrac{a}{d}z + \dfrac{b}{d}$, so that for $a \neq 0$, the points at infinity in the two planes correspond.

Again if we can write Equation (3.31) as

$$w = \frac{a}{c} + \frac{bc-ad}{c}\frac{1}{cz+d} \text{ with } ad - bc \neq 0 \text{ and } c \neq 0$$

or $\quad w = \frac{a}{c} + \frac{bc - ad}{c^2} \frac{1}{z + \frac{d}{c}}$

then it is obtained by a superimposition of successive mappings

NOTES

(*i*) $\tau = z + \frac{d}{c}$ of the type $z + \alpha$ known as **translation**

(*ii*) $\zeta = \frac{1}{\tau}$ known as **inversion**

(*iii*) $w = \frac{a}{c} + \frac{bc - ad}{c^2}\zeta$ of the type $w = \beta z$, known as **magnification and rotation**.

Note. Points which coincide with their transforms under bilinear transformation are known as **fixed points of the bilinear transformation**. Thus z is a fixed point of (5) if

$$z = \frac{az + b}{cz + d}, \text{ i.e., } cz^2 + (d - a)z - b = 0$$

It will have distinct roots ζ_1, ζ_2 if $c \neq 0$ and the discriminant $\Delta^2 = (d - a)^2 + 4bc \neq 0$.

We can take $\zeta_1 = \frac{a - d + s}{2c}, \zeta_2 = \frac{a - d - s}{2c}$.

Conversely if $c \neq 0$, each of these points is mapped onto itself, but if $\Delta = 0$, there is only one fixed point.

If $c = 0$, but $d \neq 0$, we have $w = \frac{a}{d}z + \frac{b}{d}$ ...(3.39)

In bilinear transformation ∞ is also a fixed point and another possible fixed point is given by Equation (3.39), i.e., by

$$(d - a)z = b$$

Hence, if $d - a \neq 0$, the transformation has the two fixed points ∞ and $\frac{b}{d - a}$ and if $d = a$, it has only one fixed point, *viz.*, ∞.

Conclusively, the mapping or transformation of a simple analytic (regular) function $w = f(z)$ can be performed in three manners:

1. **Translation**. $w = z + \alpha$, where α is a complex consant; or $u + iv = x + iy + (p + iq)$, where $\alpha = p + iq$

   $\Rightarrow \quad u = x + p$ and $v = y + q$

   or $\quad x = u - p$ and $y = v - q$

   Thus a point $P(x, y)$ in the z-plane is mapped onto the point $Q(x + p, y + q)$, in the w-plane, which is merely a translation of the coordinate axes.

2. **Inversion**. $w = \frac{1}{z}$ or $Re^{i\phi} = \frac{1}{re^{i\theta}} = \frac{1}{r}e^{-i\theta}$ in polar coordinates

   $\Rightarrow \quad R = \frac{1}{r}$ and $\phi = -\theta$

   Thus a point $P(r, \theta)$ in the z-plane is mapped onto the point $Q\left(\frac{1}{r}, -\theta\right)$ in the w-plane, representing a reflection into real axis. As such the interior of

NOTES

a unit circle $|z| = 1$ in the z-plane is mapped onto the exterior of the unit circle in the w-plane.

3. **Magnification and Rotation**. $w = \beta z$, where w, z are complex numbers and β is a complex constant.

   Taking $w = Re^{i\phi}$, $z = re^{i\theta}$ and $\beta = be^{i\alpha}$, we have

   $$Re^{i\phi} = bre^{i(\theta + \alpha)} \Rightarrow R = br \text{ and } \phi = \theta + \alpha$$

Thus the modulus is magnified as $R = br$ and the angle is rotated through α.

Example 3.1. *To find all the Möbius transformations which transform the half plane $I(z) \geq 0$ into the unit circle $|w| \leq 1$ (one).*

Möbius transformation is

$$w = \frac{az+b}{cz+d} = \frac{a}{c}\frac{z+b/a}{z+d/c} \qquad ...(1)$$

which transforms $I(z) = 0$ into $|w| = 1$, i.e., real axis in z-plane transforms into the unit circle in w-plane.

It is observed that points w, $\frac{1}{\overline{w}}$ inverse w.r.t. the unit circle in w-plane transform into points z, $\overline{z}$ symmetrical (inverse) w.r.t. the real axis in z-plane. In particular $w = 0, \infty$ correspond $z = \alpha$, $\overline{\alpha}$ (say).

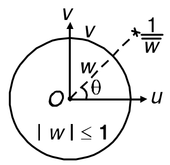

Fig. 3.5

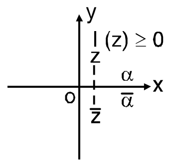

Fig. 3.6

$\therefore$ (1) $\Rightarrow$ $-\frac{b}{a} = \alpha, -\frac{d}{c} = \overline{\alpha},$, so that Equation (1) reduces to

$$w = \frac{a}{c}\frac{z-\alpha}{z-\overline{\alpha}} \quad \text{or} \quad |w| = \left|\frac{a}{c}\right|\left|\frac{z-\alpha}{z-\overline{\alpha}}\right| \qquad ...(2)$$

The point $z = 0$ must correspond to a point in the circle $|w| = 1$, so that Equation (2) gives

$$1 = \left|\frac{a}{c}\right|\left|\frac{0-\alpha}{0-\overline{\alpha}}\right| \quad \text{or} \quad 1 = \left|\frac{a}{c}\right|\left|\frac{a}{\overline{\alpha}}\right| \quad \text{or} \quad 1 = \left|\frac{a}{c}\right| \text{ as } |\alpha| = |\overline{\alpha}|$$

We may write $\frac{a}{c} = e^{i\lambda}$, λ being real, then

(2) $\Rightarrow$ $$w = \frac{z-\alpha}{z-\overline{\alpha}}e^{i\lambda} \qquad ...(3)$$

Since $z = \alpha$ gives $w = 0$, α must be a point in the *upper half plane*, i.e., $I(\alpha) > 0$. With this condition, Equation (3) gives the required transformation.

Again, we have from Equation (3)

$$w\overline{w} - 1 = \frac{z-\alpha}{z-\overline{\alpha}}e^{i\lambda} \cdot \frac{\overline{z}-\overline{\alpha}}{\overline{z}-\alpha}e^{-i\lambda} - 1 \text{ as } \overline{\overline{\alpha}} = \alpha$$

NOTES

or $$|w|^2 - 1 = \frac{(z-\alpha)(\bar{z}-\bar{\alpha})}{(z-\bar{\alpha})(\bar{z}-\alpha)} - 1 \text{ as } w\bar{w} = |w|^2$$

$$= \frac{z\bar{z} - z\bar{\alpha} - \bar{z}\alpha + \alpha\bar{\alpha} - z\bar{z} + z\alpha + \bar{z}\bar{\alpha} - \alpha\bar{\alpha}}{|z-\bar{\alpha}|^2} \text{ as } \bar{z} - \alpha = \overline{z - \bar{\alpha}}$$

$$= \frac{(z-\bar{z})(\alpha-\bar{\alpha})}{|z-\bar{\alpha}|^2} = \frac{2i\,Im(z) \times 2i\,Im(\alpha)}{|z-\bar{\alpha}|^2}$$

$$= -\frac{4Im(z)\,Im(\alpha)}{|z-\bar{\alpha}|^2} \qquad ...(4)$$

$\therefore$ (4) $\Rightarrow |w|^2 - 1 < 0$ for $I(z) > 0$

$\Rightarrow |w| < 1$ corresponds to $I(z) > 0$

$\Rightarrow$ upper half plane in the z-plane corresponds to interior of the unit circle in w-plane.

Hence Form (3) is the required transformation.

Check Your Progress

1. What are complex numbers?
2. What is an exponential function?
3. What is the polar form of exponential function?
4. What is the necessary condition for conformality?

3.4 ARGAND DIAGRAM

Consider a point P in xy-plane. Let an ordered pair of values of x and y correspond to the coordinates of the point P. Then a complex number z may be made to correspond to the point P, where

$$z = x + iy,$$

Here z is called the complex coordinate of the point P.

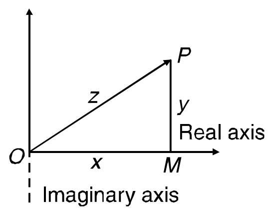

Fig. 3.7

In the adjoining figure, the x-axis is called the *real axis* or *axis of reals* and y-axis is called the *imaginary axis* or the *axis of imaginaries*.

Here $|z| = |x + iy| = \sqrt{(x^2 + y^2)}$ is the measure of length OP.

If (r, θ) be the polar coordinates of the point P, the polar form of the complex number z is

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta}.$$

NOTES

Here the number r (being taken +ve) is called the **modulus** or **absolute value** of the complex number z and θ is called the **angle** or **argument** of z and usually written as arg z, *i.e.*, $|z| = r$ and $\arg z = \theta$.

Now the coordinates of a point P' which is conjugate of z are $\bar{z} = (x, -y)$ or $(r, -\theta)$ in polars.

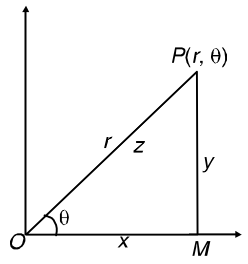

Fig. 3.8

Since $\bar{z} = r\{\cos(-\theta) + i\sin(-\theta)\}$, geometrically the points P and P' represent z and $\bar{z}$ respectively and their situations are symmetrical about the axis of reals, *i.e.*, x-axis. The conjugate of z is called the **reflection** or **image** of z in the real aixs.

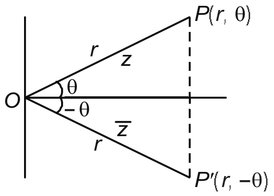

Fig. 3.9

Note 1. The plane whose points are represented by complex numbers is known as *Argand Plane* or *Argand diagram* or *Complex plane* or *Gaussian plane*.

Note 2. The complex number z representing the point (x, y) is sometimes called *Affix* of the point (x, y).

Geometrical Representation on the Argand Plane.

The *sum*, *difference*, *product* and *quotient* of complex numbers can be geometrically represented on the Argand pland as follows:

[1] Sum. Taking z_1 and z_2 two complex numbers represented by the points P and Q on Argand Plane and completing the parallelogram $OPRQ$, we observe that mid-points of its diagonals OR and PQ coincide, since they bisect each other, *i.e.*, if

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

then mid-point of PQ is $\left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right)$ which is also the mid-point of OR showing that coordinates of R are $(x_1 + x_2, y_1 + y_2)$.

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

NOTES

Therefore, the sum $z_1 + z_2$ corresponds to a vector whose components are $x_1 + x_2$ and $y_1 + y_2$. As such the sum of two complex numbers z_1 and z_2 can be represented by a vector $(z_1 + z_2)$.

i.e., if $\overrightarrow{OP} = z_1,\ \overrightarrow{OQ} = z_2$ then $\overrightarrow{OR} = \overrightarrow{OP} + \overrightarrow{PR} = \overrightarrow{OP} + \overrightarrow{OQ} = z_1 + z_2$.

Hence the point R on Argand Plane corresponds to the sum of two complex numbers z_1 and z_2 as shown in Fig. 3.10.

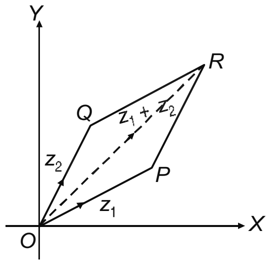

Fig. 3.10

[2] Difference. Taking $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$ two complex numbers represented by the points P and Q on Argand Plane and completing the parallelogram $OQPR$, we see that the point R represents the complex number $z_1 - z_2$, since $z_1 - z_2 = (x_1 - x_2, y_1 - y_2)$ being a complex number corresponds to a vector whose components are $x_1 - x_2$ and $y_1 - y_2$ and

$$\text{if} \quad \overrightarrow{OP} = z_1,\ \overrightarrow{OQ} = z_2$$

$$\text{then} \quad \overrightarrow{QO} = -z_2, \text{ so that}$$

$$z_1 - z_2 = \overrightarrow{OP} - \overrightarrow{OQ} = \overrightarrow{OP} + \overrightarrow{QO}$$

$$= \overrightarrow{QO} + \overrightarrow{OP} = \overrightarrow{QP} + \overrightarrow{QR}$$

i.e., the difference of two complex numbers can be represented by a vector.

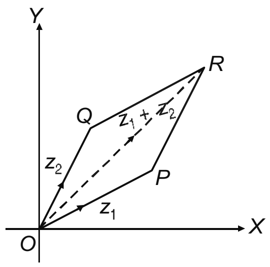

Fig. 3.11

[3] Product. If $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$ are two complex numbers, then $z_1z_2 = (x_1x_2 - y_1y_2,\ x_1y_2 + x_2y_1) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$

Changing to polars by putting

$$x_1 = r_1 \cos\theta_1,\ y_1 = r_1 \sin\theta_1$$

$$x_2 = r_2 \cos\theta_2,\ y_2 = r_2 \sin\theta_2$$

where r_1, r_2 are moduli and θ_1, θ_2 are arguments of z_1 and z_2 respectively, we have

$$z_1z_2 = r_1r_2\,[(\cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2)$$

$$+ i(\cos\theta_1 \sin\theta_2 + \cos\theta_2 \sin\theta_1)]$$

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

$$\therefore \quad |z_1z_2| = r_1r_2 = |z_1||z_2| \qquad ...(3.40)$$

and $\quad \arg(z_1z_2) = \theta_1 + \theta_2 = \arg\theta_1 + \arg\theta_2$

NOTES

i.e., the modulus of product of two complex numbers is equal to the product of their moduli and argument of the product of two complex numbers is the sum of their arguments.

In general if there are n complex numbers $z_1, z_2, ..., z_n$ with moduli $r_1, r_2, ... r_n$ and arguments $\theta_1, \theta_2, ..., \theta_n$ respectively, then repeated application of the above result yields,

$$z_1z_2...z_n = r_1r_2...r_n\,[\cos(\theta_1 + \theta_2 + ... + \theta_n) + i \sin(\theta_1 + \theta_2 + ... + \theta_n)]$$

so that $|z_1z_2...z_n| = r_1r_2...r_n = |z_1||z_2|...|z_n|$

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

i.e., the modulus of the product of any number of complex quantities is equal to the product of the moduli and the argument of the product of these complex numbers is equal to the sum of their arguments.

Geometrically represented on an Argand Plane the product of n complex quantities $z_1, z_2, ..., z_n$ as shown in Fig. 3.12 follows that the length of the vector $(z_1z_2...z_n)$ is the product of the lengths of the vectors $z_1, z_2 ....., z_n$, *i.e.*, $|z_1z_2...z_n| = |z_1||z_2|...|z_n|$ and the amplitude of $(z_1z_2...z_n)$ is equal to the sum of the amplitudes of $z_1, z_2, ..., z_n$.

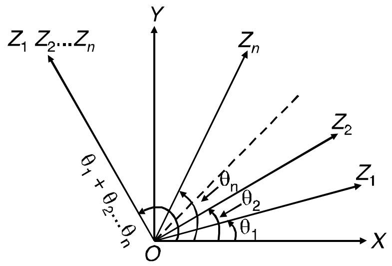

Fig. 3.12

In a particular case when $z_1 = z_2 = ... = z_n = z$ (say), the above results may be summarised as

$z^n = r^n(\cos n\theta + i \sin n\theta)$ under the assumptions

$$r_1 = r_2 = ... = r_n = r \text{ (say)}$$
$$\theta_1 = \theta_2 = ... = \theta_n = \theta \text{ (say)}$$

i.e., $\quad |z^n| = r^n = |z|^n$

and amp $z^n = n\theta = n.$ (amp z)

Also if $r = 1$, we get the *De Moivre's theorem* for positive integral exponents such as

$$z^n = (\cos\theta + i \sin\theta)^n = \cos n\theta + i \sin n\theta \qquad ...(3.41)$$

[4] Quotient. Consider two complex numbers z_1 and z_2 such that

$$z_1 = x_1 + iy_1 = r_1(\cos\theta_1 + i \sin\theta_1)$$
$$z_2 = x_2 + iy_2 = r_2(\cos\theta_2 + i \sin\theta_2)$$

NOTES

The quotient of complex numbers z_1 and z_2 is given by

$$\frac{z_1}{z_2} = \frac{r_1(\cos\theta_1 + i\sin\theta_1)}{r_2(\cos\theta_2 + i\sin\theta_2)} = \frac{r_1}{r_2}[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)]$$

$$\therefore \quad \left|\frac{z_1}{z_2}\right| = \frac{r_1}{r_2} = \frac{|z_1|}{|z_2|}$$

and $\arg\left(\frac{z_1}{z_2}\right) = \theta_1 - \theta_2 = \arg z_1 - \arg z_2$

i.e., the modulus of the quotient of two complex numbers is the quotient of their moduli and the argument of the quotient of two complex numbers is the difference of their arguments.

As a particular case defining the division as the inverse of multiplication, we have

$$\frac{1}{z} = \frac{1}{r}[\cos(-\theta) + i\sin(-\theta)] = \frac{1}{r}[\cos\theta - i\sin\theta]$$

so that $\frac{1}{z^n} = z^{-n} = \frac{1}{r^n}(\cos n\theta - i\sin n\theta) = \left(\frac{1}{z}\right)^n$ ...(3.42)

which shows that De Movire's theorem is valid when the exponent is any negative integer.

Geometrical representation of $\frac{z_1}{z_2}$ may be shown as below:

Let $\overrightarrow{OP}$ and $\overrightarrow{OQ}$ represent the vectors z_1 and z_2 in an Argand Plane such that $|z_1| = OP$, $|z_2|$ OQ and $\arg z_1 = \theta_1$, $\arg z_2 = \theta_2$.

Rotate the line OP in clockwise direction through an angle θ_2 (= arg z_2) such that its new position is OP' and $\angle POP' = \theta_2$. Take $OA = 1$ (unit length) on OX and draw a line AR to meet OP' in R such that $\angle OAR = \angle OQP$.

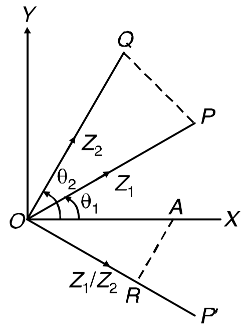

Fig. 3.13

The point R thus obtained corresponds to the quotient $\frac{z_1}{z_2}$ and it may be justified as follows:

In similar triangles OAR and OQP have

$$\frac{OR}{OA} = \frac{OP}{OQ}, \text{ i.e., } OR = \frac{OP}{OQ}, \because OA = 1$$

$$= \frac{|z_1|}{|z_2|} = \left|\frac{z_1}{z_2}\right|$$

which shows that the radius vector of the point R is $\left|\frac{z_1}{z_2}\right|$.

NOTES

Also $\angle AOR = \angle POR - \angle POX = \theta_2 - \theta_1 = -(\theta_1 - \theta_2)$

i.e., vectorial angle of R is $-(\theta_1 - \theta_2)$ which, when measured in positive sense is $\theta_1 - \theta_2$.

Hence the point R represents the quotient $\frac{z_1}{z_2}$.

Multiplication of a Complex Number by i

Let z be a complex number with its modulus r and amplitude θ

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

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

$$\therefore\ iz = \left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right) r(\cos\theta + i\sin\theta)$$

$$= r\left[\cos\left(\frac{\pi}{2} + \theta\right) + i\sin\left(\frac{\pi}{2} + \theta\right)\right] \quad ...(3.43)$$

which follows that iz represents a vector obtained by rotating the vector z through a right angle in the positive direction.

Extraction of Roots

Suppose that $z_o^n = z$, n being positive integer ...(3.44)

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

so that $z_o = r_o(\cos\theta_o + i\sin\theta_o)$ provided $z \neq 0$ and r_o, θ_o, are unknown.

$\therefore$ (3.44) gives, $r_o^n(\cos n\theta_o + i\sin n\theta_o) = r(\cos\theta + i\sin\theta)$ by (3.41)

Measuing the angles in radians, we therefore, have

$r_o^n = r$, $n\theta = \theta \pm 2m\pi$, m being zero or any positive integer which follow that r, r_o being positive, r_o is the positive nth root of r and $\theta_o = \frac{\theta}{n} \pm \frac{2m\pi}{n}$ has n distinct values for $m = 0, 1, 2...., n - 1$.

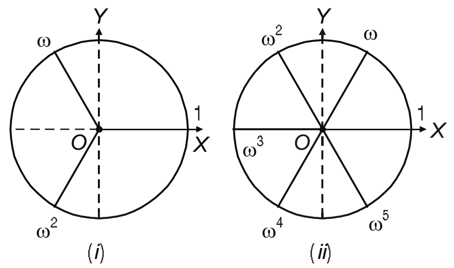

Fig. 3.14

As such there are n distinct solutions of (3.40), given by

$$z^{1/n} = z_o = r^{1/n}\left(\cos\frac{\theta+2m\pi}{n} + i\sin\frac{\theta+2m\pi}{n}\right),\ m = 0, 1, 2,....., n-1. \quad ...(3.45)$$

which are distinct values of $z^{1/n}$.

NOTES

Here the length of each of the n vectors $z^{1/n}$ is the positive number $r^{1/n}$ and argument of one of these vectors is $\frac{\theta}{n}$ while the other arguments are obtained by adding multiples of $\frac{2\pi}{n}$ to $\frac{\theta}{n}$.

In particular $z = 0$, (12) has the only solution $z_o = 0$.

But $1 = \cos 0 + i \sin 0$, then nth roots of unity are given by

$$1^{1/n} = \cos\frac{2\pi m}{n} + i\sin\frac{2\pi m}{n},\ m = 0, 1, 2,...., n-1 \quad ...(3.46)$$

Taking $m = 1$, the root of unity being a complex number and denoted by ω, is given by

$$\omega = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n} \quad ...(3.47)$$

According to De Moivre's theorem, the n, nth roots of unity are given by

$$1, \omega, \omega^2,...., \omega^{n-1} \quad ...(3.48)$$

which are the vertices of a regular polygon in complex plane of n sides inscribed in the unit circle $|z| = 1$ with one vertex at the point $z = 1$.

The case (*i*) of Fig. 3.14 shows for $n = 3$ and case (*ii*) for $n = 6$.

Now if ζ is a particular nth root of z, then we have the n roots of z as

$$\zeta, \zeta\omega, \zeta\omega^2,...., \zeta\omega^p,... \zeta\omega n^{-1} \quad ...(3.49)$$

since ζ multiplied by ω^p implies the increment of arg ζ by the angle $\frac{2p\pi}{n}$.

If m, n be two positive integers prime to each other, then (3.45) and (3.49) yield

$$(z^m)^{1/n} = (r^m)^{1/n}\left(\cos\frac{m\theta}{n} + i\sin\frac{m\theta}{n}\right)\omega^q,\ q = 1, 2,...., n-1 \quad ...(3.50)$$

$$(z^{1/n})^m = (r^{1/n})^m\left[\left(\cos\frac{\theta}{n} + i\sin\frac{\theta}{n}\right)\omega^l\right]^m$$

$$= (r^m)^{1/n}\left(\cos\frac{\theta m}{n} + i\sin\frac{\theta m}{n}\right)\omega^{lm},\ l = 0, 1, 2, ..., n-1 \quad ...(3.51)$$

The two sets of n numbers will be identical if the set ω^q and ω^{lm} coincide and then n numbers in either set can be written as $z^{m/n}$, *i.e.,*

$$\mathbf{z}^{m/n} = (r^m)^{1/n}\left[\cos\left\{\frac{m}{n}(\theta + 2\pi m)\right\} + i\sin\left\{\frac{m}{n}(\theta+2\pi m)\right\}\right]$$

$$m = 0, 1, 2,..., n-1 \quad ...(3.52)$$

We may similarly define,

$z^{-m/n} = (z^{1/n})^{-m} = (z^{-m})^{1/n}$

3.5 ANALYTICITY OF COMPLEX FUNCTION

NOTES

If there exist two functions $f_1(z)$ and $f_2(z)$, such that they are analytic (regular) in domains D_1 and D_2 respectively and that D_1 and D_2 have a common part, throughout which $f_1(z) = f_2(z)$, then the aggregate of values of $f_1(z)$ and $f_2(z)$ at the interior points of D_1 or D_2, can be regarded as a single regular function (say) $F(z)$. It is obvious that $F(z)$ is regular in the common part say Δ of the two domains and $F(z) = f_1(z)$ in domain D_1 and $F(z) = f_2(z)$ in domain D_2. We thus regard the function $f_2(z)$ as one, extending the domain in which $f_1(z)$ is defined and so it is called an **Analytic Continuation** of $f_1(z)$.

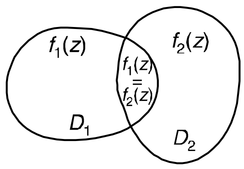

Fig. 3.15

The method of analytic continuation. Its standard method is the method of power series, which can be summarised as below.

Suppose that there is a point $P(z_0)$ in the neighbourhood of which $f(z)$ is analytic; then the function $f(z)$ can be expanded by Taylor's theorem, in a series of ascending powers of $(z - z_0)$, the coefficients of which involve the successive derivatives of $f(z)$ at z_0.

Let there be a singularity S of $f(z)$ which is nearest to P. Then a circle of centre P and radius PS is the circle of convergence within which the Taylor's expansion is valid. If we now take any point P' (not on PS) within this circle then we can find the values of $f(z)$ and all its derivatives at P', from the series by applying the method of term by term differentiation. We thus find the Taylor's series for $f(z)$ with P' as origin and this series will define a function which is regular in the circle whose centre is P. Such a circle will extend as far as the singularity of the function defined by the new series, which is nearest to P', and this may or may not be S. In either case the new circle of convergence may lie partly outside the old circle and for points in the region which is included in the new circle but not in the old one, so that the new series may be utilized in defining the values of $f(z)$ while the old series failed to do so.

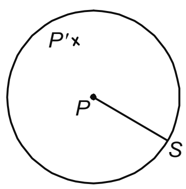

Fig. 3.16

In a similar manner, we can take another point P'' in the region for which the values of the function are known and form the Taylor's series with P'' as origin which will, in general, still further extend the domain of definition of the function and so on and so forth.

NOTES

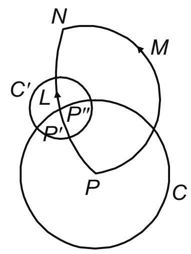

Fig. 3.17

By this method of continuation, starting from a representation of a function by any one power series, we can find any number of other power series, which between them define the value of the function at all points of a domain, any point of which can be reached from P without passing through a singularity of the function $f(z)$.

It is easy to show that continuation by two different paths PLN and PMN gives the same power series provided the function is analytic and has no singularity inside the closed curve $PLNMP$. To show it, let us suppose that S, S', S'',... be the power series with P, P', P'',... as origin. Then $S' = S''$ (for, each is equal to S) over a certain domain which contains P', when P'' is taken sufficiently near to P', and therefore S' will be the continuation of S''. Continuing this process we can deform the path PLN into PMN provided no singular point lies inside the path $PLNMP$.

Note. Weierstrass defined an **Analytic function** *of z as one power series together with all the other power series derivable from it by anlaytic continuation.*

An important remark. There must be at least one singularity of the analytic function on the circle of convergence C_0 of the power series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$.

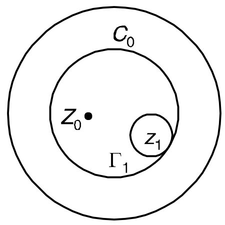

Fig. 3.18

Assuming that there is no such singularity, we can construct by the method of continuation, a function equal of $f(z)$ within C_0 and regular in a larger concentric circle Γ_0. The expansion of this function in Taylor series in powers of $(z-z_0)$ would then converge everywhere within the larger circle Γ_0. But this is not possible, since the series would be the original series which has got C_0 as its circle of convergence. Let there be a point z_1 within C_0 and let C_1 be the circle of convergence of the power series,

$$\sum_{0}^{\infty} A_n(z-z_1)^n,$$

where $A_n = \dfrac{f^n(z_1)}{n!}$ (by Taylor's expansion).

Let Γ_1 be the circle of centre z_1, touching the circle C_0 internally. Then the new power series defined by (1) is certainly convergent within Γ_1 and has the sum $f(z)$ there.

Since the radius of C_1 cannot be less than that of Γ_1, there are these possibilities:

(*i*) C_1 has a larger radius, then Γ_1, in the which case C_1 lies partly outside C_0 and the new power series provides an analytic continuation of $f(z)$. Then taking a point z_2 within C_1 and outside C_0, the process can be repeated.

(*ii*) C_0 is a natural boundary of $f(z)$. In this case we cannot continue $f(z)$ outside C_0 and the circle C_1 touches C_0 internally, wherever the point z lies within C_0.

Note. A closed curve is called a **natural boundary** *of a function if the function is such that it cannot be analytically continued to any point of the same.*

(*iii*) C_1 may touch C_0 internally, though C_0 is not a natural boundary of $f(z)$. In this case the point of contact of C_0 and C_1 is a singularity of the analytic function which has been found by continuation of the original power series: for, there is necessarily one singularity on C_1 and this cannot lie within C_0.

Natural Boundary Theorem [Lambert's Series]

If $f(z) = \sum_{n=0}^{\infty} d(n)z^n$, $(|z|<1)$, $d(n)$ *being the number of divisors of n, then the unit circle is a natural boundary of this function.* (*R.U., 1990*)

Taking the double series

$$\sum_{m=1}^{\infty}\sum_{l=1}^{\infty} z^{ml} \qquad ...(3.53)$$

We can find $f(z)$ by considering this series as a single power series and summing up by rows, *i.e.*,

$$f(z) = \sum_{m=1}^{\infty} (z^m + z^{2m} + z^{3m} + \cdots),$$

on putting $l = 1, 2, 3,...$

in (3.53) and taking $|z| < 1$

$$= \sum_{m=1}^{\infty} \frac{z^m}{1-z^m};\ |z|<1 \qquad ...(3.54)$$

The series (3.53) being absolutely convergent for $|z| < 1$, we can take the transformation,

$$z = re^{\frac{2\pi pi}{q}} \qquad ...(3.55)$$

where p, q are positive integers s.t. $p > 0$, $q > 1$ and p, q are prime to each other.

We, now claim

$$(1-r)\, f(z) \to \infty \text{ as } r \to 1$$

NOTES

Assuming

$$f(z) = \sum_1 \frac{z^m}{1-z^m} + \sum_2 \frac{z^m}{1-z^m} \qquad \text{...(3.56)}$$

where in $\sum_1$, m takes all values $\equiv 0 \pmod q$ and in $\sum_2$ all other values, and setting $m = tq$ in $\sum_1$, we obtain

$$z^m = \left[r\, e^{2\pi ip/q}\right]^{tq} = r^{tq} \quad \text{or} \quad e^{2\pi ip}$$

so that
$$(1-r)\sum_1 = (1-r)\sum_{t=1}^{\infty} \frac{r^{tq}}{1-r^{tq}} = \frac{1-r}{1-r^q}\sum_{t=1}^{\infty}\frac{1-r^q}{1-r^{tq}}\cdot r^{tq}$$

$$= \frac{1}{1+r+r^2+\cdots+r^{q-1}}\sum_{t=1}^{\infty}\frac{r^{tq}}{1+r^q+\cdots+r^{(t-1)q}}$$

$$\geq \frac{1}{q}\sum_{t=1}^{\infty}\frac{r^{tq}}{t} = \frac{1}{q}\log\frac{1}{1-r^q} \to \infty \text{ as } r \to 1$$

$$\left[\text{Here for } r \to 1, -\log(1-x^k) = \sum_t \frac{(x^k)^t}{t} = \log\frac{1}{1-x^k}\right]$$

Now, if $m \not\equiv 0 \pmod q$, then

$$|1-z^m|^2 = \left|-r^m e^{p\pi im/q}\right|^2 = \left|-r^m\left(\cos\frac{2\pi pm}{q} + i\sin\frac{2\pi pm}{q}\right)\right|^2$$

$$= \left(1 - r^m\cos\frac{2\pi pm}{q}\right)^2 + r^{2m}\sin^2\frac{2\pi pm}{q}$$

$$= 1 - 2r^m\cos\frac{\pi pm}{q} + r^{2m}$$

$$= (1-r^m)^2 + 2r^m\left(1-\cos\frac{2\pi pm}{q}\right)$$

$$= (1-r^m)^2 + 4r^m\sin^2\frac{\pi pm}{q}$$

$$\geq 4r^m\sin^2\frac{\pi pm}{q} \quad \text{as } (1-r^m)^2 \geq 0$$

$$\geq 4r^m\sin^2\frac{\pi}{q}.$$

Since the inequality does still hold as $\sin^2\frac{\pi pm}{q} \leq 1$ as well as $\sin^2\frac{\pi}{q} \leq 1$

$$\therefore \quad |1-z^m| \geq 2r^{m/2}\sin\frac{\pi}{q} \qquad \text{...(3.57)}$$

so that
$$\left|(1-r\sum_2)\right| = |(1-r)|\left|\sum_2\frac{z^m}{1-z^m}\right| \leq |(1-r)|\sum_2\frac{|z^m|}{|1-z^m|}$$

$$\leq (1-r)\sum_{m=0}^{\infty}\frac{r^m}{2r^{m/2}\sin\frac{\pi}{q}} \quad \text{by (3.57)}$$

$$\leq \frac{1-r}{2\sin\frac{\pi}{q}}\sum_{m=0}^{\infty} r^{m/2}$$

$$\le \frac{1-r}{2\sin\frac{\pi}{q}}\cdot\frac{1}{1-r^{1/2}} \text{ as } \sum_{m=0}^{\infty} r^{m/2} = 1 + r^{1/2} + r^{2/2} + \cdots \infty$$

$$= \frac{1}{1-r^{1/2}}$$

NOTES

$$\le \frac{1+r^{1/2}}{2\sin\frac{\pi}{q}} \text{ as } 1-r = 1^2 - \left(r^{1/2}\right)^2 = \left(1-r^{1/2}\right)\left(1+r^{1/2}\right)$$

$$\le \frac{1}{\sin\frac{\pi}{q}} \text{ as } r \to 1$$

Thus

$$(1-r)\,f(z) = (1-r)\{\Sigma_1 + \Sigma_2\}$$

$$= (1-r)\sum\nolimits_1 + (1-r)\sum\nolimits_2$$

$$= \frac{1}{q}\log\frac{1}{1-r^q} + \frac{1}{\sin\frac{\pi}{q}}$$

$$\to \infty \text{ as } r \to 1$$

showing that $z = e^{2\pi ip/q}$ is a singularity of $f(z)$. But the points of such types are dense everywhere round the unit circle $|z| = 1$ and hence an arc, however small, on $f(z)$ cannot be found to be regular. As such $f(z)$ cannot be continued across the unit circle so that the unit circle is a natural boundary of the function.

As an example, we claim that the power series $1+\sum_{n=0}^{\infty} z^{2^n}$ *cannot be continued analytically beyond the circle* $|z| = 1$ *which forms a natural boundary.*(*R. U., 2006*)

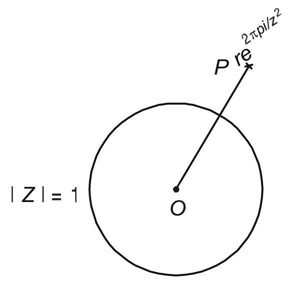

Fig. 3.19

We have

$$f(z) = 1+\sum_{n=0}^{\infty} z^{2^n} = 1+z+z^2+z^4+z^8+\cdots$$

$$= 1+z^2+z^4+\cdots+z^{2^{q-1}}+\sum_{n=q}^{\infty} z^{2^n}$$

$$= \sum_{n=0}^{q-1} z^{2^n} + \sum_{n=q}^{\infty} z^{2^n}$$

$$= f_1(z) + f_2(z) \text{ (say)}$$

But

$$\text{Lim}\,|u_n(z)|^{1/n} = \text{Lim}\left|z^{2^n}\right|^{1/n} = |z^2| = |z|^2$$

$\Rightarrow$ the series converges for $|z| < 1$

$\Rightarrow |z| = 1$ is the circle of convergence.

Now, take a point P outside the circle $|z = 1$ with complex coordinate (affix) $r e^{2\pi p i/2^q}$, with $r > 1$ and p, q are integers, also assume that P tends to move to the circle of convergence along the radius vector through P.

At P,

$$f_1(z) = \sum_{n=0}^{q-1} z^{2^n} = \sum_{n=0}^{q-1} \left(re^{2\pi pi/2^q}\right)^{2^n} = \sum_{n=0}^{q-1} r^{2^n} \left(e^{2\pi pi/2^q}\right)^{2^n}$$

which is a polynomial in r of degree 2^{q-1} and so will tend to a unique limit when $r \to 1$.

and $$f_2(z) = \sum_{n=q}^{\infty} z^{2^n} = \sum_{n=q}^{\infty} r^{2^n} \left(e^{2\pi pi/2^q}\right)^{2^n}$$

$$= \sum_{n=q}^{\infty} r^{2^n}, \text{ since } (e^{2\pi pi/2^q})^{2^n} = 1$$

for $n = q, q + 1, ...$

Hence $$f(z) = f_1(z) + f_2(z)$$

$$= \sum_{n=0}^{q-1} r^{2^n} \left(e^{2\pi pi/2^q}\right)^{2^n} + \sum_{n=q}^{\infty} r^{2^n}$$

$$\to \infty \text{ as } r \to 1.$$

Conclusively $|z| = 1$ is a singularity of $f(z)$ and it forms a natural boundary so that $f(z)$ cannot be continued analytically beyond the circle of convergence, *i.e.*, $|z| = 1$.

Problem 3.3. *Show that the two power series*

$$z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \cdots \qquad ...(1)$$

and $$i\pi - (z-2) + \frac{1}{2}(z-2)^2 - \cdots \qquad ...(2)$$

have no common region of convergence, but they are analytic continuations of the same function.

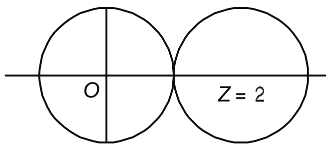

Fig. 3.20

Here the series (1) defines the function $-\log(1 - z)$, whose circle of convergence is $|z| = 1$, *i.e.*, a circle with centre $z = 0$ and radius 1. The only singularity is $z = + 1$.

The series (2) defines the function $i\pi - \log[1 + (z - 2)]$; where $|z - 2| \le 1$ is a circle with centre $z = 2$ and radius 1. It touches the first circle externally as shown in the adjoining figure.

The function $i\pi - \log[1 + (z - 2) = i\pi - \log(z - 1)$

$= i\pi - \log[-(1 - z)]$

$= i\pi - \log(-1) - \log(1 - z)$

$$= i\pi - \log e^{\pi i} - \log (1 - z)$$
$$= i\pi - \pi i - \log (1 - z)$$
$$= - \log (1 - z).$$

It is clear that the two functions defined by the given power series have no common region of convergence, but they are analytic continuations of the same function – log (1 – z).

Problem 3.4. *Show that the series*

$$\frac{1}{a} + \frac{z}{a^2} + \frac{z^2}{a^3} + \frac{z^3}{a^4} + \cdots$$

represents the function which can be continued analytically outside the circle of convergence.

Let the given series define a function $f(z)$, *i.e.*,

$$f(z) = \frac{1}{a} + \frac{z}{a^2} + \frac{z^2}{a^3} + \frac{z^3}{a^4} + \cdots$$

$$= \frac{\frac{1}{a}}{1 - \frac{z}{a}} \left[\text{being an infinite geometric series if } \left|\frac{z}{a}\right| > 1\right]$$

$$= \frac{1}{a - z},$$

only for points within the circle of convergence $|z| = |a|$.

If we consider a series

$$\frac{1}{a-b} + \frac{z-a}{(z-b)^2} + \frac{(z-a)^2}{(z-b)^3} + \cdots \text{where } \frac{b}{a} \text{ is real,}$$

then $$f(z) = \frac{1}{a-b}\left[1 + \frac{z-b}{a-b} + \left(\frac{z-b}{a-b}\right)^2 + \cdots\right]$$

$$= \frac{1}{a-b}\left[1 - \frac{z-b}{a-b}\right]^{-1}$$

provided the circle of convergence say C_1 is $|z - b| = |a - b|$, when a and b both are real and positive and also $0 < b < a$, the circle of convergence C_1 of the second series touches the first circle (say) C internally at its only singularity $z = a$. If follows that there is no analytic continuation.

In case $\frac{b}{a}$ is not real and $|b| < |a|$ and positive, the second series converges at points inside a circle which is partly inside and partly outside the circle $|z| = a$. In fact the two series represent the same function $\frac{1}{a-z}$ at points outside the circle $|z| = |a|$ and hence can be continued analytically.

Check Your Progress

5. What is Argand plane?
6. State Natural Boundary theorem.

3.6 CAUCHY–RIEMANN CONDITIONS

NOTES

Before going into the details of Regular or Analytic functions, we first define some terms which are used frequently.

Neighbourhood of a Point: Neighbourhood of a pint z_0 in the Argand diagram means the set of all points z such that $|z - z_0| < \varepsilon$, where ε is an arbitrarily chosen small positive number.

Limit Point: A point z_0 is said to be a *limit point* of a set of points S in the Argand plane, if every neighbourhood of z_0 contains a point of S other than zero.

The limit points of a set may not necessary be the points of the set. There are two types of limit points:

(*i*) **Interior Points:** A limit point z_0 of the set S is said to be the *interior or inner point* if in the neighbourhood of z_0 there exists entirely the points of the set S.

(*ii*) **Boundary Points:** A limit points z_0 is said to be the *boundary point* if all the points in the neighbourhood of z_0 do not belong to the set S.

Closed Set: If all the limit points of the set belong to the set, then the set is said to be a *closed set*.

Open Set: A set which consists entirely of interior points is known to be an *open set*.

Bounded and Unbounded Sets: A set of points is said to be *bounded* if there exists a constant number k, such that $|z| \leq k$ for all points z of the set. If there does not exist such number k the set is said to be *unbounded*.

Domain: If every pair of the points of a set of points in the Argand diagram can be joined by a polygonal arc which consists only of the points of the set, then the set of points in the Argand diagram is said to be *connex* (means connected) or *domain* or *region*.

Open Domain is an open connex set of points.

Closed Domain: When the boundary points of the set are also added to an open domain, it is then called a *closed domain*.

Functions of a Complex Variable: If $w = u + iv$ and $z = x + iy$ are two complex numbers, then w is said to be the *function* of z and written as $w = f(z)$, if to every value of z in a certain domain D, there correspond one or more values of w. If w takes only one value for each value of z in the domain D, then w is said to be *uniform* or *single-valued* function of z and if it takes more than one values for some or all values of z in the domain D, then w is known as a *many-valued* or a *multiple-valued* function of z.

Since u and v both are functions of x, y

$$\therefore\ w = f(z) = u(x, y) + iv(x, y).$$

It is however notable that the path of a complex variable z is either a straight line or a curve.

Continuity: The function $f(z)$ of a complex variable z is continuous at the point z_0 if, given a positive number $\varepsilon > 0$, a number δ can be so found that

$$|f(z) - f(z_0)| < \varepsilon$$

for all points z of the domain D satisfying $|z - z_0| < \delta$, where δ depends upon ε and also, in general, upon z_0, i.e.,

$$\delta = \phi(\varepsilon, z_0).$$

If δ is independent of z_0 or rather say that if a number $h(\varepsilon)$ can be found independent of z_0 such that $|f(z) - f(z_0)| < \varepsilon$ holds for every pair of points z, z_0 of the domain D for which $|z - z_0| < h$, then $f(z)$ is called uniformly continuous in D.

It should be noted that if a function f is continuous at $z = z_0$, i.e., *if* $f = u + iv$ *is continuous at* $z = z_0$ *then it will be so iff its real and imaginary parts are separately continuous functions of x and y at the point* $(x, y) = (x_0, y_0)$

Since if f is continuous at $z = z_0$ then $u(x_0, y_0)$ and $v(x_0, y_0)$ both are uniquely defined such that

$$0 \le |u(x, y) - u(x_0, y_0)| \le |f(z) - f(z_0)| \quad ...(3.58)$$

for, $\quad |f(z) - f(z_0)| = +\left[\{u(x, y) - u(x_0, y_0)\}^2 + \{v(x, y) - v(x_0, y_0)\}^2\right]^{1/2}$

as $\quad z \to z_0$, $u(x, y_0) \to u(x_0, y_0)$, *i.e.*, $u(x_0, y_0) = \lim\limits_{(x, y)\to(x_0, y_0)} u(x, y) \quad ...(3.59)$

This limit exists independent of the manner in which $x \to x_0$, $y \to y_0$. Equation (3.59) shows that $u(x, y)$ is continuous at (x_0, y_0).

Similarly $v(x, y)$ is continuous at (x_0, y_0).

Thus continuity of u and v for f to be continuous at $z = z_0$, is a necessary condition.

Conversely if $u(x, y)$ and $v(x, y)$ are continuous, then

$$u(x, y) \to u(x_0, y_0) \text{ and } v(x, y) \to v(x_0, y_0) \text{ as } z \to z_0$$

so that $f(z) = u(x, y) + iv(x, y) \to u(x_0, y_0) + iv(x_0, y_0) = f(z_0)$

So the condition is also sufficient.

Differentiability: If $f(z)$ be a single-valued function defined in a domain D of the Argand diagram, then $f(z)$ is said to be *differentiable* at $z = z_0$ at a point of D if $\dfrac{f(z) - f(z_0)}{z - z_0}$ tends to a unique limit when $z \to z_0$, provided that z is also a point of D.

A function $f(z)$ is said to be differentiable at a point z_0, if $\lim\limits_{z\to z_0} \dfrac{f(z) - f(z_0)}{z - z_0}$ exists and is a finite quantity provided by whatever path $z \to z_0$, then $f(z)$ is differentiable at $z = z_0$. The finite limit when exists is denoted by $f'(z_0)$ and termed as the differential coeffficient or derivative of $f(z)$ at $z = z_0$, i.e.,

$$f'(z_0) = \lim_{z\to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

Precisely if for a given $\varepsilon > 0$, there exists a number δ such that

$$\left|\frac{f(z)-f(z_0)}{z-z_0}-f'(z_0)\right| < \varepsilon \text{ whenever } 0 < |z-z_0| < \delta$$

NOTES

i.e., writing $z - z_0 = \Delta z$ if for an $\varepsilon > 0$, there exists a number δ such that $\left|\frac{f(z_0+\Delta z)-f'(z_0)}{\Delta z}-f'(z_0)\right| < \varepsilon$ whenever $0 < |\Delta z| < \delta$ then $f'(z_0)$ is known as the derivative of $f(z)$ at z_0.

Clearly the limiting value of $\frac{f(z)-f(z_0)}{z-z_0}$ is independent of the path in D along which $z \to z_0$.

Consider $f(z) = z^2$, for example, then $f'(z_0) = 2z_0$ at any point z_0

since $f'(z_0) = \lim\limits_{\Delta z\to 0}\frac{(z_0+\Delta z)^2 - z_0^2}{\Delta z} = \lim\limits_{\Delta z\to 0}(2z_0+\Delta z) = 2z_0$

In view of the relation $f'(z_0) = \lim\limits_{\Delta z\to 0}\frac{f(z_0+\Delta z)-f(z_0)}{\Delta z}$ whenever $f'(z_0)$ exists at any point z_0, we have

$$\lim_{\Delta z\to 0}[f(z_0+\Delta z)-f(z_0)] = \lim_{\Delta z\to 0}\frac{f(z_0+\Delta z)-f(z_0)}{\Delta z}\lim_{\Delta z\to 0}\Delta z = 0$$

which follows $\lim\limits_{z\to z_0} f(z) = f(z_0)$

i.e., *f is necessarily continuous at any point z_0 where its derivative exists. But the converse is not true,* i.e., *if a function is continuous, it is not necessarily differentiable as is evident from the following example:*

Consider the function $w = |z|^2$ which is differentiable at every point. It will be shown that its derivative exists only at the point $z = 0$ and nowhere else, since

$$\frac{\Delta w}{\Delta z} = \frac{|z_0+\Delta z|^2 - |z_0|^2}{\Delta z} = \frac{(z_0+\Delta z)(\bar{z}_0+\overline{\Delta z}) - z_0\bar{z}_0}{\Delta z}$$

$$\because |z|^2 = z\bar{z}$$

$$= \bar{z}_0 + \overline{\Delta z} + z_0\frac{\overline{\Delta z}}{\Delta z} \quad ...(3.60)$$

$$\lim_{\Delta z\to 0}\frac{\Delta w}{\Delta z} = \frac{dw}{dz} = \lim_{\Delta z\to 0}\left[\bar{z}_0 + \overline{\Delta z} + z_0\frac{\overline{\Delta z}}{\Delta z}\right]$$

$$= 0 \text{ when } z_0 = 0$$

But if $z_0 \neq 0$, then taking $\theta = \arg \Delta z = \arg(z - z_0)$ we have

$$\frac{\overline{\Delta z}}{\Delta z} = \frac{e^{-i\theta}}{e^{i\theta}} = e^{-2i\theta}$$

$$= \cos 2\theta - i \sin 2\theta.$$

So that $\lim\limits_{\Delta z\to 0}\frac{\Delta w}{\Delta z} = \lim\limits_{\Delta z\to 0}[\bar{z}_0 + \overline{\Delta z} + z_0(\cos 2\theta - i\sin 2\theta)]$

Here the $\lim\limits_{\Delta z\to 0}\frac{\overline{\Delta z}}{\Delta z}$ does not exist as $\Delta z \to 0$ in any manner, since if Δz is real,

$\Delta z = \Delta x$, i.e., $\overline{\Delta z} = \Delta x = \Delta z$ then limit of Equation (3.60) is $\bar{z}_0 + z_0$.

Also if Δz is imaginary, i.e., $\Delta z = i\,\Delta y$ so that $\overline{\Delta z} = -\Delta z$, then limit of Equation (3.60) is $\bar{z}_0 - z_0$. As such the limit does not exist when $z_0 \neq 0$ and hence $|z|^2$ has no derivative at z_0.

Analytic or Regular or Holomorphic or Monogesic Functions

A function $f(z)$ which is single-valued and differentiable at every point of a domain D, is said to be *regular* in the domain D.

A function may be differentiable in a domain D save possibly for a finite number of points. Such points are called *singularities* or *singular points of* $f(z)$

The Necessary and Sufficient Conditions for $f(z)$ to be Regular

Necessary Conditions: If $w = f(z)$, where $w = u + iv$ and $z = x + iy$.

As such u and v both are the functions of x and y and therefore we can write $w = f(z) = u(x, y) + iv(x, y)$.

Now if $f(z) = u(x, y) + iv(x, y)$ is differentiable at a given point z, the ratio $\dfrac{f(z+\Delta z) - f(z)}{\Delta z}$ must tend to a certain finite limit as $\Delta z \to 0$ in any manner.

From the relation $z = x + iy$, we get $\Delta z = \Delta x + i\Delta y$.

If we take Δz to be wholly real, so that $\Delta y = 0$, then

$$\lim_{\Delta x \to 0} \frac{\Delta w}{\Delta x} = \lim_{\Delta x \to 0} \left[\frac{u(x+\Delta x, y) - u(x, y)}{\Delta x} + i \frac{v(x+\Delta x, y) - v(x, y)}{\Delta x} \right]$$

must exist and tend to a definite limit.

$$\therefore \quad \frac{dw}{dx} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}$$

$$= u_x + iv_x \text{ (say)},$$

i.e., the partial derivatives u_x, v_x must exist at the point (x, y) and the limiting value is $u_x + iv_x$.

Similarly again if Δz be taken wholly imaginary, so that $\Delta x = 0$, we find that the partial derivatives u_y, v_y must exist at the point (x, y) and the limiting value is $v_y - iu_y$.

Since the function is differentiable, the two limits so obtained must be identical, i.e., $u_x + iv_x \equiv v_y - iu_y$.

Equating real and imaginary parts, we get

$$u_x = v_y \text{ and } u_y = -v_x$$

or

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

These two relations, which are necessary conditions for a function to be analytic, are called the *Cauchy Rieman Differential Equations*.

Sufficient Conditions: The continuous single-valued function $f(z)$ is regular in a domain C if the four partial derivatives u_x, u_y, v_x, v_y exist, are continuous and satisfy the Cauchy-Riemann equations at all points of the region D.

Assuming $u_x = v_y$ and $u_y = -v_x$ and these partial derivatives are continuous, we have to show that they exist and are finite.

NOTES

By the mean value theorem, we have

$$f(x + \Delta x, y) - f(x, y) = \Delta x \frac{\Delta f(x + \Delta x\theta, y)}{\Delta x},$$

where $0 < \theta < 1$

and $$f(x, y + \Delta y) - f(x, y) = \Delta y \frac{\Delta f(x, y + \theta' \Delta y)}{\Delta y}, \text{ where } 0 < \theta' < 1.$$

Now if $w = f(z)$ and $f(z) = u(x, y) + iv(x, y)$,

$$w + \Delta w = f(z + \Delta z).$$

and $$f(z + \Delta z) = u(x + \Delta x, y + \Delta y) + iv(x + \Delta x, y + \Delta y),$$

$$\therefore \quad \Delta w = f(z + \Delta z) - f(z);$$

Also if $z = x + iy$, then $\Delta z = \Delta x + i\Delta y$.

Thus

$$\frac{\Delta w}{\Delta z} = \frac{f(z + \Delta z) - f(z)}{\Delta x + i\Delta y}$$

$$= \frac{u(x + \Delta x, y + \Delta y) + iv(x + \Delta x, y + \Delta y) - u(x, y) - iv(x, y)}{\Delta x + i\Delta y}.$$

Now $u(x + \Delta x, y + \Delta y) - u(x, y)$

$$= \{u(x + \Delta x, y + \Delta y) - u(x + \Delta x, y)\} + \{u(x + \Delta x, y) - u(x, y)\}$$

$$= \Delta y \frac{\Delta u(x + \Delta x, y + \theta \Delta y)}{\Delta y} + \Delta x \frac{\Delta u(x + \theta' \Delta x, y)}{\Delta x}$$

[by mean value theorem]

$$= \Delta x \left[\frac{\Delta u(x, y)}{\Delta x} + \varepsilon_1\right] + \Delta y \left[\frac{\Delta u(x, y)}{\Delta y} + \varepsilon_2\right]$$

[$\because$ if the function is continuous
$|f(z) - f(z_0)| < \varepsilon$; $\therefore f(z) = f(z_0) + \varepsilon_1$
when $|z - z_0| < \delta$ where $|\varepsilon_1| < \varepsilon$]

$$= \Delta x \left[\frac{\partial u}{\partial x} + \varepsilon_1\right] + \Delta y \left[\frac{\partial u}{\partial y} + \varepsilon_2\right]$$

$$\left[\because \lim_{\Delta x \to 0} \frac{\Delta u(x, y)}{\Delta x} = \frac{\partial u}{\partial x} \text{ etc.}\right]$$

Hence

$$\lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} = \frac{dw}{dz} = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta x + i\Delta y}$$

$$= \frac{\Delta x\left(\frac{\partial u}{\partial x} + \varepsilon_1\right) + \Delta y\left(\frac{\partial u}{\partial y} + \varepsilon_2\right) + i\Delta x\left(\frac{\partial v}{\partial x} + \varepsilon_1'\right) + i\Delta y\left(\frac{\partial v}{\partial y} + \varepsilon_2'\right)}{\Delta x + i\Delta y}$$

$$= \frac{\Delta x\left(\frac{\partial u}{\partial x} + i\frac{\partial u}{\partial x}\right) + \Delta y\left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right) + \Delta x(\varepsilon_1 + i\varepsilon_1') + \Delta y(\varepsilon_2 + i\varepsilon_2')}{\Delta x + i\Delta y}$$

NOTES

$$= \frac{(\Delta x + i\Delta y)\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}(\Delta x + i\Delta y) + \eta\Delta x + \eta'\Delta y}{\Delta x + i\Delta y}$$

[By applying Cauchy-Riemann's equations and putting
$\eta = \varepsilon_1 + i\varepsilon_1', \ \eta' = \varepsilon_2 + i\varepsilon_2'$; also $i^2 = -1$]

$$\therefore \quad \frac{dw}{dz} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} + \eta\frac{\Delta x}{\Delta x + i\Delta y} + \eta'\frac{\Delta y}{\Delta x + i\Delta y}.$$

Now $\left|\eta\frac{\Delta x}{|\Delta x + i\Delta y|}\right| = \frac{|\eta|\,|\Delta x|}{|\Delta x + i\Delta y|}$

$$= \frac{|\eta|\,\Delta x}{\sqrt{\{(\Delta x)^2 + (\Delta y)^2\}}}$$

$$\leq |\eta| \qquad \because \Delta x < \sqrt{\{\{\Delta x\}^2 + \{\Delta y\}^2\}}$$

$$\leq \sqrt{(\varepsilon_1^2 + \varepsilon_1'^2)} \qquad \because \eta = \varepsilon_1 + i\varepsilon_1'$$

$\leq$ 0 when $\Delta x \to 0$; ε_1 and $\varepsilon_1' \to 0$.

But by definition the modulus of any quantity is always +ve or zero and it is never negative.

$$\therefore \quad \frac{\eta\Delta x}{\Delta x + i\Delta y} \to 0, \text{ when } \Delta x \to 0$$

Similarly $\frac{\eta'\Delta y}{\Delta x + i\Delta y} \to 0$, when $\Delta y \to 0$.

Hence $\frac{dw}{dz} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}$,

i.e., the limit exists and is finite and unique.

Therefore the sufficient conditions for the function $f(z)$ to be regular require the continuity of the four first partial derivatives of u and v.

Condition for a Function when it Ceases to be Analytic

If $w = F(\zeta)$ and $z = f(z)$, then w is said to be *function of a function* of z and we have $\frac{dw}{dz} = \frac{dw}{d\zeta}\frac{d\zeta}{dz}$, $F(\zeta)$ and $f(z)$ both being analytic.

Also if $w = f(z)$ be an analytic function of z such that corresponding to each point w_0 there exists a point w_0 and $z = F(w)$ is such that to each value w_0 of w there corresponds a value z_0 of z, then the function $z = F(w)$ is said to be the *Inverse function of* $w = f(z)$. Clearly, if $f'(z_0) \neq 0$, then w_0 is a regular point of $z = F(w)$ i.e. z is analytic in the neighbourhood of w_0.

On account of functions being inverse we have $F'(w_0) = \frac{1}{f'(z_0)}$ *and hence the function* $z = F(w)$ *ceases to be analytic where* $f'(z) = 0$, i.e., $\frac{dw}{dz} = 0$ *also that* $w = f(z)$ *where* $z = F(w)$ *ceases to be regular when*

$$\frac{dw}{dz} = 0.$$

Evidently when $z = x + iy$, $w = f(z)$, we have $\frac{dw}{dz} = \frac{\partial w}{\partial x}$

NOTES

So that if $w = u + iv$, $z = f(w)$, then $\frac{dz}{dw} = \frac{\partial z}{\partial u}$.

As an illustration if $w = e^{-v}(\cos u + i \sin u)$, then $\frac{dz}{dw} = \frac{\partial z}{\partial u}$ gives

$\frac{dz}{dw} = e^{-v}(-\sin u + i \cos u) = -ie^{-v}(\cos u + i \sin u) = -iz.$

$\therefore$ w ceases to be analytic when $\frac{dz}{dw} = 0$ *i.e.* $z = 0$.

To prove that if a function is regular, it is independent of $\bar{z}$ and is function of z.

$$z = x + iy,$$

$$\bar{z} = x - iy.$$

Adding and subtracting, we get

$$x = \frac{1}{2}(z+\bar{z}) \quad \text{and} \quad y = \frac{1}{2i}(z-\bar{z})..$$

$$\therefore \quad \frac{\partial x}{\partial \bar{z}} = \frac{1}{2} \quad \text{and} \quad \frac{\partial y}{\partial \bar{z}} = -\frac{1}{2i} \text{ etc.}$$

Now $\quad w = \quad f(z) = u + iv.$

$$\therefore \quad \frac{\partial x}{\partial \bar{z}} = \frac{du}{d\bar{z}} + i\frac{dv}{d\bar{z}}$$

$$= \left[\frac{\partial u}{\partial x}\cdot\frac{\partial x}{\partial \bar{z}} + \frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial \bar{z}}\right] + i\left[\frac{\partial v}{\partial x}\cdot\frac{\partial x}{\partial \bar{z}} + \frac{\partial v}{\partial y}\cdot\frac{\partial y}{\partial \bar{z}}\right]$$

$$= \frac{\partial u}{\partial x}\cdot\frac{1}{2} - \frac{1}{2i}\frac{\partial u}{\partial y} + i\left[\frac{\partial v}{\partial x}\cdot\frac{1}{2} - \frac{1}{2i}\frac{\partial v}{\partial y}\right]$$

$= 0$ by Cauchy-Riemann's equations.

Thus if the function is regular, it is independent of $\bar{z}$, as its differential is zero.

Laplace's Functions

Cauchy-Riemann's equations are

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x};$$

$$\therefore \quad \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y} \quad \text{and} \quad \frac{\partial^2 u}{\partial y^2} = -\frac{\partial^2 v}{\partial x \partial y}$$

Adding, $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$

$$\because \frac{\partial^2 v}{\partial x \partial y} = \frac{\partial^2 v}{\partial y \partial x}$$

i.e., $$\nabla^2 u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \qquad ...(3.61)$$

Similarly, $$\nabla^2 v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0 \qquad ...(3.62)$$

These are known as Laplace's equations, in which both u and v satisfy Laplace's equation in two dimensions.

NOTES

Harmonic Functions

A function of x, y is said to be a harmonic function if it possesses continuous partial derivatives of the first and second orders and satisfies Laplace's equation.

Two harmonic functions u and v as satisfying (3.61) and (3.62) are known as *Conjugate harmonic functions* or simply *conjugate functions*.

Determination of Conjugate Functions

If $f(z) = u + iv$ is an analytic function such that u and v are conjugate functions then being given one of them say u, we have to determine v.

We have

$$dv = \frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy \qquad \because\ v \text{ is a function of } x, y.$$

$$= -\frac{\partial u}{\partial y}dx + \frac{\partial u}{\partial x}dy \quad \text{by Cauchy-Riemann equations.}$$

The R.H.S. of this equation being of the form $Mdx + Ndy$ will be exact if

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \text{ where } M = -\frac{\partial u}{\partial y} \text{ and } N = \frac{\partial u}{\partial x}$$

i.e., $$\frac{\partial}{\partial y}\left(-\frac{\partial u}{\partial y}\right) = \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)$$

or $$-\frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial x^2}$$

i.e., $$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

As u satisfies Laplace's equation, it is harmonic and hence its conjugate v can be found out by integrating the equation.

$$dv = -\frac{\partial u}{\partial y}dx + \frac{\partial u}{\partial x}dy.$$

As an illustration if $u = y^3 - 3x^2y$, then $\frac{\partial u}{\partial x} = -6xy$, $\frac{\partial u}{\partial y} = 3y^2 - 3x^2$

$$\frac{\partial^2 u}{\partial x^2} = -6y, \frac{\partial^2 u}{\partial y^2} = 6y$$ so that u satisfies Laplace's equation and hence is harmonic.

Now $$v = \frac{\partial v}{\partial x}dx + \frac{\partial v}{dy}dy$$

$$= -\frac{\partial u}{\partial y}dx + \frac{\partial u}{\partial x}dy \quad \text{by Cauchy-Riemann equations}$$

$$= -(3y^2 - 3x^2)\,dx - 6xy\,dy$$

$$= -(3y^2\,dx + 6xydy) + 3x^2dx$$

Integrating $v = -3xy^2 + x^3 + c$ which is harmonic conjugate to u. Corresponding analytic function $f(z) = u + iv$

NOTES

$$= y^3 - 3x^2y + i(3xy^2 + x^3 + c)$$
$$= i(x + iy)^3 + ic$$
$$= iz^3 + c'$$

Using an alternative method, $\dfrac{\partial u}{\partial x} = -6xy = \dfrac{\partial v}{\partial y}$ by Cauchy-Riemann equations.

Integrating $\dfrac{\partial v}{\partial x} = -6xy$, we get $v = -3xy^2 + \phi(x)$, $\phi(x)$ being arbitrary.

But $\dfrac{\partial v}{\partial x} = -\dfrac{\partial u}{\partial y}$ gives

$$-3y^2 + \phi'(x) = -3y^2 + 3x^2$$

i.e., $\phi'(x) = 3x^2$ and so $\phi(x) = x^3 + c$

Construction of a function f (z) when one conjugate is given (due to **Milne-Thomson**)

If $z = x + iy$, $\bar{z} = x - iy$ so $x = \dfrac{z+\bar{z}}{2}, y = \dfrac{z-\bar{z}}{2i}$

$\therefore$ $$f(z) = u(x, y) + iv(x, y)$$
$$= u\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i}\right) + iv\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i}\right)$$

Treating it as a formal identity in two independent variables z and $\bar{z}$ and putting $z = \bar{z}$, we get $x = z$, $y = 0$, so that

$$f(z) = u(z, 0) + iv(z, 0)$$

Taking $f(z) = u + iv$ to be analytic, we have

$$f'(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}$$
$$= \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y} \text{ by Cauchy-Riemann equations.}$$

Writing $\phi(x, y) \equiv \dfrac{\partial u}{\partial x}, \psi(x, y) \equiv \dfrac{\partial u}{\partial y}$, we have

$$f'(z) = \phi(x, y) - i\psi(x, y)$$
$$= \phi(z, 0) - i\psi(z, 0)$$

Integrating, $f(z) = \int\{\phi(z,0) - i\psi(z,0)\}\,dz + c$, c being arbitrary constant.

Similarly if $v(x, y)$ is given then we can find

$$f(z) = \int\{\Phi(z,0) - i\Psi(z,0)\}\,dz + C$$

where $\Phi(x, y) = \dfrac{\partial v}{\partial y}$ and $\Psi(x, y) = \dfrac{\partial v}{\partial x}$

As an illustration if $u = e^x(x\cos y - y\sin y)$, then

$$\frac{\partial u}{\partial x} = e^x (x \cos y - y \sin y + \cos y) = \phi (x, y) \text{ (say)}$$

and $$\frac{\partial u}{\partial y} = e^x (-x \sin y - y \cos y - \sin y) = \psi (x, y) \text{ (say)}$$

NOTES

So that $\phi(z, 0) = e^z(z + 1)$ and $\psi (z, 0) = 0$

$$\therefore \quad f'(z) = \phi (z, 0) - i\psi (z, 0)$$
$$= e^z (z + 1)$$

Integrating $f(z) = ze^x + c$.

Problem 3.5. *Prove that the function $u + iv = f(z)$ where*

$$f(z) = \frac{x^3(1+i) - y^3(1-i)}{x^2+y^2} (z \neq 0),\ f(0) = 0$$

is continuous and that the Cauchy-Riemann equations are satisfied at the origin, yet $f'(0)$ does not exist.

Here $$u + iv = f(z)$$
$$= \frac{x^3(1+i) - y^3(1-i)}{x^2+y^2}$$

Equating real and imaginary parts on either side, we get

$$u = \frac{x^3 - y^3}{x^2+y^2} \text{ and } v = \frac{x^3 + y^3}{x^2+y^2} \text{ when } z \neq 0.$$

Obviously both u and v are rational and finite for all values of $z \neq 0$. Thus u and v are continuous at all those points for which $z \neq 0$. Hence $f(z)$ is continuous when $z \neq 0$.

Given that $f(0) = 0$, therefore at the origin $u = 0$, $v = 0$. Hence u and v both are continuous at the origin. As such $f(z)$ is continuous at the origin.

Conversely $f(z)$ is continuous everywhere.

Now, $$\left(\frac{\partial u}{\partial x}\right)_{\text{at } \substack{x=0\\y=0}} = \frac{\partial u(x,0)}{\partial x} \text{ at } x = 0$$

$$= \lim_{h\to 0} \frac{u(h,0) - u(0,0)}{h} \qquad \because u(h,0) = \frac{h^3 - 0}{h^2 + 0} = h \text{ etc.}$$

$$= \lim_{h\to 0} \frac{h}{h} = 1.$$

$$\left(\frac{\partial u}{\partial y}\right)_{\text{at } \substack{x=0\\y=0}} = \lim_{y\to 0} \frac{\partial u(0,y)}{\partial y} = \lim_{k\to 0} \frac{u(0,k) - u(0,0)}{k}$$

$$= \lim_{k\to 0} \frac{-k}{k} = -1 \qquad \because u(0,k) = \frac{0 - k^3}{0 + k^2} = -k \text{ etc.}$$

$$\left(\frac{\partial v}{\partial x}\right)_{\text{at } \substack{x=0\\y=0}} = \lim_{x\to 0} \frac{\partial v(x,0)}{\partial x}$$

$$= \lim_{h\to 0} \frac{v(h,0) - v(0,0)}{k} = \lim_{h\to 0} \frac{h}{h} = 1$$

and $$\left(\frac{\partial v}{\partial y}\right)_{\text{at } \substack{x=0\\y=0}} = \lim_{y\to 0} \frac{\partial v(0,y)}{\partial y}$$

$$= \lim_{k \to 0} \frac{v(0,k) - v(0,0)}{k} = \lim_{k \to 0} \frac{k}{k} = 1.$$

Thus we have found that at the origin

NOTES

$$\frac{\partial u}{\partial x} = 1 = \frac{\partial v}{\partial y} \text{ and } \frac{\partial u}{\partial y} = -1 = -\frac{\partial v}{\partial x},$$

which clearly satisfy the Cauchy-Riemann equations at $z = 0$. Again differential coefficient of $f(z)$ at $z = 0$, i.e.,

$$f'(0) = \lim_{z \to 0} \frac{f(z) - f(0)}{z - 0}$$

$$= \lim_{\substack{x \to 0 \\ y \to 0}} \frac{x^3(1+i) - y^3(1-i)}{(x^2 + y^2)(x + iy)}$$

$\because$ $$z = x + iy$$

$$= \lim_{x \to 0} \frac{x^3(1+i) - m^3(1-i)x^3}{(x^2 + m^2x^2)(x + imx)} \quad \text{by putting } y = mx$$

$$= \frac{(1+i) - m^3(1-i)}{(1+m^2)(1+im)}$$

which is not unique, as it is different for different values of m. Therefore $f(z)$ is not continuous at $z = 0$.

Hence $f'(z)$ does not exist at the origin, i.e., $z = 0$.

Problem 3.6. *Show that the function $f(z) = \sqrt{(|xy|)}$ is not regular at the origin although the Cauchy-Riemann equations are satisfied at that point.*

Let the function be

$$f(z) = \sqrt{(|xy|)} = u(x, y) + iv(x, y).$$

Equating real and imaginary parts, we have

$$u(x, y) = \sqrt{(|xy|)} \text{ and } v(x, y) = 0$$

Thus, $$\left(\frac{\partial u}{\partial x}\right)_{\substack{\text{at } x=0 \\ y=0}} = \frac{\partial u(x,0)}{\partial x} \text{ at } x = 0$$

$$= \lim_{h \to 0} \frac{u(h,0) - u(0,0)}{h} = \lim_{h \to 0} \frac{0-0}{h} = 0.$$

Similarly at $x = 0, y = 0$,

$$\frac{\partial v}{\partial x} = 0, \frac{\partial u}{\partial y} = 0, \frac{\partial u}{\partial y} = 0.$$

These values clearly satisfy the Cauchy-Riemann equations

$$\therefore \lim_{z \to 0} \frac{f(z) - f(0)}{z} = \lim_{\substack{x=0 \\ y=0}} \frac{\sqrt{(|xy|)} - 0}{x + iy}$$

Again $f'(0)$

$$= \lim_{x \to 0} \frac{x\sqrt{|m|}}{x(1+im)} \text{ by putting } y = mx$$

$$= \frac{\sqrt{|m|}}{1+im},$$

which is not unique as its values are different for different values of m. So $f(z)$ is not continuous at $z = 0$. Hence it is not regular there.

NOTES

3.6.1 Cauchy's Theorem

If $f(z)$ is a regular function of z and if $f'(z)$ is continuous at each point within and on a closed contour C, then $\int_C f(z)\,dz = 0$. (i.e., the integral of the function round a closed contour is zero).

Elementary Proof

Green's theorem states that if $P(x, y)$, $Q(x, y)$, $\frac{\partial Q}{\partial x}, \frac{\partial P}{\partial y}$ are all continuous functions of x and y in the domain D, then

$$\int_C (P\,dx + Q\,dy) = \iint_{\mathbf{D}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy \qquad \text{...(3.63)}$$

Let us now assume that $f(z) = u + iv$, where $z = x + iy$.

$$\therefore \qquad dz = x + i\,dy.$$

Substituting these values in $\int_C f(z)\,dz$, we get

$$\int_C f(z).dz = \int_C (u+iv)\,(dx+idy)$$

$$= \int_C (udx - vdy) + i\int (vdx + udy)$$

$$= \iint_D \left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dx\,dy + i\iint_D \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) dx\,dy$$

[by Result (3.63) of Green's theorem]

$$= 0$$

$$\because \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \text{ and } \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$$

(Cauchy-Riemann equations)

Rigorous Proof of Cauchy's Theorem

If $f(z)$ is analytic (regular) at all points within and on the closed contour, C *then* $\int_C f(z) = 0$.

To prove this theorem let us first consider two lemmas.

Lemma I. If C is a closed contour, then we must have

$$\int_C dz = 0 \quad \text{and also} \quad \int_C z\,dz = 0.$$

It follows from the definition of integral that

$$\int_C f(z)\,dz = \lim_{n\to\infty} \sum_{r=1}^{n} (z_r - z_{r-1})\, f(z).$$

Taking $f(z) = 1$, we have

NOTES

$$\int_C dz = \lim_{n\to\infty} \sum_{r=1}^{n} \{(z_r - z_{r-1})1\}$$

$$= 0 \text{ as max. } (z_r - z_{r-1}) \to 0, \text{ when } n \to \infty.$$

Also $$\int_C z\,dz = \lim_{n\to\infty} \sum_{r=1}^{n} \{(z_r - z_{r-1})z\}$$

$$= \lim_{n\to\infty} \sum_{r=1}^{n} \{z_{r-1}(z_r - z_{r-1})\} = \lim_{n\to\infty} \sum_{r=1}^{n} \{z_r(z_r - z_{r-1})\}$$

$$= \frac{1}{2}\left[\lim_{n\to\infty} \sum_{r=1}^{n} \{z_r(z_r - z_{r-1})\} + \lim_{n\to\infty} \sum_{r=1}^{n} \{z_{r-1}(z_r - z_{r-1})\}\right]$$

$$= \frac{1}{2}\lim_{n\to\infty} \sum \{(z_r + z_{r-1})(z_r - z_{r-1})\}$$

$$= \frac{1}{2}\lim_{n\to\infty} \sum (z_r^2 - z_{r-1}^2)$$

$$= 0 \text{ (for a closed curve).}$$

Lemma II. (Goursat's Lemma). Given ε, it is possible by suitable transversals, to divide the interior of C into a finite number of meshes, either complete squares or parts of squares, such that within each mesh there is a point z_0, such that

$$\left|\frac{f(z)-f(z_0)}{z-z_0} - f'(z_0)\right| < \varepsilon,$$

i.e., $$f(z) - f(z_0) = f'(z_0)(z - z_0) + \eta(z - z_0) \qquad ...(3.64)$$

for all values of z in the mesh when $\eta < \varepsilon$.

[*Note. Unless the contour is a square, the sum of the meshes will not be a perfect square*].

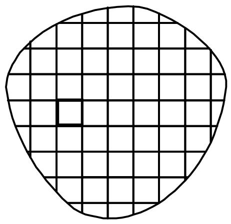

Fig. 3.21

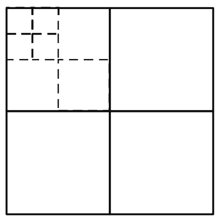

Fig. 3.22

Let us suppose that the lemma is false and however the interior of C be sub-divided. Then there will be at least one mesh for which Equation (3.46) is not true. We have to show that this necessarily implies the existence of a point within or on C at which $f(z)$ is not differentiable.

Suppose that we enclose C in a large square Γ, of area A and apply the process of repeated quadrisection. When Γ is quadrisected there is at least one of the four quarters of the square Γ for which Equation (3.64) is untrue. Let it denoted by Γ_1. We quadrisect Γ_1 and take its quarter say Γ_2 for which Equation (3.64) does not hold. This process is carried on indefinitely. Let the infinite sequence of squares so obtained be $\Gamma_1, \Gamma_2, \Gamma_3, ...\Gamma_n$.... each contained in the preceding, for which the lemma is not true. Let this sequence of squares determine a limiting point ζ which clearly lies within C.

NOTES

Now $f(z)$ being analytic everywhere and so at $z = \zeta$, it is differentiable at ζ and therefore, we have

$$\left|\frac{f(z) - f(\zeta)}{z - \zeta} - f'(\zeta)\right| < \varepsilon,$$

i.e., $f(z) - f(\zeta) = f'(\zeta)(z - \zeta) + \eta'(z - \zeta)$, where $|z - \zeta| < \delta$ ανδ $\eta' < \varepsilon$.

With centre ζ let us draw a circle of radius $\delta_1 < \delta$ and let

$$|z - \zeta| \le \delta_1 < \delta,$$

which contradicts our hypothesis, for, by taking ζ to be z_0, Equation (3.64) is satisfied and thus it follows Goursat's lemma.

Proof of the Theorem

It is obvious that some of the meshes obtained by the subdivision of the interior of C will be squares and others will not be squares. Let $C_1, C_2, ...C_m ....$ be the complete squares and $D_1, D_2, ... D_n,...$ be the partial squares, then

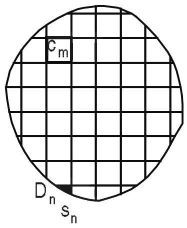

Fig. 3.23

$$\int_C f(z)\,dz = \sum \int_{C_m} f(z)\,dz + \sum \int_{D_n} f(z)\,dz \qquad ...(3.65)$$

Also we have form (3.64) of the Goursat's lemma

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \eta(z - z_0), \text{ where } |\eta| < \varepsilon, \qquad ...(3.66)$$

Now $\int_{C_n} f(z)\,dz = \int_{C_n} [f(z_0) + f'(z_0)(z - z_0) + \eta(z - z_0)]\,dz$

$$= [f(z_0) - f'(z_0)z_0]\int_{C_n} dz + f'(z_0)\int_{C_n} z\,dz + \eta\int_{C_n} (z - z_0)\,dz$$

$$= \eta\int_{C_n} (z - z_0)\,dz \text{ as } \int_{C_n} dz = \int_{C_n} z\,dz = 0 \text{ from lemma 1.}$$

$$\therefore \qquad \left|\int_{C_n} f(z)\,dz\right| \le \eta\int_{C_n} |z - z_0|\,dz$$

$\le \varepsilon\int_{C_n} \sqrt{2}\, l_n\, |dz|$, side of square being l_n; and

Max $|z - z_0| = \sqrt{2}\, l_n$, where $\sqrt{2}\, l_n$ is the length of the diagonal of square.

$\le \varepsilon\sqrt{2}\, l_n \int_{C_n} ds$, s being the entire perimeter of the square

$\le \varepsilon\sqrt{2}\, l_n \,.\, 4l_n$

$\le 4\sqrt{2}\,\varepsilon \cdot l_n^2$

$\le 4\sqrt{2}\,\varepsilon \cdot A_n, l_n^2 = A_n$, the area of the square.

NOTES

Similarly,

$$\left|\int_{D_n} f(z)\,dz\right| \le \varepsilon\sqrt{2}\,l'_n \int_{D_n} ds$$

$$\le \varepsilon\sqrt{2}\,l'_n\,(4l_n + s_n)$$

s_n being length of arc forming the curved boundary of D_n

$$\le \varepsilon\, 4\sqrt{2}\, A'_n + 2\in s_n l'_n,\ A'_n \text{ is the area of square } D_n \text{ of side } l'_n.$$

Hence Equation (3.47) gives

$$\left|\int_C f(z)\,dz\right| \le \varepsilon\sqrt{2}\sum_{n=1}^{\infty}\{4(A_n + A'_n) + s_n l'_n\}$$

$\le \varepsilon\sqrt{2}\,(4A + SL)$, S being perimeter of contour C, L is the length of a side of some square enclosing C and A the total area.

≤ 0 as $\varepsilon \to 0$, S and L both being finite.

So that $\int_C f(z)\,dz = 0.$

This proves the theorem.

Extension of Cauchy's Theorem

For this purpose, we define some terms which have not yet been introduced.

Connected Region: A region is known as a connected region if any two points of the region D can be connected by a curve lying wholly within the region.

Simply-Connected Region: A connected region is known as a simply-connected region if all the interior points of a closed curve C described in the region D, are also the points of D.

Multiply-Connected Region: A connected region is known as a multiply-connected region if all the points enclosed by two or more closed curves described in a region D are also the points of D.

Cross Cut or Simply Cut: The lines drawn in a multiply-connected region, without intersecting any curve, such that the multiply-connected region is converted to a simply-connected region, are said to be cross-cuts or cuts.

If the function $f(z)$ is not analytic in the whole region enclosed by a closed contour but it is analytic in the region enclosed between two closed contours then also Cauchy's integral theorem can be applied.

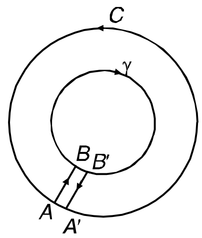

Fig. 3.24

Let the nearly equal and parallel lines AB and $A'B'$ as shown in Figure 3.24, be used as cross-cuts by connecting the points A and B (very near to each other) on outer contour C with points A' and B' on inner contour γ. Let the simply connected contour so obtained be denoted by Γ. The function $f(z)$ being analytic in this region, the Cauchy's integral formula can be applied for this contour, i.e.,

$$\int_\Gamma f(z)\,dz = 0 \text{ which gives here.}$$

$$\int_\Gamma f(z)\,dz = \int_C f(z)\,dz + \int_{AB} f(z)\,dz + \int_{A'B'} f(z)\,dz + \int_\gamma f(z)\,dz = 0$$

$$= \int_C f(z)\,dz + \int_\gamma f(z) = 0, \quad \text{other two integrals being equal and opposite in sense, cancel each other.}$$

i.e., $$\int_C f(z)\,dz = -\int_{-\gamma} f(z)\,dz$$

where minus sign shows that the integral is traversed in clockwise direction.

Therefore, taking the integral along γ in anti-clockwise direction, we get

$$\int_C f(z)\,dz = \int_\gamma f(z)\,dz.$$

Note. In general if C be a closed curve and $C_1, C_2, C_3 ..., C_n$ be the other n closed curves lying inside C and f(z) is analytic within these curves, then $\int_C f(z)\,dz = \int_{C_1} f(z)\,dz + \int_{C_2} f(z) + dz + \cdots + \int_{C_n} f(z)\,dz$ *integrals being taken in anti-clockwise direction.*

Cauchy's Integral Formula

If the function $f(z)$ is regular within and on a closed contour C and if ζ be a point within C, then

$$f(\zeta) = \frac{1}{2\pi i}\int_C \frac{f(z)\,dz}{z-\zeta}.$$

Let us describe about the point $z = \zeta$ a small circle γ of radius δ lying entirely within C. Now consider a function

$$\phi(z) = \frac{f(z)}{z-\zeta}$$

which is regular in the region between C and γ.

By making a cross-cut joining any point of γ to any point of C by two almost equal and parallel lines, let us form a closed contour $LMCM'L'\gamma L = \Gamma$ (say) within which the function $\phi(z)$ is regular, so that by Cauchy's theorem, we get

$$\int_\Gamma \phi(z)\,dz = 0, \text{ where } \phi(z) = \frac{f(z)}{z-\zeta}.$$

The function is analytic within and on the boundary of the contour and as points M, M' are very near,

$$LM = L'M' \quad \text{and} \quad LM \,||\, L'M' \text{ approximately.}$$

NOTES

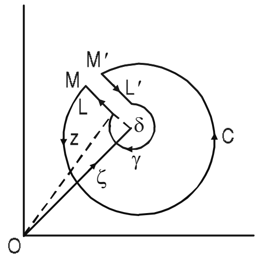

Fig. 3.25

It follows that if the contour Γ is described in anti-clockwise direction (i.e., positive sense), the cross-cut is traversed twice, once in each sense, and hence we have

$$\int_C \phi(z)\,dz - \int_\gamma \phi(z)\,dz = 0$$

or
$$\int_C \phi(z)\,dz = \int_\gamma \phi(z)\,dz,$$

i.e.,
$$\int_C \frac{f(z)}{z-\zeta}\,dz = \int_\gamma \frac{f(z)}{z-\zeta}\,dz. \qquad ...(3.67)$$

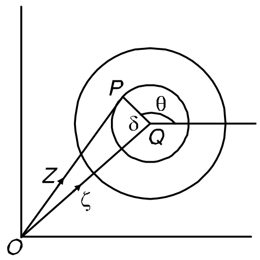

Fig. 3.26

Referring to the adjoining figure,

$$\overrightarrow{OP} = \overrightarrow{QO} + \overrightarrow{OP} = -\zeta + z = z - \zeta.$$

$\therefore$ Complex coordinate of QP is $z - \zeta = \delta e^{i\theta}$. ... (3.68)

[$\because$ if $z = x + iy$, $\therefore$ $z = r(\cos\theta + i\sin\theta) = re^{i\theta}$]

where θ is the argument and δ is the magnitude of QP which is very small.

Differentiating Equation (3.68) partially, we get $dz = i\delta e^{i\theta}\,d\theta$.

Now Equation (3.67) may be written as

$$\int_C \frac{f(z)}{z-\zeta}\,dz = \int_\gamma \frac{f(\zeta)}{z-\zeta}\,dz + \int_\gamma \frac{f(z)-f(\zeta)}{z-\zeta}\,dz$$

$$= f(\zeta)\int_0^{2\pi} \frac{i\delta\, e^{i\theta}\,d\theta}{\delta e^{i\theta}} + I \text{ (say)}$$

$$= 2\pi i f(\zeta) + I,$$

where
$$|I| = \left|\int_\gamma \frac{f(z)-f(\zeta)}{z-\zeta}\,dz\right|$$

$$\leq \int_\gamma \left|\frac{f(z)-f(\zeta)}{z-\zeta}\,dz\right|.$$

But from the definition of continuity, the function $f(z)$ is continuous at $z = \zeta$, when

$$|f(z) - f(\zeta)| < \delta_1 \text{ if } |z - \zeta| < \delta.$$

Also if
$$z = x + iy,$$
$$dz = dx + i\, dy.$$

$$\therefore \quad |dz| = \sqrt{[(dx)^2 + (dy)^2]} = ds.$$

So, $|I| \leq \dfrac{\delta_1}{\delta}\int_r ds$ where $\int_\gamma ds$ means the entire circumference of the circle γ

$$\leq \frac{\delta_1}{\delta} \times 2\pi\delta$$

$$\leq 2\pi\delta_1$$

$$\leq 0 \text{ as } \delta_1 \to 0 \text{ which is so when } \delta \to 0.$$

But we know that modulus of any quantity cannot be negative; therefore $|I| = 0$ as $\delta_1 = 0$, i.e. when z coincides with ζ.

Hence
$$\int_C \frac{f(z)}{z - \zeta} dz = 2\pi\, i f(\zeta),$$

i.e.,
$$f(\zeta) = \frac{1}{2\pi i}\int_C \frac{f(z)\, dz}{z - \zeta}.$$

Some Results Based on Cauchy's Integral Formula

1. *If (z) is regular in a domain D, then its derivative at any point $z = \zeta$ of the region D is also regular in that domain and is given by*

$$f'(\zeta) = \frac{1}{2\pi i}\int_C \frac{f(z)\, dz}{(z - \zeta)^2}.$$

where C is any single closed contour in D surrounding the point $z = \zeta$.

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

$$= \lim_{h \to 0} \frac{1}{2\pi i}\int_C \frac{f(z)\, dz}{h}\left\{\frac{1}{z - \zeta - h} - \frac{1}{z - \zeta}\right\}$$

applying Cauchy's integral formula.

$$= \lim_{h \to 0} \frac{1}{2\pi i}\int_C \frac{f(z)\, dz}{(z - \zeta)(z - \zeta - h)}$$

$$= \frac{1}{2\pi i}\lim_{h \to 0}\int_C \frac{f(z)\, dz}{(z - \zeta)(z - \zeta - h)} = \frac{1}{2\pi i}\lim_{h \to 0}\int_C \frac{(z - \zeta - h + h)\, f(z)\, dz}{(z - \zeta)^2\,(z - \zeta - h)}$$

$$= \frac{1}{2\pi i}\lim_{h \to 0}\int_C \left\{\frac{1}{(z - \zeta)^2} + \frac{h}{(z - \zeta)^2 (z - \zeta - h)}\right\} f(z)\, dz$$

$$= \frac{1}{2\pi i}\int_C \frac{f(z)\, dz}{(z - \zeta)^2} + \frac{1}{2\pi i}\lim_{h \to 0}\int_C \frac{h f(z)\, dz}{(z - \zeta)^2 (z - \zeta - h)}$$

$$= \frac{1}{2\pi i}\int_C \frac{f(z)\, dz}{(z - \zeta)^2} + I \text{ (say)},$$

NOTES

where $$I = \frac{1}{2\pi i} \underset{h\to 0}{\text{Lim}} \int_C \frac{hf(z)\,dz}{(z-\zeta)^2(z-\zeta-h)}.$$

$$\therefore \quad |I| \le \underset{h\to 0}{\text{Lim}} \frac{|h|}{2\pi} \int_C \frac{|f(z)||dz|}{|(z-\zeta)|^2\,|(z-\zeta-h)|}$$

$$\le \underset{h\to 0}{\text{Lim}} \frac{|h|}{2\pi} \cdot \frac{Ml}{d^2(d-h)}$$

$$\le 0 \text{ as } h \to 0$$

i.e. $$I = 0 \text{ as } h \to 0$$

Hence $$f'(z) = \frac{1}{2\pi i} \int_C \frac{f(z)\,dz}{(z-\zeta)^2}. \qquad \text{...(3.69)}$$

Similarly the second order derivative of $f(z)$ at $z = \zeta$ of domain D may be found as

$$f''(\zeta) = \underset{h\to 0}{\text{Lim}} \frac{f'(\zeta+h) - f'(\zeta)}{h}$$

Applying Cauchy's integral formula, in the form (3.51)

$$f''(\zeta) = \frac{1}{2\pi i} \underset{h\to 0}{\text{Lim}} \frac{1}{h} \left\{ \frac{1}{(z-\zeta-h)^2} - \frac{1}{(z-\zeta)^2} \right\} f(z)\,dz$$

$$= \frac{1}{2\pi i}\, 2 \int_C \frac{f(z)}{(z-\zeta)^3}\,dz + \underset{h\to 0}{\text{Lim}}\, h \int_C R(z)\, f(z)\,dz$$

Since $$\frac{d}{dz}\left[\frac{1}{h}\left\{\frac{1}{(z-\zeta-h)^2} - \frac{1}{(z-\zeta)^2}\right\}\right]$$

$$= \frac{2}{(z-\zeta)^3} + h\,R(z)$$ is bounded on C so that $\int_C f(z)\,R(z)$ is finite.

$$= \frac{2}{2\pi i} \int_C \frac{f(z)}{(z-\zeta)^3}\,dz$$

$$= \frac{\underline{|2}}{2\pi i} \int_C \frac{f(z)}{(z-\zeta)^3}\,dz$$

Proceeding similarly we can show that

$$f'''(\zeta) = \frac{\underline{|3}}{2\pi i} \int_C \frac{f(z)}{(z-\zeta)^4}\,dz \text{ etc.}$$

Here below the general result follows:

2. *If a function $f(z)$ is regular in a domain D then $f(z)$ has at any point $z = \zeta$ of the domain D, derivatives of all orders, values being given by*

$$f^{(n)}(\zeta) = \frac{\underline{|2}}{2\pi i} \int_C \frac{f(z)\,dz}{(z-\zeta)^{n+1}}.$$

Let us suppose that the theorem is valid for $n = m$ and then consider

$$f^{(m+1)}(\zeta) = \underset{h\to 0}{\text{Lim}} \frac{f^{(m)}(\zeta+h) - f^m(\zeta)}{h}$$

$$= \frac{(m+1)!}{2\pi i}\int_C \frac{f(z)\,dz}{(z-\zeta)^{m+2}} + I \text{ by the last result,}$$

where $$I = \frac{(m+1)!}{2\pi i}\lim_{h\to 0}\int_C \frac{hf(z)\,dz}{(z-\zeta)^{m+2}(z-\zeta-h)}.$$

NOTES

It is easy to show as in the previous result that

$$|I| \to 0 \quad \text{as} \quad |h| \to 0.$$

$$\therefore \qquad f^{(m+1)}(\zeta) = \frac{(m+1)!}{2\pi i}\int_C \frac{f(z)\,dz}{(z-\zeta)^{m+2}},$$

which shows that the result is true for $n = m + 1$ and hence we have in general

$$f^{(n)}(\zeta) = \frac{\lfloor n}{2\pi i}\int_C \frac{f(z)\,dz}{(z-\zeta)^{n+1}}.$$

Check Your Progress

7. What is a closed set?
8. What is a regular function?
9. Define harmonic function..

3.7 POLES

A pole is a certain type of singularity of a function, nearby which the function behaves relatively regularly, in contrast to essential singularities, such as 0 for the logarithm function, and branch points, such as 0 for the complex square root function. A function f of a complex variable z is meromorphic in the neighbourhood of a point z_0 if either f or its reciprocal function $1/f$ is holomorphic in some neighbourhood of z_0 (that is, if f or $1/f$ is complex differentiable in a neighbourhood of z_0).

A zero of a meromorphic function f is a complex number z such that $f(z) = 0$. A pole of f is a zero of $1/f$. This induces a duality between zeros and poles, that is obtained by replacing the function f by its reciprocal $1/f$. This duality is fundamental for the study of meromorphic functions. For example, if a function is meromorphic on the whole complex plane, including the point at infinity, then the sum of the multiplicities of its poles equals the sum of the multiplicities of its zeros.

A zero of a meromorphic function f is a complex number z such that $f(z) = 0$. A pole of f is a zero of $1/f$. If f is a function that is meromorphic in a neighbourhood of a point z_0 of the complex plane, then there exists an integer n such that $(z - z_0)^n f(z)$.

This is holomorphic and non-zero in a neighbourhood of z_0 (this is a consequence of the analytic property). If $n > 0$, then z_0 is a pole of order (or multiplicity) n of f. If $n < 0$, then z_0 is a zero of order $|n|$ of f. Simple zero and simple pole are terms used for zeroes and poles of order $|n| = 1$. Degree is sometimes used synonymously to order.

This characterization of zeros and poles implies that zeros and poles are isolated, that is, every zero or pole has a neighbourhood that does not contain any other zero and pole.

NOTES

3.8 RESIDUE AND CAUCHY'S RESIDUE THEOREM

If the function f (z) be single-valued, continuous and regular within and on a closed contour C, except a finite number of poles (singularities) within C, then

$$\int_C f(z)\,dz = 2\,\pi i \Sigma R$$

where ΣR represents the sum of the residues say R_1, R_2, R_3....R_n of f (z) at the poles α_1, α_2, α_3.....α_n (say) within A.

(*Vikram, 1969; Rohilkhand, 1985*)

Let us draw a set of circles $\gamma_1, \gamma_2....\gamma_n$ with centres $\alpha_1, \alpha_2,....\alpha_n$ and radius ρ, such that they do not intersect each other and lie entirely within the closed curve *C*. Then *f* (*z*) is regular within the region enclosed between *C* and the small circles $\gamma_1, \gamma_2,....\gamma_n$. The entire region *C* may be deformed to consist of these small circles and the polygon *P*. Now by Cauchy's theorem, we have

$$\int_C f(z)\,dz = \int_P f(z)\,dz + \sum_{r=1}^{n} \int_{\gamma_r} f(z)\,dz$$

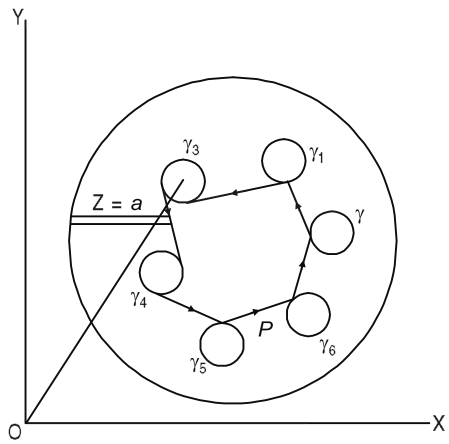

Fig. 3.27

But the integral round the polygon *P* vanishes since *f* (*z*) is regular within and on the closed contour *P*. Therefore

$$\int_C f(z)\,dz = \sum_{r=1}^{n} \int_{\gamma_r} f(z)\,dz.$$

Let us now consider *z* = *a*, a pole of order *m*; then by Laurent's expansion,

$$f(z) = \sum_{n=0}^{n} a_n (z-a)^n + \sum_{s=0}^{m} \frac{b_s}{(z-a)^s}$$

$$= \phi(z) + \sum_{1}^{m} \frac{b_s}{(z-a)^s}.$$

where $\phi(z) \left\{ = \sum_{0}^{n} a_n (z-a)^n \right\}$ is regular within and on γ_r and has no pole.

$$\therefore \qquad \int_{\gamma_r} f(z)\,dz = \int_{\gamma_r} \phi(z)\,dz + \sum_{1}^{m} b_s \int_{\gamma_r} \frac{dz}{(z-a)^s}$$

NOTES

where $\int_{\gamma_r} \phi(z)\, dz = 0$ by Cauchy's fundamental theorem.

$$\therefore \qquad \int_{\gamma_r} f(z)\, dz = \sum_{1}^{m} b_s \int_{\gamma_r} \frac{dz}{(z-a)^s}.$$

Putting $\quad z - a = \rho e^{i\theta}$, where θ varies from 0 to 2π,

$$dz = \rho i e^{i\theta} d\theta.$$

As the point z makes a circuit which consists of the circle γ_r, therefore

$$\int_{\gamma_r} f(z)\, dz = \sum_{s=1}^{m} b_s \rho i \int_0^{2\pi} \frac{e^{i\theta} d\theta}{\rho^s e^{is\theta}}$$

$$= \sum_{1}^{m} b_s \rho^{(1-s)}\, i \int_0^{2\pi} e^{i(1-s)\theta}\, d\theta.$$

Now $\quad \int_0^{2\pi} e^{i(1-s)\theta}\, d\theta = \left[\frac{e^{i(1-s)\theta}}{i(1-s)}\right]_0^{2\pi} = 0$ if $s \neq 1$.

But if $s = 1$, all the terms will be zero except one.

$$\therefore \qquad \int_{\gamma_r} f(z)\, dz = b_1 i \int_0^{2\pi} d\theta = 2\pi i b_1,$$

where b_1 is called the residue for the function.

Let the residues for $r = 1, 2, 3, \ldots n$ be respectively $R_1, R_2, R_3, \ldots, R_n$; then

$$\int_{\gamma_1} f(z)\, dz = 2\pi i R_1,$$

$$\int_{\gamma_2} f(z)\, dz = 2\pi i R_2,$$

$$\ldots \quad \ldots \quad \ldots$$

$$\ldots \quad \ldots \quad \ldots$$

$$\int_{\gamma_n} f(z)\, dz = 2\pi i R_n.$$

Hence $\quad \int_C f(z)\, dz = \sum_{1}^{m} \int_{\gamma_r} f(z)\, dz$

$$= 2\pi i \Sigma R.$$

3.9 CONTOUR INTEGRATION

The residue of a function $f(z)$ at the pole $z = a$ is defined to be the coefficient of $(z - a)^{-1}$ in the Laurent's expansion of the function $f(z)$, *i.e.*,

$$f(z) = \sum_{0}^{\infty} a_n (z-a)^n + \sum_{1}^{m} b_n (z-a)^{-n},$$

where $z = a$ is a pole of order m.

If $z = a$ be the pole of order one, then the residue is

$$b_1 = \lim_{z \to a} (z-a) f(z),$$

i.e., in case of a simple pole $z = a$.

$$\text{Residue} = \lim_{z \to a} (z-a) f(z).$$

Now consider the integral $b_n = \frac{1}{2\pi i} \int_{C_2} (z-a)^{n-1} f(z)\, dz$ which is the value of b_n in Laurent's expansion. Here the circle C_2 is arbitrary and may therefore be replaced by any closed contour C containing within it no other singularities except $z = a$. Thus

$$\frac{1}{2\pi i}\int_{C_2} f(z)\,dz = \frac{1}{2\pi i}\int_C f(z)\,dz,$$

where $n = 1$, *i.e.*, pole is of order one,

NOTES

i.e., $$b_1 = \frac{1}{2\pi i}\int_C f(z)\,dz.$$

Hence as an alternative, the residue of a funciton $f(z)$ at the pole $z = a$ is equal to $\frac{1}{2\pi i}\int_C f(z)\,dz$, where C is a closed contour containing within it the only singularity $z = a$ and integration being taken round C in anti-clockwise direction, *i.e.*, the positive sense.

Similarly the residue of $f(z)$ at infinity, *i.e.*, at the point $z = \infty$ is $\frac{1}{2\pi i}\int_C f(z)\,dz$ taken round C in clockwise direction, as it is negative w.r.t. the interior of C and positive w.r.t. the exterior of C, provided the value of this integral is definite.

Check Your Progress

10. What is pole?
11. State Cauchy's Residue theorem.

3.10 ANSWERS TO 'CHECK YOUR PROGRESS'

1. An ordered pair of real numbers such as (x, y) is termed as a complex number.
2. If $z = x + iy$ and y is used as radian measure of the angle to define cos y, sin y etc., then the exponential function in terms of real valued functions is defined by $e^x = e^{x + iy} = e^x e^{iy} = e^x (\cos y + i \sin y)$.
3. $e^z = r(\cos q + i \sin q)$.
4. The necessary condition for conformality is that $f(z)$ must be analytic.
5. The plane whose points are represented by complex numbers is known as Argand Plane or Argand diagram or Complex plane or Gaussian plane.
6. If d(n) being the number of divisors of n, then the unit circle is a natural boundary of this function.
7. If all the limit points of the set belong to the set, then the set is said to be a closed set.
8. A function $f(z)$ which is single-valued and differentiable at every point of a domain D, is said to be regular in the domain D.
9. A function of x, y is said to be a harmonic function if it possesses continuous partial derivatives of the first and second orders and satisfies Laplace's equation.
10. A pole is a certain type of singularity of a function, nearby which the function behaves relatively regularly, in contrast to essential singularities, such as 0 for the logarithm function, and branch points, such as 0 for the complex square root function.
11. If the function f(z) be single-valued, continuous and regular within and on a closed contour C, except a finite number of poles (singularities) within C, then where SR represents the sum of the residues say R1 , R2 , R3....Rn of f(z) at the poles a1 , a2 , a3.....an (say) within A.

3.11 SUMMARY

- An ordered pair of real numbers such as (x, y) is termed as a complex number. If we write $z = (x, y)$ or $x + iy$, where $i = -1$, then x is called the *real part* and y the *imaginary part* of the complex number z.
- If the function $f(z)$ is regular within and on a closed contour C and if z be a point within C, then

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(z)\,dz}{z-\zeta}.$$

- Let three be an annulus between two concentric circles C_1 and C_2 of centre $z = a$ and radii r_1 and r_2 $(r_1 > r_2)$; then if $f(z)$ be regular within the annulus between C_1 and C_2, as well as on the circles, and z be any point of the annulus,

$$f(z) = \sum_{n=0}^{\infty} a_n (\zeta - a)^n + \sum_{n=1}^{\infty} b_n (\zeta - a)^{-n},$$

$$\text{where } a_n = \frac{1}{2\pi i}\int_{C_1} \frac{f(z)\,dz}{(z-a)^{n+1}} \text{ and } b_n = \frac{1}{2\pi i}\int_{C_2} (z-a)^{n-1} f(z)\,dz.$$

- If $z = x + iy$ and y is used as radian measure of the angle to define cos y, sin y etc., then the exponential function in terms of real valued functions is defined by

  $e^x = e^{x + iy} = e^x e^{iy} = e^x (\cos y + i \sin y)$
- If $w = f(z)$ and corresponding to each point (x, y) in z-plane in a domain of function f, there is a point (u, v) in w-plane where $z = x + iy$ and $w = u + iv$, then this correspondence between the points of two planes is said to be a mapping or a transformation of points in the z-plane into points of the w-plane by the function f.
- If the sense of rotation as well as the magnitude of the angle is preserved, the mapping is conformal.
- The plane whose points are represented by complex numbers is known as Argand Plane or Argand diagram or Complex plane or Gaussian plane.
- The modulus of product of two complex numbers is equal to the product of their moduli and argument of the product of two complex numbers is the sum of their arguments.
- If there exist two functions $f1(z)$ and $f2(z)$, such that they are analytic (regular) in domains $D1$ and $D2$ respectively and that $D1$ and $D2$ have a common part, throughout which $f1(z) = f2(z)$, then the aggregate of values of $f1(z)$ and $f2(z)$ at the interior points of $D1$ or $D2$, can be regarded as a single regular function (say) $F(z)$.
- If $w = u + iv$ and $z = x + iy$ are two complex numbers, then w is said to be the function of z and written as $w = f(z)$, if to every value of z in a certain domain D, there correspond one or more values of w. If w takes only one

NOTES

value for each value of z in the domain D, then w is said to be uniform or single-valued function of z and if it takes more than one values for some or all values of z in the domain D, then w is known as a many-valued or a multiple-valued function of z.

- If $f(z)$ be a single-valued function defined in a domain D of the Argand diagram, then $f(z)$ is said to be differentiable at $z = z_0$ at a point of D if tends to a unique limit when z ® z_0, provided that z is also a point of D.
- A function $f(z)$ which is single-valued and differentiable at every point of a domain D, is said to be *regular* in the domain D.
- The two relations, which are necessary conditions for a function to be analytic, are called the Cauchy Rieman Differential Equations.
- The continuous single-valued function $f(z)$ is regular in a domain C if the four partial derivatives u_x, u_y, v_x, v_y exist, are continuous and satisfy the Cauchy-Riemann equations at all points of the region D.
- A pole is a certain type of singularity of a function, nearby which the function behaves relatively regularly, in contrast to essential singularities, such as 0 for the logarithm function, and branch points, such as 0 for the complex square root function.
- A zero of a meromorphic function f is a complex number z such that $f(z) = 0$. A pole of f is a zero of $1/f$. This induces a duality between zeros and poles that is obtained by replacing the function f by its reciprocal $1/f$. This duality is fundamental for the study of meromorphic functions.
- The residue of a function $f(z)$ at the pole $z = a$ is defined to be the coefficient of $(z - a)$–1 in the Laurent's expansion of the function $f(z)$,

3.12 KEY TERMS

- **Complex number:** Gauss introduced a number of the form a + *i*b which satisfies every algebraic equation with real coefficients. Such a number a + *i*b with $i = -1$ and a, b being real, is known as a complex number.
- **Cauchy–Riemann equations**: The Cauchy–Riemann equations consist of a system of two partial differential equations which, together with certain continuity and differentiability criteria, form a necessary and sufficient condition for a complex function to be complex differentiable, that is, holomorphic.
- **Conformal map**: A conformal map is a function that preserves orientation and angles locally. In the most common case, the function has a domain and an image in the complex plane.
- **Residue:** In mathematics, more specifically complex analysis, the residue is a complex number proportional to the contour integral of a meromorphic function along a path enclosing one of its singularities.

3.13 SELF-ASSESSMENT QUESTIONS AND EXERCISES

NOTES

Short Answer Questions

1. What are trigonometric functions?
2. What are the properties of hyperbolic functions?
3. What are conformal and isogonal transformations?
4. Write in brief about Laplace's functions.
5. State and prove Cauchy's theorem.
6. What do you mean by the term 'residue'?
7. What conditions are satisfied by an analytic function of a complex variable?

Long Answer Questions

1. Discuss necessary and sufficient conditions for conformality.
2. Illustrate the concept of linear or bilinear or Möbius' transformation.
3. Show that the transformation changes the circle $x^2 + y^2 - 4x = 0$ into the straight line $4u + 3 = 0$.
4. How can the *sum*, *difference*, *product* and *quotient* of complex numbers be geometrically represented on the Argand plane? Explain.
5. Describe the method of analytic continuation.
6. Explain the necessary and sufficient conditions for $f(z)$ to be regular.
7. Discuss polar form of Cauchy-Riemann equations.
8. Prove that the function:

   $$u = x^3 - 3xy^2 + 3x^2 - 3y^2 + 1$$

   satisfies Laplace's equation and determine the corresponding regular function $u + iv$.
9. State and prove Cauchy's Residue theorem for a complex function. Explain how it is extended for the case of an isolated first order pole lying on the contour of integration. Using this theorem show that

   $$\int_{-\infty}^{\infty} \frac{e^{ax}}{1+x^2}\,dx = \frac{\pi}{\sin \pi a};\ \text{where } 0 < a < 1.$$

10. If m > 0, a > 0, prove by contour integration

    $$\int_0^{\infty} \frac{x \sin mx}{x^4 + a^4}\,dx = \frac{\pi}{2a^2} e^{-ma/\sqrt{2}} \cdot \sin\frac{ma}{\sqrt{2}}$$

11. Show by contour integration

    $$\int_0^{\infty} \frac{e^{\alpha x}}{1+e^x}\,dx = \frac{\pi}{\sin \pi\alpha},\ 0 < \alpha < 1.$$

3.14 FURTHER READING

Dass, HK. 2008. *Mathematical Physics*. New Delhi: S. Chand

Chattopadhyay, PK. 1990. *Mathematical Physics*. New Delhi: New age International

Hassani, Sadri. 2013. *Mathematical Physics: A Modern Introduction to Its Foundations* Berlin: Springer Science & Business Media

Khanna, V.K. and S.K. Bhambri. 2008. *A Course in Abstract Algebra*, 3rd edition. New Delhi: Vikas Publishing House

Grewal, B.S. 1998. *Higher Engineering Mathematics,* 34th edition. New Delhi: Khanna Publishers

Narayan, Shanti. 1996. *Differential Calculus,* 14th edition. New Delhi: S. Chand and Company

Kreyszig, Erwin. *Advanced Engineering Mathematics,* 7th edition. Mumbai: Wiley Eastern Ltd

Bali, N.P. 2007. *A Textbook of Engineering Mathematics*. New Delhi: Laxmi Publications (P) Ltd

Narayan, Shanti. 2010. *A Text Book of Vector Analysis.* New Delhi: S. Chand and Sons

Spiegel. 1959. *Vector Analysis (Schaum'S Outline).* New York: McGraw Hill Professional

Gupta, B. D. 2009. *Mathematical Physics*, 4th Edition. Noida: Vikas Publishing House Pvt. Ltd.

NOTES

UNIT 4 SPECIAL FUNCTIONS AND SPHERICAL HARMONICS

Structure

4.0 INTRODUCTION

In mathematics, Legendre polynomials (named after Adrien-Marie Legendre) are the polynomial solutions to Legendre's differential equation. The polynomials are either even or odd functions of x for even or odd orders n. Legendre's differential equation is frequently encountered in physics and other technical fields. In particular, it occurs when solving Laplace's equation (and related partial differential equations) in spherical coordinates. The generating function is the basis for multipole expansions. The Legendre differential equation may be solved using the standard power series method. The equation has regular singular points at $x = \pm 1$ so, in general, a series solution about the origin will only converge for $|x| < 1$. Many linear differential equations having variable coefficients cannot be solved by usual methods and we need to employ series solution method to find their solutions in terms of infinite convergent series. In this unit, we introduce the Bessel differential equation and deduce from it the Bessel's function. Bessel function possesses properties of which great use can be made in the discussion of physical problems. A few of the more important of these are recurrence formulae, generating function for the Bessel coefficients, and modified Bessel function. Many facts about Bessel functions can be proved by using its generating function. 'Modified' equation plays a significant role in science and engineering. Hermite functions play significant role in equatorial dynamics and it is convenient to collect information about them from various sources. They were named after Charles Hermite, who wrote on the polynomials in 1864. Hermite was the first to define the multidimensional polynomials. In this Unit, you will study Hermite differential equation, recurrence relations and generating functions for Hermite function, Hermite polynomials and Rodrigue's formula for Hermite function. You will learn two independent solutions that the Hermite equation may have. Orthogonal properties of Hermite polynomials is explained not only for normalised Hermite functions but in terms of Kronecker delta symbol also. Laguerre functions can be exhibited in terms of the degenerate hyper geometric function or in terms of Whittaker functions. They were named after Edmond Laguerre. The Laguerre polynomials are solutions of Laguerre's equation. They are also helpful for Gaussian quadrature to numerically calculate integrals. They originate in quantum mechanics, in the radial part of the solution of the Schrödinger equation for a one electron atom. In this unit, you will study Laguerre differential equation, recurrence relations and generating functions for Laguerre functions and Laguerre polynomials. You will learn Laguerre polynomials with their representation in terms of confluent hyper geometricseries. Orthogonal properties of Laguerre polynomials are also explained.

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Explain the Legendre's differential equation
- Discuss generating functions and their recurrence relations
- Describe the orthogonal properties of Legendre's polynomials
- Discuss Bessel's differential equation and generating function
- State the orthogonal properties of Bessel's polynomials, Hermite polynomials and Laguerre polynomials
- Understand the concept of recurrence relation and generating functions for Hermite function.

4.2 LEGENDRE FUNCTION

Legendre's Differential Equation

This equation is of the form

$$(1-x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + n(n+1)y = 0. \qquad ...(4.1)$$

The equation of such type can be solved in series of ascending or descending powers of x. Suppose, we have to integrate it in a series of descending powers of x. There is no singularity at $x = 0$, so the solution of the equation can be obtained in the form of a series developed about $x = 0$.

Let us assume the solution of the given equation in the form of series.

$$y = \sum_{r=0}^{\infty} a_r x^{k-r}$$

$$\therefore \qquad \frac{dy}{dx} = \sum_{r=0}^{\infty} (k-r)\, a_r x^{k-r-1}$$

and
$$\frac{d^2y}{dx^2} = \sum_{r=0}^{\infty} (k-r)\,(k-r-1)\, a_r x^{k-r-2}.$$

Substituting these values in (4.1), we have

$$\sum_{r=0}^{\infty} [(1-x^2)\,(k-r)\,(k-r-1)x^{k-r-2} - 2x\,(k-r)\,x^{k-r-1} + n(n+1)x^{k-r}]a_r = 0$$

or
$$\sum_{r=0}^{\infty} [(k-r)\ (k-r-1)x^{k-r-2} + \{n(n+1) - 2(k-r) - (k-r)\,(k-r-1)\}\,x^{k-r}]\,a_r = 0$$

or
$$\sum_{r=0}^{\infty} [(k-r)\ (k-r-1)x^{k-r-2} + \{n(n+1) - (k-r)(k-r+1)\}\,x^{k-r}]a_r = 0. \quad ...(4.2)$$

The relation (4.2) is an identity and therefore the coefficients of various powers of x can be equated to zero.

Let us first equate the coefficient of x^k the highest power of x {by putting $r = 0$ in (4.2)} to zero; then we get

$$a_0\,\{n\,(n+1) - k(k+1) = 0,$$

where a_0 being the coefficient of the first term of the series cannot be zero, *i.e.* $a_0 \neq 0$ and thus

$$n\,(n+1) - k\,(k+1) = 0$$

or $\quad n^2 + n - k^2 - k = 0$

or $\quad n^2 - k^2 + (n-k) = 0$

or $\quad (n-k)\,(n-k+1) = 0$

which gives $\quad k = n \text{ or } -n-1 \quad$ ...(4.3)

Again equating the coefficient of x^{k-1} to zero, by putting $r = 1$ in (4.2), we have

$$\{n\,(n+1) - (k-1)\,k\}\ a_1 = 0 \quad ...(4.4)$$

From (4.3), $\{n\,(n+1) - k\,(k-1)\} \neq 0$

and therefore $\quad a_1 = 0.$

Let us now equate the coefficient of x^{k-r}, the general term in (4.2), to zero,

$$(k-r+2)\,(k-r+1)a_{r-2} + \{n\,(n+1) - (k-r)\,(k-r+1)\}\ a_r = 0$$

or
$$a_r = -\frac{(k-r+2)\,(k-r+1)}{n\,(n+1) - (k-r)\,(k-r+1)}a_{r-2}. \quad ...(4.5)$$

NOTES

Putting $k = n$, the recurrence formula is

$$a_r = -\frac{(n-r+2)(n-r+1)}{n^2+n-(n-r)(n-r+1)} a_{r-2}$$

$$= -\frac{(n-r+2)(n-r+1)}{n^2+n-n^2+nr-n+nr-r^2+r} a_{r-2}$$

$$= -\frac{(n-r+2)(n-r+1)}{r(2n-r+1)} a_{r-2}. \quad ...(4.6)$$

Again putting $k = -n-1$ in (4.5), we have the recurrence formula as

$$a_r = -\frac{(-n-r+1)(-n-r)}{n^2+n-(-n-r-1)(-n-r)} a_{r-2}$$

$$= -\frac{(n+r-1)(n+r)}{n^2+n-(n+r+1)(n+r)} a_{r-2}$$

$$= \frac{(n+r-1)(n+r)}{r(2n+r+1)} a_{r-2}. \quad ...(4.7)$$

Case I. When $k = n$, we have by putting $r = 2, 3, ...$in (4.6),

$$a_2 = -\frac{n(n-1)}{2(2n-1)} a_0$$

$$a_3 = -\frac{(n-1)(n-2)}{3(2n-2)} a_1$$

$$= 0 \text{ since } a_1 = 0.$$

Similarly a_5, a_7, a_9....etc. all the a's having odd suffixes are zero.

$$a_4 = -\frac{(n-2)(n-3)}{4(2n-3)} a_2 = \frac{n(n-1)(n-2)(n-3)}{2\cdot 4(2n-1)(2n-3)} a_0$$ (by putting value of a_2 etc.)

In general, $a_{2r} = (-1)^r \dfrac{n(n-1)(n-2)(n-3)\cdots(n-2r+1)}{2\cdot 4\cdots 2r\,(2n-1)(2n-3)\cdots(2n-2r+1)} a_0$

(by putting value of a_2 etc.)

Hence the series solution when $k = n$, is

$$y = a_0\left[x^n - \frac{n(n-1)}{2(2n-1)}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2\cdot 4(2n-1)(2n-3)}x^{n-4}\cdots\right] \quad ...(4.8)$$

where a_0 is an arbitrary constant and is equal to

$$\frac{1\cdot 3\cdot 5...(2n-1)}{\lfloor n} \quad ...(4.9)$$

where n is a positive integer.

This solution of Legendre's equation is known as $P_n(x)$, *i.e.*

$$P_n(x) = \frac{1\cdot 3\cdot 5...(2n-1)}{\lfloor n}\left[x^n - \frac{n(n-1)}{2(2n-1)}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2\cdot 4(2n-1)(2n-3)}x^{n-4} - \cdots\right] \quad ...(4.10)$$

NOTES

Case II. When $k = -n-1$, we have by putting $r = 2, 3,...$ in (4.7),

$$a_2 = \frac{(n+1)(n+2)}{2(2n+3)}a_0.$$

Now a_3 will contain a_1 and hence is zero. As such $a_5, a_7, a_9...$ all the zero.

$$a_4 = \frac{(n+3)(n+4)}{4(2n+5)}a_2$$

$$= \frac{(n+1)(n+2)(n+3)(n+4)}{2\cdot 4(2n+3)(2n+5)}a_0, \text{ by putting value of } a_2 \text{ etc.}$$

In general, $$a_{2r} = \frac{(n+1)...(n+2r)}{2\cdot 4...2r\,(2n+3)...(2n+2r+1)}a_0.$$

Hence the series solution is

$$y = a_0\left[x^{-n-1} + \frac{(n+1)(n+2)}{2(2n+3)}x^{-n-3} + \frac{(n+1)(n+2)(n+3)(n+4)}{2\cdot 4(2n+3)(2n+5)}x^{-n-5} + \cdots\right] \quad ...(4.11)$$

With $$a_0 = \frac{\lfloor n}{1\cdot 3\cdot 5...(2n+1)} \quad ...(4.12)$$

this solution is known as $Q_n(x)$.

$$\therefore \quad Q_n(x) = \frac{\lfloor n}{1\cdot 3\cdot 5...(2n+1)}\left[x^{-n-1} + \frac{(n+1)(n+2)}{2(2n+3)}x^{-n-3} + \frac{(n+1)(n+2)(n+3)(n+4)}{2\cdot 4(2n+3)(2n+5)}x^{-n-5} + \cdots\right] \quad ...(4.13)$$

which does not terminate when $n > -1$.

The most general solution of the Legendre's equation is

$$y = AP_n(x) + BQ_n(x). \quad ...(4.14)$$

where A and B are arbitrary constants.

Thus the series (4.8) or (4.11) will be convergent if $|x| > 1$ *i.e.* the above solutions for Legendre equation are not convergent in the interval $-1 < x < 1$. In order to find the convergent solution of (4.1) we seek for solution in descending powers of x.

Suppose a series solution of (4.1) is

$$y = \sum_{r=0}^{\infty} a_r x^{k+r}, \; a_0 \neq 0 \quad ...(4.15)$$

NOTES

So that $$y' = \sum_{r=0}^{\infty}(k+r)a_r x^{k+r-1}$$

$$y'' = \sum_{r=0}^{\infty}(k+r)\ (k+r-1)\ a_r x^{k+r-2}$$

Subsituting these values of y, y' and y'' in (4.1) we get the identity

$$\sum_{r=0}^{\infty}\left[(k+r)\ (k+r-1)\ x^{k+r-2} - (k+r-n)\ (k+r+n+1)x^{k+r}\right] a_r \equiv 0 \qquad ...(4.16)$$

Equating to zero the coefficient of x^{k-2} (when $r = 0$) the first term in (4.16) under the assumption $a_0 \neq 0$ yields $k(k-1) = 0$ giving $k = 0, 1$.

Now equating the coefficient of second term *i.e.* x^{k+1} to zero, we have a_1 $(k + 1)k = 0$ giving $a_1 \neq 0$ for $k = -1$ while a_1 may or may not be zero for $k = 0$.

Equating the coefficient of general term *i.e.* x^{k+r} to zero, we find the recursion formula

$$a_{r+2} = \frac{(k+r-n)\ (k+r+n+1)}{(k+r+2)\ (k+r+1)} a_r \qquad ...(4.17)$$

Putting $k = 0$, (4.17) gives $a_{r+2} = \dfrac{(r-n)\ (r+n+1)}{(r+2)\ (r+1)} a_r$ ...(4.18)

And putting $k = 1$, (4.17) gives $a_{r+2} = \dfrac{(1+r-n)\ (2+r+n)}{(3+r)\ (2+r)} a_r$ ...(4.19)

Case I. When $k = 0$, we have by putting $r = 0, 1, 2, 3, 4, 5, ...$ in (4.18)

$$a_2 = -\frac{n(n+1)}{\lfloor 2} a_0 \text{ and } a_3 = -\frac{(n-1)\ (n+2)}{\lfloor 3} a_1$$

$$a_4 = \frac{n\ (n-2)\ (n+1)\ (n+3)}{\lfloor 4} a_0 \text{ and}$$

$$a_5 = \frac{(n-1)\ (n-3)\ (n+2)\ (n+4)}{\lfloor 5} a_1$$

and in general $$a_{2r} = (-1)^r \frac{n(n-2)...(n-2r+1)}{\lfloor 2r} a_0$$

$$a_{2r+1} = \frac{(-1)^r\ (n-1)\ (n-3)...(n-2r+1)\ (n+2)...(n+2r)}{\lfloor 2r+1} a_1$$

Hence the series solution for $k = 0$, is

$$y = a_0\left[1 - \frac{n(n+1)}{\lfloor 2}x^2 + \frac{n\ (n-2)\ (n+1)\ (n+3)}{\lfloor 4}x^4 + \cdots\right]$$

$$+ a_1 x\left[1 - \frac{(n-1)\ (n+2)}{\lfloor 3}x^2 + \frac{(n-1)\ (n-3)\ (n+2)\ (n+4)}{\lfloor 5}x^4 \cdots\right]$$

...(4.19*a*)

Case II. When $k = 1$, we have from (4.19),

NOTES

$$a_2 = -\frac{(n-1)(n+2)}{\lfloor 3}a_0 \text{ and } a_1 = a_3 = a_5 = a_{2r+1} = ... = 0$$

$$a_4 = \frac{(n-1)(n-3)(n+2)(n+4)}{\lfloor 5}a_0$$

$$a_{2r} = \frac{(-1)^r (n-1)(n-3)...(n-2r+1)(n+2)...(n+2r)}{\lfloor 2r+1}a_0$$

Hence the series solution for $k = 1$ is

$$y = a_0 x\left[1 - \frac{(n-1)(n+2)}{\lfloor 3}x^2 + \frac{(n-1)(n-3)(n+2)(n+4)}{\lfloor 5}x^4 + \cdots\right] \quad ...(4.20)$$

The solution (4.20) is included in (19*a*) in the coefficient of a_1 except that a_1 is to be replaced by a_0. Hence setting $a_1 = 0$ for $k = 0$ also, the solution (4.19*a*) reduces to

$$y = a_0\left[1 - \frac{n(n+1)}{\lfloor 2}x^2 + \frac{n(n-2)(n+1)(n+3)}{\lfloor 4}x^4 + \cdots\right] \quad ...(4.21)$$

It may be shown by ratio test that the solutions (20) and (21) are convergent in the interval $-1 < x < 1$.

Calling the solution (4.21) as $S_n(x)$ and (4.20) as $T_n(x)$, the general solution of Legendre equation in ascending powers of x is

$$y = A\,S_n(x) + B\,T_n(x). \quad ...(4.22)$$

where A and B are arbitrary constants.

Legendre Polynomials

If we put $n = 2r$ (say) *i.e.* if n be taken as even positive integer then (4.21) gives the Legendre polynomial as

$$y = a_0\left[1 - \frac{n(n+1)}{\lfloor 2}x^2 + \cdots + (-1)^{n/2}\frac{n(n-2)\cdots 4.2\,(n+1)(n+3)\cdots(2n-1)}{\lfloor n}x^n\right] \quad ...(4.23)$$

While (8) gives

$$y = a_0 x^n\left[1 - \frac{n(n-1)}{2(2n-1)}x^{-2} + \cdots + (-1)^{n/2}\frac{\lfloor n}{n(n-2)...2(n+1)(n+3)...(2n-1)}x^{-n}\right] \quad ...(4.24)$$

(4.23) and (4.24) will be identical if (24) is multiplied by

$$(-1)^{n/2}\frac{n(n-2)...4\cdot 2\,(n+1)(n+3)...(2n-1)}{\lfloor n}$$

and then the solutions (4.8), (4.21) will become identical.

Again if we take, n as an odd negative integer then (4.21) is identical with (4.11). Also if n is an odd positive integer then (4.20) reduces to

$$y = a_0\left[x - \frac{(n-1)(n+2)}{\lfloor 3}x^3 + \cdots + (-1)^{(n-1)/2}\frac{(n-1)(n-3)....2(n+2)...(2n-1)}{\lfloor n}x^n\right] \quad ...(4.25)$$

and (4.8) reduces to

$$y = a_0 x^n \left[1 - \frac{n(n-1)}{2(2n-1)} x^{-2} + \cdots + (-1)^{(n-1)/2} \frac{\lfloor n}{2 \cdot 4 \ldots (n-1)\,(n-3) \ldots (2n-1)} x^{-n+1}\right] \quad \ldots(4.26)$$

which becomes identical when multiplied by the coefficient of last term in (4.26).

NOTES

Further if n is an even negative integer, (4.20) and (4.11) become identical.

These discussions follow the conclusions:

(*i*) For integral values for n, the solutions (4.8) and (4.11) have great utility of them.

(*ii*) For positive integral n, (4.8) is a polynomial but (4.11) is an infinite series and the complete integral is a linear combination of them.

(*iii*) For negative integral n, (4.8) is an infinite series and (4.11) is a polynomial.

(*iv*) For positive integral n, in (4.9) or (4.10) we have chosen

$$a_0 = \frac{1 \cdot 3 \cdot 5 \ldots (2n+1)}{\lfloor n}$$

(*v*) For negative integral n, in (12) or (13) we have chosen

$$a_0 = \frac{\lfloor n}{1 \cdot 3 \cdot 5 \ldots (2n+1)}.$$

(*vi*) For positive integral n, the polynomial $P_n(x)$ has the expansion given by (4.10) ending with term from x *i.e.*

$$P_n(x) = \sum_{r=0} (-1)^r \frac{1 \cdot 3 \cdot 5 \ldots (2n-1)}{\lfloor n} \frac{n(n-1) \ldots (n-2r+1)}{2 \cdot 4 \ldots 2r(2n-1)\,(2n-3) \ldots (2n-2r+1)} x^{n-2r}$$

$$= \sum_{n=0}^{n/2} (-1)^r \frac{1 \cdot 3 \cdot 5 \ldots (2n-2r-1)}{2^r \lfloor r \lfloor n-2r} x^{n-2r} \quad \ldots(4.27)$$

or more concisely

$$P_n(x) = \sum_{n=0}^{n/2} (-1)^r \frac{\lfloor 2n-2r}{2^n \lfloor r \lfloor n-r \lfloor n-2r} x^{2n-r} \quad \ldots(4.28)$$

4.2.1 Generating Functions

$P_n(\mu)$ is the coefficient of h^n in $(1 - 2\mu h + h^2)^{-1/2}$

We have

$$(1 - 2\mu h + h^2)^{-1/2} = [1 - h\,(2\mu - h)]^{-1/2}$$

$$= 1 + \frac{1}{2} h\,(2\mu - h) + \frac{1.3}{2.4} h^2\,(2\mu - h^2)^2$$

$$+ \frac{1.3.5}{2.4.6} h^3\,(2\mu - h)^3 + \cdots + \frac{1.3.5 \ldots (2n-1)}{2.4.6 \ldots 2n} h^n (2\mu - h)^n + \cdots$$

(by Binomial expansion)

The coefficient of h^n in the expansion

$$= \frac{1.3.5 \ldots (2n-1)}{2.4.6 \ldots 2n} (2\mu)^n + \frac{1.3.5 \ldots (2n-3)}{2.4.6 \ldots (2n-2)} (2\mu)^{n-2} \cdot {}^{n-1}C_1$$

NOTES

$$+\frac{1.3.5...(2n-5)}{2.4.6...(2n-4)}(2\mu)^{n-4}\,.^{n-2}C_2-\cdots$$

$$=\frac{1.3.5...(2n-1)}{n!}\left[\mu^n-\frac{n(n-1)}{2\cdot(2n-1)}\mu^{n-2}+\frac{n(n-1)(n-2)(n-3)}{2\cdot 2\cdot 2!(2n-1)(2n-4)}\mu^{n-4}-\cdots\right],$$

$$=\frac{1.3.5...(2n-1)}{n!}\left[\mu^n-\frac{n(n-1)}{2\cdot(2n-1)}\mu^{n-2}+\frac{n(n-1)(n-2)(n-3)}{2\cdot 4(2n-1)(2n-3)}\mu^{n-4}+\cdots\right],$$

which is $P_n(\mu)$.

Hence $$\sum_{n=0}^{\infty} h^n P_n(\mu) = (1-2\mu h+h^2)^{-1/2} \qquad ...(4.29)$$

4.2.2 Recurrence Relation

Recurrence Formulae for $\mathbf{P}_n(\mu)$.

I. $nP_n(\mu) = (2n-1)\,\mu\, P_{n-1}(\mu) - (n-1)\, P_{n-2}(\mu)$.

Suppose $V=(1-2\mu h+h^2)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} h^n P_n(\mu)$

or $V^2(1-\mu h+h^2) = 1$.

Differentiating w.r.t. h.

$$2V\cdot\frac{dV}{dh}(1-2\mu h+h^2)+V^2(2h-2\mu) = 0$$

or $$\frac{dV}{dh}(1-2\mu h+h^2)+V(h-\mu) = 0$$

or $$(1-2\mu h+h^2)\sum_{n=0}^{\infty} nh^{n-1} P_n(\mu)+(h-\mu)\sum_{n=0}^{\infty} h^n P_n(\mu) = 0$$

$$\left[\because \frac{dV}{dh}=\sum_{n=0}^{\infty} nh^{n-1}P_n(\mu)\right]$$

The coefficient of h^{n-1} equated to zero gives

$nP_n(\mu) - 2\mu\,(n-1)\,P_{n-1}(\mu) + (n-2)\,P_{n-2}(\mu) + P_{n-2}(\mu) - \mu\, P_{n-1}(\mu) = 0$

or $$nP_n(\mu) = (2n-1)\,\mu P_{n-1}(\mu) - (n-1)P_{n-2}(\mu). \qquad ...(4.30)$$

II. $(\mu^2-1)\dfrac{dP_n(\mu)}{d\mu} = n\{\mu P_n(\mu) - P_{n-1}(\mu)\}$

$= -(n+1)\,\{\mu P_n(\mu) - P_{n+1}(\mu)\}$.

We have

$\mu P_n(\mu) - P_{n+1}(\mu)$

$$=\frac{\mu}{\pi}\int_0^{\pi}\{\mu+\sqrt{(\mu^2-1)}\cos\phi\}^n\,d\phi-\frac{1}{\pi}\int_0^{\pi}\{\mu+\sqrt{(\mu^2-1)}\cos\phi\}^{n-1}d\phi$$

$$=\frac{1}{\pi}\int_0^{\pi}\{\mu+\sqrt{(\mu^2-1)}\cos\phi\}^{n-1}\left[\mu\{\mu+\sqrt{(\mu^2-1)}\cos\phi\}-1\right]d\phi$$

NOTES

$$= \frac{\mu^2-1}{\pi}\int_0^{\pi}\{\mu^2+\sqrt{(\mu^2-1)}\cos\phi\}^{n-1}\left[1+\frac{\mu\cos\phi}{\sqrt{(\mu^2-1)}}\right]d\phi$$

$$= \frac{\mu^2-1}{\pi}\frac{d}{d\mu}\int_0^{\pi}\frac{\{\mu+\sqrt{(\mu^2-1)}\cos\phi\}^n}{n}d\phi$$

$$= \frac{\mu^2-1}{n}\frac{d}{d\mu}P_n(\mu).$$

$$\therefore\ (\mu^2-1)\frac{dP_n(\mu)}{d\mu} = n\,\{\mu P_n(\mu)-P_{n-1}(\mu)\} \qquad ...(4.31)$$

Replacing n by $(-n-1)$, as $P_n(\mu) = P_{-(n+1)}(\mu)$ and as such $P_{-(n+2)}\mu = P_{n+1}(\mu)$, we get

$$(\mu^2-1)\frac{dP_n(\mu)}{d\mu} = -(n+1)\,\{\mu P_n(\mu)-P_{n+1}(\mu)\} \qquad ...(4.32)$$

III. By Legendre's equation, we have

$$(\mu^2-1)\frac{d^2P_n(\mu)}{d\mu^2}+2\mu\frac{dP_n(\mu)}{d\mu}-n(n+1)\,P_n(\mu) = 0. \qquad ...(4.33)$$

This may be written as

$$\frac{d}{d\mu}\left\{(\mu^2-1)\frac{dP_n(\mu)}{d\mu}\right\} = n\,(n+1)\,P_n(\mu) \qquad ...(4.34)$$

or $$\frac{d}{d\mu}[n\{\mu P_n(\mu)-P_{n-1}(\mu)\}] = n\,(n+1)\,P_n(\mu)$$

[From (4.31)]

or $$\mu\frac{dP_n(\mu)}{d\mu}+P_n(\mu)-\frac{dP_{n-1}(\mu)}{d\mu} = (n+1)\,P_n(\mu)$$

or $$\mu\frac{dP_n(\mu)}{d\mu}-\frac{dP_{n-1}(\mu)}{d\mu} = nP_n(\mu). \qquad ...(4.35)$$

IV. Replacing n by $-(n+1)$ and applying $P_{-n-1}(\mu) = P_n(\mu)$, $P_{-n-2}(\mu) = P_{n+1}(\mu)$ etc.., in (4.44) we get

$$-\mu\frac{dP_n(\mu)}{d\mu}+\frac{dP_{n+1}(\mu)}{d\mu} = (n+1)\,P_n(\mu). \qquad ...(4.36)$$

Addition of (4.35) and (4.36) yields

$$(2n+1)\,P_n(\mu) = \frac{d}{d\mu}P_{n+1}(\mu)-\frac{d}{d\mu}P_{n-1}(\mu). \qquad ...(4.37)$$

Now (4.37) may be written as

$$\frac{d}{d\mu}P_{n+1}(\mu) = (2n-1)\,P_n(\mu)+\frac{d}{d\mu}P_{n-1}(\mu).$$

Replacing n by $(n-1)$, we get

$$\frac{d}{d\mu}P_n(\mu) = (2n-1)\,P_{n-1}(\mu) + \frac{d}{d\mu}P_{n-2}(\mu).$$

$$= (2n-1)\,P_{n-1}(\mu) + (2n-5)\,P_{n-3}(\mu) + \frac{d}{d\mu}P_{n-4}(\mu). \quad ...(4.38)$$

NOTES

$$\left[\begin{array}{l}\text{since by putting } n-2 \text{ for } n, \\ \dfrac{dP_{n-2}(\mu)}{d\mu} = (2n-5)P_{n-3}(\mu) + \dfrac{d}{d\mu}P_{n-4}(\mu)\end{array}\right]$$

The repeated application of this replacement will give

$$\frac{d}{d\mu}P_n(\mu) = (2n-1)\,P_{n-1}(\mu) + (2n-5)\,P_{n-3}(\mu) + (2n-9)\,P_{n-5}(\mu) + \cdots$$

ending with $3P_1(\mu)$ or $P_0(\mu)$ according as n is even or odd. ...(4.39)

This is known as **Christoffel's Expansion.**

4.2.3 Orthogonality

Orthogonal Properties of Legendre's Polynomials of the First Kind

To show that

$$\int_{-1}^{1} P_m(\mu)P_n(\mu)\,d\mu = 0 \text{ unless } m = n$$

$$= \frac{2}{2n+1} \text{ if } m = n$$

or $$= \frac{2}{2n+1}\delta_{m,n}$$

in Kronecker delta symbol m, n being positive integers.

Rodrigue's formula is

$$P_n(\mu) = \frac{1}{2^n \lfloor n}\left(\frac{d}{d\mu}\right)^n (\mu^2-1)^n.$$

With its applications, we have

$$\int_{-1}^{1} P_m(\mu)\cdot P_n(\mu)d\mu = \frac{1}{2^{m+n}\lfloor m \lfloor n}\int_{-1}^{1}\left(\frac{d}{d\mu}\right)^m(\mu^2-1)^m\cdot\left(\frac{d}{d\mu}\right)^m(\mu^2-1)^n\,d\mu$$

$$= \frac{1}{2^{m+n}\lfloor m\lfloor n}\left[\left\{\left(\frac{d}{d\mu}\right)^m(\mu^2-1)^m\cdot\left(\frac{d}{d\mu}\right)^{n-1}(\mu^2-1)^n\right\}_{-1}^{1} - \int_{-1}^{1}\left(\frac{d}{d\mu}\right)^{m+1}(\mu^2-1)^m\cdot\left(\frac{d}{d\mu}\right)^{n-1}(\mu^2-1)^n\,d\mu\right]$$

(integrating by parts)

$$= -\frac{1}{2^{m+n}\lfloor m\lfloor n}\left[\int_{-1}^{1}\left(\frac{d}{d\mu}\right)^{m+1}(\mu^2-1)^m\cdot\left(\frac{d}{d\mu}\right)^{n-1}(\mu^2-1)^n\,d\mu\right]$$

(the first term vanishing for both the limits)

NOTES

Continuing the process of integration by parts m times on R.H.S., we get

$$\int_{-1}^{1} P_m(\mu) P_n(\mu)\, d\mu = \frac{(-1)^m}{2^{m+n}\lfloor m \lfloor n} \int_{-1}^{1} \left(\frac{d}{d\mu}\right)^{2m} (\mu^2-1)^m \left(\frac{d}{d\mu}\right)^{n-m} (\mu^2-1)^n\, d\mu$$

Here $$\left(\frac{d}{d\mu}\right)^{2m} (\mu^2-1)^m = \frac{d^{2m}(\mu^2-1)^m}{d\mu^{2m}} = \lfloor(2m)$$

$$\therefore \quad \int_{-1}^{1} P_m(\mu)\, P_n(\mu)\, d\mu = \frac{(-1)^m \lfloor 2m}{2^{m+n}\lfloor m \lfloor n} \int_{-1}^{1} \left(\frac{d}{d\mu}\right)^{n-m} (\mu^2-1)^n\, d\mu \qquad ...(4.40)$$

$$= \frac{(-1)^m \lfloor 2m}{2^{m+n}\lfloor m \lfloor n} \left[\left(\frac{d}{d\mu}\right)^{n-m-1} (\mu^2-1)^n\right]_{-1}^{1}$$

$$= \frac{(-1)^m \lfloor 2m}{2^{m+n}\lfloor m \lfloor n}[0] = 0 \text{ when } m \neq n \text{ and } n > m. \qquad ...(4.41)$$

Again when $m = n$, we have

$$\int_{-1}^{1} P_n^2(\mu)\, d\mu = \frac{(-1)^n \lfloor 2n}{2^n \lfloor n} \int_{-1}^{1} (\mu^2-1)^n\, d\mu$$

or $$\int_{-1}^{1} P_n^2(\mu)\, d\mu = \frac{(-1)^m \lfloor 2n}{2^n \lfloor n} (-1)^n \int_{-1}^{1} (1-\mu^2)^n\, d\mu$$

$$= \frac{1.3.5...(2n-1)}{2^n \lfloor n} \int_{-1}^{1} (1-\mu^2)^n\, d\mu \text{ put } \mu = \cos\theta,\ d\mu = -\sin\theta\, d\theta$$

$$= -\frac{1.3.5...(2n-1)}{2^n \lfloor n} \int_{\pi}^{0} \sin^{2n+1}\theta\, d\theta$$

$$= \frac{2 \times 1.3.5...(2n-1)}{2^n \lfloor n} \int_{0}^{\pi/2} \sin^{2n+1}\theta\, d\theta$$

$$= 2 \cdot \frac{1.3.5...(2n-1)}{2^n \lfloor n} \times \frac{2n\,(2n-2)...2}{(2n+1)(2n-1)....3.1}$$

$$= 2 \cdot \frac{2^n \lfloor n}{2^n \lfloor n\,(2n+1)} = \frac{2}{2n+1}. \qquad ...(4.42)$$

Check Your Progress

1. Write Legendre's differential equation.
2. What is Christoffel's Expansion?
3. What is the most general form of Legendre's equation?

NOTES

4.3 BESSEL FUNCTIONS

Bessel's Differential Equation

This equation is of the form

$$\frac{d^2y}{dx^2}+\frac{1}{x}\frac{dy}{dx}+\left(1-\frac{n^2}{x^2}\right)y = 0. \qquad ...(4.43)$$

There is singularity at $x = 0$, and this is non-essential or removable singularity and hence the given equation may be solved by the method of series integration as allowed by Fusch-theorem.

In order to integrate it in a seires of ascending powers of x, let us assume that its series solution is

$$y = \sum_{r=0}^{\infty} a_r x^{k+r}$$

$$\therefore \qquad \frac{dy}{dx} = \sum_{r=0}^{\infty} a_r\,(k+r)\,x^{k+r-1},$$

$$\frac{d^2y}{dx^2} = \sum_{r=0}^{\infty} a_r\,(k+r)\,(k+r-1)\,x^{k+r-2}.$$

Substituting these values in (4.43), we get

$$\sum_{r=0}^{\infty}\left[(k+r)\,(k+r-1)x^{k+r-2}+\frac{1}{x}(k+r)x^{k+r-1}+\left(1-\frac{n^2}{x^2}\right)x^{k+1}\right]a_r \equiv 0$$

or $$\sum_{r=0}^{\infty}\left[(k+r)\,(k+r-1)+\{(k+r)-n^2\}x^{k+r-2}+x^{k+r}\right]a_r \equiv 0$$

or $$\sum_{r=0}^{\infty}\left[\{(k+r)^2-n^2\}x^{k+r-2}+x^{k+r}\right]a_r \equiv 0. \qquad ...(4.44)$$

The relation (4.44) being an identity, let us equate the coefficients of various powers of x to zero.

Equating to zero the coefficient of lowest power of x, *i.e.*, x^{k-2} by putting $r = 0$ in (4.44), we have

$$(k^2 - n^2)\ a_0 = 0.$$

Being the coefficient of first term, $a_0 \neq 0$.

$\therefore \quad k^2 - n^2 = 0,\ i.e.,\ k = \pm n.$ ...(4.45)

Now equating to zero the coefficient of x^{k-1} by putting $r = 1$ in (4.44), we get

$$\{(k+1)^2 - n^2\}\ a_1 = 0.$$

But from (4.45), $(k + 1)^2 - n^2 \neq 0 \ \therefore\ a_1 = 0.$ ...(4.46)

Equating to zero the coefficient of general term, *i.e.* x^{k-r} in (4.44), we find

$$\{(k + r + 2)^2 - n^2\}\ a_{r+2} + a_r = 0$$

or $$a_{r+2} = -\frac{a_r}{(k+r+2-n)\,(k+r+2+n)}. \qquad ...(4.47)$$

Case I. When $k = +n$. By putting $r = 0, 1, 2,...$in (4.47), we get

$$a_2 = -\frac{a_0}{2\,(2n+2)}$$

and $a_1 = a_3 = a_5... = 0$.

$$a_4 = -\frac{a_2}{4\,(2n+4)} = \frac{a_0}{2\cdot 4\,(2n+2)\,(2n+4)},$$

$$a_6 = -\frac{a_4}{6\,(2n+6)} = -\frac{a^2}{2\cdot 4\cdot 6\,(2n+2)\,(2n+4)\,(2n+6)}$$

. . . .

$$a_{2r} = \frac{(-1)^r a_0}{2\cdot 4\cdot 6...2r\cdot(2n+2)\,(2n+4)...(2n+2r)}$$

Hence the series solution is

$$y = a_0\left[x^n - \frac{x^{n+2}}{2\,(2n+2)} + \frac{x^{n+1}}{2\cdot 4\,(2n+2)\,(2n+4)} - \cdots\right]$$

$$= a_0 x^n\left[1 - \frac{x^2}{2\,(2n+2)} + \frac{x^4}{2\cdot 4\,(2n+2)\,(2n+4)} - \cdots\right.$$

$$\left.\cdots\frac{(-1)^r x^{2r}}{2\cdot 4...2r\,(2n+2)...(2n+2r)} + \cdots\right]$$

$$= a_0 x^n \sum_{r=0}^{\infty} \frac{(-1)^r\; x^{2r}}{2^r\,(r)!\cdot 2^r\,(n+1)...(n+r)} \qquad ...(4.48)$$

If $a_0 = \dfrac{1}{2^n\,\Gamma(n+1)}$, this solution is called as $J_n(x)$.

Thus

$$J_n(x) = \frac{x^n}{2^n\,\Gamma(n+1)}\cdot\sum_{r=0}^{\infty}\frac{(-1)^r\; x^{2r}}{2^{2r}\,(r)!\,(n+1)\,(n+2)...(n+r)}$$

$$= \sum_{r=0}^{\infty}(-1)^r\left(\frac{x}{2}\right)^{n+2r}\cdot\frac{1}{r\,!\,\Gamma(n+r+1)} \qquad ....(4.49)$$

Case II. When $k = -n$.

The series solution is obtained by replacing n by $-n$ in the value of $J_n(x)$, whence, we get

$$J_{-n}(x) = \sum_{r=0}^{\infty}(-1)^r\cdot\left(\frac{x}{2}\right)^{-n+2r}\cdot\frac{1}{r\,!\,\Gamma(-n+r+1)}. \qquad ...(4.50)$$

The complete primitive of Bessel's equation is

$A\,J_n(x) + B\,J_{-n}(x)$,where n is not an integer, A, B being two arbitrary constants.

Corollary. Bessel's equation for $n = 0$ is

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

NOTES

Its series solution by the same substitution $y = \sum_{r=0}^{\infty} a_r x^{k+r}$ (as above) is obtained to be

$$y = a_0\left(1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 \cdot 4^2} - \frac{x^6}{2^2 \cdot 4^2 \cdot 6^2} + \cdots\right).$$

If $a_0 = 1$, this solution is denoted by $J_0(x)$, *i.e.*

$$J_0(x) = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 \cdot 4^2} - \frac{x^6}{2^2 \cdot 4^2 \cdot 6^2} + \cdots \qquad \text{...(4.51)}$$

where $J_0(x)$ is called *Bessel function of zeroeth order*.

In fact $J_0(x)$ is that solution of Bessel's equation for $n = 0$, which is equal to unity for $x = 0$.

Note. $J_n(x)$ is called Bessel's function of the first kind of order n.

4.3.1 Generating Functions

Generating Function for $J_n(x)$, *i.e.* to show that

$$\mathbf{e}^{(x/2)\,(t-1/t)} = \sum_{n=-\infty}^{\infty} \mathbf{t}^n \mathbf{J}_n(\mathbf{x}).$$

We know that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$= \sum_{r=0}^{\infty} \frac{x^r}{r!}.$$

$$\therefore \qquad e^{xt/2} = \sum_{r=0}^{\infty} \frac{x^r t^r}{2^r r!}. \qquad \text{...(4.52)}$$

Similarly,

$$e^{-x/2t} = \sum_{s=0}^{\infty} \frac{(-1)^s x^s}{2^s\, t^s s!}. \qquad \text{...(4.53)}$$

Multiplying (4.52) and (4.53), we get

$$e^{x/2\,(t-1/t)} = \sum_{s=0}^{\infty} \frac{x^r t^r}{2^r\, r!} \times \sum_{s=0}^{\infty} \frac{(-1)^s x^s}{2^s t^s s!}.$$

In order to find the (t^n)th term, we should replace r by $n + s$ and then coefficient of t^n is

$$\sum_{s=0}^{\infty} \frac{x^{n+s}}{2^{n+s}\,(n+s)!} \times \frac{-(1)^s x^s}{2^s\, s!} = \sum_{s=0}^{\infty} (-1)^s \left(\frac{x}{2}\right)^{n+2s} \frac{1}{(n+s)!\,s!} = J_n(x) \qquad \text{...(4.54)}$$

Again the coefficient t^{-n} is obtained by putting $s = n + r$ and then coefficient of

$$t^{-n} = \sum_{r=0}^{\infty} \frac{x^r}{2^r\, r!} \times \frac{(-1)^{n+r}\, x^{n+r}}{2^{n+r}\,(n+r)!}.$$

$$= (-1)^n \sum_{r=0}^{\infty} (-1)^r \left(\frac{x}{2}\right)^{n+2r} \frac{1}{r!\,(n+r)!}.$$

$$= (-1)^n \cdot J_n(x)$$

$$= J_{-n}(x) \qquad \text{...(4.55)}$$

NOTES

since $J_{-n}(x) = (-1)^n J_n(x)$, where n is a positive integer.

It may be shown as below:

$$J_{-n}(x) = \sum_{r=0}^{\infty} (-1)^r \left(\frac{x}{2}\right)^{-n+2r} \cdot \frac{1}{r!\,\Gamma(-n+r+1)},$$

which tends to zero if $-n + r + 1 = 0$, *i.e.*, $r = n - 1$ $(\because \Gamma 0 = \infty)$.

Hence all the terms up to nth, vanish and therefore the limit $r = 0$ may be changed to $r = n$.

$$\therefore \quad J_{-n}(x) = \sum_{r=n}^{\infty} (-1) r \left(\frac{x}{2}\right)^{-n+2r} \cdot \frac{1}{r!\,\Gamma(-n+r+1)}.$$

Now putting $r = n + s$, where s is a positive integer, we have

$$J_{-n}(x) = \sum_{s=0}^{\infty} (-1)^{n+s} \left(\frac{x}{2}\right)^{n+2s} \cdot \frac{1}{(n+s)!\,\Gamma(-n+s+n+1)}$$

$$= \sum_{s=0}^{\infty} (-1)^{n+s} \left(\frac{x}{2}\right)^{n+2s} \cdot \frac{1}{(n+s)!\,\Gamma(s+1)}$$

or

$$J_{-n}(x) = (-1)^n \sum_{s=0}^{\infty} (-1)^s \left(\frac{x}{2}\right)^{n+2s} \cdot \frac{1}{s!\,\Gamma(n+s+1)}$$

$$= (-1)^n J_n(x). \qquad \text{...(4.56)}$$

Hence from (4.54) and (4.55), we have

$$e^{x/2(t-1/t)} = \sum_{n=-\infty}^{\infty} t^n J_n(x). \qquad \text{...(4.57)}$$

4.3.2 Recurrence Relation

Recurrence Formulae for $\mathbf{J}_n(\mathbf{x})$.

I. We know that

$$J_n(x) = \sum_{r=0}^{\infty} (-1)^n \frac{1}{r!\,(n+r)!} \left(\frac{x}{2}\right)^{n+2r}$$

where n is a positive integer.

Diffentiating it w.r.t. x, we get

$$J_n'(x) = \sum_{r=0}^{\infty} (-1)^r \frac{1}{r!\,(n+r)!} (n+2r) \left(\frac{x}{2}\right)^{n+2r-1} \cdot \frac{1}{2}.$$

Multiplying both sides by x, we have

$$xJ_n'(x) = \sum_{r=0}^{\infty} (-1)^r \frac{n+2r}{r!\,(n+r)!} \cdot \left(\frac{x}{2}\right)^{n+2r}$$

$$= n\sum_{r=0}^{\infty}(-1)^r \frac{n}{r!(n+r)!}\cdot\left(\frac{x}{2}\right)^{n+2r} + x\sum_{r=0}^{\infty}(-1)^r\cdot\frac{2r}{r!(n+r)!}\cdot\left(\frac{x}{2}\right)^{n+2r-1}$$

NOTES

or $x \cdot J_n'(x) = nJ_n(x) + x\sum_{r=1}^{\infty}(-1)^r \frac{n}{(r-1)!(n+r)!}\cdot\left(\frac{x}{2}\right)^{n-1+2r}$

[since on R.H.S. the second term vanishes for $r = 1$ and hence limit of $r = 0$ may be replaced by $r = 1$.]

Putting $r - 1 = s$, we have

$$x \cdot J_n'(x) = nJ_n(x) + x\sum_{s=0}^{\infty}(-1)^{s+1}\frac{n}{s!(n+1+s)!}\cdot\left(\frac{x}{2}\right)^{n+2s-1}$$

$$= nJ_n(x) - xJ_{n+1}(x). \quad ...(4.58)$$

II. Again $\quad xJ_n'(x) = \sum_{r=0}^{\infty}(-1)^r\frac{(n+2r)}{r!(n+r)!}\left(\frac{x}{2}\right)^{n+2r-1}\quad \frac{x}{2}$ may be written as

$$xJ_n'(x) = \sum_{r=0}^{\infty}(-1)^r\frac{-n+2(n+r)}{r!(n+r)!}\left(\frac{x}{2}\right)^{n+2r}$$

$$= -n\sum_{r=0}^{\infty}(-1)^r\frac{1}{r!(n+r)!}\left(\frac{x}{2}\right)^{n+2r} + x\sum_{r=0}^{\infty}(-1)^r\frac{1}{r!(n+r-1)!}\cdot\left(\frac{x}{2}\right)^{n-1+2r}$$

$$= -nJ_n(x) + xJ_{n-1}(x). \quad ...(4.59)$$

Sum and difference of (4.58) and (4.59) give

III. $2J_n'(x) = J_{n-1}(x) - J_{n+1}(x)$. ...(4.60)

IV. $2nJ_n(x) = x\{J_{n+1}(x) + J_{n-1}(x)\}$. ...(4.61)

V. $\frac{d}{dx}\{x^n J_n(x)\} = x^n J_{n-1}(x)$. ...(4.62)

Here $\quad \frac{d}{dx}\{x^n J_n(x)\} = nx^{n-1}J_n(x) + x^n J_n'(x)$

$$= x^{n-1}\{nJ_n(x) + xJ_n'(x)\}$$

$$= x^{n-1}\{nJ_n(x) - nJ_n(x) + xJ_{n-1}(x)\} \text{ by (30)}$$

$$= x^n J_{n-1}(x).$$

VI. Similarly it is easy to show that

$$\frac{d}{dx}\{x^{-n}J_n(x)\} = -x^{-n}J_{n+1}(x). \quad ...(4.63)$$

4.3.3 Orthogonality

Orthogonal Properties of Bessel's Polynomials.

To prove that $\int_0^a \mathbf{J}_n(\mu \mathbf{r})\,\mathbf{J}_n(\mu' \mathbf{r})\,\mathbf{r}\,\mathbf{dr} = \mathbf{0}$ **where** μ **and** μ' **are different roots of** $\mathbf{J}_n(\mu\,\mathbf{a}) = 0$.

Since $J_n(x)$ is a solution of Bessel's equation,

NOTES

$$\frac{d^2y}{dx^2}+\frac{1}{x}\frac{dy}{dx}+\left(1-\frac{n^2}{x^2}\right)y = 0 \qquad ...(4.64)$$

therefore putting $x = \mu r$ and calling $y = u$ in (4.64), we get

$$\frac{1}{\mu^2}\frac{d^2u}{dr^2}+\frac{1}{\mu r}\cdot\frac{1}{\mu}\frac{du}{dr}+\left(1-\frac{n}{\mu^2 r^2}\right)u = 0,$$

$$\left[\because \frac{dy}{dx}=\frac{du}{dx}=\frac{du}{dr}\cdot\frac{dr}{dx}=\frac{1}{\mu}\cdot\frac{du}{dr} \text{ and } \frac{d^2y}{dx^2}=\frac{1}{\mu}\frac{d}{dr}\left(\frac{1}{\mu}\frac{du}{dr}\right)=\frac{1}{\mu^2}\frac{d^2u}{dr^2}\right]$$

Multiplying throughout by $\mu^2 r^2$,

$$r^2\frac{d^2u}{dr^2}+r\frac{du}{dr}+(\mu^2r^2-n^2)\,u = 0 \qquad ....(65)$$

which is **Bessel's modified equation** and its solution $J_n(\mu a) = 0$ is **Modified Bessel Function**.

Similarly putting $x = \mu' r$ and calling $y = v$ in (4.64), we have

$$r^2\frac{d^2v}{dr^2}+r\frac{dv}{dr}+(\mu'^2r^2-n^2)\,v = 0. \qquad ...(4.66)$$

If we multiply (4.65) by $\frac{v}{r}$, (4.66) by $\frac{u}{r}$ and subtract

$$r(vu'' - ruv'') + (vu' - uv') + (\mu^2 - \mu'^2)\, ruv = 0,$$

where $u' = \frac{du}{dr}$ and $v' = \frac{dv}{dr}$ etc.

or $$\frac{d}{dr}\{r\,(vu'-uv')+(\mu^2-\mu'^2)\}\,ruv = 0, \qquad ...(4.67)$$

where $u = J_n(\mu r)$ and $v = J_n(\mu' r)$.

Integrating (4.67) w.r.t. r between the limits 0 and a, we get

$$[r\,\{J_n(\mu r)\,J_n'(\mu' r)\mu' - J_n(\mu' r)\,J_n'\,(\mu r)\,\mu\}]_0^a - \int_0^a(\mu^2-\mu'^2)\,J_n(\mu r)\,J_n(\mu' r)r\,dr = 0.$$

The first term vanishes for both the limits since

$$J_n(\mu\, a) = 0, J_n(\mu'\, a) = 0.$$

Hence $\int_0^a(\mu^2-\mu'^2)\,J_n(\mu r)\,J_n(\mu' r)\,r\,dr = 0,$

i.e., $$\int_0^a J_n(\mu r)\,J_n(\mu' r)\,r\,dr = 0 \quad \text{as} \quad \mu^2-\mu'^2 \neq 0. \qquad ...(4.68)$$

Check Your Progress

4. Write Bessel's differential equation.
5. At what point there is singularity in Bessel's differential equation?
6. Write Bessel's function of zeroeth order.

4.4 HERMITE FUNCTIONS

NOTES

This equation is of the form

$$\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + 2\nu y = 0 \qquad ...(4.69)$$

where ν is a parameter.

Suppose its series solution is

$$y = \sum_{r=0}^{\infty} a_r x^{k+r} \cdot a_0 \neq 0 \qquad ...(4.70)$$

So that $$\frac{dy}{dx} = \sum_{r=0}^{\infty} a_r\,(k+r)\,x^{k+r-1}$$

and $$\frac{d^2y}{dx^2} = \sum_{r=0}^{\infty} a_r\,(k+r)\,(k+r-1)\,x^{k+r-2}.$$

Substituting the values of y, $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ in (4.69), we get the identity

$$\sum_{r=0}^{\infty} [(k+r)\,(k+r-1)\,x^{k+r-2} - 2\,(k+r-\nu)\,x^{k+r}]\,a_r \equiv 0 \qquad ...(4.71)$$

Equating the coefficient of the first term (*i.e.* x^{k-2}) (by putting $r = 0$), to zero, we get

$$a_0 k(k-1) = 0 \text{ giving } k = 0, 1 \text{ as } a_0 \neq 0 \qquad ...(4.72)$$

Now equating to zero the coefficient of second term (*i.e.* x^{k-1}) in (4.71), we get

$a_1\, k(k + 1) = 0$ *i.e.* $a_1 = 0$ when $k = -1$ and a_1 may or may not be zero when $k = 0$.

Also equating the coefficient of x^{k+r} to zero, we find

$$a_{r+2}\,(k + r + 2)\,(k + r + 1) - 2ar\,(k + r - \nu) = 0$$

giving the recurrence relation

$$a_{r+2} = \frac{2\,(k+r-\nu)}{(k+r+2)\,(k+r+1)}\,a_r \qquad ...(4.73)$$

when $k = 0$, (4.73) becomes $a_{r+2} = \dfrac{2\,(r-\nu)}{(r+2)\,(r+1)}\,a_r$ ...(4.74)

and when $k = 1$, (4.73) becomes $a_{r+2} = \dfrac{2\,(1+r-\nu)}{(r+3)\,(r+2)}\,a_r$ ...(4.75)

Case I. When $k = 0$, putting $r = 0, 1, 2, 3,...$ in (4.74) we have

$$a_2 = -\frac{2}{\lfloor 2}\nu a_0;\quad a_3 = -\frac{2(\nu-1)}{\lfloor 3}a_1$$

$$a_4 = -\frac{2^2\nu(\nu-2)}{\lfloor 4}a_0;\quad a_5 = \frac{2^2\,(\nu-1)\,(\nu-3)}{\lfloor 5}a_1$$

....

$$a_{2r} = \frac{(-2)^r \nu(\nu-2)...(\nu-2r+2)}{\lfloor 2r} a_0; \; a_{2r+1} = \frac{(-2)^r (\nu-1)(\nu-3)...(\nu-2r+1)}{\lfloor 2r+1} a_1$$

Now if $a_1 = 0$, then $a_3 = a_5 = a_7 = a_{2r+1} =...= 0$.

But if $a_1 \neq 0$, then (4.70) gives for $k = 0$, $y = \sum_{r=0}^{\infty} a_r x^r$

i.e. $y = a_0 + a_1x + a_2x^2 + a_3x^3 + ...$

$$= a_0 + a_2x^2 + a_4x^4 +...+ a_1x + a_3x^3 + a_5x^5 + ...$$

$$= a_0\left[1 - \frac{2\nu}{\lfloor 2}x^2 + \frac{2^2\nu(\nu-2)}{\lfloor 4}x^4 - ... + (-1)^r \frac{2^r}{\lfloor 2r}\nu(\nu-2)...(\nu-2r+2)x^{2r} + \cdots\right]$$

$$+ a_1 x\left[1 - \frac{2(\nu-1)}{\lfloor 3}x^2 + \frac{2^2(\nu-1)(\nu-3)}{\lfloor 5} - \cdots\right.$$

$$\left. + (-1)^r \frac{2^r}{\lfloor 2r+1}(\nu-1)(\nu-3)...(\nu-2r+1)\, x^{2r} + \cdots\right] \quad ...(4.76)$$

$$= a_0\left[1 + \sum_{r=1}^{\infty} \frac{(-1)^r 2^r}{\lfloor 2r}\nu(\nu-2)...(\nu-2r+2)\, x^{2r}\right]$$

$$+ a_0\left[x + \sum_{r=1}^{\infty} \frac{(-1)^r 2^r}{2r+1}(\nu-1)(\nu-3)...(\nu-2r+2)\, x^{2r+1}\right] \quad ...(4.77)$$

Case II. When $k = 1$, then $a_1 = 0$ and so by putting $r = 0, 1, 2, 3,...$ in (4.75) we find

$$a_2 = -\frac{2(\nu-1)}{\lfloor 3} a_0$$

$$a_4 = \frac{2^2(\nu-1)(\nu-3)}{\lfloor 5} a_0$$

....

$$a_{2r} = (-1)^r \frac{2^2(\nu-1)(\nu-3)...(\nu-2r+1)}{\lfloor 2r+1} a_0$$

Hence the solution is

$$a_0 x\left[1 - \frac{2(\nu-1)}{\lfloor 3}x^2 + \frac{2^2(\nu-1)(\nu-3)}{\lfloor 5}x^4 ... + \frac{(-1)^r 2^r (\nu-1)(\nu-3)...(\nu-2r+1)}{\lfloor 2r+1}x^{2r} + \cdots\right.$$

...(4.78)

Clealry the solution (4.78) is included in the second part of (4.76) except that a_0 is replaced by a_1 and hence in order that the Hermite equation may have two independent solutions, a_1 must be zero, even if $k = 0$ and then (4.76) reduces to

$$y = a_0\left[1 - \frac{2\nu}{\lfloor 2}x^2 + \frac{2^2\nu(\nu-2)}{\lfloor 4}x^4 - \cdots\right.$$

$$+(-1)^r \frac{2^r}{\lfloor 2r} \nu(\nu-2)...(\nu-2r+2)\, x^{2r} + \cdots \Bigg] \qquad ...(4.79)$$

NOTES

The complete integral of (4.69) is then given by

$$y = A\left[1 - \frac{2\nu}{\lfloor 2} x^2 + \frac{2^2 \nu(\nu-2)}{\lfloor 4} x^4 - \cdots\right] + B\left[1 - \frac{2(\nu-1)}{\lfloor 3} x^2 + \frac{2^2 (\nu-1)(\nu-3)}{\lfloor 5} x^4 \cdots\right] \qquad ...(4.80)$$

where A and B are arbitrary constants.

4.4.1 Recurrence Relation and Generating Functions

Recurrence Formulae for $H_n(x)$ and to show that $H_n(x)$ is a solution of Hermite Equation

Hermite equation is $y'' - 2xy' + 2ny = 0$ for integral values taking $\nu = n$.

Also, $$e^{2tx-t^2} = \sum_{n=0}^{\infty} \frac{H_n(x)\, t^n}{\lfloor n}$$

I. Differentiating partially w.r.t. x, we have

$$2t \cdot e^{-2tx-t^2} = \sum_{n=0}^{\infty} H_n'(x) \frac{t^n}{\lfloor n}$$

i.e. $$2t \sum_{n=0}^{\infty} \frac{H_n(x)\, t^n}{\lfloor n} = \sum_{n=0}^{\infty} H_n'(x) \frac{t^n}{\lfloor n}$$

which yields on equating the coefficients $\frac{t^n}{\lfloor n}$ on either side,

$$2 \frac{\lfloor n}{\lfloor n-1} H_{n-1}(x) = H_n'(x)$$

i.e. $$2n\, H_{n-1}(x) = H_n'(x) \qquad ...(4.81)$$

II. Differentiating partially w.r.t. 't', we get

$$2(x-t)\, e^{2tx-t^2} = \sum_{n=1}^{\infty} H_n(x) \frac{t^{n-1}}{\lfloor n-1} \because n = 0 \text{ corresponds to the vanishing of R.H.S.}$$

or $$2x \sum_{n=0}^{\infty} H_n(x) \frac{t^n}{\lfloor n} - 2 \sum_{n=0}^{\infty} H_n(x) \frac{t^{n+1}}{\lfloor n} = \sum_{n=1}^{\infty} H_n(x) \frac{t^{n-1}}{\lfloor n-1}$$

Equating the coefficients of t^n on either side we find

$$2x \frac{H_n(x)}{\lfloor n} - 2 \frac{H_{n-1}(x)}{\lfloor n-1} = \frac{H_{n+1}(x)}{\lfloor n}$$

i.e. $$2x\, H_n(x) = 2nH_{n-1}(x) + H_{n+1}(x) \qquad ...(4.82)$$

III. Eliminating $H_{n-1}(x)$ from (4.81) and (4.82) we get

$$2xH_n(x) = H_n'(x) + H_{n+1}(x)$$

NOTES

or $$H'_n(x) = 2x\, H_n(x) - H_{n+1}(x) \qquad ...(4.83)$$

IV. Differentiating (4.83) w.r.t. x we find

$$H_n''(x) = 2x\, H_n'(x) + 2H_n(x) - H'_{n+1}(x)$$

Putting $H'_{n+1}(x) = 2(n+1)\, H_n(x)$ obtained from (4.81) on replacing n by $n + 1$: we have

$$H_n''(x) = 2xH_n'(x) + 2H_n(x) - 2(n+1)H_n(x)$$

or $$H_n''(x) - 2x H_n'(x) + 2n\, H_n(x) = 0 \qquad ...(4.84)$$

which clearly follows that $y = H_n(x)$ is a solution of Hermite equation.

We have $$e^{tx-t^2} = e^{2tx}\cdot e^{-t^2} = \sum_{n=0}^{\infty}\frac{(2tx)^n}{\lfloor n}\sum_{m=0}^{\infty}\frac{(-t^2)^n}{\lfloor m} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^m 2^n x^n t^{n+2m}}{\lfloor n \lfloor m}$$

$$= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^m 2^{2n-2m} x^{n-2m} t^n}{\lfloor (n-2m) \lfloor m}$$

(on replacing n by $n - 2m$)

Here $n \geq 0 \Rightarrow n - 2m \geq 0 \Rightarrow m \leq \frac{n}{2}$ so that for evern n, m varies from 0 to $\frac{\pi}{2}$ and for odd n, m varies from 0 to $\frac{n-1}{2}$, m being an integer. Also if $\left[\frac{n}{2}\right]$ denotes the greater integer $\frac{n}{2}$, then coefficient of t^n in

$$e^{2tx-t^2} = \sum_{m=0}^{[n/2]}(-1)^m \frac{1}{\lfloor m \lfloor (n-2m)}\cdot(2x)^{n-2m}$$

$$H_n(x) = \sum_{m=0}^{[n/2]}\frac{(-1)^m \lfloor n\,(2x)^{n-2m}}{\lfloor m \lfloor (n-2m)}$$

$$= \frac{1}{\lfloor n}\sum_{m=0}^{[n/2]}\frac{(-1)^m \lfloor n\,(2x)^{n-2m}}{\lfloor m \lfloor (n-2m)} = \frac{1}{\lfloor n} H_n(x)$$

Hence $$e^{2tx-t^2} = \sum_{n=0}^{\infty}\frac{H_n(x)}{\lfloor t} t^n . \qquad ...(4.84a)$$

Note 1. e^{2tx-t^2} is known as **Generating function of Hermite** Polynomial.

Hermite Functions

An equation closely related to Hermite equation is

$$\frac{d^2\psi}{dx^2} + (\lambda - x^2)\psi = 0 \qquad ...(4.85)$$

If we change the dependent variable ψ to y by the substitution

$$\psi = e^{-x^2/2} y \qquad ...(4.86)$$

NOTES

So that $\dfrac{d\psi}{dx} = e^{-x^2/2}\dfrac{dy}{dx} - y e^{-x^2/2}\cdot x$

and $\dfrac{d^2\psi}{dx^2} = e^{-x^2/2}\dfrac{d^2y}{dx^2} - 2x e^{-x^2/2}\dfrac{dy}{dx} - (e^{-x^2/2} - x^2 e^{-x^2/2})y$

We get from (4.69)

$$y'' - 2xy' + (\lambda - 1)y = 0 \qquad \text{...(4.87)}$$

If we put $\lambda - 1 = 2\nu$, then (4.87) reduces to Hermite equation *i.e.*

$$y'' - 2xy' + 2\nu y = 0$$

It therefore follows that the general solution of (4.85) is given by

$$\psi = e^{-x^2/2} y$$

where y is given by (4.80).

Thus if the parameter λ be of the form $1 + 2n$, n being a positive integer, then the solution of (4.85) will be a constant multiple of the function ψ_n defined by

$$\psi_n(x) = e^{-x^2/2} H_n(x) \qquad \text{...(4.88)}$$

where $H_n(x)$ is the Hermite polynomial of degree n.

Here the function $\psi_n(x)$ is said to be the **Weber Hermite function of order n.**

Recurrence Relations for $\psi_n(x)$:

Differentiating (4.88) w.r.t. x, we have

$$\psi_n'(x) = e^{-x^2/2} H_n'(x) - e^{-x^2/2} x H_n(x)$$

$$= 2ne^{-x^2/2} H_{n-1}(x) - xe^{-x^2/2} H_n(x)$$

$$[\because H'_n(x) = 2n H_{n-1}(x) \text{ by (4.81)}]$$

$$= 2n\,\psi_{n-1}(x) - x\,\psi_n(x) \qquad \text{[using (4.88)]}$$

$$\therefore 2n\,\psi_{n-1}(x) = x\,\psi_n(x) + \psi_n'(x) \qquad \text{... (4.89)}$$

Also from (24), $2x H_n(x) = 2n H_{n-1}(x) + H_{n+1}(x)$

which may be expressed by using (4.88), as

$$2xe^{-x^2/2} H_n(x) = 2ne^{-x^2/2}(x) + e^{-x^2/2} H_{n+1}(x)$$

i.e. $$2x\psi_n(x) = 2n\,\psi_{n-1}(x) + \psi_{n+1}(x) \qquad \text{...(4.90)}$$

Eliminating $2n\,\psi n{-}1\,(x)$ from (4.89) and (4.90) we find

$$x\,\psi_n(x) + \psi_n'(x) = 2x\,\psi_n(x) - \psi_{n+1}(x)$$

i.e. $$\psi_n'(x) = x\,\psi_n(x) - \psi_{n+1}(x) \qquad \text{...(4.91)}$$

4.4.2 Orthogonality

Orthogonal Properties of Hermite Polynomials.

Now since $H_n(x)$ is a solution of Hermite equation, we have

$$H_n''(x) - 2x\,H_n'(x) + 2n\,H_n'(x) = 0 \qquad \text{by ...(4.84)}$$

If we put $y = e^{-x^2/2}\ H_n(x)$ *i.e.* $H_n(x) = ye^{-x^2/2}$

So that $H'_n(x) = y'e^{x^2/2} + xy\, e^{x^2/2}$

NOTES

and $\quad H''_n(x) = y''x^{x^2/2} + 2xy'\, e^{x^2/2} + y(1+x^2)e^{x^2/2}$

then we get $y'' + (1 - x^2 + 2n)y = 0$ ...(4.92)

Since $y = e^{-x^2/2} H_n(x) = \psi_n(x)$ by (4.89), it therefore follows that $\psi_n(n)$ satisfies (4.92) and hence

$$\psi''_n + (2n + 1 - x^2)\ \psi_n = 0 \qquad ...(4.93)$$

for a function ψ_m, this relation is

$$y''_m + (2m + 1 - x^2)\ \psi_m = 0 \qquad ...(4.94)$$

Multiplying (31) by ψ_m; (32) by ψ_n and subtracting we get

$$2(m - n)\ \psi_m\psi_n = \psi_m\psi''_n - \psi_n\psi''_m$$

Integrating over $(-\infty, \infty)$, we have

$$2(m-n)\int_{-\infty}^{-\infty} \psi_m\ \psi_n\ dx = \int_{-\infty}^{\infty} (\psi_m\ \psi''_n - \psi_n\ \psi''_m)dx \qquad ...(4.95)$$

$$= [\psi_m\psi'_n - \psi_n\psi'_m]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} (\psi'_m\psi'_n - \psi'_n\psi'_m)\,dx$$

(on integrating by parts)

$$= 0\ \because\ \psi_n(x) \to 0\ |x| \to \infty$$

for all positive integral values of n.

or $\quad \int_{-\infty}^{\infty} \psi_m\psi_n\ dx = 0$ if $m \neq n$

Symbolically $I_{m,\,n} = \int_{-\infty}^{\infty} \psi_m\psi_n\ dx = \int_{-\infty}^{\infty} e^{-x^2}\ H_m(x)\, H_n(x)\, dx = 0$ when $m \neq n$

$2x\psi_n(x) = \qquad 2n\ \psi_{n-1}(x) + \psi_{n+1}(x)$

...(4.96)

In particular $\ L_{n-1,\ n+1} = 0$ ...(4.97)

Now from (4.90) we have $2x\psi_n(x) = 2n\psi_{n-1}(x) + \psi_{n+1}(x)$

$$\therefore \int_{-\infty}^{\infty} 2x\psi_n(x)\ \psi_{n-1}\, dx = 2n\int_{-\infty}^{\infty} \psi_{n-1}(x)\ \psi_{n-1}(x)\, dx$$

$$\left[\because\ \int_{-\infty}^{\infty} \psi_{n-1}\ \psi_{n+1}\ dx = 0 \text{ by }\right]$$

$$= 2n\ I_{n-1,\ n-1} \qquad ...\ (4.98)$$

Also $\psi_n(x) = e^{-x^2/2}H_n(x)$

$$= (-1)^n e^{x^2/2}\frac{d^n}{dx^n}(e^{-x^2})$$

Thus (4.98) gives

$$-\int_{-\infty}^{\infty} 2x\, e^{x^2/2}\frac{d^n}{dx^n}(e^{-x^2/2})\frac{d^{n-1}}{dx^{n-1}}(e^{-x^2})\, dx = 2n\ I_{n-1,\, n-1}$$

NOTES

or $$2nI_{n-1,\,n-1} = -\int_{-\infty}^{\infty} d(e^{x^2})\frac{d^n}{dx^n}(e^{-x^2})\frac{d^{n-1}}{dx^{n-1}}(e^{-x^2})dx$$

$$= \left[e^{-x^2}\frac{d^n}{dx^n}(e^{-x^2})\cdot\frac{d^{n-1}}{dx^{n-1}}(e^{-x^2})\right]_{-\infty}^{\infty}$$

$$+\int_{-\infty}^{\infty} e^{-x^2}\left\{\frac{d^n}{dx^n}(e^{-x^2})\frac{d^n}{dx^n}(e^{-x^2})+\frac{d^{n+1}}{dx^{n+1}}(e^{-x^2})\cdot\frac{d^{n-1}}{dx^{n-1}}(e^{-x^2})\right\}dx$$

(on integrating by parts)

$$= 0 + I_{n,\,n} + I_{n+1,\,n-1}$$

$$= I_{n,\,n} \qquad \text{by (4.97)}$$

$$\therefore \quad I_{n,\,n} = 2n\, I_{n-1,\,n-1} \qquad ...(4.99)$$

Applying (4.99), repeatedly we have

$$\begin{aligned} I_{n,\,n} &= 2n\, I_{n-1,\,n-1} = 2n\,(2n-1)\, I_{n-2,\,n-2} \\ &= 2^2 n\,(n-1)\cdot 2(n-2)\, I_{n-3,\,n-3} \\ &= 2^3 n\,(n-1)\,(n-2)\, I_{n-3,\,n-3} \\ &= \text{..... ....} \\ &= 2^n\, n(n-1)\,(n-2)....3.2.1.\, I_{0,\,0} \end{aligned}$$

where $I_{0,\,0} = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$ (*see* Beta and Gamma functions)

$$\therefore \quad I_{n,\,n} = 2^n \underline{|n}\sqrt{\pi} \qquad ...(4.100)$$

Combining the two results (4.96) and (4.100); we have in terms of Kronecker delta symbol

$$I_{m,\,n} = \int_{-\infty}^{\infty} e^{-x^2} H_m(x) H_n(x)\, dx = 2^n \underline{|n}\sqrt{\pi}\,\delta_{m,\,n.} \qquad ...(4.101)$$

where $\delta_{m,\,n} = 0$ when $m \neq n$.

$= 1$ when $m = n$.

(4.101) may also be written as

$$I_{m,\,n} = \int_{-\infty}^{\infty} \psi_m(x)\psi_n(x)\,dx = \int_{-\infty}^{\infty} e^{-x^2} H_m(x) H_n(x)\, dx$$

$$= 2^n \underline{|n}\sqrt{\pi}\,\delta_{m,\,n} \qquad ...(4.102)$$

Again $2x\,\psi_n(x) = 2n\,\psi_{n-1}(x)\,\psi_{n+1}(x)$ gives

$$\int_{-\infty}^{\infty} x\,\psi_m(x)\,\psi_n(x)\,dx = nI_{m,\,n-1} + \frac{1}{2} I_{m,\,n+1}$$

$$= 0 \qquad (\text{for } m \neq n = 1)$$

and $$\int_{-\infty}^{\infty} x\psi_n(x)\,\psi_{n+1}(x)\,dx = n\,I_{n+1,\,n-1} + \frac{1}{2} I_{n+1,\,n+1}$$

$$= \frac{1}{2} 2^{n+1} \underline{|n+1}\sqrt{\pi} \text{ as above}$$

$$= 2^n \underline{|(n+1)}\sqrt{\pi} \qquad (\text{for } m = n)$$

Hence $\int_{-\infty}^{\infty} x\,\psi_m(x)\,\psi_n(x)\,dx = 2^n \underline{|n+1} \sqrt{\pi}\,\delta_{m,n}$ ...(4.103)

Further $2n\,\psi_{n-1}(x) = x\,\psi_n(x) + \psi'_n(x)$ gives

NOTES

$$\int_{-\infty}^{\infty} \psi_m(x)\,\psi'_n(x)\,dx = 2n\int_{-\infty}^{\infty} \psi_m(x)\,\psi_{n-1}(x)\,dx - \int_{-\infty}^{\infty} x\psi_m(x)\,\psi_n(x)\,dx$$

$$= 0 \text{ if } m \neq n = 1$$

and $$= 2n\,I_{n-1,\,n-1} - 2^{n-1}\underline{|n}\sqrt{\pi} \text{ if } m = n = 1$$

$$= 2^n\underline{|n}\sqrt{\pi} - 2^{n-1}\underline{|n}\sqrt{\pi} = 2^{n-1}\underline{|n}\sqrt{\pi}$$

Hence $\int_{-\infty}^{\infty} \psi_m(x)\,\psi'_n(x)\,dx = 2^{n-1}\underline{|n}\sqrt{\pi}\,\delta_{m,n}$ ...(4.104)

In the last if we take $m = n + 1$, then

$$\int_{-\infty}^{\infty} \psi_m(x)\,\psi'_n(x)\,dx = 2n\int_{-\infty}^{\infty} \psi_{n+1}(x)\,\psi_{n-1}(x)\,dx - \int_{-\infty}^{\infty} x\,\psi_{n+1}(x)\,\psi_n(x)\,dx$$

$$= -2^n\,\underline{|n+1}\sqrt{\pi}.$$

Check Your Progress

7. What is the form of Hermite differential equation?
8. What is the generating function of Hermite polynomial?
9. What would be the equation if Hn(x) is a solution of hermite equation?

4.5 LAGUERRE FUNCTION

This equation is of the form

$$xy'' + (1 - x)\,y' + \nu y = 0 \qquad ...(4.105)$$

Dividing by x, it is observed that $x = 0$ is a regular singularity of (4.105) and hence it has a series solution. Let its series solution be

$$y = \sum_{r=0}^{\infty} a_r\,x^{k+r} \qquad a_0 \neq 0 \qquad ...(4.106)$$

$$\therefore \quad y' = \sum_{r=0}^{\infty} a_r\,(k+r)x^{k+r-1}$$

and $$y'' = \sum_{r=0}^{\infty} a_r\,(k+r)\,(k+r-1)\,x^{k+r-2}$$

Substituting these values in (4.105), we get the identity

$$\sum_{r=0}^{\infty}[(k+r)^2 x^{k+r-1} - (k+r-\nu)\,x^{k+r}]a_r \equiv 0 \qquad ...(4.107)$$

Equating to zero the coefficient of x^{k-1} (the first and the lowest term); we get

$$k^2 = 0 \quad i.e. \quad k = 0 \quad \text{as} \quad a_0 \neq 0 \qquad ...(4.108)$$

Again equating to zero the coefficient of x^{k+r} in (4.107), we get

$$(k + r + 1)^2\,a_{r+1} - (k + r - \nu)\,a_r = 0$$

which gives the recurrence relation

$$a_{r+1} = \frac{k+r-\nu}{(k+r+1)^2} a_r \qquad ...(4.109)$$

For $k = 0$, this yields, $a_{r+1} = \frac{r-\nu}{(r+1)^2} a_r$

$\therefore \quad a_1 = -\nu a_0 = (-1)\ \nu a_0$

$$a_2 = \frac{1-y}{2^2} = (-1)^2 \frac{\nu(\nu-1)}{(\lfloor 2)^2} a_0$$

Similarly $a_3 = (-1)^3 \frac{\nu(\nu-1)\,(\nu-2)}{(\lfloor 3)^2} a_0$

$$a_r = (-1)^r \frac{\nu(\nu-1)...(\nu-r+1)}{(\lfloor r)^2} a_0$$

Hence the solution is

$$y = a_0 \left[1 - \nu x + \frac{\nu(\nu-1)}{(\lfloor 2)^2} x^2 - \cdots + (-1)^r \frac{\nu(\nu-1)...(\nu-r+1)}{(\lfloor r)^2} x^r + \cdots\right] \qquad ...(4.110)$$

$$= a_0 \sum_{k=0}^{\infty} \frac{(-1)^k \lfloor \nu}{(\lfloor k)^2 \lfloor \nu - k} \qquad ...(4.111)$$

In case ν is a positive integer put $\nu = n$, so that Laguerre equation becomes

$xy'' + (1-x)y' + ny = 0$ for positive integeral n. ...(4.112)

When $\nu = n$ (a positive integer) and $a_0 = \lfloor n$ then solution for (4.112) is said to be the Laguerre polynomial of degree n and denoted by $L_n(x)$ *i.e.*

$$L_n(x) = (-1)^n \left[x^n - \frac{n^2}{\lfloor 1} x^{n-1} + \frac{n^2(n-1)^2}{\lfloor 2} x^{n-2} + \cdots + \cdots (-1)^n \lfloor n\right] \qquad ...(4.113)$$

Then the solution of Laguerre solution for ν to be a positive integer is

$$y = AL_n(x) \qquad ...(4.114)$$

From (4.113), it is easy to show that

$L_n(0) = \lfloor n$, $\quad L_2(x) = x^2 - 4x + 2$

$L_0(x) = 1$, $\quad L_3(x) = -x^3 + 9x^2 - 18x + 6$ ...(4.115)

$L_1(x) = 1 - x$, $L_4(x) = x^4 - 16x^3 + 7x^2 - 97x + 48$

Also $L_n(x)$ being the solution of (8), we should have

$$xL''_n + (1-x)\,L'_n(x) + nL_n(x) = 0 \qquad ...(4.116)$$

4.5.1 Generating Functions

Laguerre Polynomials with Their Representation in Terms of Confluent Hypergeometric Series

The Laguerre Polynomials $L_n(x)$ are given by the relation

$$(1-t)\sum_{n=0}^{\infty}\frac{L_n(x)}{\lfloor n}t^n = e^{\frac{-xt}{1-t}} \text{ or } \frac{1}{1-t}e^{-xt/(t-0)} = \sum_{n=0}^{\infty}\frac{L_n(x)t^n}{\lfloor n} \qquad ...(4.117)$$

NOTES

where n is a positive integer and x is a positive real number, and $\frac{1}{1-t}e^{-xt/(1-t)}$ is known as **Generating function** for Laguerre polynomial.

(4.117) can be written as

$$\sum_{n=0}^{\infty}\frac{L_n(x)t^n}{\lfloor n} = \frac{1}{1-t}e^{\frac{-xt}{1-t}}$$

$$= \frac{1}{1-t}\left[1-\frac{xt}{1-t}+\frac{x^2t^2}{\lfloor 2\,(1-t)^2}-\cdots+\frac{(-1)^k\,x^k\,t^k}{\lfloor k\,(1-t)^k}+\cdots\right]$$

$$= \sum_{k=0}^{\infty}-\frac{(-1)^k\,x^kt^k}{\lfloor k\,(1-t)^{k+1}}$$

$$= \sum_{k=0}^{\infty}-\frac{(-1)^k\,x^kt^k}{\lfloor k}(1-t)^{-(k+1)}$$

$$= \sum_{k=0}^{\infty}-\frac{(-1)^k\,x^k}{\lfloor k}t^k\left[1+(k+1)t+\frac{(k+1)(k+2)}{\lfloor 2}t^2+\cdots\right.$$

$$\left.\cdots+\frac{(k+1)(k+2)...(k+l)}{\lfloor l}t^l+\cdots\right]$$

$$= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{(-1)^k\,(k+1)}{\lfloor k\lfloor l}x^kt^{k+1} \text{ where } (k+1)_t = \frac{\Gamma(k+1+l)}{\Gamma(k+1)}$$

Equating the coefficients of t^n on either side (coefficient of t^n on R.H.S. being obtained by putting $l = n - k$), we get

$$\frac{L_n(x)}{\lfloor n} = \sum_{k=0}^{\infty}\frac{(-1)^k\,(k+1)_{n-k}}{\lfloor r\lfloor n-k}x^k \text{ or } L_n(x) = \sum_{r=0}^{n}\frac{(-1)^r}{\lfloor(n-r)}\left(\frac{\lfloor n}{\lfloor r}\right)^2 x^r \qquad ...(4.118)$$

Here $(k+1)_{n-k} = \frac{\Gamma(k+1+n-k)}{\Gamma(k+1)}+\frac{\Gamma(n+1)}{\Gamma(k+1)}=\frac{\lfloor n}{\lfloor k}$

$$\frac{(-1)^k}{\lfloor n-k} = \frac{(-1)^k\,n(n-1)...(n-k+1)}{\lfloor n}$$

$$= \frac{(-n)(-n+1)(-n+2)\cdots(-n+k-1)}{\lfloor n} = \frac{(-n)_k}{\lfloor n}$$

$\therefore$ (4.118) yields

$$L_n(x) = \lfloor n\sum_{k=0}^{\infty}\frac{(-n)_k}{\lfloor n}\cdot\frac{\lfloor n}{(\lfloor k)^2}x^k = \lfloor n\sum_{k=0}^{\infty}\frac{(-n)_k}{(\lfloor k)^2}x^k$$

$$= \lfloor n\left[1+\frac{(-n)}{\lfloor 1\,\lfloor 1}x+\frac{(-n)(-n+1)}{\lfloor 2\,\lfloor 2}x^2+\frac{(-n)(-n+1)(-n+2)}{\lfloor 3\,\lfloor 3}x^3+\cdots\right]$$

$$= \lfloor n \, F(-n, 1; x) \qquad ...(4.119)$$

From which it follows that $L_n(x)$ is a polynomial of degree n in x and that the coefficient of x^n is $(-1)^n$.

NOTES

4.5.2 Recurrence Relation

Recurrence formulae for Laguerre Polynomials

The generating function for Laguerre Polynomial is

$$e^{\frac{-xt}{1-t}} = (1-t)\sum_{n=0}^{\infty} \frac{L_n(x)}{\lfloor n} t^n \qquad \text{[by (4.117)]} \quad ...(4.120)$$

I. Differentiating w.r.t. 't' it gives

$$-\frac{x}{(1-t)^2} e^{-\frac{xt}{1-t}} = (1-t)\sum_{n=0}^{\infty} \frac{L_n(x)t^{n-1}}{\lfloor n-1} - \sum_{n=0}^{\infty} \frac{L_n(x)t^n}{\lfloor n}$$

Using (4.120), we have

$$-\frac{x}{(1-t)^2}(1-t)\sum_{n=0}^{\infty} \frac{L_n(x)\,t^n}{\lfloor n} = (1-t)\sum_{n=0}^{\infty} \frac{L_n(x)\,t^{n-1}}{\lfloor n-1} - \sum_{n=0}^{\infty} \frac{L_n(x)\,t^n}{\lfloor n}$$

or $$x\sum_{n=0}^{\infty} \frac{L_n(x)\,t^n}{\lfloor n} + (1-t)^2 \sum_{n=0}^{\infty} \frac{L_n(x)\,t^{n-1}}{\lfloor n-1} - (1-t)\sum_{n=0}^{\infty} \frac{L_n(x)t^n}{\lfloor n} = 0$$

Equating to zero the coefficient of t^n, we find

$$x\frac{L_n(x)}{\lfloor n} + \frac{L_{n+1}(x)}{\lfloor n} - 2\frac{L_n(x)}{\lfloor n-1} + \frac{L_{n-1}(x)}{\lfloor n-2} - \frac{L_n(x)}{\lfloor n} + \frac{L_{n-1}(x)}{\lfloor n-1} = 0$$

i.e. $$L_{n+1}(x) + (x - 2n - 1)\, L_n(x) + n^2 L_{n-1}(x) = 0 \qquad ...(4.121)$$

II. Again differentiating (4.120) w.r.t. x, we get

$$-\left(\frac{t}{1-t}\right)e^{\frac{-xt}{1-t}} = (1-t)\sum_{n=0}^{\infty} \frac{L_n'(x)t^n}{\lfloor n}$$

or $$-\left(\frac{t}{1-t}\right)(1-t)\sum_{n=0}^{\infty} \frac{L_n(x)}{\lfloor n} t^n = (1-t)\sum_{n=0}^{\infty} \frac{L_n'(x)}{\lfloor n} t^n \quad \text{[by (4.120)]}$$

or $$t\sum_{n=0}^{\infty} \frac{L_n(x)\,t^n}{\lfloor n} + (1-t)\sum_{n=0}^{\infty} \frac{L_n'(x)\,t^n}{\lfloor n} = 0$$

Equating to zero the coefficient of t^n, we get

$$\frac{L_n'(x)}{\lfloor n} - \frac{L_{n-1}'(x)}{\lfloor n-1} + \frac{L_n(x)}{\lfloor n-1} = 0$$

i.e. $$L_n'(x) - nL_{n-1}'(x) + nL_{n-1}(x) = 0 \qquad ...(4.122)$$

III. Now differentiating (4.121) w.r.t. x, we find

$$L_{n+1}'(x) - (x - 2n - 1)\, L_n'(x) + L_n(x) + n^2 L_{n-1}'(x) = 0.$$

Differentiating it again w.r.t. x, we have

$$L_{n+1}''(x) + (x - 2n - 1)\, L_n''(x) + 2L_n'(x) + n^2 L_{n-1}''(x) = 0$$

NOTES

Replacing n by $n + 1$, this yields

$$L''_{n+2}(x) + (x - 2n - 3)\, L''_{n-1}(x) + (n+1)^2 L''_n(x) + 2L'_{n+1}(x) = 0 \quad ...(a)$$

Whence from (4.122),

$$L'_n(x) = n\{L'_{n-1}(x) - L_{n-1}(x)\}$$

or $L'_{n+1}(x) = (n + 1)\ \{L'_n(x) - L_n(x)\}$ = (on replacing n by $n + 1$) ...(b)

$\therefore$ $L''_{n+1}(x) = (n + 1)\ \{L''_{n+1}(x) - L'_{n+1}(x)\}$ (on differentiating) ...(c)

or $L''_{n+2}(x) = (n + 2)\ \{L''_{n+1}(x) - L'_{n+1}(x)\}$ (on replacing n by $n + 1$)

Thus we have from (a)

$$(n+2)\ \{L''_{n+1}(x) - L'_{n+1}(x)\} + (x - 2n - 3)\ L''_{n+1}(x) + (n+1)^2\ L''_n(x) + 2L''_{n+1}(x) = 0$$

or $(x - n - 1)\ L''_{n+1}(x) - nL'_{n+1}(x) + (n+1)^2\ L''_n(x) = 0$

Eliminating $L''_{n+1}(x)$ and L'_{n+1} by (b) and (c), we get

$$xL''_n(x) + (1-x)\, L'_n(x) + n\, L_n(x) = 0 \quad ...(4.123)$$

which clearly shows that $y = AL_n(x)$ is a solution of Laguerre equation.

4.5.3 Orthogonality

Orthogonal Properties of Laugerre Polynomials

The Laguerre polynomials themselves do not form an orthogonal set since the Laguerre equation is not self-adjoint. We, therefore, introduce a function

$$\phi_n(x) = \frac{1}{\lfloor n} e^{-x/2} L_n(x) \quad ...(4.124)$$

and then will show that ϕ's form an orthogonal set *i.e.*

$$\int_0^\infty \phi_m(x)\, \phi_n(x)\, dx = \int_0^\infty e^{-x} \frac{L_m(x)}{\lfloor m} \frac{L_n(x)}{\lfloor n} dx = \delta_{m,\, n} \quad ...(4.125)$$

Over the interval $0 \le x \le \infty$, when $\delta_{m,\, n}$ $= 0$ for $m \neq n$.
$= 1$ for $m = n$.

Since $L_n(x) = e^x \dfrac{d^n}{dx^n}(x^n e^{-x})$, therefore, we have

$$\int_0^\infty e^{-x} x^m L_n(x)\, dx = \int_0^\infty e^{-x} x^m e^x \frac{d^n}{dx^n}(x^n e^{-x})\, dx$$

$$= \int_0^\infty x^m \frac{d^n}{dx^n}(x^n e^{-x})\, dx$$

Integrating the R.H.S. by parts, we get

$$\int_0^\infty e^{-x} x^m L_n(x)\, dx = \left[x^m \frac{d^{n-1}}{dx^{n-1}} x^n e^{-x} \right]_0^\infty - \int_0^\infty m x^{m-1} \frac{d^{n-1}}{dx^{n-1}}(x^n e^{-x})\, dx$$

$$= (-1)m \int_0^\infty x^{m-1} \frac{d^{n-1}}{dx^{n-1}}(x^n e^{-x})\, dx$$

$$= (-1)^2 m(m-1)\int_0^\infty x^{m-2} \frac{d^{n-2}}{dx^{n-2}}(x^n e^{-x})\, dx$$

$$= \text{...}$$

$$= (-1)^m \lfloor m \int_0^\infty \frac{d^{n-m}}{dx^{n-m}}(x^n e^{-x})\, dx$$

$$= 0 \text{ if } n > m.$$

Similarly, $\int_0^\infty e^{-x} x^n L_m(x)\, dx = 0$ for $m < n$.

Now $L_m(x)$ being a polynomial of degree m in x and and $L_n(x)$ that of degree n in x, we have

$$\int_0^\infty e^{-x} L_m(x)\, L_n(x)\, dx = 0 \text{ if } m \neq n.$$

i.e. $$\int_0^\infty e^{-x} \frac{L_m(x)\, L_n(x)}{\lfloor m \; \lfloor n} dx = 0 \text{ if } m \neq n. \qquad ...(4.126)$$

In case $m = n$, then the term of degree n in $L_n(x)$ is $(-1)^n x^n$,

$$\therefore \int_0^\infty e^{-x}\{L_n(x)\}^2\, dx = (-1)^n \int_0^\infty e^{-x} x^n L_n(x)\, dx$$

$$= (-1)^n \int_0^\infty e^{-x} x^n e^x \frac{d^n}{dx^n}(x^n e^{-x})\, dx$$

$$= (-1)^n \int_0^\infty x^n \frac{d^n}{dx^n}(x^n e^{-x})\, dx$$

$$= (-1)^{2n} \lfloor n \int_0^\infty x^n e^{-x}\, dx \quad \text{(on integrating by parts } n \text{ times)}$$

$$= (\lfloor n)^2$$

or $$\int_0^\infty e^{-x/2} \frac{L_n(x)}{\lfloor n} e^{-x/2} \frac{L_n(x)}{\lfloor n} dx = 1 \qquad ...(4.127)$$

Combining (4.126) and (4.127) we have

$$\int_0^\infty \phi_m(x)\, \phi_n(x)\, dx = \int_0^\infty e^{-x/2} \frac{L_m(x)}{\lfloor m} e^{-x/2} \frac{L_n(x)}{\lfloor n} dx = \delta_{m,n}$$

Note. $\phi_n(x)$ satisfies the equation

$$x\phi_n''(x) + \phi_n'(x) + \left(n + \frac{1}{2} - \frac{x}{4}\right)\phi_n(x) = 0$$

4.6 SPHERICAL HARMONICS

One of the varieties of special functions which are encountered in the solution of physical problems is the class of functions called spherical harmonics.

NOTES

Table 4.1

ℓ	m_ℓ	$Y_{\ell m_l}(\theta,\phi) = \Theta_{\ell m_l}(\theta)\Phi_{m_l}(\phi)$
0	0	$(1/4\pi)^{1/2}$
1	0	$(3/4\pi)^{1/2}\cos\theta$
1	± 1	$\mp(3/8\pi)^{1/2}\sin\theta e^{\pm i\phi}$
2	0	$(5/16\pi)^{1/2}(3\cos^2\theta - 1)$
2	± 1	$\mp(15/8\pi)^{1/2}\sin\theta\cos\theta e^{\pm i\phi}$
2	± 2	$(15/32\pi)^{1/2}\sin^2\theta e^{\pm 2i\phi}$

$$\Phi_{m_\ell}(\phi) = \frac{1}{\sqrt{2\pi}} e^{im_\ell \phi}$$

$$\Theta_{\ell m_\ell}(\theta) = \left[\frac{2\ell+1}{2}\frac{(\ell-m_\ell)!}{(\ell+m_\ell)!}\right]^{1/2} P_\ell^{m_\ell}(\theta)$$

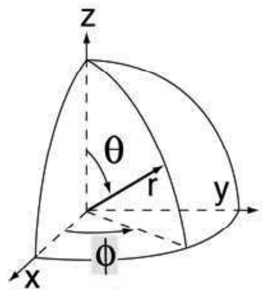

Fig. 4.1

The functions in this table are placed in the form appropriate for the solution of the Schrodinger equation for the spherical potential well, but occur in other physical problems as well. The dependence upon the colatitude angle q in spherical polar coordinates is a modified form of the associated Legendre functions.

Spherical Potential Well

The idealized infinite-walled one-dimensional and three-dimensional square-well potentials can be solved by the Schrodinger equation to give quantized energy levels. For the case of a nucleus, a useful idealization is an infinite-walled spherical potential. That is, we model the nucleus with a potential which is zero inside the nuclear radius and infinite outside that radius.

In spherical polar coordinates, the Schrodinger equation is separable in the general form $\Psi(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi)$, as it is in the case of the hydrogen atom solution. In this case with zero potential, the separation of the azimuthal (ö) and colatitude (è) equations requires

NOTES

$$\frac{d^2\Phi}{d\phi^2} + m_\ell^2\Phi = 0 \quad \textit{with solution} \quad \Phi_{m_\ell}(\phi) = \frac{1}{\sqrt{2\pi}}e^{i\,m_\ell\phi}$$

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left[\sin\theta\frac{d\Theta}{d\theta}\right] + \left[\ell(\ell+1) - \frac{m_\ell^2}{\sin^2\theta}\right]\Theta = 0$$

$$\ell = 0,1,2,3... \qquad m_\ell = 0,\pm1,\pm2,...,\pm\ell$$

The solutions for Θ and Φ, when normalized, give a standard set of functions called spherical harmonics.

$$\Theta_{\ell,m_\ell}(\theta)\Phi_{m_\ell}(\phi) = Y_{\ell,m_\ell}(\theta,\phi)$$

The radial equation is

$$\frac{-\hbar^2}{2m}\left[\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr}\right] + \left[V(r) + \frac{\ell(\ell+1)\hbar^2}{2mr^2}\right]R = ER$$

and the solution of this equation can be expressed in terms of another set of functions called spherical Bessel functions.

4.7 SERIES SOLUTIONS OF HERMITE AND LAGUERRE POLYNOMIALS

Let us first consider the Hermite polynomials.

Hermite Polynomials

The Hermite polynomials play an important role in problems involving Laplace's equations in cylinderical coordinates, in various problems in quantum mechanics and in probability theory.

The Hermite polynomials are polynomial solutions to **Hermite's equation**

$$y'' - 2xy' + 2n\,y = 0 \qquad (4.128)$$

where n is a constant.

Using the power series method it can be easily shown that $y(x) = a_0y_1(x) + a_1y_2(x)$ is the general solution of (4.128), where

$$y_1(x) = 1 - \frac{2n}{2!}x^2 + 2^2n\frac{(n-2)}{4!}x^4 - \frac{2^3n(n-2)\,(n-4)}{6!}x^6 + \cdots \qquad (4.129)$$

NOTES

And

$$y_2(x) = x - \frac{2(n-1)}{3!}x^3 + \frac{2^2(n-1)(n-3)}{5!}x^5 - \frac{2^3(n-1)(n-3)(n-5)}{7!}x^7 + \cdots$$

(4.130)

and both the series converge for all x .

If n is a non-negative integer, then one of these series terminates and is thus a polynomial- y (x) 1 if n is even, and y (x) 2 if n is odd, while the other remains an infinite series. It can be easily verified that for n = 0,1, 2,3, 4, 5, these polynomials are

$$1, x, 1 - 2x^2, x - \frac{2}{3}x^3, 1 - 4x^2 + \frac{4}{3}x^4, x - \frac{4}{3}x^3 + \frac{4}{15}x^5,$$

respectively

The polynomial solutions of Hermite's Eqn.(63) are constant multiples of these polynomials. The constant multiples with the property that the terms containing the highest powers of x are of the form $2^n x^n$ are denoted by H_n (x) and called the **Hermite polynomials.**

We define the Hermite polynomials H_n (x) by means of the relation

$$e^{(2xt-t^2)} = \sum_{n=0}^{\infty} H_n(x)\frac{t^n}{n!}, \quad \text{for all finite } x \text{ and } t \qquad (4.131)$$

The function on the left hand side is called the generating function of the Hermite polynomials. This definition of Hermite polynomials is used often in statistical applications.

The Hermite polynomials H_n (x) satisfy the following orthogonality property

$$\int_{-\infty}^{\infty} e^{-x^2} H_n(x)\, H_m(x)\,dx = \begin{cases} 0, & \text{if } m \neq n \\ 2^n n!\sqrt{\pi}, & \text{if } m = n \end{cases} \qquad (4.132)$$

Laguerre Polynomials

The Laguerre polynomials L (x) n are polynomial solutions to Laguerre's equation

$$xy'' + (1 - x)\, y' + ny = 0 \qquad (4.133)$$

where n is a constant. These polynomial solutions are obtained when n is a nonnegative integer. An important application involving Laguerre polnomials in quantum mechanics is to find the wave function associated with the electron in a hydrogen atom.

We define the Laguerre polynomials L_n (x) by means of the relations

$$(1-t)^{-1} e^{\frac{-xt}{1-t}} = \sum_{n=0}^{\infty} L_n(x)\, t^n, \; |t| < 1, \; 0 \le x \le \infty \qquad (4.134)$$

NOTES

The function on the left hand side is called the generating function of the Laguerre polynomials.

The Rodrigue's formula for Laguerre polynomials $L_n(x)$ is given by

$$L_n(x) = \frac{e^x}{n!}\frac{d^n}{dx^n}(x^n e^{-x}), \quad n = 0, 1, 2, \ldots \tag{4.135}$$

The Laguerre polynomials $L_n(x)$ satisfy the orthogonality property

$$\int_0^\infty e^{-x} L_n(x) L_m(x)\, dx = 0, \text{ if } m \neq n \tag{4.136}$$

4.8 ASSOCIATED LAGUERRE POLYNOMIALS

Solutions to the associated Laguerre differential equation with $\nu \neq 0$ and k an integer are called associated Laguerre polynomials $L_n^k(x)$. Associated Laguerre polynomials are implemented in the Wolfram Language as LaguerreL[*n*, *k*, *x*]. In terms of the unassociated Laguerre polynomials, $L_n(x) = L_n^0(x)$.

The Rodrigues representation for the associated Laguerre polynomials is

$$\begin{aligned} L_n^k(x) &= \frac{e^x x^{-k}}{n!}\frac{d^n}{dx^n}\left(e^{-x} x^{n+k}\right) \\ &= (-1)^k \frac{d^k}{dx^k}[L_{n+k}(x)] \\ &= \frac{(-1)^n x^{-(k+1)/2}}{n!} e^{x/2} W_{k/2+n+1/2,k/2}(x) \\ &= \sum_{m=0}^{n} (-1)^m \frac{(n+k)!}{(n-m)!\,(k+m)!\,m!} x^m, \end{aligned}$$

where

$W_{k,m}(x)$ is a Whittaker function.

The associated Laguerre polynomials are a Sheffer sequence with

$$\begin{aligned} g(t) &= (1-t)^{-k-1} \\ f(t) &= \frac{t}{t-1}, \end{aligned}$$

giving the generating function

$$\begin{aligned} g(x, z) &= \frac{\exp\left(-\frac{xz}{1-z}\right)}{(1-z)^{k+1}} \\ &= 1 + (k+1-x)\, z\, \tfrac{1}{2}\left[x^2 - 2(k+2)x + (k+1)(k+2)\right] z^2 + \ldots \end{aligned}$$

where the usual factor of in the denominator has been suppressed. Many interesting properties of the associated Laguerre polynomials follow from the fact that $f^{-1}(t) = f(t)$.

NOTES

The associated Laguerre polynomials are given explicitly by the formula

$$L_n^k(x) = \frac{1}{n!}\sum_{i=0}^{n}\frac{n!}{i!}\binom{k+n}{n-i}(-x)^i,$$

where $\binom{n}{k}$ is a binomial coefficient, and have Sheffer identity

$$n!\,L_n^k(x+y) = \sum_{i=0}^{n}\binom{n}{i} i!\,L_i^k(x)\,(n-i)!\,L_{n-i}^{-1}(y)$$

The associated Laguerre polynomials are orthogonal over $[0, \infty)$ with respect to the weighting function $x^k e^{-x}$,

$$\int_0^\infty e^{-x} x^k L_n^k(x) L_m^k(x)\,dx = \frac{(n+k)!}{n!}\delta_{mn},$$

where δ_{mn} is the Kronecker delta. They also satisfy

$$\int_0^\infty e^{-x} x^{k+1}\left[L_n^k(x)\right]^2 dx = \frac{(n+k)!}{n!}(2n+k+1).$$

$$\begin{aligned}
L_0^k(x) &= 1\\
L_1^k(x) &= -x+k+1\\
L_2^k(x) &= \tfrac{1}{2}\left[x^2 - 2(k+2)x + (k+1)(k+2)\right]\\
L_3^k(x) &= \tfrac{1}{6}\left[-x^3 + 3(k+3)x^2 - 3(k+2)(k+3)x + (k+1)(k+2)(k+3\right.
\end{aligned}$$

A generalization of the associated Laguerre polynomial to k not necessarily an integer is called a Laguerre function or a generalized Laguerre function.

4.9 HYPERGEOMETRIC FUNCTION

[A] 'Hypergeometric or Gauss' Differential Equation

This equation is of the form

$$x(1-x)\frac{d^2y}{dx^2} + \{\gamma - (\alpha+\beta+1)x\}\frac{dy}{dx} - \alpha\beta\, y = 0 \qquad ...(4.137)$$

where α, β, γ are parametric constants.

Here $x = 0$, $x = 1$ and $x = \infty$ are the singularities, since on dividing (4.137) by $x(x-1)$ we observe that coefficients of $\frac{dy}{dx}$ and y become infinite when $x =$ 0, 1 or ∞. Thus we can integrate (4.137) in series about $x = 0$ or $x = 1$ or $x =$ ∞. We therefore discuss the series integration in three cases.

Case [a$_1$]. When $x = 0$, then taking the series solution of (4.137) as

$$y = \sum_{r=0}^{\infty} a_r\, x^{k+r} \qquad ...(4.138)$$

We already have a solution as

$$y = AF\,(\alpha, \beta, \gamma, x) + Bx^{1-\gamma}\, F(\alpha + 1 - \gamma, \beta + 1 - \gamma, 2 - \gamma, x) \qquad ...(4.139)$$

where A and B are arbitrary constants, and

$$F(\alpha, \beta, \gamma, x) = 1 + \frac{\alpha \cdot \beta}{1 \cdot \gamma} x + \frac{\alpha(\alpha+1)\,(\beta+1)}{1\cdot 2\cdot \gamma \cdot (\gamma+1)} x^2 + \frac{\alpha(\alpha+1)\,(\alpha+2)\beta\,(\beta+1)\,(\beta+2)}{1\cdot 2\cdot 3\cdot \gamma\,(\gamma+1)\,(\gamma+2)} x^3 + \cdots$$

Case [$\mathbf{a_2}$]. When $x = 1$, is the singularity, then the series solution is obtained by developing the series about $x = 1$, by making a substitution

$$X = 1 - x \text{ in } (4.137) \qquad ...(4.140)$$

$$\therefore \quad \frac{dy}{dx} = \frac{dy}{dX}\frac{dX}{dx} = -\frac{dy}{dX}, \qquad \text{since by (1) } \frac{dX}{dx} = -1$$

and $$\frac{d^2y}{dx^2} = \frac{d}{dX}\left(-\frac{dy}{dX}\right)\frac{dX}{dx} = \frac{d^2y}{dX^2}$$

Substituting these values in (4.137) and arranging, we get

$$(X^2 - X)\frac{d^2y}{dX^2} + [(1+\alpha+\beta)X + (\gamma - 1 - \alpha - \beta)]\frac{dy}{dX} + \alpha\beta\, y = 0. \qquad ...(4.141)$$

which is similar to (4.137) except that γ is replaced by $1 + \alpha + \beta - \gamma$ and x by $1 - x$ and hence by (4.139) the solution in this case becomes,

$$y = AF\,(\alpha, \beta, 1 + \alpha + \beta - \gamma, 1 - x) + B(1 - x)^{\gamma-\alpha-\beta}$$
$$F(\gamma - \alpha, \gamma - \beta, \gamma - \alpha - \beta + 1, 1 - x). \qquad ...(4.142)$$

Case [$\mathbf{a_3}$]. When $x = \infty$ gives a singularity, then the series solution of (4.137) is obtained by developing the series about $x = \infty$, by making a substitution.

$$x = \frac{1}{X} \text{ in } (4.137) \qquad ...(4.143)$$

So that $$\frac{dy}{dx} = \frac{dy}{dX}\frac{dX}{dx} = -\frac{1}{x^2}\frac{dy}{dX} = -X^2\frac{dy}{dX} \text{ etc.}$$

$\therefore$ (4.137) becomes

$$X^2(1 - X)\frac{d^2y}{dX^2} + [2X(1-X) - (1+\alpha+\beta)X + \gamma X^2]\frac{dy}{dX} + \alpha\beta\, y = 0 \qquad ...(4.144)$$

Let its series solution be

$$y = \sum_{r=0}^{\infty} a_r X^{k+r}, \quad a_0 \neq 0 \qquad ...(4.145)$$

So that $$\frac{dy}{dX} = \sum_{r=0}^{\infty} a_r\,(k+r)\,X^{k+r-1}$$

$$\frac{d^2y}{dX^2} = \sum_{r=0}^{\infty} a_r\,(k+r)\,(k+r-1)\,X^{k+r-2}$$

Substituting these values in (4.144) we get the identity

$$\sum_{r=0}^{\infty} [\{(k+r)(k+r-1)+(1-\alpha-\beta)(k+r)+\alpha\beta\}X^{k+r}$$

NOTES

$- \{(k + r)(k + r - 1) + (2 - \gamma)(k + r)\} X^{k+r+1}] a_r \equiv 0$ ...(4.146)

Equating to zero the coefficient of X^k (the first term) in (10), we get

$[k(k - 1) + (1 - \alpha - \beta) k + \alpha\beta] a_0 = 0$

$\because a_0 \neq 0, \therefore k(k - 1) + (1 - \alpha - \beta) k + \alpha\beta = 0$

or $k^2 - (\alpha + \beta) k + \alpha\beta = 0$

or $(k - \alpha)(k - \beta) = 0$ giving $k = \alpha, \beta$ ...(4.147)

Again equating to zero the coefficient of x^{k+r} in (4.146), we find the recurrence relation as

$\{k + r)(k + r - 1) + (1 - \alpha - \beta)(k + r) + \alpha\beta\} a_r - \{(k + r - 1)(k + r - 2)$

$+ (2 - \gamma)(k + r - 1)\} a_{r-1} = 0$

i.e. $(k + r - \alpha)(k + r - \beta) a_r - (k + r - 1)(k + r - \gamma) a_{r-1} = 0$

or $$a_r = \frac{(k+r-1)(k+r-\gamma)}{(k+r-\alpha)(k+r-\beta)} a_{r-1}$$ ...(4.148)

when $k = \alpha$, (4.146) gives $a_r = \dfrac{(\alpha+r-1)(\alpha+r-\gamma)}{r(\alpha+r-\beta)} a_{r-1}$

So that $$a_1 = \frac{\alpha(\alpha+1-\gamma)}{1\cdot(\alpha+1-\beta)} a_0$$

$$a_2 = \frac{\alpha(\alpha+1)\cdot(\alpha+1-\gamma)(\alpha+2-\gamma)}{1\cdot 2\cdot(\alpha+1-\beta)(\alpha+2-\beta)} a_0,$$

$$a_3 = \frac{\alpha(\alpha+1)(\alpha+2)\cdot(\alpha+1-\gamma)(\alpha+2-\gamma)(\alpha+3-\gamma)}{1\cdot 2\cdot(\alpha+1-\beta)(\alpha+2-\beta)(\alpha+3-\gamma)} a_0 \text{ etc.}$$

$\therefore$ Solution is $y = a_0 X^{\alpha} F(\alpha, \alpha + 1 - \gamma, \alpha + 1 - \beta, X)$

$$= a_0 x^{-\alpha} F\left(\alpha, \alpha+1-\gamma, \alpha+1-\beta, \frac{1}{x}\right)$$ ...(4.149)

and similarly when $k = \beta$, the solution is

$$y = a_0 x^{-\beta} F\left[\beta, \beta+1-\gamma, \beta+1-\alpha, \frac{1}{x}\right]$$ ...(4.150)

Hence the complete integral is

$$y = Ax^{-\alpha} F(\alpha, \alpha+1-\gamma, \alpha+1-\beta, \frac{1}{x} + Bx^{-\beta} F\left[\beta, \beta+1-\gamma, \beta+1-\alpha, \frac{1}{x}\right]$$ ...(4.151)

where A and B are arbitrary constants.

We thus have found all the three possible solutions of Hypergeometric or Gauss equation, *i.e.*,

(*i*) for $x = 0$, exponents are $0, 1 - \gamma$ by (4.193)

(*ii*) for $x = \infty$, exponents are $-\alpha, -\beta$ by (4.49) (*iii*) for $x = 1$, exponents are $0, \gamma - \alpha - \beta$ by (4.142).

These results may be shown by a scheme as follows:

$$y = P\begin{bmatrix} 0 & \infty & 1 & \\ 0 & \alpha & 0 & x \\ 1-\gamma & \beta & \gamma-\alpha-\beta & \end{bmatrix} \quad ...(4.52)$$

NOTES

where the R.H.S. is said to be the **Reimann-P Function of the equation.**

In symbolic form, $F(\alpha, \beta; \gamma; x) = F\begin{bmatrix} \alpha, & \beta & \\ & & x \\ & \gamma & \end{bmatrix}$ ...(4.153)

which is known as **Hypergeometric function**.

We also denote $F(\alpha, \beta; \gamma; x)$ by $\sum_{k=0}^{\infty} \frac{(\alpha)_k (\beta)_k}{(\gamma)_k \lfloor k} x^k$ ...(4.154)

[B] Paritcular Cases of Hypergeometric Series

(*i*) we have $(1 + x)^n = 1 + nx + \frac{n(n-1)}{1 \cdot 2} x^2 + \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3} x^3 + \cdots$

$$= F(-n, 1, 1, -x)$$

(*ii*) $\log(1 + x) = x - \frac{1}{2}x^2 + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$

$$= x\left[1 - \frac{1}{2}x + \frac{1}{3}x^2 - \frac{1}{4}x^3 + \cdots\right]$$

$$= x\, F(1, 1, 2, -x)$$

(*iii*) $\sin^{-1} x = x + \frac{x^3}{\lfloor 3} + \frac{3^2 \cdot x^5}{\lfloor 5} + \frac{3^2 \cdot 5^2}{\lfloor 7} x^7 + \cdots$

$$= x\left[1 + \frac{\frac{1}{2} \cdot \frac{1}{2}}{1 \cdot \frac{3}{2}} x^2 + \frac{\frac{1}{2} \cdot \frac{3}{2} \cdot \frac{1}{2} \cdot \frac{3}{2}}{1 \cdot 2 \cdot \frac{3}{2} \cdot \frac{5}{2}} x^4 + \cdots\right]$$

$$= xF\left[\frac{1}{2}, \frac{1}{2}; \frac{3}{2}; x^2\right]$$

(*iv*) $\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$

$$= x\left[1 - \frac{x^2}{3} + \frac{x^4}{5} - \frac{x^6}{7} + \cdots\right]$$

$$= x\left[1 + \frac{1 \cdot \frac{1}{2}}{1 \cdot \frac{3}{2}}(-x^2) + \frac{1 \cdot 2 \cdot \frac{1}{2} \cdot \frac{3}{2}}{1 \cdot 2 \cdot \frac{3}{2} \cdot \frac{5}{2}}(-x^2)^2 + \cdots\right]$$

$$= xF\left[1, \frac{1}{2}; \frac{3}{2}; -x^2\right]$$

NOTES

(v) $e^x = 1 + x + \frac{x^2}{\lfloor 2} + \frac{x^3}{\lfloor 3} + \cdots$

$$= \lim_{n\to\infty}\left\{1 + \frac{n\cdot 1\, x}{1\cdot 1\, n} + \frac{n\,(n+1)\,1\cdot 2}{1\cdot 2\cdot 1\cdot 2}\left[\frac{x}{n}\right]^2 + \cdots\right\}$$

$$= \lim_{n\to\infty} F\left[n, 1; 1; \frac{x}{n}\right].$$

[C] Simple Properties of Hypergeometric Function

(1) Symmetry property. The value of a hypergeometric function does not change by the interchange of parameters α and β.

$$\because \quad F(\alpha, \beta; \gamma; x) = \sum_{k=0}^{\infty} \frac{(\alpha)_k\,(\beta)_k}{(\gamma)_k\,\lfloor k} x^k$$

$$= \sum_{k=0}^{\infty} \frac{(\beta)_k\,(\alpha)_k}{(\gamma)_k\,\lfloor k} x^k$$

$$= F(\beta, \alpha; \gamma; x)$$

$$\therefore \quad F\begin{bmatrix}\alpha, & \beta & \\ & & x \\ & \gamma & \end{bmatrix} = F\begin{bmatrix}\beta, & \alpha & \\ & & x \\ & \gamma & \end{bmatrix} \qquad \text{...(4.155)}$$

(2) Differentiation of Hypergeometric Functions. We have

$$F(\alpha, \beta, \gamma, x) = \sum_{k=0}^{\infty} \frac{(\alpha)_k\,(\beta)_k}{(\gamma)_k\,\lfloor k} x^k$$

Its differentiation w.r.t. x, yields,

$$\frac{d}{dx}(F(\alpha, \beta, \gamma, x)) = \sum_{k=0}^{\infty} k\,\frac{(\alpha)_k(\beta)_k}{k\,\lfloor k-1\,(\gamma)_k}\,x^{k-1} = \sum_{k=1}^{\infty} \frac{(\alpha)_k\,(\beta)_k}{\lfloor k-1\,(\gamma)_k} x^{k-1}$$

($\because$ it vanishes for $k = 0$)

$$= \sum_{p=0}^{\infty} \frac{(\alpha)_{p+1}\,(\beta)_{p+1}}{\lfloor p\,(\gamma)_{p+1}} \qquad \text{(on putting } k - 1 = p\text{)}$$

$$= \sum_{p=0}^{\infty} \frac{\alpha\cdot(\alpha+1)_p\cdot\beta\,(\beta+1)_p}{\lfloor p\,\gamma(\gamma+1)_p} \qquad \left[\begin{array}{ll}\because (\alpha)_{p+1} & = \alpha(\alpha+1)\,(\alpha+2)...(\alpha+p). \\ & = \alpha[(\alpha+1)\,(\alpha+2),...(\alpha+1+p-1)] \\ & = \alpha(\alpha+1)_p \text{ etc.}\end{array}\right]$$

$$= \frac{\alpha\beta}{\gamma}\sum_{p=0}^{\infty} \frac{(\alpha+1)_p\,(\beta+1)_p}{\lfloor p\,(\gamma+1)_p}$$

$$= \frac{\alpha\beta}{\gamma}\, F(\alpha + 1; \beta + 1; \gamma + 1; x) \qquad \text{...(4.156)}$$

NOTES

$$\frac{d^2}{dx^2}F(\alpha, \beta; \gamma; x) = \frac{d}{dx}\left[\frac{\alpha\beta}{\gamma}F(\alpha+1, \beta+1, \gamma+1, x)\right]$$

$$= \frac{\alpha\beta}{\gamma}\frac{d}{dx}F(\alpha+1, \beta+1, \gamma+1, x)$$

$$= \frac{\alpha\beta}{\gamma}\cdot\frac{(\alpha+1)(\beta+1)}{\gamma+1}F(\alpha+1+1, \beta+1+1, \gamma+1+1, x)$$

(by applying (4.156))

$$= \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)}F(\alpha+2, \beta+2, \gamma+2, x) \quad ...(4.157)$$

Repeating the process m times, we may have

$$\frac{d^m}{dx^m}F(\alpha, \beta; \gamma; x) = \frac{\alpha(\alpha+1)...(\alpha+m-1)\cdot\beta(\beta+1)...(\beta+m-1)}{\gamma(\gamma+1)...(\gamma+m-1)}$$

$F(\alpha + m, \beta + m, \gamma + m, x)$

$$= \frac{(\alpha)_m(\beta)_m}{(\gamma)_m}F(\alpha + m, \beta + m, \gamma + m, x) \quad ...(4.158)$$

In symbolic form (4.158) can be written as

$$\frac{d^m}{dx^m}F\begin{bmatrix}\alpha, & \beta & \\ & & x\\ & \gamma & \end{bmatrix} = \frac{(\alpha)_m(\beta)_m}{(\gamma)_m}\begin{bmatrix}\alpha+m, & \beta+m & \\ & & x\\ & \gamma+m & \end{bmatrix}$$

Corollary 1. When $x = 0$,

$$F(\alpha, \beta; \gamma; 0) = \lim_{x\to 0}\sum_{k=0}^{\infty}\frac{(\alpha)_k(\beta)_k}{\lfloor k\,(\gamma)_k}x^k$$

$$= \lim_{x\to 0}\left[1+\frac{\alpha\cdot\beta}{\gamma}x+\frac{\alpha(\alpha+1)\beta(\beta+1)}{1\cdot 2\cdot\gamma(\gamma+1)}x^2+\cdots\right]$$

= 1, since all the terms except the first vanish.

Similarly $F(\alpha + 1, \beta + 1; \gamma + 1; 0) = 1$

Hence from (4.154), it follows that

$$\left[\frac{d}{dx}F(\alpha, \beta; \gamma; x)\right]_{x=0} = \frac{\alpha\beta}{\gamma}. \quad ...(4.159)$$

Corollary 2. Had the parameter α been a negative integer say $-N$, then we should have

$$F(-N, B; \gamma; x) = \sum_{k=0}^{N}\frac{(-N)_k(\beta)_k}{\lfloor k\,(\gamma)_k}x^k \quad ...(4.160)$$

since it vanishes for $k = N, N + 1, N + 2,$...etc.

Similarly if β were a negative integer say $-M$, then

$$F(\alpha, -M; \gamma; x) = \sum_{k=0}^{M}\frac{(\alpha)_k(-M)_k}{\lfloor k\,(\gamma)_k}x^k \quad ...(4.161)$$

In case $\alpha = -N$ and $\gamma = -(N + M)$, N, M being positive integers, we have

$$F(-N, \beta, -N-M, x) = \sum_{k=0}^{\infty}\frac{(-N)_k\beta_k}{(-N-M)_k\lfloor k}x^k \quad ...(4.162)$$

NOTES

where $(-N)_k = (-N)(-N+1)\ldots(N+k-1) = (-1)^k \frac{\lfloor N}{\lfloor N-k}$.

Similarly $(-N-M)_k = (-1)^k \frac{\lfloor N+M}{\lfloor N+M-k}$

So that $\frac{(-N)_k}{(-N-M)_k} = \frac{\lfloor N}{\lfloor N-k} \times \frac{\lfloor N+M-k}{\lfloor N+M} = \frac{\lfloor N}{\lfloor N+M}$
$[(N+M-k)(N+M-k-1)\ldots(N-k+1)]$

$$= \left[\left(1-\frac{k}{N+M}\right)\left(1-\frac{k}{N+M-1}\right)\ldots\left(1-\frac{k}{N+1}\right)\right]$$

Hence (4.162) yields

$$F(-N, \beta; -N-M; x) = \sum_{k=0}^{\infty}\left[\left(1-\frac{k}{N+M}\right)\left(1-\frac{k}{N+M-1}\right)\ldots\left(1-\frac{k}{N+1}\right)\right]\frac{(\beta)_k}{\lfloor k} x^k \qquad \ldots(4.163)$$

Here it is notable that on the R.H.S. the terms do not vanish for $k = 0, 1, 2, \ldots N$ and they vanish for $k = N+1, N+2, \ldots, N+M$; but the terms do not vanish for $k = N+M+1, N+M+2, \ldots$, since the term corresponding to $k = N+M+1$ is

$$\left(1-\frac{N+M+1}{N+M}\right)\ldots\left(1-\frac{N+M+1}{N+1}\right)\cdot\frac{(\beta)_{N+M+1}}{\lfloor N+M+1} x^{N+M+1}$$

Conclusively the series which stopped for $\alpha = -N$ or $\beta = -M$ at Nth and Mth terms respectively, starts again when $\alpha = -N$ and $\gamma = -(N+M)$ or likewise when $\beta = -M$ and $\gamma = -(N+M)$.

(3) Integral Formula for the Hypergeometric Functions (*R.U., 1992, 93*)

We have $F(\alpha, \beta; \gamma; x) = \sum_{k=0}^{\infty} \frac{(\alpha)_k (\beta)_k}{\lfloor k (\gamma)_k} x^k$

where $(\alpha)_k = \alpha(\alpha+1)(\alpha+2)\ldots(\alpha+k-1) = \frac{\lfloor \alpha+k-1}{\lfloor \alpha-1} = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}$.

Similarly $(\beta)_k = \frac{\Gamma(\beta+k)}{\Gamma\beta}, (\gamma)_k = \frac{\Gamma(\gamma+k)}{\Gamma\gamma}$.

Also we have $B(m, n) = \frac{\Gamma m \, \Gamma n}{\Gamma(m+n)}$ (Beta function)

$$\therefore \frac{(\beta)_k}{(\gamma)_k} = \frac{\Gamma(\beta+k)}{\Gamma\beta} \times \frac{\Gamma\gamma}{\Gamma(\gamma+k)} = \frac{\Gamma\gamma}{\Gamma\beta}\cdot\frac{\Gamma(\beta+k)}{\Gamma(\gamma+k)} \times \frac{\Gamma(\gamma-\beta)}{\Gamma(\gamma-\beta)}$$

$$= \frac{\Gamma\gamma}{\Gamma\beta\Gamma(\gamma-\beta)}\cdot\frac{\Gamma(\beta+k)\,\Gamma(\gamma-\beta)}{\Gamma(\gamma+k)}$$

$$= \frac{\Gamma\gamma}{\Gamma\beta\Gamma(\gamma-\beta)}\,\frac{\Gamma(\beta+k)\,\Gamma(\gamma-\beta)}{\Gamma\{(\beta+k)+(\gamma-\beta)\}} = \frac{\Gamma\gamma}{\Gamma\beta\Gamma(\gamma-\beta)}\cdot B(\gamma-\beta, \beta+k)$$

$$= \frac{\Gamma\gamma}{\Gamma\beta\Gamma(\gamma-\beta)}\int_0^1 (1-t)^{\gamma-\beta-1} t^{\beta+k-1} dt \quad \because\ B(P, Q) = \int_0^1 (1-x)^{P-1} x^{Q-1} dx$$

$$= \frac{1}{B(\beta, \gamma-\beta)}\int_0^1 (1-t)^{\gamma-\beta-1} t^{\beta+x-1} dt\,.$$

NOTES

With these substitutions, we therefore, have

$$F(\alpha,\ \beta,\ \gamma,\ x) = \frac{1}{B(\beta, \gamma-\beta)} \sum_{k=0}^{\infty} \frac{(\alpha)_k}{\lfloor k} x^k \int_0^1 (1-t)^{\gamma-\beta-1} t^{\beta+k-1} dt$$

$$= \frac{1}{B(\beta, \gamma-\beta)} \int_0^1 (1-t)^{\gamma-\beta-1} t^{\beta-1} \left\{ \sum_{k=0}^{\infty} \frac{(\alpha)_k}{\lfloor k} (xt)^k \right\} dt$$

(on interchanging the order of integration and summation)

$$= \frac{1}{B(\beta, \gamma-\beta)} \int_0^1 (1-t)^{\gamma-\beta-1} t^{\beta-1} (1-xt)^{-\alpha}\, dt$$

$$\because\ \sum_{k=0}^{\infty} \frac{(\alpha)_k}{\lfloor k} (xt)^k = 1 + \frac{\alpha\,(xt)}{\lfloor 1} + \frac{\alpha\,(\alpha+1)}{\lfloor 2} (xt)^2 + \cdots = (1-xt)^{-\alpha}$$

We thus get the integral formula for hypergeometric series as

$$F(\alpha,\ \beta,\ \gamma,\ x) = \frac{1}{B(\beta, \gamma-\beta)} \int_0^1 (1-t)^{\gamma-\beta-1} t^{\beta-1} (1-xt)^{-\alpha}\ dt \qquad ...(4.164)$$

which is valid for $|x| < 1$ and $\gamma > \beta > 0$.

Corollary 3. Gauss Theorem or Gauss Formula

If we put $x = 1$ in (4.161), we have

$$F(\alpha,\ \beta;\ \gamma;\ 1) = \frac{1}{B(\beta, \gamma-\beta)} \int_0^1 (1-t)^{\gamma-\beta-1} t^{\beta-1} (1-t)^{-\alpha} dt$$

$$= \frac{1}{B(\beta, \gamma-\beta)} \int_0^1 (1-t)^{\gamma-\beta-\alpha-1} t^{\beta-1}\, dt = \frac{B(\beta, \gamma-\alpha-\beta)}{B(\beta, \gamma-\beta)} \qquad ...(4.165)$$

Using $B(P, Q) = \dfrac{\Gamma P \Gamma Q}{\Gamma(P+Q)}$, (27) yields

$$F(\alpha,\ \beta;\ \gamma;\ 1) = \frac{\Gamma\gamma\, \Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\, \Gamma(\gamma-\beta)} \qquad ...(4.166)$$

which is known as **Gauss' theorem**.

Corollary 4. Vandermond's theorem.

In (4.166) if we put $\alpha = -n$, we get

$$F(-n,\ \beta;\ \gamma;\ 1) = \frac{\Gamma\gamma\, \Gamma(\gamma-\beta+n)}{\Gamma(\gamma+n)\Gamma(\gamma-\beta)}$$

$$= \frac{1.2...(\gamma-1)\,.\,1.2...(\gamma-\beta+n-1)}{1.2...(\gamma+n-1)\,.\,1.2...(\gamma-\beta-1)}$$

$$= \frac{(\gamma-\beta)\,(\gamma-\beta+1)...(\gamma-\beta+n-1)}{\gamma(\gamma+1)...(\gamma+n-1)}$$

$$= \frac{(\gamma-\beta)_n}{(\gamma)_n} \qquad \text{...(4.167)}$$

This is known as **Vandermond's theorem**.

NOTES

Corollary 5. Kummer's theorem. (*R.U. 1991, 96*)

In (4.163) if we put $x = -1$ and $\gamma = \beta - \alpha + 1$, we get

$F(\alpha, \beta, \beta - \alpha + 1; -1)$

$$= \frac{1}{B(\beta, \gamma-\beta)} \int_0^1 (1-t^2)^{-\alpha} t^{\beta-1} dt$$

Put $t^2 = y$ so that $2t\, dt = dy$

$$= \frac{\Gamma(\beta-\alpha+1)}{\Gamma\beta\Gamma(\gamma-\beta)} \int_0^1 (1-y)^{-\alpha} y^{(\beta-1)/2} \frac{dy}{2\sqrt{y}}$$

$$= \frac{1}{2} \frac{\Gamma(\beta-\alpha+1)}{\Gamma\beta\Gamma(\gamma-\beta)} \int_0^1 (1-y)^{-\alpha} y^{\beta/2-1} dy$$

$$= \frac{1}{2} \cdot \frac{\Gamma(\beta-\alpha+1)}{\Gamma\beta\Gamma(\gamma-\beta)} B\left(\frac{\beta}{2}, 1-\alpha\right)$$

$$= \frac{1}{2} \cdot \frac{\Gamma(\beta-\alpha+1)}{\Gamma\beta\Gamma(1-\alpha)} \cdot \frac{\Gamma\frac{\beta}{2}\Gamma(1-\alpha)}{\Gamma\left(\frac{\beta}{2}+1-\alpha\right)}$$

$\because\ \gamma = \beta - \alpha + 1$ gives, $\gamma - \beta = 1 - \alpha$

$$= \frac{\Gamma(\beta-\alpha+1)}{\Gamma(\beta+1)} \cdot \frac{\Gamma\left(\frac{\beta}{2}+1\right)}{\Gamma\left(1-\alpha+\frac{\beta}{2}\right)} \qquad \text{...(4.168)}$$

(on dividing and multiplying by β)

which is known as **Kummer's theorem.**

[D] Linear Relationships of Hypergeometric Functions.

If we put $1 - t = p$ in (4.165), we get

$$F(\alpha, \beta; \gamma; x) = \frac{1}{B(\beta, \gamma-\beta)} \int_0^1 p^{\gamma-\beta-1}(1-p)^{\beta-1}\{1-x(1-p)\}^{-\alpha}\, dp$$

$$= \frac{(1-x)^{-\alpha}}{B(\beta, \gamma-\beta)} \int_0^1 (1-p)^{\beta-1} p^{\gamma-\beta-1} \left\{1-\frac{xp}{x-1}\right\}^{-\alpha} dp$$

or

$$F(\alpha, \beta; \gamma; x) = \frac{(1-x)^{-\alpha}}{B(\beta, \gamma-\beta)} B(\gamma-\beta, \beta)\, F\left(\alpha, \beta-\gamma; y; \frac{x}{x-1}\right)$$

[by using (4.164)]

$$= (1-x)^{-\alpha} F\left(\alpha, \gamma-\beta; \gamma; \frac{x}{x-1}\right) \qquad \text{...(4.169)}$$

NOTES

$\because$ by symmetry property $B(\beta, \gamma - \beta) = B(\gamma - \beta, \beta)$.

Similarly it may be shown that

$$F(\alpha, \beta; \gamma; x) = (1-x)^{-\beta} F\left(\gamma-\alpha, -\beta; \gamma; \frac{x}{x-1}\right) \qquad ...(4.170)$$

With $x = \frac{1}{2}$, (4.169) yields

$$F\left(\alpha, \beta; \gamma; \frac{1}{2}\right) = 2^{\alpha} F(\alpha, \gamma - \beta; \gamma; -1) \qquad ...(4.171)$$

If $\beta = 1 - \alpha$, this gives

$$F\left(\alpha, 1-\alpha; \gamma; \frac{1}{2}\right) = 2^{\alpha} F(\alpha, \gamma + \alpha - 1; \gamma; -1)$$

$$= \frac{2^{\alpha}\,\Gamma\gamma\,\Gamma\left(\frac{\gamma+\alpha-1}{2}+1\right)}{\Gamma(\gamma+\alpha)\,\Gamma\left(1-\alpha+\frac{\gamma+\alpha-1}{2}\right)} \qquad \text{using (4.168) } i.e. \text{ Kummer's theorem}$$

$$= \frac{2^{\alpha}\,\Gamma\gamma\,\Gamma\left(\frac{\gamma}{2}+\frac{1}{2}+\frac{\alpha}{2}\right)}{\Gamma(\gamma+\alpha)\,\Gamma\left(\frac{1}{2}-\frac{\alpha}{2}+\frac{\gamma}{2}\right)}$$

$$= \frac{2^{\alpha}\,\Gamma\left(\frac{\gamma}{2}\right)\cdot\frac{\gamma}{2}\left(\frac{\gamma}{2}+1\right)\left(\frac{\gamma}{2}+2\right)...(\gamma-1)\cdot\Gamma\left(\frac{\gamma}{2}+\frac{1}{2}\right)\cdot\left(\frac{\gamma}{2}+\frac{1}{2}\right)...\left(\frac{\gamma}{2}-\frac{1}{2}+\frac{\alpha}{2}\right)}{\Gamma\left(\frac{\gamma}{2}+\frac{\alpha}{2}\right)\left(\frac{\gamma}{2}+\frac{\alpha}{2}\right)...(\gamma+\alpha-1)\,\Gamma\left(\frac{1}{2}-\frac{\alpha}{2}+\frac{\gamma}{2}\right)}$$

$$= \frac{2^{\alpha}\,\Gamma\left(\frac{\gamma}{2}\right)\cdot\frac{1}{2^{\gamma/2}}\gamma(\gamma+2)...(2\gamma-2)\cdot\frac{1}{2^{\alpha/2}}\Gamma\left(\frac{\gamma}{2}+\frac{1}{2}\right)(\gamma+1)\,(\gamma+3)...(\gamma-1+\alpha)}{\frac{1}{2^{\gamma/2-\alpha/2}}\Gamma\left(\frac{\gamma}{2}+\frac{\alpha}{2}\right)\cdot(\gamma+\alpha)\,(\gamma+\alpha+2)...(2\gamma+2\alpha-2)\,\Gamma\left(\frac{1}{2}-\frac{\alpha}{2}+\frac{\gamma}{2}\right)}$$

$$= \frac{\Gamma\left(\frac{\gamma}{2}\right)\Gamma\left(\frac{\gamma+1}{2}\right)}{\Gamma\left(\frac{\gamma+\alpha}{2}\right)\Gamma\left(\frac{1-\alpha+\gamma}{2}\right)} \qquad ...(4.172)$$

(on simplifying)

Further, we have shown the solution of the hypergeometric series for $x = 0$ as

$$y = AF(\alpha, \beta, \gamma, x) + Bx^{1-\gamma} F(\alpha + 1 - \gamma, \beta + 1 - \gamma, 2 - \gamma, x)$$

by (4.139)

which is convergent for $|x| \le 1$, *i.e.*, in the interval $(-1, 1)$, and the solution for $x = 1$ by (4.142) is

$$y = AF(\alpha, \beta, 1 + \alpha + \beta - \gamma, 1 - x) + B(1-x)^{\gamma-\alpha-\beta} F(\gamma - \alpha, \gamma - \beta, \gamma - \alpha - \beta + 1, 1 - x)$$

which is convergent for $|1 - x| \le 1$, *i.e.*, in the interval $(0, 2)$.

NOTES

Clearly the solutions (4.139) and (4.142) are convergent in the common interval (0, 1) in which there exists a linear relationship between different hypergeometric functions.

Let the relation be

$$F(\alpha, \beta, \gamma, x) = AF(\alpha, \beta, 1 + \alpha + \beta - \gamma, 1 - x)$$
$$+ B(1 - x)^{\gamma-\alpha-\beta} F(\gamma - \alpha, \gamma - \beta, \gamma - \alpha - \beta + 1, 1 - x) \quad ...(4.173)$$

If we put $x = 0$ in (37) we get by Cor. 1 of §8.5 (C),

$$F(\alpha, \beta, \gamma, 0) = 1 = AF(\alpha, \beta, 1 + \alpha + \beta - \gamma, 1) + BF(\gamma - \alpha, \gamma - \beta, \gamma - \alpha - \beta + 1, 1)$$

$$= A\frac{\Gamma(1+\alpha+\beta-\gamma)\,\Gamma(1-\gamma)}{\Gamma(1+\beta-\gamma)\,\Gamma(1+\alpha-\gamma)} + B\frac{\Gamma(\gamma-\alpha-\beta+1)\,\Gamma(1-\gamma)}{\Gamma(1-\beta)\,\Gamma(1-\alpha)} \quad ...(4.174)$$

[by Gauss theorem (4.165)]

Putting again, $x = 1$ in (4.173) we get with the help of Gauss theorem,

$$F(\alpha, \beta, \gamma, 1) = AF(\alpha, \beta, 1 + \alpha + \beta - \gamma, 0) = \frac{\Gamma\gamma\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}$$

$$i.e.\ A = \frac{\Gamma\gamma\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}$$

$$F(\alpha, \beta, 1 + \alpha + \beta - \gamma, 0) = 1.$$

Substituting this value of A in (4.174), we find

$$1 = \frac{\Gamma\gamma\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)} \cdot \frac{\Gamma(1+\alpha+\beta-\gamma)\,\Gamma(1-\alpha)}{\Gamma(1+\beta-\gamma)\,\Gamma(1+\alpha-\gamma)} + B\frac{\Gamma(\gamma-\alpha-\beta+1)\Gamma(1-\gamma)}{\Gamma(1-\beta)\,\Gamma(1-\alpha)}$$

$$= \frac{\sin \pi(\gamma-\alpha)\sin \pi(\gamma-\beta)}{\sin \pi\gamma \sin(\gamma-\alpha-\beta)} + B\frac{\Gamma(1-\gamma)\,\Gamma(1+\gamma-\alpha-\beta)}{\Gamma(1-\alpha)\Gamma(1-\beta)}$$

[by using $\Gamma p\Gamma(1-p) = \dfrac{\pi}{\sin p\pi}$]

$$\text{or}\quad B = \frac{\Gamma(1-\alpha)\Gamma(1-\beta)}{\Gamma(1-\gamma)\,\Gamma(1+\gamma-\alpha-\beta)}\left[\frac{\sin \pi\gamma \sin \pi(\gamma-\alpha-\beta) - \sin \pi(\gamma-\alpha)\sin \pi(\gamma-\beta)}{\sin \pi\gamma \sin \pi(\gamma-\alpha-\beta)}\right]$$

$$= \frac{\Gamma(1-\alpha)\Gamma(1-\beta)\Gamma(\alpha+\beta-\gamma)}{\Gamma(1-\gamma)\cdot\dfrac{\pi}{\sin \pi(\alpha+\beta-\gamma)}} \cdot \frac{-\sin \pi\alpha \sin \pi\beta}{\sin \pi\gamma \sin \pi(\gamma-\alpha-\beta)}$$

$$= \frac{\Gamma(1-\alpha)\,\Gamma(1-\beta)\,\Gamma(\alpha+\beta-\gamma)}{\Gamma(1-\gamma)\,\pi} \cdot \frac{\dfrac{\pi}{\Gamma\alpha\,\Gamma(1-\alpha)} \cdot \dfrac{\pi}{\Gamma\beta\,\Gamma(1-\beta)}}{\dfrac{\pi}{\Gamma\gamma\,\Gamma(1-\gamma)}}$$

[by using $\Gamma p\,\Gamma(1-p) = \dfrac{\pi}{\sin p\pi}$]

$$= \frac{\Gamma\gamma\,\Gamma(\alpha+\beta-\gamma)}{\Gamma\alpha\,\Gamma\beta} x^{\gamma-\alpha-\beta}$$

Now (4.173) yields with the substitutions for A and B

NOTES

$$F(\alpha, \beta, \gamma, x) = \frac{\Gamma\gamma\,\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)} F(\alpha, \beta, 1+\alpha+\beta-\gamma, 1-x)$$

$$+\frac{\Gamma\gamma\,\Gamma(\alpha+\beta-\gamma)}{\Gamma\alpha\,\Gamma\beta} x^{\gamma-\alpha-\beta}\; F(\gamma-\alpha, \gamma-\beta, \gamma-\alpha-\beta+1, 1-x) \quad ...(4.175)$$

which is the required relationship.

If now replace x by $\frac{1}{x}$ in (4.175), we get

$$F\left(\alpha, \beta, \gamma, \frac{1}{x}\right) = \frac{\Gamma\gamma\,\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\,\Gamma(\alpha-\beta)} F\left(\alpha, \beta, 1+\alpha+\beta-\gamma, 1-\frac{1}{x}\right)$$

$$+\frac{\Gamma\gamma\Gamma(\alpha+\beta-\gamma)}{\Gamma\alpha\,\Gamma\beta}\left(1-\frac{1}{x}\right)^{\gamma-\beta-\alpha} F\left(\gamma-\alpha, \gamma-\beta, x-\alpha-\beta+1, 1-\frac{1}{x}\right)$$

But from (4.169) we have

$$F\left(\alpha, \beta, \gamma, 1-\frac{1}{x}\right) = \left(\frac{1}{x}\right)^{-\alpha} F(\alpha, \gamma-\beta, \gamma, 1-x)$$

$$= x^{\alpha} F(\alpha, \gamma-\beta, \gamma, 1-x),$$

$$\therefore \quad F\left(\alpha, \beta, \gamma\,\frac{1}{x}\right) = \frac{\Gamma\gamma\,\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)} x^{\alpha} F(\alpha, \alpha-\gamma+1, \alpha+\beta-\gamma+1, 1-x)$$

$$+\frac{\Gamma\gamma\,\Gamma(\alpha+\beta-\gamma)}{\Gamma\alpha\,\Gamma\beta} x^{\beta}\,(x-1)^{\gamma-\alpha-\beta}$$

$$F = (\gamma-\alpha, 1-\alpha, \gamma-\alpha-\beta+1, 1-x) \quad ...(4.176)$$

where $1 < x < 2$ and $1 > \gamma > \alpha + \beta$.

[E] Various Representations in Terms of Hypergeometric Functions

[e_1] If we put $x = (r_2 - r_1)\,u + r_1$ *i.e.* $u = \frac{x - r_1}{r_2 - r_1}$ ...(4.177)

i.e. $$\frac{du}{dx} = \frac{1}{r_2 - r_1}$$

So that $$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{1}{r_2 - r_1}\frac{dy}{du}$$

and $$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{1}{r_2 - r_1}\frac{dy}{du}\right) = \frac{1}{(r_2 - r_1)^2}\frac{d^2y}{du^2}, r_1 \neq r_2$$

in the equation $(x - r_1)\,(x - r_2)\,\frac{d^2y}{dx^2} + \frac{b}{a}(x - r_3)\frac{dy}{dx} + \frac{c}{a}y = 0.$ ...(4.178)

Then we find the Gauss equation or hypergeometric equation,

$$u(1-u)\frac{d^2y}{du^2} + \left[\frac{b}{a}\frac{r_1 - r_3}{r_1 - r_2} - \frac{b}{a}u\right]\frac{dy}{du} - \frac{c}{a}y = 0 \quad ...(4.179)$$

NOTES

Its series solution by usual method or comparison with §8.3 F [f_3], is

$y = AF\,(\alpha, \beta, \gamma, u) + Bu^{1-\gamma}\,F\,(\alpha - \gamma + 1, \beta - \gamma + 1, 2 - \gamma, u)$ ...(4.180)

where $\gamma = \dfrac{b\,(r_1 - r_3)}{a\,(r_1 - r_2)}, \alpha + \beta + 1 = \dfrac{b}{a}$ and $\alpha\beta = \dfrac{c}{a}$ ...(4.181)

$$\therefore \quad y = AF\left(\alpha, \beta, \gamma; \frac{x - r_1}{r_2 - r_1}\right) + B\left(\frac{x - r_1}{r_2 - r_1}\right)^{1-\gamma} F\left(\alpha - \gamma + 1, \beta - \gamma + 1, 2 - \gamma, \frac{x - r_1}{r_2 - r_1}\right)$$

...(4.182)

Corollary. Tschebycheff's equation is

$$(x - 1)\,(x + 1)\frac{d^2 y}{dx^2} + x\frac{dy}{dx} - n^2 y = 0 \qquad ...(4.183)$$

Comparing it with (42), we get

$r_1 = 1, r_2 = -1, r_3 = 0, a = 1, b = 1, c = -n^2, x = -2u + 1, \alpha = n, \beta = -n, \gamma = \dfrac{1}{2}$

and hence by (4.182), the solution is

$$y = AF\left(n, -n; \frac{1}{2}, \frac{1 - x}{2}\right) + B\left(\frac{1 - x}{2}\right)^{1/2} F\left(n + \frac{1}{2}, -n + \frac{1}{2}, \frac{3}{2}; \frac{1 - x}{2}\right) \qquad ...(4.184)$$

with the help of (4.184) the *Tschebycheff's Polynomials* may be found as

$$\left.\begin{array}{l} T_0(x) = 1, T_1(x) = x, T_2(x) = 2x^2 - 1, \\ T_3(x) = 4x^3 - 3x, T_4(x) = 8x^4 - 8x^2 + 1 \end{array}\right\} \qquad ...(4.185)$$

[e_2] Legendre's Polynomials have already been discussed in problems on Legendre's Polynomials.

[e_3] **Elliptical integrals** of the first and second kind are

$$k(x) = \int_0^{\pi/2} \frac{d\theta}{(1 - x^2 \sin^2 \phi)^{1/2}} \qquad ...(4.186)$$

and
$$E(x) = \int_0^{\pi/2} (1 - x^2 \sin^2 \phi)^{1/2}\, d\phi \qquad ...(4.187)$$

respectively.

We have $k(x) = \displaystyle\int_0^{\pi/2} (1 - x^2 \sin^2 \phi)^{-1/2}\, d\phi$

$$= \int_0^{\pi/2} \left(1 + \frac{1}{2} \cdot x^2 \sin^2 \phi + \frac{\frac{1}{2}\left(\frac{1}{2} + 1\right)}{\lfloor 2} x^4 \sin^4 \phi + \cdots\cdots \right) d\phi$$

$$= \int_0^{\pi/2} \sum_{k=0}^{\infty} \frac{\frac{1}{2}\left(\frac{1}{2} + 1\right) \cdots\cdots \left(\frac{1}{2} + k - 1\right)}{\lfloor k} x^{2k} \sin^{2k} \phi\, d\phi$$

$$= \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\right)_k x^{2k}}{\underline{|k}} \int_0^{\pi/2} \sin^{2k} \phi \, d\phi$$

NOTES

$$= \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\right)_k x^{2k}}{\underline{|k}} \cdot \frac{\Gamma\left(2k+\frac{1}{2}\right)}{\Gamma(k+1} \frac{\sqrt{\pi}}{2} \qquad \text{[by Gamma integrals]}$$

$$= \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\right)_k x^{2k}}{\underline{|k}} \cdot \frac{\left(\frac{1}{2}\right)_k}{\underline{|k}} \frac{\pi}{2}$$

$$= \frac{\pi}{2} \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\right)_k \left(\frac{1}{2}\right)_k}{(1)_k \, \underline{|k}} \cdot (x^2)^k \text{ since } \underline{|k} = 1 \cdot 2 \ldots\ldots (k-1) = (1)_k$$

$$= \frac{\pi}{2} F\left(\frac{1}{2}, \frac{1}{2}, 1; x^2\right) \qquad \ldots(4.188)$$

We can similarly show that

$$E(x) = \frac{\pi}{2} F\left(-\frac{1}{2}, \frac{1}{2}, 1; x^2\right) \qquad \ldots(4.189)$$

[G] Various Transformations

The Gauss hypergeometric equation is

$$x(1-x)\frac{d^2y}{dx^2} + \{\gamma - (\alpha+\beta+1)x\}\frac{dy}{dx} - \alpha\beta y = 0 \qquad \ldots(4.190)$$

Whose solution is $y = F(\alpha, \beta; \gamma; x)$ ...(4.191)

For $\gamma = 2\beta$, (54) $\Rightarrow x(1-x)\frac{d^2y}{dx^2} + \{2\beta - (\alpha+\beta+1)x\}\frac{dy}{dx} - \alpha\beta y = 0$ ...(4.192)

With its solution $y = F(\alpha, \beta; 2\beta; x)$ ...(4.193)

If we make the transformation

$$x = \frac{4u}{(1+u)^2} \; i.e. \; \frac{dx}{du} = \frac{4(1-u)}{(1+u)^3}, \text{ so that}$$

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = \frac{(1+u)^3}{4(1-u)} \frac{dy}{du}$$

and $$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{(1-u)^6}{16(1-u)^2} \cdot \frac{d^2y}{du^2} + \frac{(1+u)^5(4-2u)}{16(1-u)^3} \frac{du}{dx}$$

then (4.192) transforms to

$$u(1-u)(1+u)^2 \frac{d^2y}{du^2} + 2(1+u)(b - 2\alpha u + \beta u^2)\frac{dy}{du} - 4\alpha\beta(1-u)y = 0 \qquad \ldots(4.194)$$

with its solution $$y = F\left[\alpha, \beta; 2\beta; \frac{4u}{(1+u)^2}\right] \qquad \ldots(4.195)$$

If $y = (1+u)^{2\alpha} \cdot v$, then (4.194) $\Rightarrow \frac{dy}{du} = (1+u)^{2a}\frac{dv}{du} + 2\alpha(1+u)^{2\alpha-1} \cdot v$

and $$\frac{d^2y}{du^2} = (1+u)^{2\alpha}\frac{d^2v}{du^2} + 4\alpha(1+u)^{2\alpha-1}\frac{dv}{du} + 2\alpha\,(2\alpha-1)\,(1+u)^{2\alpha-2}v$$

so that the transformed equation is

NOTES

$$u(1-u^2)\frac{d^2v}{du^2} + 2[\beta-(2\alpha-\beta+1)\,u^2]\frac{dv}{du} - 2\alpha u\,(1+2\alpha-2\beta)v = 0 \quad ...(4.196)$$

with its solution

$$v(1+u)^{-2\alpha}\; F\left[\alpha, \beta; 2\beta; \frac{4u}{(1+u)^2}\right] \quad ...(4.197)$$

Putting $u^2 = w$ in (4.196) which is unchanged on replacing u by $-v$, we get

$$w(1-w)\frac{d^2v}{dw^2} + \left[\beta+\frac{1}{2}-\left(2\alpha-\beta+\frac{3}{2}\right)w\right]\frac{du}{dw} - \alpha\left(\alpha-\beta+\frac{1}{2}\right)v = 0 \quad ...(4.198)$$

which is identical with (4.190) and have the general solution for $|w| < 1$,

$$v = AF\left(\alpha, \alpha-\beta+\frac{1}{2}; \beta+\frac{1}{2}; w\right) + Bw^{\frac{1}{2}-\beta}F\left(\alpha-\beta+\frac{1}{2}, \alpha+1-2\beta; \frac{3}{2}-\beta; w\right)$$
... (4.199)

by comparison of coefficients in the solution (4.189)

Conclusively, for $|u| < 1$ and $\left|\frac{4u}{(1+u)^2}\right| < 1$, provided 2β is neither zero nor a negative number, we have from (4.197) and (4.199), the identity

$$(1+u)^{-2\alpha}F(\alpha, \beta; 2\beta; \frac{4u}{(1+u)^2}) = AF\left(\alpha, \alpha-\beta+\frac{1}{2}; \beta+\frac{1}{2}; u^2\right)$$
$$+Bu^{1-2\beta}F\left(\alpha-\beta+\frac{1}{2}, \alpha+1-2\beta; \frac{3}{2}-\beta; u^2\right) \quad ...(4.200)$$

which yields $A = 1$ when $u = 0$.

The last term in (64) being not analytic due to the presence of the factor $u^{1-2\beta}$ as $u \to 0$, we have $B = 0$, since $F\left(\alpha-\beta+\frac{1}{2}, \alpha+1-2\beta; \frac{3}{2}-\beta; 0\right) \neq 0$.

$$\therefore\ (4.200) \Rightarrow (1-u)^{-2\alpha}\; F\left(\alpha, \beta; 2\beta; \frac{4u}{(1+u)^2}\right) = F\left(\alpha, \alpha-\beta+\frac{1}{2}; \beta+\frac{1}{2}; u^2\right)$$
...(4.201)

Replacing γ by $\alpha+\beta+\frac{1}{2}$ in (4.190) and (4.191), we find

$$x(1-x)\frac{d^2y}{dx^2} + \left\{\left(\alpha+\beta+\frac{1}{2}\right) - (\alpha+\beta+1)\,x\right\}\frac{dy}{dx} - \alpha\beta y = 0 \quad ...(4.202)$$

With its solution $y = F\left(\alpha, \beta; \alpha+\beta+\frac{1}{2}; x\right)$ ...(4.203)

Setting $x = 4z(1-z)$ so that $\frac{dy}{dx} = \frac{dy}{dz}\cdot\frac{dz}{dx} = \frac{1}{4(1-2z)}\cdot\frac{dy}{dz}$

NOTES

and $\dfrac{d^2y}{dx^2} = \dfrac{1}{16(1-2z)^2}\dfrac{d^2y}{dz^2} + \dfrac{1}{8(1-2z)^3}\cdot\dfrac{dy}{dz}$,

$$(4.202) \Rightarrow z(1-z)\frac{d^2y}{dz^2} + \left\{\left(\alpha+\beta+\frac{1}{2}\right) - (2\alpha+2\beta+1)z\right\}\frac{dy}{dz} - 4\alpha\beta y = 0 \quad ...(4.204)$$

For $|x| < 1$ *i.e.* $|4z(1-z)| < 1$, its solution is

$$y = AF\left(2\alpha, 2\beta; \alpha+\beta+\frac{1}{2}; z\right) + Bz^{\frac{1}{2}-\alpha-\beta}F\left(\frac{1}{2}+\alpha-\beta, \frac{1}{2}+\beta-\alpha; \frac{3}{2}-\alpha-\beta; z\right) \quad ...(4.205)$$

(4.203) and (4.205) lead to the identity.

$$F\left(\alpha, \beta; \alpha+\beta+\frac{1}{2}; 4z(1-z)\right) = AF\left(2\alpha, 2\beta; \alpha+\beta+\frac{1}{2}; z\right) + Bz^{\frac{1}{2}-\alpha-\beta}F\left(\frac{1}{2}+\alpha-\beta, \frac{1}{2}+\beta-\alpha; \frac{3}{2}-\alpha-\beta; z\right) \quad ...(4.206)$$

With the same argument as applied above, for $x \to 0$, $A = 1$, $B = 0$ so that (4.206) reduces to

$$F\left[\alpha, \beta; \alpha+\beta+\frac{1}{2}; 4z(1-z)\right] = F\left[2\alpha, 2\beta; \alpha+\beta+\frac{1}{2}; z\right] \quad ...(4.207)$$

Which yields on replacing α by $\dfrac{\gamma}{2}-\dfrac{\alpha}{2}$ and β by $\dfrac{\gamma}{2}+\dfrac{\alpha}{2}-\dfrac{1}{2}$,

$$F\left(\frac{\gamma}{2}-\frac{\alpha}{2}, \frac{\gamma}{2}+\frac{\alpha}{2}-\frac{1}{2}; \gamma; 4z(1-z)\right) = F(\gamma-\alpha, \gamma+\alpha-1; \gamma; z) \quad ...(4.208)$$

Again if we put $\beta = l$, $a = l + \dfrac{1}{2}$ in (4.201), then we find

$$F\left(l, l+\frac{1}{2}; 2l; \frac{4u}{(1+u)^2}\right) = (1+u)^{2l+1} \; F\left(l+\frac{1}{2}, 1; l+\frac{1}{2}; u^2\right)$$

$$= (1+u)^{2l+1}\sum_{m=0}^{\infty}u^{2m} = (1+u)^{2l+1}(1-u^2)^{-1} = (1+u^2)^{2l-1}\left(\frac{1+u}{1-u}\right)$$

Putting $x = \dfrac{4u}{(1+u)^2}$ *i.e.* $1-x = \left(\dfrac{1-u}{1+u}\right)^2$ and $1+u = \dfrac{2}{1+\sqrt{(1-x)}}$, the last result reduces to

$$F\left(l, l+\frac{1}{2}; 2l; x\right) = \left\{1+\frac{2}{1+\sqrt{(1-x)}}\right\}^{2l-1}(1-x)^{-1/2} \quad ...(4.209)$$

Also (4.201) yields for $\alpha = l - \dfrac{1}{2}, \beta = l$,

$$F\left(l, l-\frac{1}{2}; 2l; \frac{4u}{(1+u)^2}\right) = (1+u)^{2l-1}F(l, 0; l; u^2)$$

$$= (1+u)^{2l-1}$$

NOTES

and hence for $x = \dfrac{4u}{(1+u)^2}$, we find

$$F\left(l, l-\frac{1}{2}; 2l; x\right) = \left\{\frac{2}{1+\sqrt{(1-x)}}\right\}^{2l-1}. \quad ...(4.210)$$

Problem 4.1. *Solve in series*

$$x(1-x)y'' + 4(1-x)y' - 2y = 0.$$

Comparing it with the hypergeometric series, we get

$$\alpha + \beta + 1 = 4, \gamma = 4, \alpha\beta = 2$$

i.e. $\alpha = 1, \beta = 2, \gamma = 4$ or $\alpha = 2, \beta = 1, \gamma = 4$

For the first choice the solution is $y_1\, F(1, 2, 4, x)$

$= F(2, 1, 4, x)$ by symmetry property

$$= 1 + \frac{x}{2} + \frac{3x^2}{10} + \frac{x^3}{5} + \frac{x^4}{7} + \cdots$$

For the second choice the solution is

$y_2 = x^{-3} F(-2, -1, -2, x)$

$= x^{-3}(1-x)$ since the third term vanishes as one of $\alpha - \gamma + 2$ or $\beta - \gamma + 2$ is zero. Moreover fourth term has zero for its denominator as $\gamma = 4$.

Hence the complete solution is

$$y = AF(1, 2, 4, x) + B\frac{1-x}{x^3}.$$

Problem 4.2. *Solve in series*

$$(x - x^2)\frac{d^2y}{dx^2} + \left(\frac{3}{2} - 2x\right)\frac{dy}{dx} - \frac{1}{4}y = 0.$$

Comparing with Gauss equation, we have

$$\alpha + \beta + 1 = 2, \alpha\beta = \frac{1}{4}, \gamma = \frac{3}{2} \text{ giving } \alpha = \beta = \frac{1}{2}, \gamma = \frac{3}{2}.$$

So the solution is

$y = AF(\alpha, \beta, \gamma, x) + Bx^{1-\gamma} F(\alpha - \gamma + 1, \beta - \gamma + 1, 2 - \gamma, x)$

$$= AF\left(\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, x\right) + Bx^{-\frac{1}{2}} F\left(0, 0, \frac{1}{2}, x\right)$$

$$= AF\left(\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, x\right) + Bx^{-\frac{1}{2}} \text{ since } F\left(\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, x\right)$$

$$= 1 + \frac{x}{6} + \frac{3x^2}{40} + \frac{5x^3}{112} + \cdots \text{ and } F\left(0, 0, \frac{1}{2}, x\right) = 1.$$

Problem 4.3. *Transform $y'' + n^2y = 0$ to hypergeometric form by the substitution $u = \sin^2 x$ and prove that*

NOTES

(*i*) $cos\ nx = F\left(\frac{n}{2}, -\frac{n}{2}, \frac{1}{2}, sin^2 x\right)$

(*ii*) $sin\ nx = n\ sin\ x\ F\left(\frac{1}{2}-\frac{1}{2}n, \frac{1}{2}+\frac{1}{2}n, \frac{3}{2}; sin^2 x\right)$ for $-\frac{\pi}{2} \le x \le \frac{\pi}{2}$.

Given equation is $y'' + n^2 y = 0$ ...(1)

Also given that $u = \sin^2 x$ ...(2)

which gives $\frac{du}{dx} = 2 \sin x \cos x = \sin 2x$.

So that $\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \sin 2x \frac{dy}{du}$ and $\frac{d^2 y}{dx^2} = 2 \cos 2x \frac{dy}{du} + \sin^2 2x \frac{d^2 y}{du^2}$

Substituting these values in (1),

$$4 \sin^2 x \cos^2 x \frac{d^2 y}{du^2} + 2(1 - 2\sin^2 x)\frac{dy}{du} + n^2 y = 0$$

or $$u(1-u)\frac{d^2 y}{du^2} + \left(\frac{1}{2} - u\right)\frac{dy}{du} + \frac{n^2}{4} y = 0 \quad ...(3)$$

which is hypergeometric form, with $\gamma = \frac{1}{2}$, $\alpha + \beta + 1 = 1$, $\alpha\beta = -\frac{n^2}{4}$.

i.e. $\alpha = \frac{n}{2}$, $\beta = -\frac{n}{2}$, $\gamma = \frac{1}{2}$ and hence the solution is

$$y = AF\left(\frac{n}{2}, -\frac{n}{2}, \frac{1}{2}, u\right) + Bu^{1/2} F\left(\frac{1+n}{2}, \frac{1-n}{2}, \frac{3}{2}, u\right)$$

$$= AF\left(\frac{n}{2}, -\frac{n}{2}, \frac{1}{2}, \sin^2 x\right) + B \sin x F\left(\frac{1+n}{2}, \frac{1-n}{2}, \frac{3}{2}, \sin^2 x\right) \quad ...(4)$$

where $|\sin x| < 1$ *i.e.* $-\frac{\pi}{2} \le x \le \frac{\pi}{2}$.

Bus sin *nx* and cos *nx* are the solutions of (1) and so of (3).We can therefore take from (4).

$$\sin nx = AF\left(\frac{n}{2}, -\frac{n}{2}; \frac{1}{2}; \sin^2 x\right) + B \sin x F\left(\frac{1+n}{2}, \frac{1-n}{2}; \frac{3}{2}; \sin^2 x\right) \quad ...(5)$$

When $x = 0$, $A = 0$ and so $\frac{\sin nx}{\sin x} = BF\left(\frac{1+n}{2}, \frac{1-n}{2}, \frac{3}{2}, \sin^2 x\right)$

by (5)

Which follows that when $x = 0$, $B = n$ and hence

$$\sin nx = n \sin xF\left(\frac{1-n}{2}, \frac{1+n}{2}; \frac{3}{2}; \sin^2 x\right) \quad ...(6)$$

Also, $\cos nx = AF\left(\frac{n}{2}, -\frac{n}{2}, \frac{1}{2}, u\right) + B\sqrt{u}\, F\left(\frac{1+n}{2}, \frac{1-n}{2}, \frac{3}{2}, u\right)$ ...(7)

NOTES

where $u = \sin^2 x$.

When $x = 0$, $A = 1$ and differentiating (7) w.r.t. x, we find

$$-n \sin nx = A(-n^2)\, F\left(\frac{n+2}{2}, \frac{2-n}{2}, \frac{3}{2}, u\right)\frac{du}{dx} + B\sqrt{u}\,\frac{1-n^2}{6}$$

$$F\left(\frac{3+n}{2}, \frac{3-n}{2}, \frac{5}{2}, u\right)\frac{du}{dx} + BF\left(\frac{1+n}{2}, \frac{1-n}{2}, \frac{3}{2}, u\right)\frac{1}{2\sqrt{u}}\frac{du}{dx}$$

by §8.5 [C] (2)

where $u = \sin^2 x$ and $\frac{du}{dx} = \sin 2x$

When $x = 0$, $B = 0$, so that

$$\cos nx = F\left(\frac{n}{2}, -\frac{n}{2}; \frac{1}{2}; \sin^2 x\right) \qquad ...(8)$$

4.9.1 Representation of Bessel Functions In Terms Of Hypergeometric Functions

The Bessel functions have been known since the 18th century when mathematicians and scientists started to describe physical processes through differential equations. Many different looking processes satisfy the same partial differential equations. These equations were named Laplace, d'Alembert (wave), Poisson, Helmholtz, and heat (diffusion) equations. Different methods were used to investigate these equations. The most powerful was the separation of variables method, which in polar coordinates often leads to ordinary differential equations of special structure:

$$w''(z)\,z^2 + w'(z)\,z + \left(z^2 - \nu^2\right) w(z) = 0.$$

This equation with concrete values of the parameter ν appeared in the articles by F. W. Bessel (1816, 1824) who built two partial solutions $w_1(z)$ and $w_2(z)$ of the previous equation in the form of series:

$$w(z) = z^{\nu} \sum_{j=0}^{\infty} a_j\, z^j + z^{-\nu} \sum_{j=0}^{\infty} b_j\, z^j = z^{\nu}\left(\sum_{k=0}^{\infty} a_{2k}\, z^{2k} + \sum_{k=0}^{\infty} a_{2k+1}\, z^{2k+1}\right) + z^{-\nu}\left(\sum_{k=0}^{\infty} b_{2k}\, z^{2k} + \sum_{k=0}^{\infty} b_{2k+1}\, z^{2k+1}\right).$$

Substituting the series into the differential equation produces the following solutions:

$$w_1(z) = z^{\nu} \sum_{k=0}^{\infty} A_k\, z^{2k} \;/;\; A_0 = \frac{2^{-\nu}}{\Gamma(\nu+1)} \bigwedge A_1 = -\frac{2^{-\nu-2}}{\Gamma(\nu+2)} \bigwedge A_k = a_{2k} = \frac{(-1)^k\, 2^{-\nu-2k}}{\Gamma(k+\nu+1)\, k!}$$

$$w_2(z) = z^{-\nu} \sum_{k=0}^{\infty} B_k\, z^{2k} \;/;\; B_0 = \frac{2^{\nu}}{\Gamma(1-\nu)} \bigwedge B_1 = -\frac{2^{\nu-2}}{\Gamma(2-\nu)} \bigwedge B_k = b_{2k} = \frac{(-1)^k\, 2^{\nu-2k}}{\Gamma(k-\nu+1)\, k!}.$$

O. Schlömilch (1857) used the name Bessel functions for these solutions, E. Lommel (1868) considered as an arbitrary real parameter, and H. Hankel (1869) considered complex values for . The two independent solutions of the differential equation were notated as and .

For integer index , the functions and coincide or have different signs. In such cases, the second linear independent solution of the previous differential equation

NOTES

was introduced by C. G. Neumann (1867) as the limit case of the following special linear combination of the functions $J_\nu(z)$ and $J_{-\nu}(z)$:

$$Y_\nu(z) = \lim_{\mu \to \nu} \frac{\cos(\mu \pi) J_\mu(z) - J_{-\mu}(z)}{\sin(\mu \pi)} /; \nu \in \mathbb{Z}.$$

The Bessel functions $J_\nu(z)$, $I_\nu(z)$, $K_\nu(z)$, and are particular cases of more general functions: hypergeometric and Meijer G functions.

In particular, the functions $J_\nu(z)$ and $I_\nu(z)$ can be represented through the regularized hypergeometric functions (without any restrictions on the parameter ν):

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu {}_0\tilde{F}_1\left(; \nu + 1; -\frac{z^2}{4}\right) \quad I_\nu(z) = \left(\frac{z}{2}\right)^\nu {}_0\tilde{F}_1\left(; \nu + 1; \frac{z^2}{4}\right).$$

Similar formulas, but with restrictions on the parameter , represent $J_\nu(z)$ and $I_\nu(z)$ through the classical hypergeometric function ${}_0F_1$:

$$J_\nu(z) = \frac{1}{\Gamma(\nu+1)} \left(\frac{z}{2}\right)^\nu {}_0F_1\left(; \nu + 1; -\frac{z^2}{4}\right) /; -\nu \notin \mathbb{N}^+ \quad I_\nu(z) = \frac{1}{\Gamma(\nu+1)} \left(\frac{z}{2}\right)^\nu {}_0F_1\left(; \nu + 1; \frac{z^2}{4}\right) /; -\nu \notin \mathbb{N}^+.$$

The functions and can also be represented through the hypergeometric functions by the following formulas:

$$J_\nu(z) = \frac{z^\nu}{2^\nu e^{i z} \Gamma(\nu+1)} {}_1F_1\left(\nu + \frac{1}{2}; 2\nu + 1; 2 i z\right) \quad I_\nu(z) = \frac{z^\nu}{2^\nu e^{z} \Gamma(\nu+1)} {}_1F_1\left(\nu + \frac{1}{2}; 2\nu + 1; 2 z\right)$$

$$J_\nu(z) = \frac{z^\nu}{2^\nu \Gamma(\nu+1)} \lim_{a \to \infty} {}_1F_1\left(a; \nu + 1; -\frac{z^2}{4a}\right) \quad I_\nu(z) = \frac{z^\nu}{2^\nu \Gamma(\nu+1)} \lim_{a \to \infty} {}_1F_1\left(a; \nu + 1; \frac{z^2}{4a}\right).$$

4.9.2 Representation of Laguerre Function In Terms Of Hypergeometric Functions

By using the isomorphism technique it is possible to define in general Laguerre-type special functions, and in particular, the 1st order Laguerre-type hypergeometric functions.

In fact, starting from the Gauss' hypergeometric equation:

x (1- x)y''+[c- (a+b+1)x]y'- aby = 0

and applying the isomorphism Tx, we find the equation

x(1- x)D2Ly+[c- (a+b+1)x]DLy- aby=0, (4.211)

that is:

[x(1- x)](x2yiv+4xy'''+2y'')+[c- (a+b+1)x](y'+xy'')- aby=0. (4.212)

The solution of Equation (4.211), corresponding to the Gauss' hypergeometric equation F(a,b,c;x), , is given by

LF(a,b,c;x)=1+? n=18 a(n)b(n)c(n)xn(n!)2 (4.213)

where the symbol a(n) denotes the rising factorial.

The rth order Laguerre-type hypergeometric functions are obtained by applying to both sides of the hypergeometric equation the iterated isomorphism of order r, but the corresponding differential equation becomes more and more complicated as r increases.

NOTES

The generalized hypergeometric functions have their 1st order Laguerre-type counterpart, which are given by:

LpFq(*a*1,⋯,*ap*;*b*1,⋯,*bq*,;*x*)=∑*n*=0∞|*a*(*n*)1⋯*a*(*n*)*pb*(*n*)1⋯*b*(*n*)*qxn*(*n*!)2, (4.214)

and those of higher order immediately follow.

Please note that the function in (4.214) can be viewed as a generalized hypergeometric function of the form *pFq*+1, by moving one of the *n*! in the first fraction under the sum and considering “(*n*+1)/“(1)=1(*n*)=*n*!=*b*(*n*)*q*+1 as the (*q*+1)-th term.

Check Your Progress

10. What is the form of Laguerre differential equation?
11. What the Laguerre polynomial of degree n is denoted by?
12. What is the solution of Laguerre solution for v to be a positive integer?
13. Write the relation that we use to give the Laguerre polynomial Ln(x).
14. What is the generating function for Laguerre polynomial?

4.10 ANSWERS TO ‘CHECK YOUR PROGRESS’

1. $(1 - x^2)\dfrac{d^2y}{dx^2} - 2x\dfrac{dy}{dx} + n(n+1) = 0.$
2. $\dfrac{d}{d\mu}P_n(\mu)$ = (2*n* – 1) *Pn*–1(µ) + (2*n* – 5) *Pn*–3(µ) + (2*n* – 9) *Pn*–5(µ) + · · ·
3. y = $AP_n(x) + BQ_n(x)$ is the most general form of Legendre’s equation where and are arbitrary constants.
4. $\dfrac{d^2y}{dx^2} + \dfrac{1}{x}\dfrac{dy}{dx} + \left(1 - \dfrac{n^2}{x^2}\right)y = 0$
5. At $x = 0$.
6. $J_o(x) = \sum_{r=o}^{\infty} \dfrac{(-1)^r\left(\frac{x}{2}\right)^{2r}}{r\,!\,\Gamma(r+1)} = 1 - \dfrac{x^2}{2^2} + \dfrac{x^4}{2^2 4^2} - \dfrac{x^6}{2^2.4^2.6^2} + \cdots$
7. This equation is of the form

   $\dfrac{d^2y}{dx^2} - 2x\dfrac{dy}{dx} + 2\nu y = 0$

   where ν is a parameter.
8. Hermite equation is $y'' - 2xy' + 2ny = 0$ for integral values taking $\nu = n$.
9. $H_n''(x) - 2x\, H_n'(x) + 2n\, H_n'(x) = 0$

10. This equation is of the form

$$xy'' + (1 - x)\,y' + \nu y = 0$$

11. The Laguerre polynomial of degree n is denoted by $L_n(x)$ *i.e.*

$$L_n(x) = (-1)^n\left[x^n - \frac{n^2}{\lfloor 1}x^{n-1} + \frac{n^2(n-1)^2}{\lfloor 2}x^{n-2} + \cdots + \cdots (-1)^n \lfloor n\right]$$

12. The solution of Laguerre solution for ν to be a positive integer is

$$y = AL_n(x)$$

13. The Laguerre Polynomials $L_n(x)$ are given by the relation

$$(1-t)\sum_{n=0}^{\infty}\frac{L_n(x)}{\lfloor n}t^n = \frac{-xt}{e^{1-t}} \quad \text{or} \quad \frac{1}{1-t}e^{-xt/(t-0)} = \sum_{n=0}^{\infty}\frac{L_n(x)t^n}{\lfloor n}$$

14. $\frac{1}{1-t}e^{-xt/(1-t)}$ is known as **Generating function** for Laguerre polynomial.

4.11 SUMMARY

- $(1 - x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + n\,(n+1)y = 0$ is the Legendre's equation and this equation can be solved in series of ascending or descending powers of.
- $y = AP_n(x) + BQ_n(x)$ = is the most general form of Legendre's equation where and are arbitrary constants.
- $\sum_{n=0}^{\infty} h^n P_n(\mu) = (1 - 2\mu h + h^2)^{-1/2}$ is the generating function for
- $P_n(\mu) = (2n-1)\,\mu\, P_{n-1}(\mu) - (n-1)P_{n-2}(\mu)$ is the recurrence formula for $P_n(\mu)$.
- $\frac{d}{d\mu}P_n(\mu) = (2n-1)\,Pn{-}1(\mu) + (2n-5)\,Pn{-}3(\mu) + (2n-9)\,Pn{-}5(\mu) + \cdots$ this is known as Christoffel's Expansion.
- Legendre's polynomials of first kind possess orthogonal properties.
- Bessel's differential equation $\frac{d^2y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + \left(1 - \frac{n^2}{x^2}\right)y = 0$
- Bessel's function of the first kind of order n

$$J_o(x) = \sum_{r=o}^{\infty}\frac{(-1)^r\left(\frac{x}{2}\right)^{2r}}{r\,!\,\Gamma(r+1)} = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 4^2} - \frac{x^6}{2^2.4^2.6^2} + \cdots$$

- Bessel's function of zeroth order $e^{\frac{x}{2}\left(t-\frac{1}{t}\right)} = \sum_{n=-\infty}^{\infty} t^n J_n(x)$
- Generating function for $e^{\frac{x}{2}\left(t-\frac{1}{t}\right)} = \sum_{n=-\infty}^{\infty} t^n J_n(x)$

- Recurrence relation for $J_n(x) = \sum_{r=0}^{\infty} \frac{(-1)^r \left(\frac{x}{2}\right)^{n+2r}}{r!\,(n+r)!}$

NOTES

- $e^{2tx-t^2} = \sum_{n=0}^{\infty} \frac{H_n(x)\,t^n}{\lfloor n}$
- $y = H_n(x)$ is a solution of Hermite equation.
- $y = a_0\left[1 - \nu x + \frac{\nu(\nu-1)}{(\lfloor 2)^2}x^2 - \dots + (-1)^r \frac{\nu(\nu-1)\dots(\nu-r+1)}{(\lfloor r)^2}x^r + \dots\right] = a_0 \sum_{k=0}^{\infty} \frac{(-1)^k \lfloor \nu}{(\lfloor k)^2 \lfloor \nu - k}$
- $L_n(x) = (-1)^n\left[x^n - \frac{n^2}{\lfloor 1}x^{n-1} + \frac{n^2(n-1)^2}{\lfloor 2}x^{n-2} + \dots + \dots (-1)^n \lfloor n\right]$
- The solution of Laguerre solution for ν to be a positive integer is

  $y = AL_n(x)$
- The generating function for Laguerre Polynomial is

  $e^{\frac{-xt}{1-t}} = (1-t)\sum_{n=0}^{\infty} \frac{L_n(x)}{\lfloor n} t^n$
- Solutions to the associated Laguerre differential equation with and an integer are called associated Laguerre polynomials . Associated Laguerre polynomials are implemented in the Wolfram Language as LaguerreL[*n*, *k*, *x*]. In terms of the unassociated Laguerre polynomials,

4.12 KEY TERMS

- **Differential equations**: A differential equation is a mathematical equation that relates some function with its derivatives.
- **Functions**: Function is an expression, rule, or law that defines a relationship between one variable (the independent variable) and another variable (the dependent variable).
- **Polynomial**: Polynomial is an expression of more than two algebraic terms, especially the sum of several terms that contain different powers of the same variable(s).
- **Arbitrary constant**: It is a symbol to which various values may be assigned but which remains unaffected by the changes in the values of the variables of the equation.
- **Coefficient**: It refers to a number by which another number or symbol is multiplied.
- **Finite series**: A finite series is a summation of a finite number of terms.
- **Infinite series**: An infinite series has an infinite number of terms and an upper limit of infinity.

- **Recurrence Relation**: A recurrence relation is an equation that defines a sequence based on a rule that gives the next term as a function of the previous term(s).
- **Singularity**: A singularity is a point at which a function does not possess a derivative.
- **Method of series integration**: This is a general method used for summing up power series. Differentiate the power series, find the sum of that, and then integrate the function obtained, choosing the antiderivative whose value at 0 equals the constant term of the power series.
- **Relation:** It is a relationship between sets of values. The relation is between the x-values and y-values of ordered pairs.
- **Equation:** It is a statement of an equality containing one or more variables.
- **Laguerre polynomials:** The Laguerre polynomials are named after Edmond Laguerre. They are solutions of Laguerre's equation in mathematics.
- **Orthogonality**: It is the generalization of the notion of perpendicularity to the linear algebra of bilinear forms.

4.13 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. Write a short note on Legendre's differential equation.
2. What are Legendre's polynomials?
3. What is the general solution of Legendre's equation in ascending powers of x?
4. Write a short note on Bessel's differential equation.
5. What is Bessel's function of the first kind of order n?
6. What is modified Bessel's equation?
7. Derive differentiation formula for Laguerre polynomials.
8. What are the solutions of Laguerre equation of various degrees for v to be a positive integer?

Long Answer Questions

1. Find a convergent solution of Legendre's equation in descending powers of x.
2. Drive recurrence formula for $P_n(\mu)$.
3. Prove the orthogonality of Legendre's polynomials of the first kind.
4. Drive Bessel's function of the first kind of order n from Bessel's differential equation.
5. Discuss generating function for Bessel's function.
6. Explain recurrence relation for Bessel's function.
7. Describe modified Bessel's function.

8. Show that $H_n(x)$ is a solution for Hermite equation.
9. Derive the two independent solutions that a Hermite equation may have.
10. Describe orthogonal properties of Hermite polynomials.
11. Describe orthogonal properties of Laguerre polynomials.
12. Show that $L_n(x)$ is a polynomial of degree n in x and that the coefficient of x^n is $(-1)^n$.
13. Explain associated Laguerre polynomials.

4.14 FURTHER READING

Dass, HK. 2008. *Mathematical Physics*. New Delhi: S. Chand

Chattopadhyay, PK. 1990. *Mathematical Physics*. New Delhi: New age International

Hassani, Sadri. 2013. *Mathematical Physics: A Modern Introduction to Its Foundations* Berlin: Springer Science & Business Media

Khanna, V.K. and S.K. Bhambri. 2008. *A Course in Abstract Algebra*, 3rd edition. New Delhi: Vikas Publishing House

Grewal, B.S. 1998. *Higher Engineering Mathematics,* 34th edition. New Delhi: Khanna Publishers

Narayan, Shanti. 1996. *Differential Calculus,* 14th edition. New Delhi: S. Chand and Company

Kreyszig, Erwin. *Advanced Engineering Mathematics,* 7th edition. Mumbai: Wiley Eastern Ltd

Bali, N.P. 2007. *A Textbook of Engineering Mathematics*. New Delhi: Laxmi Publications (P) Ltd

Narayan, Shanti. 2010. *A Text Book of Vector Analysis.* New Delhi: S. Chand and Sons

Spiegel. 1959. *Vector Analysis (Schaum'S Outline).* New York: McGraw Hill Professional

Gupta, B. D. 2009. *Mathematical Physics*, 4th Edition. Noida: Vikas Publishing House Pvt. Ltd.

UNIT 5 INTEGRAL TRANSFORM

NOTES

Structure

5.0 INTRODUCTION

In this unit, you will learn how Fourier and Laplace transformations are an important operational method for solving linear differential equations. These are of practical use in solving initial value problems connected with linear differential equations. The theory of Laplace transforms will help you to understand graphically how a continuous function will be continuous except for a finite number of jump-type discontinuities in the interval in which it is defined. You will also study about the applications of Laplace transform to potential and oscillatory problems and the evaluation of simple integrals using Fourier and Laplace transforms.

5.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe Fourier and Laplace transforms along with their properties
- Discuss the application of Fourier transform to Dirac delta function and potential problems
- Explain the applications of Laplace transform to potential and oscillatory problems
- Evaluate simple integrals using both Fourier and Laplace transforms
- Discuss the applications of Laplace transform to potential and oscillatory problems
- Evaluate simple integrals using Fourier and Laplace transforms

5.2 INTRODUCTION TO INTEGRAL TRANSFORM

NOTES

There are many classes of problems which are not easy to work out in their original representation. The fundamental idea behind integral transforms is to map these functions from their original domain to some other domain. Integral transform is a type of mathematical operator that results when a given function, *f(x)*, is multiplied by a kernel function, *K(p,x)*, and the product is integrated between suitable limits.

$$F(p)=\int_{t_1}^{t_2} K(p,x)f(x)dx,$$

It converts the function into sums of much simpler functions that are much easy to solve. After solving the problem the solution is mapped back to its original domain.

Transform of Integrals

The following results show that the Laplace transforms of the derivatives and integrals of a function $f(t)$ can be expressed in terms of the Laplace transform of $f(t)$. These results are important in solving differential equations using the methods of Laplace transformation.

Theorem 5.1: If $f(t)$ is continuous and $f'(t)$ is piecewise continuous in the interval $0 \le t \le T$ for any finite T, and $f(t)$ and $f'(t)$ are of exponential order as $t \to \infty$, then,

$$L[f'(t)] = sL[f(t)] - f(0).$$

Proof: Under the conditions stated in the theorem, the Laplace transforms of $f(t)$ and $f'(t)$ exist and,

$$L[f'(t)] = \int_0^\infty e^{-st} f'(t)dt$$

$$= \int_0^\infty e^{-st} d[f(t)]$$

$$= \left[e^{-st} f(t)\right]_0^\infty - \int_0^\infty f(t)\cdot(-s)e^{-st}\,dt$$

On integration by parts, we get $L[f'(t)]$ to be,

$$= \lim_{t\to\infty} e^{-st} f(t) - f(0) + s\int_0^\infty e^{-st} f(t)dt = -f(0) + sL[f(t)]$$

Since $\lim_{t\to\infty} e^{-st} f(t) = 0$, as shown below (see Notes).

$$\therefore \quad L[f'(t)] = sL[f(t)] - f(0)$$

Notes:

1. As $f(t)$ is of exponential order at $t \to \infty$, there exist constants α and M such that,

$$|f(t)| \le Me^{\alpha t}, \text{ for } t \ge t_0$$

$$[|f(t)|/e^{st}] \le Me^{[(s-\alpha)t]}, \text{ for } t \ge t_0$$

Now, $e^{-(s-\alpha)t} \rightarrow 0$, as $t \rightarrow \infty$, if $s > \alpha$

$\therefore$ $\lim_{t\to\infty} \frac{|f(t)|}{e^{st}}$, as also $\lim_{t\to\infty} \frac{f(t)}{e^{st}}$ vanishes if, $s \rightarrow \infty$

NOTES

2. Although $f(t)$ is of exponential order, it cannot be said that the derivatives of $f(t)$ will also be of exponential order. However, in most practical cases, the functions considered and their derivatives are all of exponential order.

Theorem 5.2: If $f(t)$ and $f'(t)$ are continuous and $f''(t)$ is piecewise continuous in $0 \le t \le T$, for any finite T, and $f(t)$ and $f'(t)$ are of exponential order as $t \rightarrow \infty$, then

$$L[f''(t)] = s^2 L[f(t)] - sf(0) - f'(0)$$

Proof: By applying the previous theorem to the function $f'(t)$, we have,

$$L[f''(t)] = s^2 L[f(t)] - sf(0) - f'(0)$$

Again applying the previous theorem to write $L[f''(t)]$ in terms of $L[f'(t)]$, we get,

$$L[f''(t)] = s[sL\{f(t)\} - f(0)] - f''(0)$$

$$L[f''(t)] = sL[f'(t)] - f'(0)$$

Notes:

1. Conditions stated in the above theorem ensure that the Laplace transforms of $f(t), f'(t)$ and f''(t) exist.
2. The result concerning the Laplace transforms of $f^n(t)$ is as follows:

   If $f(t), f'(t),..., f^{(n-1)}(t)$ are continuous and $f^n(t)$ is piecewise continuous in the interval $0 \le t \le T$ for any finite T and all these functions are of exponential order as $t \rightarrow \infty$, then,

$$L[f^{(n)}(t)] = s^n L[f(t)] - s^{(n-1)} f(0) - s^{(n-2)} f'(0) - s^{(n-3)} f''(0) - \ldots - sf^{(n-2)}(0) - f^{(n-1)}(0)$$

   This result is obtained by successive application of the result,

$$L[f'(t)] = sL[f(t)] - f(0)$$

Example 5.1: Given that $L[t \sin at] = [(2as)/(s^2 + a^2)^2]$, find $L[at \cos at + \sin at]$.

Solution:

$$L[at \cos at + \sin at] = L\left[\frac{d}{dt}(t \sin at)\right]$$

$$= sL[t \sin at] - [t \sin at]_{t=0} = s\frac{2as}{(s^2 + a^2)^2}$$

$$\therefore \quad L[at \cos at + \sin at] = \frac{2as^2}{(s^2 + a^2)^2}$$

Example 5.2: Using the result $L[f'] = sL[f] - f(0)$ and $L[f''] = s^2L[f] - sf(0) - f'(0)$, find $L[e^{at}]$, $L[\sin at]$ and $L[\cos at]$.

Solution:

To find $L(e^{at})$, take $f(t) = e^{at}$ in the result $L(f') = sL(f) - f(0)$. Then, $L(e^{at}) = sL(e^{at}) - 1$ i.e., $aL(e^{at}) - sL(e^{at}) = -1$

NOTES

$\therefore \quad L(e^{at}) = \dfrac{1}{s-a}$

Taking $f(t) = \sin at$ in the result $L[f''] = s^2L[f] - sf(0) - f'(0)$, we get,

$$L[-a^2 \sin at] = s^2L[\sin at] - s(0) - a(1)$$

i.e., $\quad -a^2L(\sin at) - s^2\text{L}(\sin at) = -a$

$$\therefore \quad L(\sin at) = \frac{a}{s^2+a^2}$$

Finding L(cos at) in a similar manner is left as an exercise for the students.

If $L[f(t)] = \bar{f}$ (s), prove that,

$$L[tf'(t)] = -[sf'(s) + f(s)]$$

Theorem 5.3: If $f(t)$ is of exponential order as $t \to \infty$ and piecewise continuous in the interval $0 \le t \le T$, for any finite T, then,

$$L\left[\int_0^t f(u)du\right] = \frac{1}{s}L[f(t)]$$

Proof: This result can be proved using the result,

$$L[g'(t) = sL[g(t)] - g(0) \qquad ...(5.1)$$

Let $\int_0^t f(u)du$ be denoted as $g(t)$. Then,

$$g'(f) = f(t) \quad \text{and} \quad g(0) = \int_0^0 f(u)du = 0$$

It can be shown that $g(t)$ is continuous in $0 \le t \le T$ and is of exponential order as $t \to \infty$. Therefore, Laplace transforms of both $f(t)$ and $g(t)$ exist and by equation (5.1),

$$L[f(t)] = sL\left[\int_0^t f(u)du\right] - 0$$

$$\therefore \quad L\left[\int_0^t f(u)du\right] = \frac{1}{s}L[f(t)]$$

Notes:

1. Replacing the dummy variable u by t, the above result is written in form

$$L\left[\int_0^t f(t)dt\right] = \frac{1}{s}L[f(t)]$$

2. Under the conditions stated in the theorem,

$$L\left[\int_0^t\int_0^t\cdots\int_0^t f(t)(dt)^n\right] = \frac{1}{s^n}L[f(t)], \text{ for any positive integer } n.$$

3. $\int_a^t f(u)du = \int_a^t f(u)du - \int_a^t f(u)du$

$$\therefore \quad L\left[\int_a^t f(u)du\right] = L\left[\int_0^t f(u)du\right] - L\left[\int_0^a f(u)du\right]$$

$$= \frac{1}{s}L[f(t)] - \frac{1}{s}\int_0^a f(u)du$$

Since, $\int_0^a f(u)du$ is a constant.

$$\therefore \qquad L\left[\int_a^t f(u)du\right] = \frac{1}{s}L[f(t)] + \frac{1}{s}\int_0^a f(u)du$$

NOTES

Example 6.3: Find $L\int_0^t [(\sin x)/x]dx$

Solution: $$L\left[\frac{\sin t}{t}\right] = \int_s^\infty L(\sin t)ds = \int_s^\infty \frac{1}{s^2+1}ds = \left[\tan^{-1}(s)\right]_s^\infty$$

$$= \frac{\pi}{2} - \tan^{-1} s = \cot^{-1} s$$

$$\therefore \qquad L\left[\int_0^t \frac{\sin x}{x}dx\right] = \frac{1}{s}L\left[\frac{\sin t}{t}\right] = \frac{1}{s}\cot^{-1} s$$

5.3 FOURIER TRANSFORM AND ITS PROPERTIES

In this section, we will study about Fourier transforms and its properties.

[A] **Fourier Sine Transforms.**

They can be subdivided in two, namely, the infinite Fourier sine transform and the Finite Fourier sine transforms.

[a_1] *The Infinite Fourier sine Transform* of a function $F(x)$ of x such that $0 < x < \infty$ is denoted by $f_s(n)$, n being a positive integer and is defined as

$$f_s(n) = \int_0^\infty F(x)\sin nx\, dx \qquad \text{... (5.2)}$$

Here $F(x)$ is called as the *Inverse Fourier sine transform of* $f_s(n)$ and defined as

$$F(x) = \frac{2}{\pi}\int_0^\infty f_s(n)\sin nx\, dx \qquad \text{... (5.3)}$$

Thus if $f_s(n) = f_s[F(x)]$, then $F(x) = f_s^{-1}[f_s(n)]$ ... (5.4)

where f is the symbol for Fourier transform and f^{-1} for its inverse.

Problem 5.1. *Find the sine transform of* e^{-x}.

We have

$$f_s(n) = \int_0^\infty e^{-x}\sin nx\, dx = \left[\frac{e^{-x}}{1+n^2}(-\sin nx - n\cos nx)\right]_0^\infty = \frac{n}{1+n^2}.$$

Problem 5.2. *Find the inverse sine transform of* $e^{-\lambda n}$.

We have

$$f_s^{-1}\left[e^{-\lambda n}\right] = \frac{2}{\pi}\int_0^\infty e^{-\lambda n}\sin nx\, dx = \frac{2}{\pi}\left[\frac{e^{-\lambda n}}{\lambda^2+x^2}(-\lambda\sin nx - x\cos nx)\right]_0^\infty$$

$$= \frac{2}{\pi}\cdot\frac{x}{\lambda^2+x^2}.$$

NOTES

[a_2] *The Finite Fourier sine transform* of a function $F(x)$ of x such that $0 < x < l$ is denoted by $f_s(n)$, n being a positive integer and is defined as

$$f_s(n) = \int_0^l F(x) \sin \frac{n\pi x}{l}\, dx \qquad \text{... (5.5)}$$

In case $l = \pi$, this becomes

$$f_s(n) = \int_0^\pi F(x) \sin nx\, dx \qquad \text{... (5.6)}$$

and the inversion formula is

$$F(x) = \frac{2}{\pi} \sum_{n=1}^{\infty} f_s(n) \sin nx \qquad \text{... (5.7)}$$

whence a_n is the coefficient of sin nx in the expansion of $F(x)$ in a sine series and is given by

$$a_n = \frac{2}{\pi} \int_0^\pi F(x) \sin nx\, dx$$

$$= \frac{2}{\pi} f_s(n) \text{ by (5)} \qquad \text{... (5.8)}$$

Problem 5.3. *Find the Fourier sine transform of $F(x) = x$ such that $0 < x < 2$.*

We have $f_s(n) = \int_0^2 F(x) \sin \frac{n\pi x}{2}\, dx$

$\because\ l = 2$ in the existing case.

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

$$= \left[x \cdot \frac{-\cos \frac{n\pi x}{2}}{\frac{n\pi}{2}} \right]_0^2 + \int_0^2 \frac{2}{n\pi} \cos \frac{n\pi x}{2}\, dx$$

(on integrating by parts)

$$= \left[\frac{-2x}{n\pi} \cos \frac{n\pi x}{2} + \frac{4}{n^2\pi^2} \sin \frac{n\pi x}{2} \right]_0^2 = -\frac{4}{n\pi} \cos n\pi.$$

[**B**] **Fourier Cosine Transforms.**

They can also be subdivided into two, namely, Infinite and Finite cosine transforms.

[b_1] *The Infinite Fourier Cosine Transform* of $F(x)$ for $0 < x < \infty$, is defined as

$$f_c(n) = \int_0^\infty F(x) \cos nx\, dx, n \text{ being a positive integer.} \qquad \text{... (5.9)}$$

Here the function $F(x)$ is called as the *Inverse cosine transform* of $f_c(n)$ and is defined as

$$F(x) = \frac{2}{\pi} \int_0^\infty f_c(n) \cos nx\, dx \qquad \text{... (5.10)}$$

Thus if $f_c(n) = f_c[F(x)]$, then $F(x) = f_c^{-1}[f_c(n)]$ ... (5.11)

Problem 5.4. *Find the cosine transform of $x^n e^{-ax}$.*

We have $\int_0^\infty e^{-ax} \cos nx\, dx = \frac{a}{a^2 + n^2}$ and $f_c(n) = \int_0^\infty x^n e^{-ax} \cos nx\, dx$

Differentiating the first relation n times w.r.t. 'a' we find

$$\int_0^\infty x^n e^{-ax} \cos nx\, dx = (-1)^n \frac{d^n}{da^n} \left(\frac{a}{a^2 + n^2} \right)$$

$$= \frac{\lfloor n \cos\left\{(n+1)\tan^{-1}\frac{n}{a}\right\}}{\left(a^2+n^2\right)^{(n+1)/2}} \text{ by usual method.}$$

NOTES

Hence $$f_c(n) = \frac{\lfloor n \cos\left\{(n+1)\tan^{-1}\frac{n}{a}\right\}}{\left(a^2+n^2\right)^{(n+1)/2}}$$

Problem 5.5. *Find* $f_c^{-1}\ \{e^{-\lambda n}\}$

We have $f_c^{-1}\left\{e^{-\lambda n}\right\} = \frac{2}{\pi}\int_0^{\infty} e^{-\lambda n} \cos nx\, dx$

$$= \frac{2}{\pi}\left[\frac{e^{-\lambda n}}{\lambda^2+x^2}(-\lambda\cos nx + x\sin nx)\right]_0^{\infty} = \frac{2}{\lambda^2+x^2}$$

[b_2] *The Finite Fourier cosine transform* of $F(x)$ for $0 < x < l$ is defined as

$$f_c(n) = \int_0^l F(x)\cos\frac{n\pi x}{l}\,dx \quad \text{... (5.12)}$$

When $l = \pi$, this becomes

$$f_c(n) = \int_0^{\pi} F(x)\cos nx\,dx \quad \text{... (5.13)}$$

and the inversion formula is

$$F(x) = \frac{1}{\pi} f_c(0) + \frac{2}{\pi}\sum_{n=1}^{\infty} f_c(n)\cos nx \quad \text{... (5.14)}$$

when $$f_c(0) = \int_0^{\pi} F(x)\,dx \quad \text{... (5.15)}$$

Also b_n the coefficient of cos nx in the expansion of $F(x)$ in a cosine series is given by

$$b_n = \frac{2}{\pi}\int_0^{\pi} F(x)\cos nx\,dx = \frac{2}{\pi} f_c(n) \text{ by (12)} \quad \text{... (5.16)}$$

Problem 5.6. *Find the finite Fourier cosine transform of x.*

We have $f_c(n) = \int_0^{\pi} x\cos nx\,dx$

$$= \left[\frac{x\sin nx}{n}\right]_0^{\pi} - \frac{1}{n}\int_0^{\pi}\sin nx\,dx \text{ (on integrating by parts)}$$

$$= 0 - \frac{1}{n}\left[\frac{\cos nx}{-n}\right]_0^{\pi}$$

$$= \frac{1}{n^2}\left\{(-1)^n - 1\right\},$$

$n = 1, 2, 3,...$

But if $n = 0$, $f_c(0) = \int_0^{\pi} x\,dx = \left[\frac{x^2}{2}\right]_0^{\pi} = \frac{\pi^2}{2}.$

Note. On the next page are tabulated some useful Fourier sine and cosine transforms in a concise form.

[**C**] **The Complex Fourier Transforms.**

The Complex Fourier Transform of a function $F(x)$ for $-\infty < x < \infty$, is defined as

$$f(n) = \int_{-\infty}^{\infty} F(x)\,e^{inx}\,dx \quad \text{... (5.17)}$$

where e^{inx} is said to be the *Kernel* of the transform.

The inversion formula is $F(x) = \dfrac{1}{2\pi}\int_{-\infty}^{\infty} f(n)\, e^{-inx}\, dn$ ... (5.18)

Problem 5.7. *Find the Fourier Complex Transform of*

NOTES

$$F(x) = \begin{cases} 1 - x^2, & |x| < 1 \\ 0, & |x| < 1 \end{cases}$$

We have $f(n) = \int_{-1}^{1} (1 - x^2)\, e^{inx}\, dx = \left[(1 - x^2)\dfrac{e^{inx}}{in} + \dfrac{2}{in}\int_{-1}^{1} x\, e^{inx}\, dx \right]$

Infinite Fourier Sine Transforms $f_s(n) = \int_0^\infty F(x)\sin nx\,dx,\ F(x) = \frac{2}{\pi}\int_0^\infty f_s(n)\sin x\,dn$		*Finite Fourier Sine Transforms* $f_s(n) = \int_0^\pi F(x)\sin nx\,dx,\ F(x) = \frac{2}{\pi}\sum_{n=1}^{\infty} f_s(n)\sin nx$	
$F(x)$	$f_c(n)$	$F(x)$	$f_c(n)$
e^{-x}	$n/(1+n^2)$	1	$\frac{1}{n}\left[1+(-1)^{n+1}\right]$
$xe^{-x^2}/2$	$\sqrt{\pi/2}\, ne^{-n^2/2}$	x	$(-1)^{n+1}\cdot \pi/n$
$\frac{\sin x}{x}$	$\frac{1}{2}\log\left\|\frac{1+n}{1-n}\right\|$	$\left.\begin{matrix} x, 0 < x \le \pi/2 \\ \pi - x, \pi/2 \le x < \pi \end{matrix}\right\}$......	$\frac{2}{n^2}\sin\frac{n\pi}{2}$
$\left.\begin{matrix} x\left(1-x_0^2\right)2, 0 < x < 1, v > -1 \\ 0 \quad x > 1 \end{matrix}\right\}$	$\sqrt{\pi}\,2^{v-1/2}\,\Gamma(v+1)\, n^{-1/2^{-v}} J_{v+3/2}(n)$	$1 - \frac{x}{\pi}$	$\frac{1}{n}$
$x^{\mu-1}, 0 < \mu < 1$	$\sqrt{\mu}\sin\left(\frac{\mu\pi}{2}\right) n^{-\mu}$	x^2.......	$\frac{\pi^2}{n}(-1)^{n-1} - \frac{2}{n^3}\left[1-(-1)^n\right]$
$x^{v'} e^{-\mu_x}$	$\frac{2^v \mu^v (\Gamma v) n}{(\mu^2+n^2)^{v+1}}$	$x(\pi - x)$	$\frac{2}{n^3}\left[1-(-1)^n\right]$
		x^3...	$(-1)^n \pi[6/n^3 - \pi^2/n]$
$\left.\begin{matrix} 0, \quad 0 < x < a \\ \{x^2 - a^2\}^{-1/2}, x > a \end{matrix}\right\}$	$J_0(an)$	$x(\pi^2 - x^2)$	$(-1)^{n+1}\, 6\pi/n^3$
		$e^{\lambda x}$.......	$\frac{n}{n^2+\lambda^2}\left[1+(-1)^n e^{\lambda\pi}\right]$
		$\cos \lambda x$	$\frac{n}{n^2-\lambda^2}\left[1-(-1)^n \cos\lambda\pi\right]$
		$\cos mx,\ m = 1, 2, 3,$......	$\begin{cases} \frac{n}{n^2-m^2}\left[1-(-1)^{n+m}\right], & n \ne m \\ 0, & n = m \end{cases}$
		$\sin mx,\ m = 1, 2, 3,$......	$\begin{cases} 0, & n \ne m \\ \pi/2, & n = m \end{cases}$
		$\frac{\sin\lambda(\lambda - x)}{\sin\lambda\pi}$	$\frac{n}{n^2-\lambda^2}, \lambda \ne 0, 1, 2,$...
		$\frac{\sinh\lambda(\pi - x)}{\sinh\lambda\pi}$	$\frac{n}{\lambda^2+n^2}$

NOTES

Infinite Fourier Cosine Transforms $f_c(n) = \int_0^{\infty} F(x)\cos nx\, dx,\ F(x) = \frac{2}{\pi}\int_0^{\infty} f_c(n)\cos nx\, dn$		*Finite Fourier Cosine Transforms* $f_c(n) = \int_0^{\pi} F(x)\cos nx\, dx,\ F(x) = \frac{1}{\pi} f_c(0) + \sum_{n=1}^{\infty} f_c(n)\cos nx$	
$F(x)$	$f_c(n)$	$F(x)$	$f_c(n)$
$\left.\begin{matrix}1, 0 < x < a \\ 0,\ x > a\end{matrix}\right\}$......	$\frac{\sin na}{n}$	1	$\begin{cases}\pi, & n = 0 \\ 0 & n = 1, 2, 3, \ldots\end{cases}$
$x^{\mu-1}$, $0 < \mu < 1$	$\sqrt{\mu}\, n^{-\mu} \sin \frac{\mu\pi}{2}$	$\left.\begin{matrix}1, & 0 < x < \pi/2 \\ -1, & \pi/2 < x < \pi\end{matrix}\right\}$......	$\begin{cases}0, n = 0 \\ 2/\pi \sin n\pi/2, n = 1, 2, 3, \ldots\end{cases}$
e^{-x}......	$\frac{1}{1+n^2}$	x	$\begin{cases}\pi^2/2, n = 0 \\ 1/n^2\left[(-1)^n - 1\right], n = 1, 2, 3, \ldots\end{cases}$
Sech $x\pi$	$\sqrt{\frac{\pi}{2}}\frac{1}{1+n^4}$	x^2......	$\begin{cases}\pi^3/3, n = 0 \\ 2\pi/n^2\,(-1)^n, n = 1, 2, 3, \ldots\end{cases}$
e^{-x^2}......	$\sqrt{\frac{\pi}{2}}\, e^{-n^2/2}$	x^3	$\begin{cases}\pi^4/4, n = 0 \\ 3\pi^2/n^2(-1)^n + 6/n^4\left[(-1)^n - 1\right]. \\ \quad n = 1, 2, 3\ldots\ldots\end{cases}$
$\cos x,\ \left.\begin{matrix}0 < x < a \\ 0, x > a\end{matrix}\right\}$	$\frac{1}{2}\left[\sin\frac{a(1-n)}{1-n} + \sin\frac{a(1+n)}{1+n}\right]$	$\left(1 - \frac{x}{\pi}\right)^2$......	$\begin{cases}\pi/3, & n = 0 \\ 2/\pi n^2, & n = 1, 2, 3, \ldots\end{cases}$
$\sin\frac{x^2}{2}$......	$\sqrt{\frac{\pi}{2}}\left[\cos\frac{n^2}{2} - \sin\frac{n^2}{2}\right]$	$e^{\lambda x}$	$\frac{\lambda}{\lambda^2+n^2}\left[(-1)^n e^{\lambda x} - 1\right]$
$\cos\frac{x^2}{2}$	$\sqrt{\frac{\pi}{2}}\left[\cos\frac{n^2}{2} + \sin\frac{n^2}{2}\right]$	$\sin \lambda x$......	$\frac{\lambda}{n^2-\lambda^2}\left[(-1)^n \cos\lambda\pi - 1\right], n \neq \lambda$
$\left.\begin{matrix}(1-x^2)^{\nu}, 0 < x < 1 \\ 0, x > 1 \text{ and } \nu > -3/2\end{matrix}\right\}$......	$\sqrt{\pi}\, 2^{\nu-1/2}\, \Gamma(\nu+1)\cdot n^{-\nu-1/2}\, J_{\nu+1/2}(n)$	$\sin mx$, $m = 1, 2, 3, \ldots\ldots$	$\begin{cases}0, & n = m \\ m/(n^2 - m^2)\left[(-1)^{n+m} - 1\right], & n \neq m\end{cases}$
		$\frac{\cos\lambda(\pi - x)}{\sinh \pi\lambda}$	$\frac{\lambda}{\lambda^2+n^2}$

(on integrating by parts)

$$= 0 + \frac{2}{in}\left[\frac{x\,e^{inx}}{in}\right]_{-1}^{1} - \frac{2}{(in)^2}\int_{-1}^{1} e^{inx}\, dx$$

$$= \frac{2}{-n^2}\left[e^{in} + e^{-in}\right] + \frac{2}{n^2}\left[\frac{e^{inx}}{in}\right]_{-1}^{1} = -\frac{2}{n^2}\,(e^{in} + e^{-in}) + \frac{2}{in^3}\,(e^{in} - e^{-in})$$

$$= -\frac{4}{n^2}\cos n + \frac{4}{n^3}\sin n = -\frac{4}{n^3}(n\cos n - \sin n).$$

Problem 5.8. *Find the Complex Fourier transform of* $e^{-|x|}$ *and then invert it.*

NOTES

(*Rohilkhand, 1982, 86*)

We have $f(n) = \int_{-\infty}^{\infty} e^{-|x|}\, e^{inx}\, dx = \int_{-\infty}^{0} e^{(1+in)x}\, dx + \int_{0}^{\infty} e^{-(1-in)x}\, dx$

$$= \frac{1}{1+in} + \frac{1}{1-in} = \frac{2}{1+n^2}$$

so that the inversion formula gives,

$F(n) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{2}{1+n^2}\, e^{inx}\, dx = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{e^{-inx}}{1+n^2}\, dx$ which may be integrated by contour Integration.

Note. Several other **Complex Fourier Transforms** have been tabulated on the next page.

[D] Parseval's Identity for Fourier Integrals.

It is stated as $\int_{-\infty}^{\infty} |F(x)|^2\, dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} |f(n)|^2\, dn$ ...(5.19)

where $f(n)$ is the Fourier transform of $F(x)$.

Problem 5.9. *Find the Fourier transform of*

$$F(x) = \begin{cases} 1, & |x| < a \\ 0, & |x| > a \end{cases}$$ (*Agra, 1982; Kanpur, 1970*)

Hence or otherwise evaluate $\int_0^{\infty} \frac{\sin^2 nx}{n^2}\, dx.$

We have $f(n) = \int_{-\infty}^{\infty} F(x)\, e^{-inx}\, dx = \int_{-a}^{a} 1\cdot e^{-inx}\, dx$

$$= \left[\frac{e^{-inx}}{-in}\right]_{-a}^{a} = \frac{e^{+ina} - e^{-ina}}{in} = \frac{2\sin na}{n} \text{ for } n \neq 0$$

For $n = 0, f(n) = 2\lim_{n\to 0}\frac{\sin na}{n} = 2\lim_{n\to 0}\frac{1}{n}\left[na - \frac{1}{\lfloor 3}n^3a^3 + ...\right] = 2a$

Now using Parseval's identity, we find

$$\int_{-a}^{a} 1^2\cdot dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{4\sin^2 na}{n^2}\, dn \text{ when } n \neq 0$$

i.e., $$\frac{1}{2\pi}\cdot 2\int_0^{\infty}\frac{4\sin^2 na}{n^2}\, dn = [x]_{-a}^{a} = 2a$$

$$\therefore \quad \int_0^{\infty}\frac{\sin^2 nx}{n^2}\, dx = \frac{\pi x}{2}.$$

[E] Relation between the Fourier Transform of the Derivatives of a Function.

If $f(n)$ be the Fourier transform of $F(x)$, then we have to express the Fourier transform of the function $\frac{d^mF}{dx^m}$ in terms of $f(n)$.

NOTES

$F(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\, e^{-inx}\, dx$	$f(n)=\int_{-\infty}^{\infty} F(x)\, e^{inx}\, dx$	$F(x)$	$f(n)$
$\frac{1}{\lvert x\rvert}$	$\frac{1}{\lvert n\rvert}$	$\frac{\sinh ax}{\sin \pi x}, -\pi < a < \pi$	$\frac{\sinh a}{\cosh n + \cos a}$
$\left.\begin{matrix} e^{i\omega t}, a < x < b \\ 0, x < a, x > b \end{matrix}\right\}$......	$i\,\frac{e^{ia(\omega+n)} - e^{ib(\omega+n)}}{n}$	$\frac{\cosh ax}{\cosh \pi x}, -\pi < a < \pi$	$2\,\frac{\cos a/2 \cosh a/2}{\cosh n + \cos a}$
$\left.\begin{matrix} e^{-\lambda x + i\omega x}, x > 0 \\ 0 \quad, x < 0 \end{matrix}\right\}$......	$\frac{i}{\omega + n + i\lambda}$	$\frac{\sin\{b(a^2+x^2)^{1/2}\}}{(a^2+x^2)^{1/2}}$......	$\left\{\begin{matrix} 0, \lvert n\rvert > b \\ \pi J_0\left[a\sqrt{b^2-n^2}\right], \lvert n\rvert < b \end{matrix}\right.$
$\frac{e^{-\lambda}\lvert x\rvert}{\lvert x\rvert^{1/2}}$......	$\frac{\left[(\lambda^2+n^2)^{1/2}+\lambda\right]^{1/2}}{(\lambda^2+n^2)^{1/2}}$	$\left.\begin{matrix} \frac{\cosh\left\{b(a^2-x^2)^{1/2}\right\}}{(a^2-x^2)^{1/2}}, \lvert x\rvert < a \\ 0 \quad , \quad \lvert x\rvert > a \end{matrix}\right]$	$\pi J_0\left[a\sqrt{n^2+b^2}\right]$
$e^{-\lambda x^2}, R(\lambda) > 0$	$\sqrt{\frac{\pi}{\lambda}}, e^{-n^2/4\lambda}$	$\begin{matrix} \frac{\cos\left\{b(a^2-x^2)^{1/2}\right\}}{\sqrt{a^2-x^2}}, \lvert x\rvert < a \\ 0 \quad , \quad \lvert x\rvert > a \end{matrix}$	$\pi J_0\left[a\sqrt{n^2-b^2}\right]$
$\sin \lambda x^2$...	$\sqrt{\frac{\pi}{\lambda}} \sin\left(\frac{n^2}{4\lambda}+\frac{\pi}{4}\right)$		
$\cos \lambda x^2$	$\sqrt{\frac{\pi}{\lambda}} \cos\left(\frac{n^2}{4\lambda}-\frac{\pi}{4}\right)$	$\left.\begin{matrix} P_\nu(x), \lvert x\rvert < 1 \\ 0 \quad, \lvert x\rvert > 1 \end{matrix}\right\}$......	$i^\nu \sqrt{2}\, J_{\nu+1/2}(n)$
$\left.\begin{matrix} \left(a^2-x^2\right)^{-1/2} \lvert x\rvert < a \\ 0, \quad \lvert x\rvert > a \end{matrix}\right\}$......	$\pi J_0 (an)$		

We have by the definition of Fourier-transform,

$$f\left[\frac{d^m F}{dx^m}\right] = \int_{-\infty}^{\infty} \frac{d^m F}{dx^m}\, e^{inx}\, dx = f^m(n) \text{ (say)} \qquad \ldots (5.20)$$

so that $\qquad f^m(n)=$

$$\int_{-\infty}^{\infty} \frac{d^m F}{dx^m}\, e^{inx}\, dx = \left[\frac{d^{m-1} F}{dx^{m-1}}\, e^{inx}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} (in)\, e^{inx}\, \frac{d^{m-1}F}{dx^{m-1}}\, dx$$

(on integrating by parts)

$$= - in \int_{-\infty}^{\infty} \frac{d^{m-1} F}{dx^{m-1}} d^{inx}\ dx,$$

NOTES

under the assumption $\frac{d^{m-1}F}{dx^{m-1}} \to 0$ as $|x| \to \infty$

$= \quad - \text{in} f^{n-1}(n)$ by (5.20) ... (5.21)

Repeating the same process under the assumption

$$\frac{d^r F}{dx^r} \to 0 \text{ as } |x| \to \infty,\ r = 1, 2, 3,\ldots (m-1)$$

we get after $(m - 1)$ operations,

$$f^m(n) = (-in)^m f(n) \qquad \ldots (5.22)$$

which follows that the Fourier transform of $\frac{d^m F}{dx^m}$ *is* $(-in)^m$ *times the Fourier transform of* $F(x)$ *subject to the condition that* $\frac{d^r F}{dx^r} \to 0$ *when* $|x| \to \infty$, *for* $r = 1, 2, 3,\ldots(m - 1)$.

By similar procedure we can find a relation between the sine and cosine Fourier transforms of the derivatives of a function, such as

$$f_c^m(n) = \int_0^{\infty} \frac{d^m F}{dx^m} \cos nx\, dx = \left[\frac{d^{m-1}F}{dx^{m-1}} \cos nx \right]_0^{\infty} + n \int_0^{\infty} \frac{d^{m-1}F}{dx^{m-1}} \sin nx\, dn$$

integrating by parts

$$= - \alpha_{n-1} + f_s^{m-1}(n) \qquad \ldots (5.23)$$

Under the assumptions,

$$\frac{d^{m-1}F}{dx^{m-1}} \to 0 \text{ as } x \to \infty \text{ and } \frac{d^{m-1}F}{dx^{m-1}} \to \alpha_{n-1} \text{ as } x \to 0.$$

Similarly, integrating, $f_s^m(n) = \int_0^{\infty} \frac{d^m F}{dx^m} \sin nx\, dx$

$$= - n f_c^{m-1}(n) \qquad \ldots (5.24)$$

(5.23) and (5.24) yield,

$$f_c^m(n) = - \alpha_{n-1} - n^2 f_c^{m-2}(n) \qquad \ldots (5.25)$$

Repeating the procedure $f_c^m(n)$ may be expressed as the sum of α'^s and either $f_c(n)$ or $f_c'(n)$ or $f_c'(n)$ or $f_c'(n)$. $f_c(n)$ will occur when x is odd and in that case we can write $\alpha_0 + n f_s(n)$ place of $f_c'(n)$. We thus have

$$f_c^{2m}(n) = -\sum_{r=0}^{m-1} (-1)^r \alpha_{2m-2r-1} n^{2r} + (-1)^m n^{2m} f_c(n) \qquad \ldots (5.26)$$

and
$$f_c^{2m+1}(n) = -\sum_{r=0}^{m} (-1)^r \alpha_{2m-2r} n^{2r} + (-1)^m n^{2m+1} f_s(n) \qquad \ldots (5.27)$$

Similar procedure with the help of (5.23) and (5.24), will yield

$$f_s^m(n) = n\alpha_{n-2} - n^2 f_s^{m-2}(n) \qquad \ldots (5.28)$$

$$f_s^{2m}(n) = -\sum_{r=1}^{m} (-1)^r n^{2r-1} \alpha_{2n-2r} + (-1)^{m+1} n^{2m} f_s(n) \qquad \ldots (5.29)$$

and
$$f_s^{2m+1}(n) = -\sum_{r=1}^{m} (-1)^r n^{2r-1} \alpha_{2n-2r+1} + (-1)^{m+1} n^{2m+1} f_s(n) \qquad \ldots (5.30)$$

NOTES

Note 1. The following results are easily deducible

(*i*) $\int_0^\infty \frac{d^2F}{dx^2} \cos nx\, dx = -n^2 f_c(n)$

(*ii*) $\int_0^\infty \frac{d^4F}{dx^4} \cos nx\, dx = -n^4 f_c(n)$

(*iii*) $\int_0^\infty \frac{d^2F}{dx^2} \sin nx\, dx = n^2 f_s(n)$

(*iv*) $\int_0^\infty \frac{d^4F}{dx^4} \sin nx\, dx = n^4 f_c(n)$

(*v*) $\int_0^\infty \frac{\partial F}{\partial t} \sin nx\, dx = \frac{\partial}{\partial t}\int_0^\infty F \sin nx\, dx = \frac{\partial f}{\partial t}$ ... (5.31)

Note 2. In case the transforms are finite, then consider

$\int_0^\pi \frac{\partial F}{\partial x} \sin nx\, dx = [F(x)\sin nx]_0^\pi - n\int_0^\pi F(x)$ $\cos nx\, dx,$ integrating by parts

$$= -n f_c(n) \qquad \text{... (5.32)}$$

under the assumption that $F(0)$ and $F(\pi)$ both are finite.

Similarly, $\int_0^\pi \frac{\partial F}{\partial x} \cos nx\, dx = [F(x)\cos nx]_0^\pi + n\int_0^\pi F(x)\sin nx\, dx$

$$= (-1)^n F(\pi) - F(0) + n f_s(n) \qquad \text{... (5.33)}$$

Assuming that $F(x) \to 0$ at $x = \pi$ and at $x = 0$, (5.33) reduces to

$$\int_0^\pi \frac{\partial F}{\partial x} \cos nx\, dx = n f_s(n) \qquad \text{... (5.34)}$$

and (5.32) reduces to

$$\int_0^\pi \frac{\partial^2 F}{\partial x^2} \sin nx\, dx = -n\int_0^\pi \frac{\partial F}{\partial x} \cos nx\, dx$$

$$= n[(-1)^{n+1} F(\pi) + F(0)] - n^2 f_s(n) \text{ by (5.33)} \qquad \text{... (5.35)}$$

If $F(0) = F(\pi) = 0$, then (5.35) yields,

$$\int_0^\pi \frac{\partial^2 F}{\partial x^2} \sin nx\, dx = -n^2 f_s(n) \qquad \text{... (5.36)}$$

Similarly (5.33) yields

$$\int_0^\pi \frac{\partial^2 F}{\partial x^2} \cos nx\, dx = (-1)^n F'(\pi) - F'(0) - n^2 f_c(n) \qquad \text{... (5.37)}$$

In case $\frac{\partial^2 F}{\partial x^2}$ vanishes at $x = 0$ and at $x = \pi$, it is easy to see that

$$\int_0^\pi \frac{\partial^4 F}{\partial x^4} \sin nx\, dx = -n^2 \int \frac{\partial^2 F}{\partial x^2} \sin nx\, dx = n^4 f_s(n) \qquad \text{... (5.38)}$$

and when $\frac{\partial F}{\partial x}, \frac{\partial^3 F}{\partial x^3}$ vanish at $x = 0$ and at $x = \pi$, (5.37) gives

$$\int_0^\pi \frac{\partial^2 F}{\partial x^2} \cos nx\, dx = -n^2 f_c(n) \qquad \text{... (5.39)}$$

So that $\int_0^\pi \frac{\partial^4 F}{\partial x^4} \cos nx\, dx = -n^4 f_c(n)$ ... (5.40)

NOTES

Problem 5.10. *Determine the function F such that*

$$\frac{\partial^2 F}{\partial x^2}+\frac{\partial^2 F}{\partial y^2} = 0,\ 0 < x < \pi,$$

with the boundary condition $F = 0$ *when* $x = 0$ *and* $x = \pi$

$= 0$ *when* $y = 0$

$= F_0$ *(const.) when* $y = \pi$.

F being given to be zero when $x = 0$ and $x = \pi$, we have to use the finite sine transform, *i.e.*, $f(n) = \int_0^{\pi} F(x) \sin nx\, dx$

Applying it to the given differential equation we have

$$\int_0^{\pi} \frac{\partial^2 F}{\partial x^2} \sin nx\, dx + \int_0^{\pi} \frac{\partial^2 F}{\partial y^2} \sin nx\, dx = 0$$

with the condition, $f = 0$ when $y = 0$ and $f = \int_0^{\pi} F_0 \sin nx\, dx$ when $y = \pi$

By (5.36) we have, $\int_0^{\pi} \frac{\partial^2 F}{\partial x^2} \sin nx\, dx = -n^2 f$

$$\therefore\ -n^2 f + \frac{\partial^2 F}{\partial y^2} = 0 \text{ where } \frac{\partial^2 f}{\partial y^2} = \int_0^{\pi} \frac{\partial^2 F}{\partial y^2} \sin nx\, dx$$

or $\frac{\partial^2 f}{\partial y^2} - n^2 f = 0.$

Its general solution is $f = A \sinh ny$

But $f = F_0 \int_0^{\pi} \sin nx\, dx$ when $y = \pi$

$$= F_0 \left[-\frac{\cos nx}{n}\right]_0^{\pi} = 0 \text{ when } n \text{ is even}$$

$= -2F_0/n$ when n is odd.

So that considering the two solutions for f we conclude

$f = 0$ when n is even and $f = \frac{2F_0}{n}$ cosech $n\pi$ sinh ny when n is odd.

Hence the inversion formula will give on replacing n by $2m + 1$,

$$F = \frac{4F_0}{\pi} \sum_{m=0}^{\infty} (2m+1)^{-1} \operatorname{cosech} (2m+1)\pi \sinh (2m+1) y \sin (2m+1) x$$

[F] Multiple Fourier Transforms.

If $F(x, y)$ be a function of two variables x and y, then assuming it to be the function of x only, its fourier transform $\phi(n, y)$ is given by

$$\phi(n, y) = \int_{-\infty}^{\infty} F(x, y)\, e^{inx}\, dx \qquad \text{... (5.41)}$$

Now if $f(n, l)$ be the Fourier complex transform of $f(n, y)$ which is regarded as function of y only then

$$f(n, l) = \int_{-\infty}^{\infty} \phi(n, y)\, e^{ily}\, dy \qquad \text{... (5.42)}$$

These two results when combined, give

$$f(n, l) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, e^{i(nx+ly)}\, dx\, dy \qquad \text{... (5.43)}$$

and the inversion formula is

$$f(x, y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(n, l)\, e^{-i(nx+ly)}\, dn\, dl \qquad \text{... (5.44)}$$

NOTES

Similarly in case of three variables x, y, z, we have

$$f(n, l, m) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y, z)\, e^{i(nx+ly+mz)}\, dx\, dy\, dz \quad \text{... (5.43 A)}$$

and $$f(x, y, z) = \frac{1}{8\pi^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(n, l, m)\, e^{-i(nx+ly+mz)}\, dn\, dl\, dm \quad \text{... (5.44 A)}$$

Note 1. The result may be generalized for any number of variables.

Note 2. In case the Fourier transforms are finite such that $F(x, y)$ is a function of two independent variables x, y where $0 \le x \le \pi$ and $0 \le y \le \pi$, then the sine transform of $F(x, y)$ is given by

$$f_s(n, l) = \int_0^{\pi}\int_0^{\pi} F(x, y) \sin nx \sin ly\, dx\, dy \quad \text{... (5.45)}$$

and the inversion formula is

$$F(x, y) = \frac{4}{\pi^2}\sum_{n=1}^{\infty}\sum_{l=1}^{\infty} f_s(n, l) \sin nx \sin ly \quad \text{... (5.46)}$$

[G] Convolution or Faltung Theorem for Fourier Transforms.

If $F(x)$ and $G(x)$ are two functions such that $-\infty < x < \infty$ then their Faltung or Convolution F^*G is defined as

$$H(x) = F^*G = \int_{-\infty}^{\infty} F(n)\, G(x-n)\, dn \quad \text{... (5.47)}$$

It is worth noting that the Fourier Transform of the Convolution of $F(x)$ and $G(x)$ is the product of their Fourier transforms, *i.e.*,

$$f[F^*G] = f[F]\, f[G] \quad \text{... (5.48)}$$

Since $f[F^*G] = \int_{-\infty}^{\infty} H(x)\, e^{-inx}\, dx$ by definition

$$= \int_{-\infty}^{\infty} F(x)\, e^{-inx}\, dx \int_{-\infty}^{\infty} G(x)\, e^{-inx}\, dx$$

$$= f[F].\, f[G].$$

[H] Evaluation of Integrals with the help of Fourier Inversion Theorem.

Let $I_1 = \int_0^{\infty} e^{-ax} \cos nx\, dx$ and $I_2 = \int_0^{\infty} e^{-ax} \sin nx\, dx.$

Integrating by parts, we have

$$I_1 = \left[\left[-\frac{1}{a}e^{-ax}\cos nx\right]_0^{\infty} - \frac{n}{a}\int_0^{\infty} e^{-ax} \sin nx\, dx\right] = \frac{1}{a} - \frac{n}{a} I_2.$$

Similarly $I_2 = \dfrac{n}{a} I_1$

These give on solving $I_1 = \dfrac{a}{a^2+n^2}$ and $I_2 = \dfrac{n}{a^2+n^2}$

Thus taking $F(x) = e^{-ax}$, its sine and cosine Fourier transforms are $\dfrac{a}{a^2+n^2}$ and $\dfrac{n}{a^2+n^2}$ respectively, so that the inversion formula gives

$$e^{-ax} = \frac{2}{\pi}\int_0^{\infty} \frac{a}{a^2+n^2} \cos nx\, dn \quad \text{... (5.49)}$$

and $$e^{-ax} = \frac{2}{\pi}\int_0^{\infty} \frac{a}{a^2+n^2} \sin nx\, dn \quad \text{... (5.50)}$$

i.e., $$\int_0^{\infty} \frac{\cos nx}{a^2+n^2}\, dn = \frac{\pi}{2a} e^{-ax} \text{ and } \int_0^{\infty} \frac{n \sin nx}{a^2+n^2}\, dn = \frac{\pi}{2} e^{-ax} \quad \text{... (5.51)}$$

NOTES

5.4 APPLICATION OF FOURIER TRANSFORM TO DIRAC DELTA FUNCTION AND POTENTIAL PROBLEMS

Here, we will discuss the application of Fourier transform to Dirac delta function and potential problems.

The Dirac delta function

Let us insert the expression $c_m = \frac{1}{L}\int_{\text{one period}} e_m(x)^* f(x)\,dx$. for the Fourier coefficients back into the Fourier series $f(x) = \sum_{n=-\infty}^{\infty} c_n e_n(x)$ for a periodic function $f(x)$:

Interchanging the order of summation and integration, we find

$$f(x) = \int_{\text{one period}} \left[\sum_{n=-\infty}^{\infty} \frac{1}{L} e_n(y)^* e_n(x)\right] f(y)\,dy = \int_{\text{one period}} \delta(x-y)\, f(y)\,dy$$

where we have introduced the Dirac delta function

$$\delta(x-y) \equiv \sum_{n=-\infty}^{\infty} \frac{1}{L} e_n(y)^* e_n(x) = \sum_{n=-\infty}^{\infty} \frac{1}{L} e^{i(x-y)2\pi n/L}$$

Despite its name, the delta function is *not* a function, even though it is a limit of functions. Instead it is a distribution. Distributions are only well-defined when integrated against sufficiently well-behaved functions known as test functions. The delta function is the distribution defined by the condition:

$$\int_{\text{one period}} \delta(x-y)\, f(y)\,dy = f(x)$$

In particular,

$$\int_{\text{one period}} \delta(y)\,dy = 1$$

Hence it depends on the region of integration. This is clear from the above expression which has an explicit dependence on the period L. In the following section, we will see another delta functions adapted to a different region of integration: the whole real line.

5.5 LAPLACE TRANSFORM

If $f(t)$ is a function of t defined for $t \geq 0$ and if the integral $\int_0^{\infty} e^{-st} f(t)dt$ exists, then it is a function of the parameter s.

This function of s, denoted as $\bar{f}(s)$ is called the *Laplace Transform* of $f(t)$ over the range of values of s for which the integral exists.

Laplace transform of $f(t)$ is also denoted as $L[f(t)]$ or simply $L(f)$ or $F(s)$ (i.e., using the upper case letter corresponding to the letter used to denote the function).

$$\therefore \qquad L[f(t)] = \int_0^\infty e^{-st} f(t)dt = F(s)$$

When the integral exists.

The function $f(t)$ is called the *Inverse Laplace Transform* of $\bar{f}(s)$ or simply the *Inverse Transform* of $\bar{f}(s)$ and is denoted by,

$$L^{-1}\{\bar{f}(s)\} \quad \text{or} \quad L^{-1}[F(s)]$$

The symbol L is called the *Laplace Transformation Operator* and the determination of $L[f(t)]$ is called *Laplace Transformation* of $f(t)$.

Note: Let $f(t)$ be a function defined in the interval $a \le t \le b$. Let us associate with $f(t)$, a function $\bar{f}(s)$ by the formula,

$$\bar{f}(s) = \int_a^b f(t)K(s,t)\,dt \qquad ...(5.52)$$

The function $\bar{f}(s)$ is called the *integral transform* of $f(t)$ over the interval (a, b) corresponding to the function $K(s, t)$. The function $K(s, t)$, is called the *kernel* of the integral transform as defined in equation (5.52).

The function $f(t)$ is called the inverse transform of $\bar{f}(s)$.

It may be noted that Laplace transform is an integral transform over the interval $0 \le t \le \infty$ with kernel e^{-st}.

5.6 PROPERTIES OF LAPLACE TRANSFORM

Laplace transform is one of the most widely used integral transform.

Change of Scale Property of Laplace Transforms

If $L[f(t)] = \bar{f}(s)$, then

$$L[f(at)] = \frac{1}{a}\bar{f}\left(\frac{s}{a}\right) \quad \text{and} \quad L\left[f\left(\frac{t}{a}\right)\right] = a\bar{f}(as)$$

This simple property is proved as,

By definition, $L[f(t)] = \int_0^\infty e^{-st} f(t)dt = \bar{f}(s)$ ...(5.53)

$$\therefore \qquad L[f(at)] = \int_0^\infty e^{-st} f(at)dt \qquad ...(5.54)$$

Putting $at = u$, we have $dt = (1/a)du$.

NOTES

Changing the limits of the integral in equation (5.54), we have

$$L[f(at)] = \int_0^\infty e^{-(s/a)u} f(u) \cdot \frac{1}{a} du$$

On replacing the dummy variable u by t, we get

$$= \frac{1}{a}\int_0^\infty e^{-(s/a)t} f(t)\, dt \quad ...(5.55)$$

Now, comparing the integrals in equations (5.53) and (5.55), we find that the integral in equation (5.55) is similar to the integral in (5.53) with (s/a) appearing in the place of s. Therefore, the integral in (5.55) must be $\bar{f}(s/a)$

Thus,

$$L[f(at)] = \frac{1}{a}\bar{f}\left(\frac{s}{a}\right) \quad ...(5.56)$$

Changing a to ($1/a$) in equation (5.56), we have

$$L\left[f\left(\frac{t}{a}\right)\right] = a\bar{f}\,(as) \quad ...(5.57)$$

Resultant equations (5.56) and (5.57) are called the change of scale property of the Laplace transform.

Existence of Laplace Transforms

We now consider a sufficiently general set of conditions for the existence of Laplace transforms.

A function $f(t)$ is said to be *sectionally continuous* or *piecewise continuous* in the finite interval, $a \le t \le b$, if this interval can be divided into a finite number of subintervals such that,

(i) $f(t)$ is continuous at every point inside each of the subintervals.

(ii) $f(t)$ has finite limits as t approaches the end points of each subinterval from the interior of the subinterval.

i.e., for each subinterval $\alpha \le t \le \beta$, $\lim_{h\to 0} f(\alpha + h)$ and $\lim_{h\to 0} f(\beta - h), (h > 0)$ both exist finitely.

The graph of a piecewise continuous function will be continuous except for a finite number of jump-type discontinuities in the interval in which it is defined.

A graph of a piecewise continuous function is given below.

A function which is continuous in $a \le t \le b$, is piecewise continuous in this interval.

A function $f(t)$ is said to be of *exponential order* $\alpha(>0)$ as $t \to \infty$, if there exist finite positive constants t_0 and M such that,

$$|f(t)| \le Me^{\alpha t} \quad \text{or} \quad |e^{\alpha t} f(t)| \le M, \text{ for all } t \ge t_0$$

NOTES

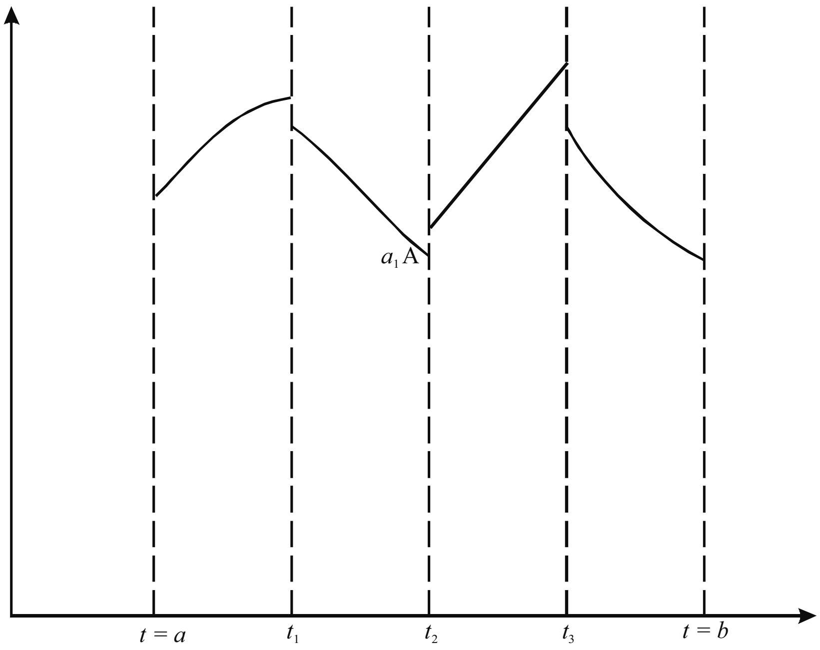

Fig. 5.1 *Graph of a Piecewise Continuous Function*

When $f(t)$ is of exponential order α as $t \to \infty$, we write

$$f(t) = O(e^{\alpha t})$$

If $\lim_{t\to\infty} e^{-s_0 t} \, | f(t) |$ exists finitely, then it can be proved that $f(t)$ is of exponential order s_0 at $t \to \infty$.

If $|f(t)| \le K$, for $0 \le t \le N$, for any large N and if $f(t)$ is of exponential order as $t \to \infty$, then it can be proved that,

$$|f(t)| \le Pe^{\alpha t}, \quad \text{for} \quad t \ge 0$$

Where, P is a positive constant.

Sufficient Conditions for the Existence of Laplace Transforms

Theorem 5.4: If the function $f(t)$ defined for $t \ge 0$ is,

(*i*) piecewise continuous in the interval $0 \le t \le N$, for every finite $N(> 0)$, and

(*ii*) of exponential order $\alpha (> 0)$ at $t \to \infty$,

then the Laplace transform of $f(t)$ exists for $s > \alpha$.

Proof: Since $f(t)$ is piecewise continuous in $0 \le t \le N$, $e^{-st} f(t)$ is integrable in $0 \le t \le N$ and $|f(t)| \le K$, for $0 \le t \le N$. Further, $f(t)$ being a function of exponential order as $t \to \infty$, there exists a positive constant P such that, $|f(t)| \le Pe^{\alpha t}$, for $t \ge 0$.

$$\therefore \qquad \left| \int_0^N e^{-st} f(t) dt \right| \le \int_0^N e^{-st} \, | f(t) | \, dt$$

$$\le \int_0^N e^{-st} P e^{\alpha t} dt$$

$$\leq P\int_0^\infty e^{-(s-\alpha)t}dt$$

NOTES

As the integral is positive for $t \geq 0$

$$= P\left[\frac{e^{-(s-\alpha)t}}{s-\alpha}\right]_0^\infty = \frac{P}{s-\alpha} \quad \text{if } s > \alpha \qquad ...(5.58)$$

$\therefore \int_0^N e^{-st} f(t)dt$ exists, however large N might be,

$$\therefore \int_0^\infty e^{-st} f(t)dt \text{ exists and } \left|\int_0^\infty e^{-(s-\alpha)t}dt\right| \frac{P}{s-\alpha}, s > \alpha \qquad ...(5.59)$$

Thus, Laplace transform of $f(t)$ exists for $s > \alpha$.

Notes: 1.The conditions stated above are only the sufficient conditions, *i.e.*, if these conditions are satisfied by a function, then its Laplace transform exists. However, the conditions are not necessary for the existence of Laplace transform, *i.e.*, Laplace transform of a function may exist even if these conditions are not satisfied by the function.

2. The function e^{-at}, e^{at}, cos h at, sin h at, t^n (where n is a positive integer), sin at, cos at are all continuous functions in $0 \leq t \leq N$, for any large N and all of them are of exponential order as $t \to \infty$. Therefore, their Laplace transforms exist and we have already found them.

3. By far most of the functions that represent physical quantities (variables) satisfy the conditions stated in the theorem and therefore have Laplace transforms.

Other Properties of Laplace Transforms

We now consider some of the properties of Laplace transform using which the Laplace transform of the product $f(t)$ $g(t)$ can be expressed in terms of the Laplace transform of $f(t)$ when $g(t)$ is a simple exponential function or a polynomial in t.

The first theorem called the first shifting property of Laplace transform enables us to find the Laplace transform of $e^{at} f(t)$ and $e^{-at} f(t)$ in terms of the Laplace transform of $f(t)$.

Theorem 5.5: (First Shifting Property)

If $L[f(t)] = F(s)$, then $L[e^{-at} f(t)] = F(s + a)$

Proof: By definition, $L[f(t)] = \int_0^\infty e^{-st} f(t)dt = F(s)$ ...(5.60)

Also, $L[e^{-at} f(t)] = \int_0^\infty e^{-st} e^{-at} f(t)dt$

$$= \int_0^\infty e^{-(s+a)} f(t)dt = \int_0^\infty e^{-pt} f(t)dt, \quad \text{Where} \quad p = s + a = F(p) \qquad ...(5.61)$$

Comparing the integrals on the RHS of equations (5.60) and (5.61), we get, $F(s + a)$

Thus, $\quad L[e^{-at} f(t)] = F(s + a)$

Notes:

1. The first shifting property stated above may also be expressed as below:

$$L\{e^{-at}f(t)\} = [L\{f(t)\}]_{s\to(s+a)}$$

Where, $s \to (s + a)$ means that s is replaced by $(s + a)$ in $L\{f(t)\}$.

2. Changing a to $-a$ in the above result, we have,

$$L\{e^{at}f(t)\} = [L\{f(t)\}]_{s\to(s-a)}$$

3. The results,

$$L\{e^{-at}f(t)\} = [L\{f(t)\}]_{s\to(s+a)}$$

$$L\{e^{at}f(t)\} = [L\{f(t)\}]_{s\to(s-a)}$$

These are called *shifting properties,* because the multiplication of $f(t)$ by e^{-at} (or e^{at}) shifts the argument s by a (or $-a$).

Example 5.4: Find the Laplace transforms of

(*i*) $e^{-at} \sin bt$ (*ii*) $e^{at} t^n$ (*iii*) $e^t \sin t \cos t$

Solution:

(i) $$L[e^{-at} \sin bt] = [L(\sin bt)]_{[s\to(s+a)]}$$

$$= \left[\frac{b}{s^2+b^2}\right]_{[s\to(s+a)]} = \frac{b}{(s+a)^2+b^2}$$

(ii) $$L[e^{at}t^n] = [L(t^n)]_{[s\to(s-a)]}$$

$$= \left[\frac{n!}{s^{(n+1)}}\right]_{[s\to(s-a)]} = \frac{n!}{(s-a)^{(n+1)}}$$

(iii) $$L[e^t \sin t \cos t] = [L(\sin t \cos t)]_{[s\to(s-1)]}$$

$$= \left[L\left(\frac{1}{2}\sin 2t\right)\right]_{[s\to(s-1)]}$$

$$= \left[\frac{1}{2}\cdot\frac{2}{s^2+4}\right]_{[s\to(s-1)]} = \frac{1}{(s-1)^2+4}$$

Example 5.5: Find the Laplace transforms of,

(*i*) $\cos h\, at \sin bt$ (*ii*) $e^{at} \sin^3 bt$ (*iii*) $e^{3t} (2t + 3)^3$

Solution:

(i) $$L[\cos h\, at \sin bt] = L\left[\frac{1}{2}\left(e^{at}+e^{-at}\right)\sin bt\right]$$

NOTES

$$= \frac{1}{2} L\left[e^{at} \sin bt\right] + \frac{1}{2} L\left[e^{-at} \sin bt\right]$$

$$= \frac{1}{2}\left[L(\sin bt)\right]_{[s \to (s-a)]} + \frac{1}{2}\left[L(\sin bt)\right]_{[s \to (s+a)]}$$

$$= \frac{1}{2}\left[\frac{b}{s^2 + b^2}\right]_{[s \to (s-a)]} + \frac{1}{2}\left[\frac{b}{s^2 + b^2}\right]_{[s \to (s+a)]}$$

$$= \frac{1}{2} \frac{b}{(s-a)^2 + b^2} + \frac{1}{2} \frac{b}{(s+a)^2 + b^2}$$

(ii) $L\left[e^{-at} \sin^3 bt\right] = L\left[e^{-at}\left(\frac{3}{4}\sin bt - \frac{1}{4}\sin 3bt\right)\right]$

$$= \left[L\left(\frac{3}{4}\sin bt - \frac{1}{4}\sin 3bt\right)\right]_{[s \to (s+a)]}$$

$$= \left[\frac{3}{4}\frac{b}{s^2 + b^2} - \frac{1}{4}\frac{3b}{s^2 + 9b^2}\right]_{[s \to (s+a)]}$$

$$= \frac{3}{4}\frac{b}{(s+a)^2 + b^2} - \frac{3}{4}\frac{b}{(s+a)^2 + 9b^2}$$

$$= \frac{3b}{4}\left[\frac{1}{(s+a)^2 + b^2} - \frac{1}{(s+a)^2 + b^2}\right]$$

(iii) $L[e^{3t}(2t+3)^3] = L[e^{3t}(8t^3 + 36t^2 + 54t + 27)]$

$$= [L(8t^3 + 36t^2 + 54t + 27)]_{[s \to (s-3)]}$$

$$= \left[\frac{48}{s^4} + \frac{72}{s^3} + \frac{54}{s^2} + \frac{27}{s}\right]_{[s \to (s-3)]}$$

$$= \frac{48}{(s-3)^4} + \frac{72}{(s-3)^3} + \frac{54}{(s-3)^2} + \frac{27}{(s-3)}$$

Example 5.6: Find the Laplace transform of $[(2e^t + e^{-2t})t]^2$

Solution: The given function of t is $(4e^{2t} + e^{-4t} + 4e^{-t})t^2$

The Laplace transform of the given function,

$$= 4L\left(e^{2t}t^2\right) + L\left(e^{-4t}t^2\right) + 4L\left(e^{-t}t^2\right)$$

$$= 4 \cdot \frac{2}{(s-2)^3} + \frac{2}{(s+4)^3} + 4 \cdot \frac{2}{(s+1)^3}$$

$$= \frac{8}{(s-2)^3} + \frac{2}{(s+4)^3} + \frac{8}{(s+1)^3}$$

NOTES

The following theorem enables us to express the Laplace transform of $t^n f(t)$ for any positive integer n, in terms of the Laplace transform of $f(t)$.

Theorem 5.6: If, $L[f(t)] = F(s)$, then

$$L[t\,f(t)] = \frac{d}{ds}F(s), \quad \text{and} \quad L[t^n f(t)] = (-1)^n \frac{d^n}{ds^n}F(s)$$

Proof: $F(s)$ is the Laplace transform of $f(t)$ and so,

$$F(s) = \int_0^\infty e^{-st} f(t)dt \qquad ...(5.62)$$

Differentiating equation (5.62) with respect to s,

$$\frac{d}{ds}F(s) = \frac{d}{ds}\int_0^\infty e^{-st} f(t)dt$$

Assuming that the conditions for interchanging the order of two operations of integration with respect to t and differentiation with respect to s are satisfied for the integrand $e^{-st} f(t)$, we have,

$$\frac{d}{ds}F(s) = \int_0^\infty \frac{d}{ds}[e^{-st}]\, f(t)dt = \int_0^\infty e^{-st}(-t)\, f(t)dt$$

$$= -\int_0^\infty e^{-st}\, tf(t)dt = -L[tf(t)]$$

$$\therefore \qquad L[tf(t)] = -\frac{d}{ds}F(s)$$

To prove the second result, we differentiate equation (5.62) n times with respect to s. Then,

$$\frac{d^n}{ds^n}F(s) = \int_0^\infty \frac{d^n}{ds^n}\left(e^{-st}\right) f(t)dt = \int_0^\infty (-t)^n e^{-st} f(t)dt$$

$$= (-1)^n \int_0^\infty e^{-st} t^n f(t)dt = (-1)^n L[t^n f(t)]$$

$$\therefore \qquad L[t^n f(t)] = (-1)^n \frac{d^n}{ds^n}F(s)$$

Note: The results of the theorem can also be written as,

$$L[tf(t)] = -\frac{d}{ds}L[f(t)]$$

$$L[t^n f(t)] = (-1)^n \frac{d^n}{ds^n}L[f(t)]$$

Example 5.7: Find the Laplace transform of,

(i) $t \sin at$ (ii) $t \cos at$ (iii) $\sin at - at \cos at$ (iv) $\sin at + at \cos at$.

Solution:

NOTES

(i) $L[t \sin at] = -\frac{d}{ds}\left(\frac{a}{s^2+a^2}\right) = \frac{2as}{(s^2+a^2)^2}$

(ii) $L[t \cos at] = -\frac{d}{ds}L(\cos at) = -\frac{d}{ds}\frac{s}{s^2+a^2}$

Completing differentiation and simplifying,

$$L[t \cos at] = \frac{s^2-a^2}{(s^2+a^2)^2}$$

(iii) $L[\sin at - at \cos at] = L(\sin at) - aL(t \cos at)$

$$= \frac{a}{s^2+a^2} - a\cdot\frac{s^2-a^2}{(s^2+a^2)^2}\text{, from equation } (ii)$$

$$\therefore L[\sin at - at \cos at] = \frac{2a^3}{(s^2+a^2)^3}$$

(iv) $L[\sin at + at \cos at] = L(\sin at) + aL(t \cos at)$

$$= \frac{a}{(s^2+a^2)} + a\cdot\frac{s^2-a^2}{(s^2+a^2)^2}\text{, from equation } (ii)$$

Notes: The first two results may also be obtained as follows:

$$L[te^{iat}] = L(t)_{[s \to s-ia]}$$

$$= \frac{1}{(s-ia)^2} = \frac{(s+ia)^2}{(s^2+a^2)^2}$$

$$= \frac{s^2-a^2}{(s^2+a^2)^2} + i.\frac{2as}{(s^2+a^2)^2}$$

$$L[te^{iat}] = L[t(\cos at + i \sin at)]$$

$$= L[t \cos at] + iL[t \sin at)]$$

$$\therefore \quad L[t \cos at] + iL[t \sin at)] = \frac{s^2-a^2}{(s^2+a^2)^2} + i\frac{2as}{(s^2+a^2)^2} \qquad ...(5.63)$$

Equating real parts on both sides of equation (5.63), we get,

$$L[t \cos at] = \frac{s^2-a^2}{(s^2+a^2)^2}$$

Equating imaginary parts on both sides of equation (5.63), we get,

$$L[t \sin at)] = \frac{2as}{(s^2+a^2)^2}$$

Note: Results of equations *(i)*, *(ii)*, *(iii)*, and *(iv)* above are often taken as formulae and are used particularly in finding the inverse Laplace transforms.

The following results which can all be proved using the first shifting property, are often remembered as formulae and are used in finding inverse Laplace transforms.

$$L[t \sin h\, at)] = \frac{2as}{(s^2 - a^2)^2}$$

$$L[t \cos h\, at] = \frac{s^2 - a^2}{(s^2 - a^2)^2}$$

$$L[at \cos h\, at - \sin h\, at] = \frac{2a^3}{(s^2 - a^2)^2}$$

$$L[\sin h\, at - at \cos h\, at] = \frac{2as^2}{(s^2 - a^2)^2}$$

NOTES

Example 5.8: Find the Laplace transforms of $(t \cos 2t)^2$.
Solution:

$$L[(t \cos 2t)^2] = L[t^2 \cos^2 2t]$$

$$= (-1)^2 \frac{d^2}{ds^2} L[\cos^2 2t] = \frac{d^2}{ds^2} L\left[\frac{1}{2}(1 + \cos 4t)\right]$$

$$= \frac{1}{2}\frac{d^2}{ds^2}\left[\frac{1}{s} + \frac{s}{s^2 + 16}\right] = \frac{1}{2}\frac{d}{ds}\left[-\frac{1}{s^2} + \frac{16 - s^2}{(s^2 + 16)^2}\right]$$

$$= \frac{1}{2}\left[\frac{2}{s^3} + \frac{2s(s^2 - 48)}{(s^2 + 16)^3}\right]$$

$$\therefore \quad L[(t \cos 2t)^2] = \frac{1}{s^3} + \frac{(s^3 - 48s)}{(s^2 + 16)^3}$$

Example 5.9: Find the Laplace transform of $e^t(t^2 - 2t + 4) \sin t$.
Solution:

$$L[e^t(t^2 - 2t + 4) \sin t] = [L(t^2 - 2t + 4) \sin t]_{[s \to (s-1)]}$$

Now, $$[L(t^2 - 2t + 4) \sin t] = (-1)^2\left[\frac{(d^2)}{(ds^2)}\right]$$

$$L(\sin t) - 2.(-1)\left[\frac{d}{ds}\right] L(\sin t) + 4.L(\sin t)$$

$$= \frac{d^2}{ds^2}\frac{1}{(s^2 + 1)} + 2\frac{d}{ds}.\frac{1}{(s^2 + 1)} + \frac{4}{s^2 + 1}$$

$$= \frac{d}{ds}\left[-\frac{2s}{(s^2 + 1)^2}\right] - 2.\frac{2s}{(s^2 + 1)^2} + \frac{4}{s^2 + 1}$$

$$= \frac{6s^2 - 4}{(s^2 + 1)^3} - \frac{4s}{(s^2 + 1)^2} + \frac{4}{s^2 + 1}$$

$$\therefore \quad L[e^t(t^2 - 2t + 4) \sin t] = \frac{6(s-1)^2 - 4}{[(s-1)^2 + 1]^3} - \frac{4(s-1)}{[(s-1)^2 + 1]^2} + \frac{4}{(s-1)^2 + 1}$$

The following theorem is concerned with the Laplace transform of $[f(t)/t]$

NOTES

Theorem 5.7: If, $L[f(t)] = F(s)$ and $\underset{t\to 0}{\text{Lim}}[f(t)/t]$ exists, then

$$L\left[\frac{f(t)}{t}\right] = \int_s^\infty F(s)ds$$

Proof:

$$L[f(t) = F(s)\int_0^\infty e^{-st} f(t)dt \qquad ...(5.64)$$

Integrating the above equation with respect to s between the limits s and ∞, we have,

$$\int_0^\infty F(s)ds = \int_s^\infty \left[\int_0^\infty e^{-st} f(t)dt\right] ds \qquad ...(5.65)$$

Assuming that the condition for changing the order of integration on the right side of equation (6.65) is satisfied, we have,

$$\int_s^\infty F(s)ds = \int_s^\infty \left[\int_0^\infty e^{-st} f(t)ds\right] dt = \int_s^\infty \left[f(t).\int_0^\infty e^{-st} ds\right] dt$$

On removing $f(t)$ outside the inner integral which is with respect to s, we get,

$$= \int_0^\infty f(t)\left[-\frac{e^{-st}}{t}\right]_{(s=s)}^{(s=\infty)} dt = \int_s^\infty f(t)\left[0 + \frac{e^{-st}}{t}\right] dt$$

Since, $\underset{s\to\infty}{\text{Lim}}[(e^{-st})/t] = 0$, if $s > 0$,

$$= \int_0^\infty e^{-st}\frac{f(t)}{t} dt$$

$$\therefore \qquad \int_0^\infty F(s)\,ds = L\left[\frac{f(t)}{t}\right]$$

Notes:

1. We may state the above result as follows:

$$L\left[\frac{f(t)}{t}\right] = \int_s^\infty L[f(t)]\,ds$$

2. Division of $f(t)$ by t corresponds to the integration of Laplace transform of $f(t)$ with respect to s between the limits s and a. By repeated application of the result, we have,

$$L\left[\frac{f(t)}{t^n}\right] = \int_s^\infty\int_s^\infty\int_s^\infty \cdots \int_s^\infty F(s)(ds)^n$$

or,
$$L\left[\frac{f(t)}{t^n}\right] = \int_s^\infty\int_s^\infty\int_s^\infty \cdots \int_s^\infty L[f(t)](ds)^n$$

provided that $\underset{t\to\infty}{\text{Lim}}[f(t)/t^n]$ exists, n being a positive integer.

Example 5.10: Find the Laplace transforms of,

(i) $[(e^{-bt} - e^{-at})/t]$, where $(a \neq b)$ (ii) $[(1 - e^{-at})/t]$ (iii) $[(\sin \omega)t/t]$

Solution:

NOTES

(i)
$$L\left(\frac{e^{-bt}-e^{-at}}{t}\right)=\int_s^\infty L[e^{-bt}-e^{-at}]ds=\int_s^\infty\left[\frac{1}{s+b}-\frac{1}{s+a}\right]ds$$
$$=\left[\log(s+b)-\log(s+a)\right]_{s=s}^{s=\infty}=\left[\log\frac{s+b}{s+a}\right]_s^\infty$$
$$=\lim_{s\to\infty}\log\frac{s+b}{s+a}-\log\frac{s+b}{s+a}$$

Now, $$\lim_{s\to\infty}\log\left(\frac{s+b}{s+a}\right)=\lim_{s\to\infty}\log\left(\frac{1+b/s}{1+a/s}\right)=\log 1=0$$

$\therefore$ $$L\left(\frac{e^{-bt}-e^{-at}}{t}\right)=\log\frac{s+a}{s+b}$$

(ii)
$$L\left(\frac{1-e^{-at}}{t}\right)=\int_s^\infty L[1-e^{-at}]ds$$
$$=\int_s^\infty\left[\frac{1}{s}-\frac{1}{s+a}\right]ds=\left[\log s-\log(s+a)\right]_{s=s}^{s=\infty}$$
$$=\left[\log\frac{s}{s+a}\right]_{s=s}^{s=\infty}=\lim_{s\to\infty}\log\frac{s}{s+a}-\log\frac{s}{s+a}$$
$$=\lim_{s\to\infty}\log\frac{1}{1+(a/s)}+\log\frac{s+a}{s}$$
$$=0+\log\frac{s+a}{s}=\log\frac{s+a}{s}$$

Note: Putting, $b=0$ in the first example, we get the Laplace transform of $[(1-e^{-at})/t]$.

(iii)
$$L\left[\frac{\sin\omega t}{t}\right]=\int_s^\infty L(\sin\omega t)ds$$
$$=\int_s^\infty\frac{\omega}{s^2+\omega^2}ds=\left[\tan^{-1}\left(\frac{s}{\omega}\right)\right]_s^\infty$$
$$=\tan^{-1}(\infty)-\tan^{-1}\left(\frac{s}{\omega}\right)=\frac{\pi}{2}-\tan^{-1}\left(\frac{s}{\omega}\right)$$

$\therefore$ $$L\left[\frac{\sin\omega t}{t}\right]=\cot^{-1}\left(\frac{s}{\omega}\right)\text{ or }\tan^{-1}\left(\frac{\omega}{s}\right)$$

Using the results, we get,

$\tan^{-1}x+\cot^{-1}x=(\pi/2)$ and $\tan^{-1}x=\cot^{-1}(1/x)$

Example 5.11: Evaluate $L[(\cos at-\cos bt)/t]$, where $(a\neq b)$.

Solution: $$L\left[\frac{\cos at-\cos bt}{t}\right]=\int_s^\infty(\cos at-\cos bt)ds=\int_s^\infty\left(\frac{s}{s^2+a^2}-\frac{s}{s^2+b^2}\right)ds$$
$$=\left[\frac{1}{2}\log(s^2+a^2)-\frac{1}{2}\log(s^2+b^2)\right]_s^\infty$$

NOTES

$$= \left[\frac{1}{2}\log\frac{(s^2+a^2)}{(s^2+b^2)}\right]_s^\infty$$

$$= \frac{1}{2}\operatorname*{Lim}_{s\to\alpha}\log\frac{(s^2+a^2)}{(s^2+b^2)} - \frac{1}{2}\log\frac{(s^2+a^2)}{(s^2+b^2)}$$

$$= \frac{1}{2}\log\left[\frac{1+[(a^2)/(s^2)]}{1+[(b^2)/(s^2)]}\right] + \frac{1}{2}\log\frac{(s^2+b^2)}{(s^2+a^2)}$$

$$= \frac{1}{2}\log 1 + \frac{1}{2}\log\frac{(s^2+b^2)}{(s^2+a^2)}$$

$$\therefore \qquad L\left\{\frac{\cos at - \cos bt}{t}\right\} = \frac{1}{2}\log\frac{(s^2+b^2)}{(s^2+a^2)}$$

Example 5.12: Find the Laplace transforms of $[(1 - \cos at)/t]$ and $[(1 - \cos at)/t^2]$.

Deduce the Laplace transform of $((2t + 3)/t^2)\sin^2 t$.

Solution: *(i)* $$L\left[\frac{1-\cos at}{t}\right] = \int_s^\infty (1-\cos at)ds = \int_s^\infty \left[\frac{1}{2} - \frac{s}{s^2+a^2}\right]ds$$

$$= \left[\log s - \frac{1}{2}\log(s^2+a^2)\right]_s^\infty = \left[\log\frac{s}{\sqrt{s^2+a^2}}\right]_s^\infty$$

$$= \operatorname*{Lim}_{s\to\infty}\log\frac{s}{\sqrt{s^2+a^2}} - \log\frac{s}{\sqrt{s^2+a^2}}$$

$$= \operatorname*{Lim}_{s\to\infty}\log\frac{1}{\sqrt{1+a^2/s^2}} + \log\frac{\sqrt{s^2+a^2}}{s}$$

$$= \log 1 + \log\frac{\sqrt{s^2+a^2}}{s}$$

$$\therefore \qquad L\left[\frac{1-\cos at}{t}\right] = \log\frac{\sqrt{s^2+a^2}}{s} \qquad ...(1)$$

$$L\left[\frac{1-\cos at}{t^2}\right] = \int_s^\infty L\left[\frac{1-\cos at}{t}\right]ds$$

$$= \int_s^\infty \log\frac{\sqrt{s^2+a^2}}{s}ds, \text{ using equation (1)}$$

On integrating by parts, we get,

$$= \left[s\log\frac{\sqrt{s^2+a^2}}{s}\right]_s^\infty - \int_s^\infty s\frac{d}{ds}\left[\log\frac{\sqrt{s^2+a^2}}{s}\right]ds$$

$$= \operatorname*{Lim}_{s\to\infty} s\log\frac{\sqrt{s^2+a^2}}{s} - s\log\frac{\sqrt{s^2+a^2}}{s}$$

$$-\int_s^\infty s\frac{d}{ds}\left[\frac{1}{2}\log(s^2+a^2)-\log s\right]ds$$

$$= A + s\log\frac{s}{\sqrt{s^2+a^2}} - \int_s^\infty s\left[\frac{s}{s^2+a^2}-\frac{1}{s}\right]ds \qquad ...(2)$$

Where, $A = \underset{s\to\infty}{\text{Lim}}\, s\log[s/\sqrt{s^2+a^2}\,]$

$$= \underset{s\to\infty}{\text{Lim}}\left[\frac{\log s-(1/2)\log(s^2+a^2)}{(1/s)}\right]$$

$$= \underset{s\to\infty}{\text{Lim}}\left[\frac{(1/s)-[s/(s^2+a^2)]}{-(1/s^2)}\right]$$

Again by applying L'Hospital's rule,

$$= \underset{s\to\infty}{\text{Lim}}\left[-\frac{a^2 s}{s^2+a^2}\right] = \underset{s\to\infty}{\text{Lim}}\left[-\frac{a^2}{2s}\right]$$

Therefore, from equation (2) above, we get,

$$L\left[\frac{1-\cos at}{t^2}\right] = s\log\frac{s}{\sqrt{s^2+a^2}} + \int_s^\infty \frac{a^2}{s^2+a^2}ds$$

$$= s\log\frac{s}{\sqrt{s^2+a^2}} + \frac{a^2}{a}\left[\tan^{-1}\left(\frac{s}{a}\right)\right]_s^\infty$$

$$= s\log\frac{s}{\sqrt{s^2+a^2}} + a\left[\frac{\pi}{2}-\tan^{-1}\left(\frac{s}{a}\right)\right]$$

$$\therefore \quad L\left[\frac{1-\cos at}{t^2}\right] = s\log\frac{s}{\sqrt{s^2+a^2}} + a\cot^{-1}\left(\frac{s}{a}\right) \qquad ...(3)$$

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

$$= L\left[\frac{2}{t}\cdot\frac{1}{2}(1-\cos 2t)\right] + L\left[\frac{3}{t^2}\cdot\frac{1}{2}(1-\cos 2t)\right]$$

$$= L\left[\frac{1-\cos 2t}{t}\right] + \frac{3}{2}L\left[\frac{1-\cos 2t}{t^2}\right]$$

$$= \log\frac{\sqrt{s^2+4}}{s} + \frac{3}{2}\left[s\log\frac{s}{\sqrt{s^2+4}} + 2\cot^{-1}\left(\frac{s}{2}\right)\right]$$

Using equations (1) and (3),

$$= \left(\frac{3s-2}{2}\right)\log\frac{s}{\sqrt{s^2+4}} + 3\cot^{-1}\left(\frac{s}{2}\right)$$

Laplace Transforms and Results

For ready reference for students, the Laplace transforms of some useful functions and important results on Laplace transforms are tabulated below.

NOTES

Table 5.1 *Laplace Transforms of Some Useful Functions*

$f(t)$	$L[f(t)]$
1	$\frac{1}{s}$
t	$\frac{1}{s^2}$
t^n (n is a positive integer)	$\frac{n!}{s^{(n+1)}}$
e^{-at}	$\frac{1}{s+a}$
e^{at}	$\frac{1}{s-a}$
$\sin at$	$\frac{a}{s^2+a^2}$
$\cos at$	$\frac{s}{s^2+a^2}$
$\sinh at$	$\frac{a}{s^2-a^2}$
$\cosh at$	$\frac{s}{s^2-a^2}$
$t \sin at$	$\frac{2as}{(s^2+a^2)^2}$
$t \cos at$	$\frac{s^2-a^2}{(s^2+a^2)^2}$
$\sin at + at \cos at$	$\frac{2as^2}{(s^2+a^2)^2}$
$\sin at - at \cos at$	$\frac{2a^3}{(s^2+a^2)^2}$
$t \sin h\, at$	$\frac{2as}{(s^2-a^2)^2}$
$t \cos h\, at$	$\frac{s^2+a^2}{(s^2-a^2)^2}$
$\sin h\, at + at \cos h\, at$	$\frac{2a^2}{(s^2-a^2)^2}$
$at \cos h\, at - \sin h\, at$	$\frac{2a^3}{(s^2-a^2)^2}$

Table 5.2 *Important Results of Laplace Transform*

If $f(t)$ is periodic function of period p,

$$L\left[e^{-at} f(t)\right] = \left[Lf(t)\right]_{s\to(s+a)}$$

$$L\left[e^{at} f(t)\right] = \left[Lf(t)\right]_{s\to(s-a)}$$

$$L\left[tf(t)\right] = -\frac{d}{ds} L[f(t)]$$

$$L\left[t^n f(t)\right] = (-1)^n \frac{d^n}{ds^n} L[f(t)]$$

$$L\left[\frac{f(t)}{t}\right] = \int_s^\infty L[f(t)]\,dx$$

$$L\left[\frac{f(t)}{t^n}\right] = \int_s^\infty \int_s^\infty \cdots \int_s^\infty L[f(t)](ds)^n$$

$$L[f'(t)] = sL[f(t)] - f(0)$$

$$L[f^{(n)}(t)] = s^n L[f(t)] - s^{(n-1)} f(0) - s^{(n-2)} f'(0) - \cdots - sf^{(n-2)}(0) - f^{(n-1)}(0)$$

$$L\left[\int_0^t f(u)\,du\right] = \frac{1}{s} L[f(t)]$$

$$L\left[\int_0^t \int_0^t \cdots \int_0^t f(t)(dt)^n\right] = \frac{1}{s^n} L[f(t)]$$

$$L[f(t)] = \frac{1}{1-e^{-sp}} \int_0^p e^{-st} f(t)\,dt$$

NOTES

5.7 APPLICATIONS OF LAPLACE TRANSFORM TO POTENTIAL AND OSCILLATORY PROBLEMS

In this section, we will discuss the applications of Laplace transform to potential and oscillatory problems.

The damped harmonic oscillator

Stability is not the only property of a system that can be detected by studying the poles of the transfer function. With some experience one can detect change in the qualitative behaviour of a system by studying the poles. A simple example is provided by the damped harmonic oscillator. This system is defined by two parameters m and w, both positive real numbers. The differential equation which governs this system is $(D^2 + 2\mu\, D + \omega^2)\, f(t) = u(t)$

NOTES

The transfer function is $H(s) = \frac{1}{s^2 + 2\mu s + \omega^2}$

which has poles at $s_{\pm} = -\mu \pm \sqrt{\mu^2 - \omega^2}$

We must distinguish three separate cases:

(a) (overdamped) $\mu > \omega$

In this case the poles are real and negative:

$$s_{\pm} = -\mu \left(1 \mp \sqrt{1 - \frac{\omega^2}{\mu^2}}\right)$$

(b) (critically damped) $\mu = \omega$

In this case there is a double pole, real and negative: $s+ = s_{j} = -\mu$.

(c) (underdamped) $\mu = \omega$

In this case the poles are complex:

$$s_{\pm} = -\mu \pm i\omega\sqrt{1 - \frac{\mu^2}{\omega^2}}$$

Hence provided that m is positive, the system is stable.

Suppose that we start with the system being overdamped so that the ratio w/m is less than 1. As we increase this ratio either by increasing w or decreasing m, the poles of the transfer function, which start in the negative real axis, start moving towards each other, coinciding when $\omega/\mu = 1$.

If we continue increasing the ratio, so that it becomes greater than 1, the poles move vertically away from each other keeping their real parts constant. It is the transition from real to complex poles which offers the most drastic qualitative change in the behaviour of the system.

5.8 EVALUATION OF SIMPLE INTEGRALS USING FOURIER TRANSFORMS

If $f(x)$ satisfies the Dirichlet's condition in $-\pi \le x \le \pi$, and $\int_{-\infty}^{\infty} f(x)\, dx$ converges, *i.e.*, is integrable in $-\infty < x < \infty$, then we have the Fourier series expansion for $f(x)$ as

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx, \qquad \text{... (5.66)}$$

where $a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\, dx$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos nx\, dx, \qquad \text{... (5.67)}$$

and $b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin nx \, dx.$

This function of x may be developed into a trigonometric series for all values of x between $x = -c$ and $x = c$ by putting

NOTES

$$z = \frac{\pi}{c}x, \text{ where } z = -\pi \quad \text{when} \quad x = -c$$

and $z = \pi$ when $x = c$,

i.e., $f(x) = f\left(\frac{c}{\pi}z\right)$

Then the series (5.66) may be developed in terms of z as

$$f\left(\frac{c}{\pi}z\right) = a_0 + a_1 \cos z + a_2 \cos 2z + a_3 \cos 3z + ...$$

$$+ b_1 \sin z + b_2 \sin 2z + b_3 \sin 3z + ... \quad ... (5.68)$$

where $a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f\left(\frac{c}{\pi}z\right) dz,$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{c}{\pi}z\right) \cos nz \, dz, \quad ... (5.69)$$

and $b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{c}{\pi}z\right) \sin nz \, dz.$

Now if we replace z by $\frac{\pi}{c}x,$ then (5.68) becomes

$$f(x) = a_0 + a_1 \cos \frac{\pi x}{c} + a_2 \cos \frac{2\pi x}{c} + a_3 \cos \frac{3\pi x}{c} + ...$$

$$+ b_1 \sin \frac{\pi x}{c} + b_2 \sin \frac{2\pi x}{c} + b_3 \sin \frac{3\pi x}{c} + \quad ... (5.70)$$

Its coefficients being the same as those of (5.68) is therefore valid from $x = -c$ to $x = c$, where

$$a_0 = \frac{1}{2\pi}\int_{-c}^{c} f(x)\cdot\frac{\pi}{c}dx = \frac{1}{2c}\int_{-c}^{c} f(x)\,dx \quad \because \text{ when } z = \frac{\pi}{c}x, dz = \frac{\pi}{c}dx$$

$$= \frac{1}{2c}\int_{-c}^{c} f(t)\,dt \text{ (say)}$$

$$a_n = \frac{1}{\pi}\int_{-c}^{c} f(x)\cos\frac{n\pi x}{c}\cdot\frac{\pi}{c}dx$$

$$= \frac{1}{c}\int_{-c}^{c} f(t)\cos\frac{n\pi t}{c}dt \quad ... (5.71)$$

and $b_n = \frac{1}{\pi}\int_{-c}^{c} f(x)\sin\frac{n\pi x}{c}\cdot\frac{\pi}{c}dx$

$$= \frac{1}{2c}\int_{-c}^{c} f(t)\sin\frac{n\pi t}{c}dt.$$

Now if we substitute the values of the coefficients $a_0, a_1, a_2, a_3 ... b_1, b_2, b_3...$ given by (5.71) in (5.70), then we get

$$f(x) = \frac{1}{2c}\int_{-c}^{c} f(t)\,dt + \frac{1}{c}\int_{-c}^{c} f(t)\cos\frac{\pi x}{c}dt + \frac{1}{c}\int_{-c}^{c} f(t)\cos\frac{2\pi t}{c}\cos\frac{2\pi x}{c}dx + ...$$

$$+ \frac{1}{c}\int_{-c}^{c} f(t)\sin\frac{\pi t}{c}\sin\frac{\pi x}{c}dt + \frac{1}{c}\int_{-c}^{c} f(t)\sin\frac{2\pi t}{c}\sin\frac{2\pi x}{c}dt + ...$$

$$= \frac{1}{c}\int_{-c}^{c} f(t)\left[\frac{1}{2} + \cos\frac{\pi t}{c}\cos\frac{\pi x}{c} + \cos\frac{2\pi t}{c}\cos\frac{2\pi x}{c} + ...\right.$$

$$\left. + \sin\frac{\pi t}{c}\sin\frac{\pi x}{c} + \sin\frac{2\pi t}{c}\sin\frac{2\pi x}{c} + ...\right]dt$$

NOTES

$$= \frac{1}{c}\int_{-c}^{c} f(t)\left[\frac{1}{2}+\left\{\cos\frac{\pi t}{c}\cos\frac{\pi x}{c}+\sin\frac{\pi t}{c}\sin\frac{\pi x}{c}\right\}\right.$$
$$\left.+\left\{\cos\frac{2\pi t}{c}\cos\frac{2\pi x}{c}+\sin\frac{2\pi t}{c}\sin\frac{2\pi x}{c}\right\}+\ldots\right]dt$$

$$= \frac{1}{c}\int_{-c}^{c} f(t)\left[\frac{1}{2}+\cos\frac{\pi}{c}(x-t)+\cos\frac{2\pi}{c}(x-t)+\ldots\right]dt$$

$$= \frac{1}{2c}\int_{-c}^{c} f(t)\left[1+2\cos\frac{\pi}{c}(x-t)+2\cos\frac{2\pi}{c}(x-t)+\ldots\right]dt$$

$$= \frac{1}{2\pi}\int_{-c}^{c} f(t)\left[\frac{\pi}{c}\cos\frac{0\pi}{c}(x-t)+\frac{\pi}{c}\cos\frac{2\pi}{c}(x-t)\right.$$
$$\left.+\frac{\pi}{c}\cos\left(\frac{-\pi}{c}\right)(x-t)+\frac{\pi}{c}\cos\frac{2\pi}{c}(x-t)+\frac{\pi}{c}\cos\left(\frac{-2\pi}{c}\right)(x-t)+\ldots\right]dt$$

since $2\cos\phi = \cos\phi + \cos(-\phi)$

$$= \frac{1}{2\pi}\int_{-c}^{c} f(t)\left[\ldots\frac{\pi}{c}\cos\left(-\frac{n\pi}{c}\right)(x-t)+\ldots+\frac{\pi}{c}\cos\left(-\frac{2\pi}{c}\right)(x-t)+\frac{\pi}{c}\cos\left(-\frac{\pi}{c}\right)(x-t)\right.$$
$$+\frac{\pi}{c}\cos\frac{0.\pi}{c}(x-t)+\frac{\pi}{c}\cos\frac{\pi}{c}(x-t)+\frac{\pi}{c}\cos\frac{2\pi}{c}(x-t)+\ldots$$
$$\left.+\frac{\pi}{c}\cos\frac{n\pi}{c}(x-t)+\ldots\right]dt$$

$$= \frac{1}{2\pi}\int_{-c}^{c} f(t)\left[\lim_{n\to\infty}\sum_{r=-n}^{r=n}\frac{\pi}{c}\cos\frac{r\pi}{c}(x-t)\right]dt$$

$$= \frac{1}{2\pi}\int_{-c}^{c} f(t)\left[\lim_{n\to\infty}\sum_{r=-n}^{r=n}\frac{1}{c/\pi}\cos\frac{r}{c/\pi}(x-t)\right]dt.$$

If c becomes indefinitely large, *i.e.*, as $c\to\infty$, $\frac{c}{\pi}\to\infty$, we have

$$\lim_{c\to\infty}\sum_{r=-\infty}^{\infty}\frac{1}{c/\pi}\cos\frac{r}{c/\pi}(x-t)=\int_{-\infty}^{\infty}\cos u\,(x-t)\,du$$

(by the definition of integral as the limit of a sum).

Hence $f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)\,dt\int_{-\infty}^{\infty}\cos u\,(x-t)\,du.$ ... (5.72)

This double integral is known as *Fourier's Integral* and holds if x is a point of continuity of $f(x)$.

Aliter. We have $f(x) = \frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi x}{l}+b_n\sin\frac{n\pi x}{l}\right)$ ... (5.73)

where $a_n = \frac{1}{l}\int_{-l}^{l} f(u)\cos\frac{n\pi u}{l}\,du$ and $b_n = \frac{1}{l}\int_{-l}^{l} f(u)\sin\frac{n\pi u}{l}\,du$ .. (5.74)

so that $\quad a_n\cos\frac{n\pi x}{l}+b_n\sin\frac{n\pi x}{l}=\frac{1}{\pi}\int_{-l}^{l} f(u)\cos\frac{n\pi}{l}(u-x)\,dx$

and $\quad \frac{a_0}{2}=\frac{1}{2l}\int_{-l}^{l} f(u)\,du$

$\therefore$ (5.73) gives

$$f(x) = \frac{1}{2l}\int_{-l}^{l} f(u)\,du + \frac{1}{l}\sum_{n=1}^{\infty} f(u)\cos\frac{n\pi}{l}(u-x)\,du \qquad \text{... (5.75)}$$

NOTES

Assuming that $\int_{-\infty}^{\infty} |f(u)|\,du$ converges, the first term on R.H.S. of (5.75) approaches zero as $l \to \infty$ and hence (5.75) yields

$$f(x) = \lim_{l\to\infty}\frac{1}{l}\sum_{n=1}^{\infty}\int_{-\infty}^{\infty} f(u)\cos\frac{n\pi}{l}(u-x)\,dx \qquad \text{... (5.76)}$$

Putting $\frac{\pi}{l} = \Delta t$, (11) can be written as

$$f(x) = \lim_{\Delta t\to 0}\sum_{n=1}^{\infty}\Delta t\, F(n\Delta t) \qquad \text{... (5.77)}$$

where $F(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(x)\cos t(u-x)\,du$ ... (5.78)

Thus (5.77) gives

$$f(x) = \int_0^{\infty} F(t)\,dt = \frac{1}{\pi}\int_0^{\infty} dt \int_{-\infty}^{\infty} f(u)\cos t(u-x)\,du \qquad \text{... (5.79)}$$

which is Fourier's Integral formula.

Note. The complex form of Fourier's integral is

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{itx}\,dt \int_{-\infty}^{\infty} f(u)\,e^{itu}\,du$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u)\,e^{it(x-u)}\,du\,dt \qquad \text{... (5.80)}$$

5.9 EVALUATION OF SIMPLE INTEGRALS USING LAPLACE TRANSFORMS

By definition, we have

$$L\{F(t)\} = \int_0^{\infty} e^{-st}F(t)\,dt = f(s) \qquad \text{... (5.81)}$$

Assuming that the integral is convergent and proceeding to the limit $s \to 0$, this reduces to

$$\int_0^{\infty} F(t)\,dt = f(0) \qquad \text{... (5.82)}$$

(5.81) and (5.82) are sometimes used to evaluate integrals.

Problem 5.11. *Evaluate the following integrals*

(a) $\int_0^{\infty} t^2 e^{-t}\sin t\,dt$ (b) $\int_0^{\infty} J_0(t)\,dt$

(c) $\int_0^{\infty} e^{-t}\operatorname{erf}\sqrt{t}\,dt.$

(a) We have $L\{\sin t\} = \frac{1}{s^2+1}$

$$\therefore \quad L\{t^2\sin t\} = (-1)^2\frac{d^2}{ds^2}L\{\sin t\}$$

$$= \frac{d^2}{ds^2}\left(\frac{1}{s^2+1}\right) = \frac{d}{dx}\left(\frac{-2s}{(s^2+1)^2}\right) = -\frac{2(1-s^2)}{(1+s^2)^3}$$

so that $\int_0^{\infty} e^{-st}\cdot t^2\sin t\,dt = -\frac{2(1-s^2)}{(1+s^2)^3}.$

Putting $s = 1$, we find $\int_0^\infty e^{-t} t^2 \sin t\, dt = 0$

(*b*) We have $L\ \{J_0(t)\} = \dfrac{1}{\sqrt{s^2+1}}$

NOTES

i.e., $$\int_0^\infty e^{-st} J_0(t)\, dt = \frac{1}{\sqrt{s^2+1}}$$

Proceeding to the limit as $s \to 0$, we get

$$\int_0^\infty J_0(t)\, dt = 1.$$

(*c*) We have, $L\ \left\{\text{erf}\ \sqrt{t}\right\}$

i.e., $$\int_0^\infty e^{-st} \text{erf}\ \sqrt{t}\ dt = \frac{1}{s\sqrt{s+1}}.$$

Proceeding to the limit $s \to 1$, we find

$$\int_0^\infty e^{-x} \text{erf}\ \sqrt{t}\ dt = \frac{1}{\sqrt{2}}.$$

Check Your Progress

1. Which transformation is used to solve linear differential equations?
2. Find the Laplace transform of 1.
3. If $f(t)$ is of exponential order, what is the order of $f'(t)$.
4. In the equation $L[f(t)] = F(s)$, what is the symbol L called?
5. Write the equation for change of scale property of Laplace transforms.
6. Write the condition for the function $f(t)$ to be sectionally continuous.
7. Write the sufficient condition for the existence of Laplace transforms.
8. State the first shifting property of Laplace transforms

5.10 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Laplace transformation is an important operational method for solving linear differential equations.
2. Transform of $f(t) = 1,\ t \geq 0$

   $$L\{f(t)\} = \int_0^\infty e^{-st} dt = \left[-\frac{e^{-st}}{s}\right]_0^\infty$$

   $e^{-st} \to 0$ as $t \to \infty$, if $s > 0$

   $$\therefore \quad L(1) = \frac{1}{s}, s > 0$$
3. Nothing can be said about the order of $f'(t)$.
4. Laplace transformation operator.
5. $L[f(\text{at})] = \dfrac{1}{a}\bar{f}\left(\dfrac{s}{a}\right)$

   $$L\left[f\left(\frac{t}{a}\right)\right] = a\bar{f}(as)$$
6. A function $f(t)$ is said to be *sectionally continuous* or *piecewise continuous* in the finite interval, $a \leq t \leq b$, if this interval can be divided into a finite number of sub-intervals such that,

(i) $f(t)$ is continuous at every point inside each of the subintervals.

(ii) $f(t)$ has finite limits as t approaches the end points of each subinterval from the interior of the subinterval.

7. If the function $f(t)$ defined for $t \geq 0$ is,

   (*i*) piecewise continuous in the interval $0 \leq t \leq N$ for every finite $N(> 0)$, and

   (*ii*) of exponential order $\alpha(> 0)$ at $t \to \infty$,

   then the Laplace transform of $f(t)$ exists for $s > \alpha$.

8. If $L[f(t)] = F(s)$,

   then $L[e^{-at} f(t)] = F(s+a)$

NOTES

5.11 SUMMARY

- Integral transform is a type of mathematical operator that results when a given function, *f(x)*, is multiplied by a kernel function, *K(p,x)*, and the product is integrated between suitable limits.
- If $f(t)$ is a function of t defined for $t \geq 0$ and if the integral $\int_0^\infty e^{-st} f(t)dt$ exists, then it is a function of the parameter s. This function of s, denoted as $\bar{f}(s)$ is called the *Laplace Transform* of $f(t)$ over the range of values of s for which the integral exists.
- A function $f(t)$ is said to be *sectionally continuous* or *piecewise continuous* in the finite interval, $a \leq t \leq b$, if this interval can be divided into a finite number of subintervals such that, *(i)* $f(t)$ is continuous at every point inside each of the subintervals. *(ii)* $f(t)$ has finite limits as t approaches the end points of each subinterval from the interior of the subinterval.
- If $f(x)$ satisfies the Dirichlet's condition in $-\pi \leq x \leq \pi$, and converges, i.e., is integrable in $-\infty < x < \infty$, then we have the Fourier series expansion for f(x) as.
- Despite its name, the delta function is *not* a function, even though it is a limit of functions. Instead it is a distribution. Distributions are only well-defined when integrated against sufficiently well-behaved functions known as test functions.
- Stability is not the only property of a system that can be detected by studying the poles of the transfer function. With some experience one can detect change in the qualitative behaviour of a system by studying the poles. A simple example is provided by the damped harmonic oscillator. This system is defined by two parameters m and w, both positive real numbers.

5.12 KEY TERMS

- **Fourier transformation:** A Fourier transform is a mathematical transform that decomposes functions depending on space or time into functions depending on spatial or temporal frequency, such as the expression of a musical chord in terms of the volumes and frequencies of its constituent notes.

- **Laplace transformation:** Laplace transformation is used for solving linear differential equations, both ordinary and partial. If $f(t)$ is a function of t defined for $t \geq 0$ and if the integral exists, then it is a function of the parameter s.

NOTES

5.13 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. Write the definition of Laplace transformation.
2. What are the Laplace transforms of simple functions?
3. Define the change of scale property of Laplace transforms.
4. Mention the silent conditions for the existence of Laplace Transform.
5. What do you mean by shifting property of inverse transform?
6. State the theorem which proves the property of transforms of integrals.
7. Mention the the application of Fourier transform to Dirac delta function.

Long Answer Questions

1. Discuss various Fourier transforms with examples.
2. If F1 (k) and F2 (k) are Fourier transforms of f1 (x) and f2 (x) respectively, show that the Fourier Transform of f1 (x) f2 (x) is given by

   $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_1(k') \, F_2(k - k') \, dk'$.

3. Apply Fourier series solution method to solve the wave equation

   $\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 y}{\partial t^2}$ for a stretched string fastened to fixed supports at its two

   ends and initially plucked at its midpoint, giving it initial displacement h. If initial velocity be zero at all points of the string, prove that the displacement d at a point distant a from a fixed end is given by

   $\delta = \frac{2h}{\pi^2} \left[\sin \frac{\pi \alpha}{l} \cos \frac{\pi c t}{l} - \frac{1}{9} \sin \frac{3\pi \alpha}{l} \cos \frac{3\pi c t}{l} + ... \right]$, where l denotes the length of the string.

4. Find the Laplace transforms of the following:

   (i) $2t^3 + 3t^2 - 5t + 2$ *(ii)* $\sqrt{e^{3(t+1)}}$

   (iii) $(e^{3t} + e^{-2t})^2$ *(iv)* $\sin at \cos at$

   (v) $\sin^3 bt$ *(vi)* $3t^2 + \cos^3 bt$

   (vii) $\sin at \cos bt$

5. Find the Laplace transforms of the following:

   (i) $t^3 e^5 t$ *(ii)* $e^{-t} \sin(2t + 3)$

   (iii) $\cosh at \cos bt$ *(iv)* $\sinh at \sin bt$

   (v) $3t^2 e^{-3t} + 5e^{3t} \cos 2t$

6. Find the Laplace transforms of the following:

NOTES

(i) $(2t+1)\sin 2t$ *(ii)* $(t+2)\cos 3t$
(iii) $t^2 \sin at$ *(iv)* $t^2 \cos at$
(v) $te^{-t} \cos 2t$ *(vi)* $te^{-at} \sin at$

7. Find the Laplace transforms of:

(i) $\frac{e^{-at} - \cos at}{t}$ *(ii)* $\frac{\sin^2 t}{t}$

(iii) $\left[\frac{\sin 2t}{\sqrt{t}}\right]^2$ *(iv)* $\left(\frac{\sin at}{at}\right)^2$

(v) $\frac{1-e^{-at}}{t}$

8. Find

(i) $L\left[\int_0^t \frac{\sin at}{t} dt\right]$ *(ii)* Prove that $L\left[\int_0^t \frac{f(t)}{t} dt\right] = \frac{1}{s} \cdot \int_s^\infty L[f(t)ds$

(iii) Find $L\left[\int_0^t \frac{\sin^2 t}{t} dt\right]$ *(iv)* Find $L\left[\int_0^t \frac{t - e^{-at}}{t} dt\right]$

(v) Find $L[e^{-t}\int_0^t t\cos t\, dt]$

9. Find the Laplace transform of the following periodic functions:

(i) $f(t) = E \sin \omega t$, for $0 \le t \le (\pi / \omega)$ and $f(t + (\pi / \omega) = f(t)$, for all t.

(ii) $f(t) = |\cos \omega t|$

(iii) $f(t) = 0$, for $0 < t < (\pi / \omega) = -\sin \omega t$, for $(\pi / \omega) < t < \frac{2\pi}{\omega}$ and $f(t + 2(\pi / \omega))$
$= f(t)$, for all t.

10. Find the Laplace transforms of the following functions given that $f(t)$ is a periodic function of period 2π.

(i) $f(t) = e^t$, for $0 < t < 2\pi$ *(ii)* $f(t) = \pi - t$, for $0 < t < 2\pi$
(iii) $f(t) = t^2$, for $0 < t < 2\pi$ *(iv)* $f(t) = t$, for $0 < t < \pi = 0$, for $\pi < t < 2\pi$
(v) $f(t) = t$, for $0 < t < \pi = 2\pi - t$, for $\pi < t < 2\pi$

11. Find the inverse transform of the following:

(i) $\frac{1}{(3p-4)^5}$ *(ii)* $\frac{1}{(2-3s)^3}$

(iii) $\frac{2s^2+5s+2}{(s-1)^4}$ *(iv)* $\frac{2s+1}{(s^2-4)}$

(v) $\frac{4s-1}{(s+1)^2}$ *(vi)* $\frac{s}{(s^2+4)^2}$

(vii) $\frac{s^2+3s-2}{(s^2+4)^2}$

12. Find the inverse transform of the following functions:

(i) $\frac{1}{s(s^2+4)}$ *(ii)* $\frac{1}{s^2(s^2+a^2)}$

(iii) $\frac{1}{s(s+2)(s-2)}$ *(iv)* $\frac{s}{(s+2)(s+7)}$

(v) $\dfrac{s}{(2s+3)(3s+5)}$

NOTES

13. Find the Laplace transforms of:

(i) $L(2t^2 - e^{-t})$ (ii) $L(t^2 + 1)^2$

(iii) $L(\sin t - \cos t)^2$ (iv) $L(\cosh^2 4t)$

(v) $L\{f(t)\}$ if $f(t) = \begin{cases} 0 & \text{when } 0 < t < 2 \\ 4 & \text{when } t > 2 \end{cases}$

(vi) $L\{t^3 e^{-3t}\}$ (vii) $L\{(t+2)^2 e^t\}$

14. Show that the Laplace transform of $\int_0^\infty te^{-3t} \sin t \, dt = \dfrac{3}{50}$

15. Solve using $L\left\{\dfrac{f(t)}{t}\right\} = \int_s^\infty F(u)\, du$

(i) Find $L\left[\dfrac{\sin^2 t}{t}\right]$ (ii) Find $L\left[\dfrac{1-e^t}{t}\right]$

(iii) Evaluate $\int_s^\infty te^{-3t} \cos t \, dt$

16. Explain Fourier's integral.

5.14 FURTHER READING

Dass, HK. 2008. *Mathematical Physics*. New Delhi: S. Chand

Chattopadhyay, PK. 1990. *Mathematical Physics*. New Delhi: New age International

Hassani, Sadri. 2013. *Mathematical Physics: A Modern Introduction to Its Foundations* Berlin: Springer Science & Business Media

Khanna, V.K. and S.K. Bhambri. 2008. *A Course in Abstract Algebra*, 3rd edition. New Delhi: Vikas Publishing House

Grewal, B.S. 1998. *Higher Engineering Mathematics,* 34th edition. New Delhi: Khanna Publishers

Narayan, Shanti. 1996. *Differential Calculus,* 14th edition. New Delhi: S. Chand and Company

Kreyszig, Erwin. *Advanced Engineering Mathematics,* 7th edition. Mumbai: Wiley Eastern Ltd

Bali, N.P. 2007. *A Textbook of Engineering Mathematics*. New Delhi: Laxmi Publications (P) Ltd

Narayan, Shanti. 2010. *A Text Book of Vector Analysis.* New Delhi: S. Chand and Sons

Spiegel. 1959. *Vector Analysis (Schaum'S Outline).* New York: McGraw Hill Professional

Gupta, B. D. 2009. *Mathematical Physics*, 4th Edition. Noida: Vikas Publishing House Pvt. Ltd.

NOTES

NOTES